OXFORD

MICROECONOMICS

Competition, Conflict, and Coordination

SAMUEL BOWLES & SIMON D. HALLIDAY

ENDORSEMENTS

"I envy the students who will have the opportunity to take a microeconomics course based on this brilliant textbook. Not only will they find it fascinating. It will change their lives, in every way, for the better."

George Akerlof, Georgetown University,
Nobel Laureate in Economics

"In a thick wall of textbooks about rational agents trading in perfect markets, Bowles and Halliday open up a window through which students can see economists at work as they seek answers to market failures, behavioral biases and all the obstacles that must be overcome to build a society that is fair and efficient. This book can change how economics is understood by students who will go on to help us find the answers."

Oriana Bandiera, Sir Anthony Atkinson
Professor of Economics, LSE,
winner of the Yrjo Jahnsson Award

"This text will make for an exciting course—and one especially relevant to contemporary problems like inequality and climate change. Normally, students don't see recent economic ideas until they reach the end of the book. Here such ideas are introduced starting in the first chapter."

Eric Maskin, Harvard University,
Nobel Laureate in Economics

"Teaching from Bowles and Halliday is very rewarding. Students are intrigued by coordination problems and enjoy the visual illustrations."

Avanti Mukherjee, State University of New York
at Cortland

"Bowles' and Halliday's textbook unusually puts at its core the key concepts of social sciences: the interactions (competition, conflict, and coordination) among individuals, groups, and firms. You will come away from this riveting reading understanding how economists deploy theory to help design impactful public policies, and why economics is essential to making this world a better place."

Jean Tirole, Toulouse School of Economics,
Nobel Laureate in Economics

"Bowles and Halliday is pure fun to teach and highly motivating for students, a true gem in the universe of microeconomics textbooks. It applies economic theory to the most pressing challenges of our time, including poverty, inequality, and climate change."

Martin Leroch, Pforzheim University

"The best possible textbook for intermediate microeconomics. It deals with important real-world issues such as inequality, incorporates relevant political, sociological, and behavioural insights, and appropriately places the topics within their historical intellectual roots, while providing rigorous economic analysis."

Giorgos Galanis, Goldsmiths, University of
London

"I congratulate the authors on a job well done! Bowles and Halliday integrate recent economic insights into a classic curriculum of intermediate microeconomics without sacrificing on the formalism. I particularly liked the efforts they have gone to to make the book as pedagogical as possible."

Amrish Patel, University of East Anglia

MICROECONOMICS

MICROECONOMICS

Competition, Conflict, and Coordination

SAMUEL BOWLES & SIMON D. HALLIDAY

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark
of Oxford University Press in the UK and certain other countries

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021936241

ISBN 978-0-19-884320-7 (pbk.)

Printed in Great Britain by
Bell & Bain Ltd., Glasgow

ABOUT THE AUTHORS

Santa Fe Institute

Samuel Bowles (PhD, economics, Harvard University) heads the Behavioral Sciences Program at the Santa Fe Institute. He has taught microeconomic theory to undergraduates and PhD candidates at Harvard University, the University of Massachusetts, and the University of Siena. His research has appeared in the *American Economic Review*, *Nature*, *Science*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Econometrica*.

Textbooks that he has written or coauthored include *Notes and Problems in Microeconomic Theory* (1980), *Microeconomics: Behavior, Institutions and Evolution* (2005), *Understanding Capitalism* (2018), and with the global CORE team, *The Economy* (2017), and *Economy, Society and Public Policy* (2019), both open-access introductions to economics (for majors and nonmajors respectively).

Simon D. Halliday (PhD, economics, University of Siena, Italy) is an Associate Professor in the School of Economics at the University of Bristol. He has also taught game theory, microeconomics, industrial organization, and behavioral economics to graduate and undergraduate students at Smith College, the University of Cape Town, and Royal Holloway, University of London.

His research in experimental economics, behavioral economics, and economics education has been published in the *Journal of Economic Behavior and Organization*, *Journal of Behavioral and Experimental Economics*, the *Journal of Economic Education*, and elsewhere.

PREFACE

To its eighteenth- and early nineteenth-century founders, the subject of economics was the wealth of nations and people. This was no less true of Karl Marx, the most famous critic of capitalism, than it was of Adam Smith, whose *The Wealth of Nations* is considered the most powerful defense of the then emerging capitalist economic system.

Economics was at the time called political economy, and it sought to understand how and why society was being transformed as a result of capitalism, a novel way of organizing how people produce, exchange, and distribute the things we live on. Capitalism continues to change the world, and the task of economics is to understand this process, and how our economies might be made to work better for people today and in the future.

Welcome to *Microeconomics: Competition, Conflict, and Coordination*, and best wishes for your journey through its content. Let's begin by saying how we came to think that economics is important and then explaining our strategy for how you can best learn to do economics.

ECONOMICS ENGAGED IN THE WORLD

Contrary to its reputation among students for being remote from reality, economics has always been about changing the way the world works. The earliest economists—the physiocrats in late eighteenth-century France and the mercantilists before them—were advisers to kings and queens of Europe. This tradition of real-world engagement is continued by today's central bank macroeconomic managers addressing the economic shock of the Covid-19 pandemic, the economic development advisers and advocates of competing policies concerning intellectual property rights, or the global movements of goods and people. Economists have never been strangers to policymaking, constitution building, and attempts at economic reform for the betterment of people's living conditions.

Alfred Marshall's (1842–1924) *Principles of Economics*, initially published in 1890, was the first great text in what came to be called neoclassical economics. It opens with these lines:

> Now at last we are setting ourselves seriously to inquire whether . . . there need be large numbers of people doomed from their birth to hard work in order to provide for others the requisites of a refined and cultured life, while they themselves are prevented by their poverty and toil from having any share or part in that life. . . . [T]he answer depends in a great measure upon facts and inferences, which are within the province of economics;

> and this is it which gives to economic studies their chief and their highest interest.

The hope that economics might assist in alleviating poverty and securing the conditions under which free people might flourish in a sustainable global environment is at once economics' most inspiring calling and its greatest challenge. Like many, both of us were drawn to economics by this hope.

One of us (Simon) grew up in Cape Town, South Africa, under the system of racial segregation called *apartheid*. He vividly remembers the demonstrations that finally brought that system down and the long lines of people waiting to vote in South Africa's first democratic elections in 1994. He volunteered in the poor townships surrounding Cape Town teaching critical thinking and debating, skills required to make the new democracy work. Having initially followed his passion for theater and poetry, he switched in to economics to gain the analytical tools to understand and address his country's challenges.

✓ **FACT CHECK** In a 1999 article in the *Quarterly Journal of Economics*, economists Robert Hall and Charles Jones document this and other productivity differences among countries, attributing them to what they term differences in "social infrastructure" including "institutions and public policies." [1]

The other of your authors (Sam), having been a schoolboy in India and a secondary school teacher in Nigeria before turning to economics, naturally came to the field expecting that it would address the enduring problem of global poverty and inequality.

At age 11 Sam had noticed how very average he was among his classmates at the Delhi Public School—in sports, in school work, in just about everything. A question that he then asked his mother has haunted him since: "How does it come about that Indians are so much poorer than Americans, given that as people we are so similar in our abilities?" And so he entered graduate school hoping that economics might, for example, explain why workers in the US produced as much in a month as those in India produce in a year, and why the Indian population was correspondingly poor.

We now know that the many conventional economic explanations for the gap in standards of living between the two countries are part of the answer but far from all of it: by any reasonable accounting, the difference in the amount of machinery, land, and other capital goods per worker and in the level of schooling of the US and Indian workforces explain much less than half of the difference in output per hour of work.

It seems likely that much of the unexplained difference results from causes that until recently have been less studied by economists but which are a central theme of this book. Chief among these are differences in institutions, that is, differences in how the activities of the millions of actors in the two economies are coordinated by some combination of markets, private property, social norms, and governments.

WHAT SHOULD ECONOMICS BE ABOUT?

We do not think that we are atypical—either among our economics colleagues, or our students, or for that matter among people generally—in our hope that economics can contribute to improving the way these institutions work. The CORE Team—a global group of economic researchers and teachers who have created an open access introductory economics course (www.core-econ.org)—posed the following question to students around the world on the first day of their introductory classes: "What is the most pressing problem economists today should be addressing?" The results are summarized in the word cloud in Figure 1.

The themes are remarkably consistent across universities and countries. Unemployment, inflation, and growth, all important topics in most macroeconomics courses, are on the minds of students. But inequality (along with "poverty") is a much greater concern, as is environmental sustainability (and "climate change"). The future of work (robots, digitalization), globalization and migration, innovation, financial instability, and how governments work ("corruption," "war") are also present.

The microeconomic theory that you will learn has a lot to say about these issues. Included are tried-and-true workhorse concepts that you have probably already encountered, like opportunity costs, mutual gains from exchange, constrained optimization, and trade-offs. Also essential in understanding issues like those in the word cloud are concepts that have more recently risen to prominence among economists. Examples include the importance of cooperation and social (rather than entirely selfish) motivations and modeling strategic interactions among people, including conflicts over the distribution of the mutual gains from exchange.

"IF YOU ARE NOT DOING SOMETHING, YOU ARE NOT LEARNING ANYTHING!"

This phrase is our motto when it comes to learning. Economics is not just something you learn. It is something you do. Think of studying economics as learning a new language. Mastering a large vocabulary and the grammatical rules is essential, but it is not the same as speaking the language.

The test of what you have learned after studying this book is not just what you know, but what you can do with it. Doing economics is what you can say or write—the case you can make for or against a proposed economic policy, the analysis of the reasons for some new development in the global economy—in other words what you can do as a result of what you know.

Like mastering a new language, doing economics is essential to learning the subject. And also like a language, you will learn to do economics more readily if you have a clear need to know.

We begin each chapter with a real-world problem or example that can be better understood using the concepts and models to be introduced in

Figure 1 Student replies to the question "What is the most pressing problem economists today should be addressing?" The size of the font is proportional to the frequency with which subjects mentioned the word or term. The top panel records 3769 student responses from 10 countries and 20 universities. The bottom panel is from 2019 based on 807 students in four universities in Colombia, the UK, and the US. Surprisingly, professional economists at the New Zealand Treasury and central bank and new hires at the Bank of England responded very similarly to students. The less frequently mentioned—smaller font—topics are more readable in the individual word clouds from each of the 25 samples of students that you can access at https://tinyco.re/6235473

With permission from CORE Economics Education.

(a) Word cloud with a 2020 sample of students

(b) Word Cloud with topics consolidated over time

the chapter. These opening paragraphs suggest the need to know what is to follow. The empirical examples also serve as a reminder that the point to the model is to understand the world; and as we proceed through chapters we will ask: How good a job does this particular model do in that respect?

You may be curious about the names we have given to the economic actors in our models. Many are the actual names of members of the team that worked with us to bring this book to you, from around the world including China, India, Chile, Mexico, the US, Germany, and South Africa.

At the beginning of each chapter is a set of learning objectives phrased as new capacities to do things that most likely you were unable to do before. We place great emphasis on your ability to solve problems in which there are right and wrong answers. But it is also important to learn how to formulate arguments and hypotheses about questions that are thus far unanswered, some of which may remain so, and to express economically

informed opinions on issues that will continue to be debated due to the fact that people's values differ.

Interspersed with the contents of the chapters, but offset by boxes, are two important resources:

Mathematics notes M-notes contain the details of mathematical derivations and other analyses as well as worked examples that illustrate the mathematical models in the text. Many of the M-Notes present analysis using calculus of points made in the body of the text using verbal or graphical reasoning.

Checkpoints are self-tests to confirm that you understand the content of the section. The first step in "doing economics" is by checking your understanding of the passage you have just read.

At the end of each chapter, you will find the following:

Important ideas The main ideas in each chapter are provided in a list. At the end of the book, you will also find that all the definitions of the book are included in the Glossary at the end of the book for you to consult and improve your understanding. Mastering the use of these terms is essential to doing economics. Try using each of them in a complete sentence of your own.

Making connections Provides some guidance in seeing how the ideas in each chapter are connected to each other and to other themes in the book, so that you will be able to draw together the 'big picture' about the main messages and themes of the book. Try restating these connections making use of the terms in important ideas. Or better yet: make a mind map using the important ideas and making connections features.

Mathematical notation The book contains a variety of important mathematical tools to help model the various economic ideas in the book. To assist you with your reading of each chapter and to understand better each model you encounter, we provide a table of the mathematical notation you will encounter in that chapter.

We use the margins of the book for a variety of purposes:

Definition We define important terms in the margins where they first are introduced. All of the definitions are collected in the Glossary.

Reminder We put reminders in the text often to help you to see the connections of ideas throughout the book.

Example An example will often illustrate an idea with a relevant example of a person, firm, or country making decisions that are similar to those described in the text.

Fact check When we need to verify or illustrate an idea with data or an empirical example we will do so with a fact check.

History These introduce you to some of those people who have contributed to economics or to relevant historical facts.

M-Check If an idea requires a brief mathematical clarification that does not require its own M-Note, then we may convey that in a margin note.

As is the case with any first edition of a text there inevitably will typographical errors and other things we would like to correct, and that others using the book should know about. Refer to our list of errata at https://tinyurl.com/bhmicro for the current list. If you find what you think is a mistake, do please add the error you've found and your name. If we add your suggested error to our list we will acknowledge the first person to point it out to us.

Economics is an integrated body of knowledge, and it is best learned in a cumulative way, mastering a set of concepts and going on to use those concepts in mastering additional concepts. What this means, practically, is that it is best to study earlier chapters before moving on to later ones. Sections labeled "application" however provide illustrations of how the ideas and models being taught in a particular chapter can be used, and these do not introduce new material that is essential to the chapters that follow.

Microeconomics is waiting for you. Just do it!

Samuel Bowles and Simon D. Halliday
Santa Fe Institute, Santa Fe, New Mexico, US, and
University of Bristol, Bristol, UK.

GUIDE TO THE ONLINE RESOURCES

As well as the boxes and features presented in the chapters to aid you in doing economics, we have a wealth of online resources to support your learning.

Our **interactive graphs** allow you to explore key models in a dynamic way, and we have also provided **video material** and a **Mathematics Appendix** to further explain figures and mathematics.

Test your knowledge with interactive **multiple-choice questions** and push your understanding of economic problems further with **mathematical questions**.

Discussion questions and **further-reading recommendations** prompt you to think around the issues.

Access the online resources by going to: www.oup.com/he/bowles-halliday1e

ACKNOWLEDGMENTS

Our most heartfelt thanks go to Duncan Foley, who first had the idea of this text and initiated work on it with us, contributing important elements of its content. His commitment to rethinking microeconomics has been an inspiration.

We have also borrowed and learned from the pedagogical and content innovations of The CORE Team's introductory open access texts, *The Economy*; *Economy, Society and Public Policy*; and *Doing Economics*. We have especially benefited from the advice of Wendy Carlin, who heads the CORE Team. A conversation with Margaret Stevens was the inspiration for Chapter 4. Weikai Chen and Martin Leroch provided detailed comments on the entire manuscript, proposing substantial improvements. Daniele Girardi and Avanti Mukherjee taught drafts of the book over the past many years; their comments have made the book much better. Sahana Subramanyam prepared the index and helped us correct errors and ambiguities in the text.

Many other test teachers have contributed to the project as it has progressed: Elizabeth Anat, Xiao Jiang, Rishabh Kumar, Sai Madhurika Mamunuru, Lisa Saunders, Markus Schneider, Gregor Semeniuk, Daniele Tavani, and Tai Young-Taft in the US, Paul Cowell, Giorgos Galanis, Thaana Ghalia, and Sebastian Ille in the UK, Ihsaan Bassier and Justine Burns in South Africa, Martin Leroch in Germany, Arjun Jayadev and Anand Srivastava in India, Seçil Akin and Erkan Gurpinar in Turkey, Marcelo Caffera in Uruguay, Hernán Bejarano in Mexico, Stephen Kinsella in Ireland, and Mark Levin in Russia.

Many people generously shared data and helped us create compelling economic narratives, including The CORE Team, John Adams, Karishma Ajmera, Robert Axtell, Doyne Farmer, Nigel Franks, Anders Fremstad, Mauro Gallegati, Gianfranco Giulioni, Diana Greenwald, Arjun Jayadev, Chris Kempes, Alan Kirman, Francois Lafond, Seung-Yun Oh, Mark Paul, Giacomo Piccoli, William A. Pizer, Steven Sexton, Eileen Tipoe, and Jessika Trancik.

We were fortunate to be able to engage an outstanding team of researchers in creating the book: Morgan Barney, Nicolas Bohme-Oliver, Harriet Brookes-Gray, Weikai Chen, Scott Cohn, Bridget Diana, Jesus Lara Jauregui, and Anoushka Sharma. Bridget Diana produced the interactive figures for the enhanced e-text with essential help from Chris Makler. Riley Boeth and Madeleine Wettach also provided valuable research assistance. The "Microeconomics Revolution" working group suggested many improvements: Nicolas Bohme-Oliver, Joshua Budlender, Pedro De Almeida, Bridget Diana, Kuo-Chih Huang, Chirag Lala, and Lisa Saunders. Anmei Zhi created the cartoons.

The Behavioral Sciences Program at the Santa Fe Institute was an optimal environment for the development of our ideas over many years, hosted meetings of our team, and provided financial support to the project. Caroline Seigel of SFI Library provided essential help. Smith College provided additional support for our research team.

We are also grateful to Felicity Boughton, Jonathan Crowe, Keith Faivre, Judith Lorton, Jon McGreevy, Amber Stone-Galilee, and their colleagues at Oxford University Press for their part in bringing our book to you.

OUTLINE CONTENTS

DETAILED CONTENTS

"The man . . . enamored of . . . his own ideal plan of government, . . . seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board . . . but . . . in the great chess-board of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it.

If those two principles coincide and act in the same direction, the game of human society will go on easily and harmoniously, and is very likely to be happy and successful. If they are opposite or different, the game will go on miserably, and the society must be at all times in the highest degree of disorder."

Adam Smith,

Theory of Moral Sentiments (1759, Part VI, Section 2, Ch 2)[1]

PART 1

PEOPLE, ECONOMY, AND SOCIETY

As people, our physical capacities are hardly remarkable compared to other animals. But by coordinating with others—finding ways that our individual efforts can add up to a whole that is more than the sum of its parts—humans are unique as a species, engaging in common pursuits on a global scale and, for better or worse, transforming nature and inventing previously unimagined devices and ways of life. Economics provides a lens for studying this social aspect of human uniqueness by analyzing how people interact with each other and with our natural surroundings to produce and acquire our livelihoods.

We begin (in Chapter 1) by developing a common framework for studying the various types of social interactions, using game theory to pose a question older than economics. This is: How can a society's institutions—its laws, unwritten rules, and social norms—harness people's pursuit of their own objectives to generate common benefits and to avoid outcomes that none would have chosen? The challenge is how to combine freedom—individuals' pursuit of their own objectives—with the common good, improving the livelihoods of all members of society.

This is called the problem of societal coordination: How can we coordinate—that is, jointly with others *organize*—our actions to yield desirable results for all the members of society? The example of societal coordination we use in Chapter 1 to illustrate this challenge is about the other aspect of economics: how we relate to our natural surroundings, illustrated by a problem of overexploiting an environmental resource.

Adam Smith, considered by many to be the founder of economics, understood the challenge well. And he understood that economics—or "political economy" as it was then called—is fundamentally a *social* science: it is about how people interact. Smith warned (in the head quote) about the disastrous consequences of treating people as if they were simply chess pieces who could be moved around on the chessboard of life at the will of a government, more or less like an engineer might design a machine.

An adequate response to the challenge of combining freedom and the common good must therefore be based on knowledge of how people act depending on the situation they are in, and how changing the situation will change how they act. We therefore (Chapter 2) turn to people and their motives—whether self-regarding or generous, opportunistic or ethical—and we use the game theory concepts you will have learned in Chapter 1 to illustrate some of the challenges that we face in coordinating our actions.

A key idea in these first two chapters is that people do the best they can in given situations. In Chapter 3 we introduce the mathematics of constrained optimization as a method to better understand this process. In this chapter we consider individuals in situations where they act in isolation rather than interacting with other individuals.

HISTORY What makes humans unique among all the animals is our capacity to cooperate in very large numbers and to adjust the ways that we cooperate to changing circumstances. Here (https://tinyurl.com/y3bpy4px) the Israeli historian Yuval Noah Harari explains why this is so.

But people rarely act in isolation: economics allows us to understand the sometimes surprising or unintended society-wide effects of when we interact with others, whether it be directly with our own employer or indirectly with literally millions of people involved in producing and distributing the goods making up our livelihoods.

A basic insight for this understanding is that we are better off by interacting with others. But our interactions also give rise to conflicts. When people engage with others in buying and selling, working and investing there are mutual benefits potentially available to all parties involved. This must be the case if participation in these and other economic activities is voluntary. But unavoidably there are also conflicts over how these mutual gains are divided (Chapters 4 and 5).

In our interactions with each other and with nature we frequently fail to exploit all of the potential mutual gains. An example is when a person with the capacity and desire to produce goods and services needed by others cannot find a job. Another is overexploitation of a fishery or some other environmental resource. These are called *coordination failures* because they result from inadequacies in the ways that our institutions allow us to coordinate the ways that we interact.

Coordination failures are often due to our conflicts over the distribution of potential mutual gains or to the fact that when we act we do not take account of the effects of our actions on others. In Chapter 5 we show how differing institutions—differing rules of the game—can help address these coordination failures so that no potential mutual gains remain unexploited. We also show how differing rules of the game, by conferring differential advantages on people, will result in differing levels of inequality.

The problems of inequality and environmental impacts will be taken up to illustrate the concepts we teach throughout the book.

CHAPTER

1

SOCIETY

COORDINATION PROBLEMS AND ECONOMIC INSTITUTIONS

Two neighbors may agree to drain a meadow, which they possess in common; because 'tis easy for them to know each others mind; and each must perceive, that the immediate consequence of his failing in his part, is the abandoning of the whole project.

But 'tis very difficult and indeed impossible, that a thousand persons shou'd agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expense, and wou'd lay the whole burden on others.

David Hume
A Treatise of Human Nature ([1742] 1967)[1]

DOING ECONOMICS

This chapter will enable you to:

- Use game theory to analyze how people interact in the economy, each affecting the conditions under which the others decide how to act.
- Understand why the outcomes of interactions are often worse than they could be when people fail to coordinate with each other and to take account of the effect of their own actions on others.
- Explain how problems like environmental damage and global poverty can be the result of failed coordination.
- Represent institutions as "the rules of the game" and see how changing these rules will change outcomes.
- See that economic institutions determine incentives for people's behavior and can affect how successfully we address coordination problems.
- Explain why when people have limited information and conflicts of interest they often fail to implement "win-win" outcomes.

Figure 1.1 Palanpur farmers threshing and winnowing grain (separating grain from chaff).

Photo courtesy of the Palanpur project photo archives. First published in London School of Economics (LSE). India's economic "revolution": a perspective from six decades of economic development in Palanpur, a north Indian village. London School of Economics (LSE), London, UK (2011) 428 pp.

1.1 INTRODUCTION: POOR ECONOMICS

At the turn of the present century, the process of economic development had bypassed almost all of the 200 or so families that made up the village of Palanpur in the Indian state of Uttar Pradesh. But for the occasional watch, bicycle, or irrigation pump, Palanpur appeared to be a timeless backwater, untouched by India's cutting-edge software industry and booming agricultural regions. Less than one-third of the adults were literate, and most had endured the loss of a child to malnutrition or to illnesses that had long been forgotten in other parts of the world.

A visitor to the village approached a farmer and his three daughters weeding a small plot of land. The conversation turned to the fact that Palanpur farmers plant their winter crops several weeks after the date that would maximize the amount of grain they could get at harvest time. The farmers knew that planting earlier would produce larger harvests, but no one, the farmer explained, wants to be the first farmer to plant, as the seeds on any lone plot would be quickly eaten by birds.

Curious, the visitor asked if a large group of farmers, perhaps members of the same extended family, had ever agreed to plant their seeds earlier, perhaps on the same day to minimize the individual losses. The farmer looked up from his hoe and made eye contact with the visitor for the first time "If we knew how to do that," he said, addressing the visitor as "bhai" or brother, "we would not be poor."[3]

HISTORY In his address accepting the Nobel Prize for economics in 1979, University of Chicago economist T. W. Schultz said: "Most of the people in the world are poor, so if we knew the economics of being poor, we would know much of the economics that really matters." He was right then and he is right now. What is called the Nobel Prize in economics, officially is the Nobel Memorial Prize in Economic Sciences.[2]

1.2 SOCIETAL COORDINATION: THE CLASSICAL INSTITUTIONAL CHALLENGE

For the Palanpur farmers, the decision when to plant is a **coordination problem**. A coordination problem is a situation in which people could all be better off, or at least some be better of and none be worse off, if they all jointly decided how to act—that is, if they coordinated their actions—than if they act individually.

The planting choice is a coordination problem because:

- the farmer does better or worse depending on what other farmers do;
- all the farmers would do better if they could *coordinate* their actions by jointly agreeing to all do what would be mutually beneficial namely, planting early; but

EXAMPLE In this video (tinyurl.com/y5fg3whx) Thomas Piketty and James Heckman explain why data is fundamental to their work (from the CORE project. www.core-econ.org).

COORDINATION PROBLEM A coordination problem is a situation in which people could all be better off (or at least one be better of and none be worse off) if they jointly decide how to act—that is, if they coordinate their actions—than if they act independently.

- it is a *problem* because the farmers may not be able to coordinate; and as a result
- if they do *not* coordinate and plant late, then all of the farmers will do worse than they all could have done (that is, had they all planted early).

To stress the fact that coordination problems often affect an entire population (even though we explain them using two-person examples) we sometimes use the expression *societal* coordination problems. Notice that one farmer cannot dictate the actions of the other farmers, nor can they come to a common agreement about what to do ("if we knew how to do that, we would not be poor")—the inability to come together and coordinate is at the heart of coordination problems.

David Hume (the eighteenth-century British philosopher and economist quoted at the start of this chapter) used an example—two landowners considering draining a meadow—to pose what he considered the most important problem facing society, namely, devising institutions that would reconcile the pursuit of individual objectives (avoiding the "trouble and expense" in his example of the meadow) with getting desired societal outcomes (improving the value of the meadow by draining it). His simple two-person example was meant to illustrate the need (in a society of "a thousand persons") for a government to address the broader societal coordination problems of his day.

Though the term was invented only two centuries after Hume, he was using what we now call game theory to make his case. Let's apply his reasoning to the farmers of Palanpur. Like Hume we will consider just two farmers as a way of representing the institutional challenge faced by the entire village.

Figure 1.3 shows the outcomes for two players, Aram and Bina, choosing when to plant their grain. The figure illustrates the values of the farmers' crops, whether they consume the crop themselves, or sell it for money to spend on other things.

Each farmer can either plant early or plant late, and while (also as in Hume's example) two people could probably come to some agreement about what to do, remember that we are using this two-person example to illustrate the entire village of about 200 families of farmers. So we assume that they cannot coordinate on some agreed upon actions for the two jointly. There are four possible outcomes:

- If both players plant early, they each achieve their *best* possible harvest, because they grow the most grain through sharing the risk of having their seeds eaten by birds (outcome (c) in Figure 1.3).
- If Aram plants early while Bina plants late, Aram has his seeds eaten by birds and gets no harvest (the *worst* outcome for him), whereas the late planter gets a *good* (but not the best) harvest. While none of her seeds are eaten by the birds, planting late is not the best for growing the most

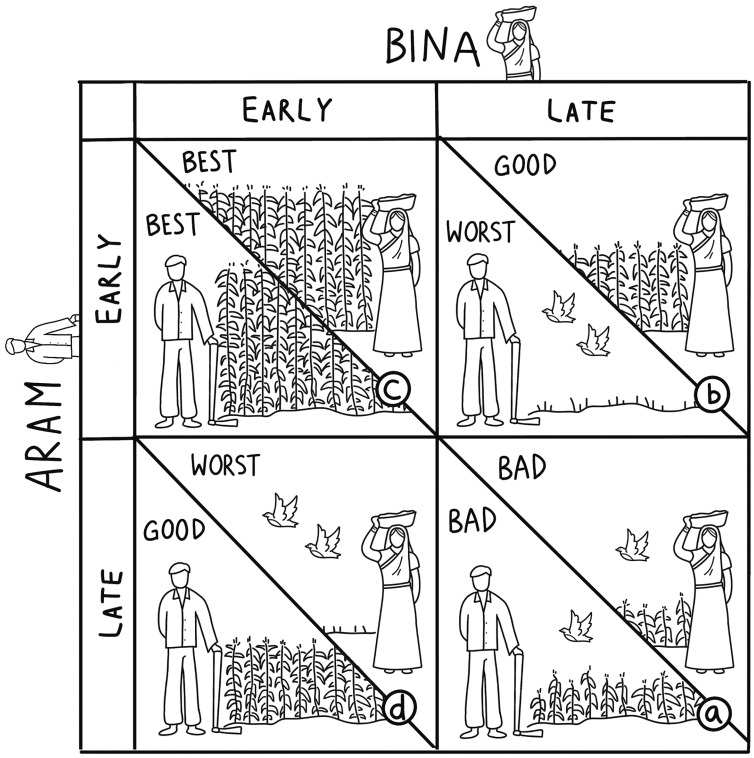
Figure 1.2 Poor economics. Esther Duflo and Abhijit Banerjee founded the Massachusetts Institute of Technology's Poverty Action Lab to bring the best minds in economics to bear on eradicating global poverty. Their 2011 book is titled *Poor Economics*.[4] In 2019 the two MIT professors were awarded the Nobel Prize in economics along with Michael Kremer for their research on the causes of poverty and methods to raise the living standards of poor people.

Photo © Bryce Vickmark

› EXAMPLE In this video (tinyurl.com/yxpf72hm) Esther Duflo explains what happened when it was mandated that randomly selected villages elect a woman to head their local council (from the CORE project. www.core-econ.org).

Figure 1.3 Planting in Palanpur. This figure shows "what-if" outcomes for planting in Palanpur. Each column represents a possible combination of Aram planting early or late and Bina planting early or late with the corresponding outcomes being worst, bad, good, or best in terms of how much grain they grow.

Illustration by Anmei Zhi.



grain (outcomes (b) and (d) in Figure 1.3). The same is true if Aram planted late when Bina planted early.

- If both plant late, they harvest a smaller crop while also sharing the risk of their seeds being eaten, a *bad* outcome, but not the worst (outcome (a) in Figure 1.3).

The people of Palanpur are stuck in the bad outcome even though they would all be better off if they all planted early (they would both move from a "bad" outcome to the "best" outcome in the figure). They are experiencing a **coordination failure**, namely a coordination problem that is not addressed

COORDINATION FAILURE A coordination failure occurs when the non-cooperative interaction of two or more people results in an outcome that is worse for at least one of those involved and not better for any.

by appropriate institutions. A modern-day David Hume would point out that a government could simply impose a sufficient tax on those planting late to ensure that most farmers would plant early.

Adam Smith, a generation after Hume, would stress the value of the exchange of privately owned goods on competitive markets as a way of coordinating the actions of large numbers of people, who would be guided (even without knowing it) by what he termed "an invisible hand." Hume, Smith, and the other founders of European political philosophy and political economy posed what we call the classical institutional challenge.

These philosophers and economists wanted to know how to *design* institutions *so that people could be left free to make their own decisions, and at the same time avoid outcomes that were inferior for everyone.* More precisely, how do we design institutions which encourage coordination by free choice while avoiding poor outcomes such as planting late in Palanpur? The eighteenth- and nineteenth-century political economists and philosophers who founded the field of economics were attempting to provide solutions to coordination problems.

HISTORY Adam Smith wrote: "[E]very individual [...], indeed, neither intends to promote the public interest, nor knows how much he is promoting it [...] he intends only his own security; ... he intends only his own gain, and he is in this ... led by an invisible hand to promote an end which was no part of his intention ... By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it."[5]

CHECKPOINT 1.1 Planting in Palanpur: a coordination problem Imagine that you are Bina in the figure above, and that you did not know whether Aram would plant early or late. What would you do? Suppose, contrary to what we have assumed, you and Aram were neighbors and you could talk with him. What would you say?

1.3 THE INSTITUTIONAL CHALLENGE TODAY

The classical institutional challenge remains with us, although some of the forms that it takes today were unknown to the great eighteenth- and nineteenth-century thinkers.

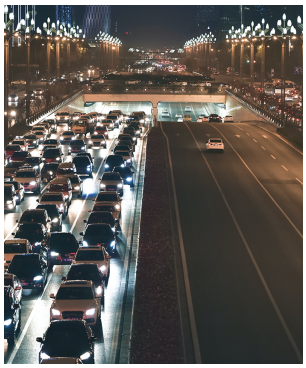
Consider the following coordination problems:

- How do we sustain the global environment? To avoid damaging climate change we need to coordinate our reduction of emissions. Many people and firms would prefer that someone *else* reduce *their* carbon footprint. How can we address climate change in a way that is both fair and imposes the least possible costs?
- How do we make the best use of our ability to create and use knowledge? If we all agree to share the knowledge we have with others we may all be better off: when I transfer my knowledge to you I do not lose the ability to continue using it. But each of us may profit by restricting others' use of our knowledge by means of patents, copyrights, and other intellectual property rights.
- How do we move around a city without overcrowding streets and causing delays? My decision whether to drive, walk, or take public transport affects not only my own travel time, but also the degree of traffic

HISTORY Adam Smith, David Hume, and the other founders of classical economics sought to solve coordination problems in ways that were fair according to their values, and respected the liberty (freedom of choice) and dignity (self-respect and social esteem) of all economic actors.

Figure 1.4 Traffic headed out of a major city.

Photo by Preillumination SeTh on Unsplash.



congestion and delays experienced by everyone else. Everyone might be better off if the use of private vehicles were substantially reduced, but few will reduce their driving unless some way is found to implement a general reduction by everyone.

These are all coordination problems because an outcome that is better for all is possible if people find a way to jointly agree to a course of action. But, for reasons we will explain in detail, people routinely fail to coordinate and suffer bad consequences as a result, including the following:

- *over*use of some resources illustrated by pollution, overgrazing, traffic congestion, and climate change; and
- *under*use of other resources such as the productive capacities and creativity of people and the knowledge that we have created, illustrated by unemployment and the enduring poverty of the people of Palanpur and villages like it around the world.

CHECKPOINT 1.2 Coordination problems you have known Think of a social interaction in which you have been involved that was a *coordination problem*, and, using the description of why planting in Palanpur is a coordination problem (the bulleted points above), explain why it was a problem and how coordination might have (or did) address the problem.

1.4 ANATOMY OF A COORDINATION PROBLEM: THE TRAGEDY OF THE COMMONS

The overuse of environmental resources provides a good illustration of why coordination problems arise.

In 1968, Garrett Hardin, an ecologist, famously described what he called the **tragedy of the commons**, an example of a coordination failure.[6] He told a story about a group of herders who share a pasture. The pasture was common land—hence a "commons"—shared by many herders. But why was his story a tragedy?

Each herder could put as many animals in the pasture as they wished, and overgrazing will lead to erosion and the ruin of the pasture. Hardin reasoned that if the land is common to all and no one herder owns it, each herder has no interest in limiting how many animals they put in the common pasture. A ruined pasture is of no value to any of the herders. But each herder's self-interest leads them to neglect the effect their actions have on others. The outcome is a tragedy.

TRAGEDY OF THE COMMONS The tragedy of the commons is a term used to describe a coordination problem in which self-interested individuals acting independently deplete a common property resource, lowering the payoffs of all.

With the term *tragedy of the commons*, Hardin gave social science one of the most evocative metaphors since Adam Smith's "invisible hand." Indeed Hardin called his tragedy a "rebuttal to the invisible hand." The two metaphors are powerful because they capture two essential yet contrasting social insights. When guided by an invisible hand, social interactions reconcile individual choice and socially desirable outcomes. By contrast, the actors in the tragedy of the commons pursue their private objectives to tragic consequences for themselves and others.

The natural setting for Hardin's tragedy was chosen for its imagery. The underlying problem applies to many situations where people typically cannot or do not take account of the effects of their actions on the well-being of others. You can think of a city's streets as a commons, and people deciding to drive rather than walk, bike, or use public transport as similar to the herders putting cattle on the common. A modern-day "tragedy of the roadways" is a traffic jam.

What are the common elements in Hume's drain-the-meadow problem, the farmers in Palanpur planting late, Hardin's herders overgrazing their pasture, and our modern city dwellers clogging the streets with their vehicles?

In each of these three cases, the reason why uncoordinated activities of people pursuing their own ends produce outcomes that are worse for all is that each participant's actions affect the well-being of others, but these effects are not taken into account by the individual actors when they decide how to act. These impacts of our actions on others that we do not take account of in deciding what to do are termed **external effects**.

Here are the external effects (italicized in the list below) that actors in our four examples do not take into account when deciding what to do:

- The person who lives in a city who drives to work, *adds congestion* to the streets, and therefore increases the travel time of others.
- Hume's farmer who does not drain the swamp and *imposes the cost* of doing so on the other farmer.
- The Palanpur farmer who plants late, *imposes a cost* on the other farmer who will have his seeds devoured by birds if he plants early.
- Likewise the farmer who plants early *confers a benefit* on the other farmer who can benefit by planting at the right time (early) without severe losses of seed to the birds.

EXTERNAL EFFECT An external effect occurs when a person's action confers a benefit or imposes a cost on others and this cost or benefit is not taken into account by the individual taking the action. External effects are also called externalities.

- The herder who places additional cattle on the common pasture *reduces the grass available* to the other herders' stock.

Addressing coordination problems by internalizing external effects

Simply abolishing these and other external effects that are the root of coordination problems is not an option. There is no way to organize society so that nothing that we do would affect others, each person on his or her self-sufficient island.

Apart from not being much fun, life would be impossible in a society of total social isolates (just think about how the next generation would be born and raised!). So, to address the classical institutional challenge to prevent or at least minimize coordination failures we need to find ways of inducing each participant *to take adequate account of the effects of their actions on others.*

This is called *internalizing* an *external effect*. We use the term *external effect* because the effect is *outside* of the individual's process of decision-making when taking the action. To internalize the external effect, you ensure that the person who acts bears the costs of their negative effects on others and reaps the rewards of their positive effects on others. In this way the otherwise "external" costs and benefits become part of the individual's decision-making process, leading them to "take adequate account of the effects of her actions on others."

HISTORY The "golden rule" is "to do unto others as you would have them do unto you" (Matthew, 7: 12). Or, treat others as you would like to be treated yourself. The same ethical principle is found in Islamic scriptures and in the teaching of other religions.

If the "others" are our family, our neighbors, or our friends, our concern for their well-being or our desire to be well regarded by others might get us to take account of the effects of our actions on them. Reflecting this fact, an important response to the classical institutional challenge—one that long predates the classical economists—is that caring for the well-being of others need not be confined to friends and relatives but may extend to all of those with whom we interact. Ethical guides such as the "golden rule" are ways that people often internalize the effects of our actions on others, even when the others are total strangers to us.

But, over the past five centuries, people have come to interact not with just a few dozen people, as humans have for most of our history and prehistory, but directly with hundreds and indirectly with millions of strangers. The classical economists in the eighteenth century were responding to the fact that the generosity or ethical motivations that one might feel towards one's family or neighbors would not be sufficient to induce people to take account of the effect of their actions on others, once these external effects spread across the entire network of global interactions.

From its eighteenth-century origins up to today, an objective that economics has set for itself, therefore, has been to design and implement institutions that would induce people to act *as if* they cared about those who were affected by their actions even when that was not literally true.

CHECKPOINT 1.3 External effects

a. Provide an example of a *negative external effect* that occurs in a social interaction. Explain *why* it is negative and why it is external.

b. Provide an example of a *positive external effect* that occurs in a social interaction. Explain *why* it is positive and why it is external.

1.5 INSTITUTIONS: GAMES AND THE RULES OF THE GAME

Institutions

Institutions are the laws, informal rules, and mutual expectations which regulate *social interactions* among people and between people and the biosphere. Think about driving on the right or on the left as a coordination problem (not a very challenging one). People adopt the behaviors prescribed by institutions (e.g. drive on the right if you are in the US) because of some combination of:

- *laws* enforced by a government (you will be arrested and fined for driving on the left in Brazil, the US, France, and other countries where driving on the right is the law);
- *social pressures*—sometimes termed informal rules because they are not enforced by governments (your friends and neighbors will disapprove and think less of you if you drive on the left); and
- *mutual expectations* that you have about what others will do and have about what you will do (you expect others to drive on the right because they expect you to drive on the right, so you will avoid accidents by doing the same).

We refer to institutions as the rules of the game. To see what this means we now introduce an important conceptual approach for understanding society. **Game theory** uses mathematical models and verbal arguments to analyze how the outcomes of the interaction for the participants will depend on the rules of the game and the objectives of the players. It has been used extensively in economics and the other social sciences, biology, and computer science.

INSTITUTIONS Institutions are the laws, informal rules, and mutual expectations which regulate social interactions among people and between people and the biosphere.

GAME THEORY Game theory is the branch of applied mathematics that studies strategic interactions.

Game theory focuses on **strategic interactions** where participants are interdependent and are aware of this interdependence: one player's outcome depends on their own and other players' actions and all players know this. We can contrast strategic with *nonstrategic* situations in which the effect of your actions on the outcomes you will experience is independent of what others do. An example: your enjoyment of the program you are streaming at home alone is substantially independent of what others may be doing.

But many of our economic and social interactions are strategic:

- those considering driving to work know that their travel time will depend on how others decided to get to work that morning;
- the Palanpur farmer knows that how his crop will fare if he plants early will depend on how many others planted early.

CHECKPOINT 1.4 Institutions

a. Give an example of a strategic and a nonstrategic interaction.

b. Which of the three items on the list of reasons why people coordinate on the side of the road on which to drive—laws, social pressures, and mutual expectations—explain why the farmers in Palanpur plant late?

Figure 1.5 John von Neumann (1903–1957) was a Hungarian-American mathematician, computer scientist, and physicist who is regarded as the father of game theory,[7] which he hoped would allow us to better understand the anti-Semitism and fascist political upheavals that he had witnessed in the early 20th century and provide the basis for understanding how groups interact.

Games

When we model strategic interactions using game theory we call the actors *players*. Players can be people, owners of firms, social movements, governments, or a variety of other entities. In biology, where game theory has been extensively used, even sub-individual entities are "players" such as viruses "trying to" spread in a pandemic or genes "trying to" get as many copies of themselves made as possible. Players may choose from a list of possible strategies (called a strategy set). For example, a strategy set might include "Purchase a bicycle for $350." But the rules of the game reflect institutions: if private property is an institution that is present and

STRATEGIC INTERACTION An interaction is strategic when participants' outcomes—their profit, standard of living, or some other measure of their well-being—depend on the actions that both they and others choose, and this interdependence is known to the actors. A shorthand expression for the term strategic is: mutual dependence, recognized.

SET A set (in mathematics) is a collection of objects defined either by enumerating the objects, or by a rule for deciding whether any particular object is in the set or not. For example, the set of positive, even integers less than or equal to 10 is, $\{2, 4, 6, 8, 10\}$.

enforced, then the strategy set would not include "Pick up any available bicycle," without specifying the possible penalties for stealing.

The Palanpur farmers' strategies are "Plant Early" or "Plant Late." The strategies could also include a strategy based on what others did in the past (called a contingent strategy) such as: "Plant early as long as at least five others planted early last season." The description of a *game* requires us to identify the following:

- *Players*: a list of every player in the game whether they be individuals (like the farmers in Palanpur), an organization such as Amazon or Alibaba, or some other entity that can be represented as a single actor choosing between alternative courses of action.
- *Strategy sets*: a list for each player of every course of action available to them at each point where they must make a choice (including actions that depend on the actions taken by other players, or on chance events). The strategies selected by each of the players—the outcome of the game—is called the strategy profile.
- *Order of play*: a game can be *simultaneous* such that players make their choices without knowing the choices of others, as in the game of rock-paper-scissors. Or a game can be *sequential* such that players move in sequence, one after the other, as in chess, so that each player knows and responds to the choices of the previous players.
- *Information*: A game also specifies
  - who "knows" what,
  - when do they "know" it,
  - if what they "know" is known to others as well,
  - if what they "know" can be used in a court of law to enforce a contract, and
  - if what they "know" is true (this is why we use the quotation marks).
- *Payoffs*: Numbers are assigned to each possible outcome of the game (each strategy profile) for each player; a player chooses a strategy with the intention of bringing about the strategy profile with the highest number.

It is often useful to consider payoffs as something that the players actually get. For example, considering the farmers in Palanpur again, an outcome of the game is a strategy profile indicating who plants early and who plants late, and the payoffs could be the amount of grain each farmer harvests. We say that the payoff associated with a particular outcome of a game is how much the player *values* that outcome. But that means nothing more than that a player will choose a strategy resulting in an outcome with a higher payoff number if possible.

EXAMPLE People can *change* these rules, so institutions can themselves be outcomes of games that govern how the rules of the game can be changed. FIFA governs how football (soccer) can be played by what are called the Laws of the Game. These institutions also change: the corner kick was introduced in 1872 when the UK Football Association changed the rules.

EXAMPLE In 2020 under the pressure of the popular protests, the government of Chile established a set of rules governing how the constitution of Chile would be amended.[8]

EXAMPLE When we model the coordination problem of the Palanpur farmers as a game we assume they plant simultaneously. But when we model the interaction between a bank and a borrower we assume that the banks first makes an offer (the loan size, interest rate, and schedule of repayment) and the prospective borrower responds.

EXAMPLE Watch this video (tinyurl.com/y2cosf8v) to see how a sophisticated game theorist brings an episode of the show Golden Balls to a surprising close (from the CORE project www.core-econ.org).

An important distinction concerning strategy sets is whether or not one of the strategies open to the players is to jointly agree on a strategy profile—that is to deliberately coordinate their actions. This is possible in what is called a **cooperative game.**

We use the set of *players*, their *strategy sets*, their *payoffs*, the *order of play*, and the *information* the players have to describe the institutions governing some economic interaction, whether it is between an employer and an employee, or a central bank like the US Federal Reserve and a commercial bank. But even this detailed description of the interaction does not give us enough information to predict how the game will be played.

The outcome of a game—how it will be played, resulting in a particular strategy profile—is called a solution. To determine the solution as a way of predicting the outcome of a game we need what is called a **solution concept**. A solution concept for a cooperative game would include some rule for deciding on what the coordination would be, for example allowing one player selected at random to dictate the outcome, or a particular system of voting.

But by positing some way that people could jointly implement some outcome, cooperative game theory *assumes away* the problem of coordination. And the problem of how coordination is to be achieved is at the heart of the classical institutional challenge whether it takes the form of climate change or traffic jams.

So we need to see how players might coordinate in what is initially a noncooperative setting—one in which coordination is not assumed at the outset—let's take a concrete example: people interacting in a way that results in the over-exploitation of an environmental resource. We will use this example to illustrate a basic solution concept for noncooperative games: the Nash equilibrium.

CHECKPOINT 1.5 Games

a. What is a game?

b. How do you describe the *outcome* of a game?

COOPERATIVE GAME A strategic interaction in which the players' choice of a strategy is subject to a binding (enforceable) agreement.

SOLUTION CONCEPT A solution concept is a rule for predicting the outcome of a game, that is, how a game will be played.

1.6 OVEREXPLOITING NATURE: ILLUSTRATING THE BASICS OF GAME THEORY

People who fish for a living interact with each other regularly. Each of them is aware that how much they benefit from fishing depends not only on their own actions, but on the actions of others. This is because the more others fish, the more difficult it will be for each to catch fish. The fishermen therefore impose negative external effects on each other. And this, along with the difficulty they face in agreeing on a common course of action, is why they face a coordination problem. Given that they cannot jointly decide on how much to fish, each faces a basic question: How much fishing to do given the strategies adopted by others who are fishing the same waters?

The game setup

Specifically, we consider two fictional fishermen, Alfredo and Bob, who share access to a lake, and catch fish, which they eat. There are no other people affected by their actions.

Here we illustrate the basic concepts of game theory in a game we call the Fishermen's Dilemma. We chose the name because it is an example of what is probably the most famous game, the Prisoners' Dilemma.

The Fishermen's Dilemma game is noncooperative, which for two people fishing in the same lake may seem unrealistic because as neighbors they might be able to come to some kind of agreement about what each will do. We do not consider this option in the two-person case because the model illustrates a large number of people interacting. When many people interact, arriving at and enforcing such a cooperative agreement would present serious challenges.

Here is the game.

- *Players*: Alfredo and Bob, two fishermen.
- *Strategy sets*: Each may fish for either 10 or 12 hours.
- *Order of play*: They simultaneously select a strategy, resulting in the game's strategy profile
- *Payoffs*: The players each catch and eat the amount of fish they caught, given by the strategy profile they have implemented.

This ends the game.

Payoffs

The payoff of each player is composed of two parts:

- The amount of fish they are able to catch and consume, which they value and would like to increase; and
- The amount of time they spend fishing, which they find tiring and would like to decrease.

Figure 1.6 Elinor Ostrom (1933–2012). Elinor Ostrom was an American political scientist who won the Nobel Prize in economics for her contributions to understanding coordination problems, such as that encountered by Alfredo and Bob in the Fishermen's Dilemma, and on the institutions that promote cooperation in groups. Especially important are her empirical studies showing how people address the coordination problems that they encounter in seeking to maintain their livelihoods, such as grazing herds of cattle, fishing, or managing shared forests.

Figure 1.7 Alfredo's payoffs to fishing more or less depend on how much Bob fishes. Alfredo's payoffs are described using the words we used for the coordination problem: Planting in Palanpur. Alfredo ranks his outcomes from best to worst: Best > Good > Bad > Worst. Alfredo's strategies and outcomes are highlighted in blue. Bob's strategies and outcomes are highlighted in red (but we have not put the words to describe Bob's outcomes in the figure).

		Bob	
		10 hours	12 hours
Alfredo	10 hours	Good	Worst
	12 hours	Best	Bad

We can describe the fishermen's interaction in the form of a *payoff matrix*.

M-CHECK A *matrix* is a rectangular array of quantities or other quantitative information.

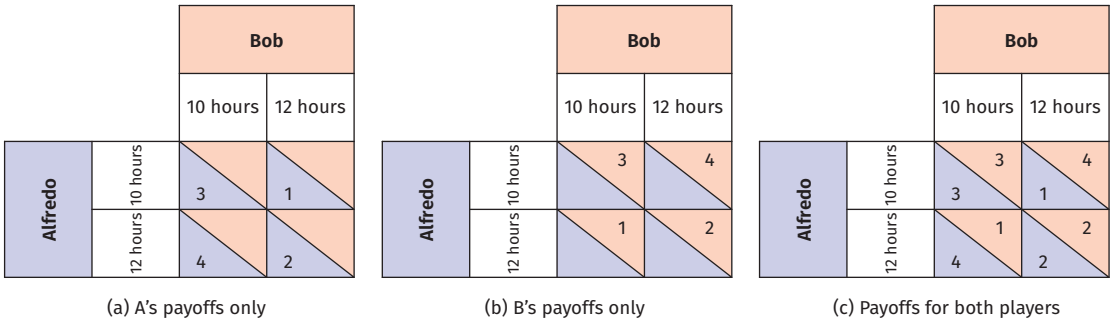
We first present a version of the payoff matrix with words to represent Alfredo's payoffs (but not yet Bob's) in Figure 1.7. Read the table this way: if Bob fishes 12 hours (the right-hand column) and Alfredo fishes 10 hours (top row) this is the *worst* outcome for Alfredo. A payoff matrix presents hypothetical 'if–then' information; it presents all of the possible sets of payoffs, whether or not each is likely ever to occur.

The complete payoff matrix for the Fishermen's Dilemma is represented in Figure 1.8 with numbers indicating the two fishermen's evaluation of how good the outcome indicated is. So for example the payoff to each if they both fish 10 hours (3) is 50 percent greater than if they both fish 12 hours (2).

The convention we will use throughout this book is to list the row player's payoffs first and in the bottom-left corner of the cell and the column player's payoffs second in the top-right corner. So, in the Fishermen's Dilemma game, we list Alfredo's payoffs first and Bob's payoffs second. We shade each player's payoffs to make them easier to differentiate: blue for the row player (Alfredo) and red for the column player (Bob).

Many of the games in this book involve two players and each player has two possible strategies. We often call a game like this a "2 × 2" game (a "two-by-two" game). We now have all the elements we need for the

Figure 1.8 Payoffs of players in the Fishermen's Dilemma. Alfredo's payoffs are in the bottom-left corner of each cell and are shaded blue. We include Alfredo's payoffs in panel (a) and panel (c). Bob's payoffs are in the top-right corner of each cell and are shaded red. We include Bob's payoffs in panel (b) and panel (c).



(a) A's payoffs only (b) B's payoffs only (c) Payoffs for both players

complete description of the Fishermen's Dilemma and its strategy profiles and associated payoffs.

- *Alfredo fishes 12 hours, Bob fishes 12 hours*: When both fishermen fish 12 hours, they each catch fewer fish per hour of work, while they also have a higher cost of effort because they've spent a lot of time fishing. Each fisherman ends up with 2.
- *Alfredo fishes 10 hours, Bob fishes 10 hours*: When both fishermen spend less time fishing they catch a decent amount of fish and they haven't fished so long that the other fisherman catches fewer fish. They also benefit from a lower cost of time spent fishing. Each gets a net benefit of 3.
- *Alfredo fishes 10 hours, Bob fishes 12 hours*: Because Bob fishes 12 hours, Al catches many fewer fish and because Bob still fishes for another two hours, he catches a lot of fish while Al doesn't fish. Consequently, with the cost of time and catching fewer fish, Al ends up with net benefits of 1 and Bob ends up with net benefits of 4.
- *Alfredo fishes 12 hours, Bob fishes 10 hours*: This is symmetrical to the previous description, so now Al gets net benefits of 4 and Bob gets net benefits of 1.

CHECKPOINT 1.6 Payoff matrixes

a. Fill in the blank red triangles showing Bob's payoffs in Figure 1.7 using the payoffs shown in Figure 1.8.

b. You can now see that the cartoon in Figure 1.3 is a payoff matrix. What are the main differences in the payoffs of the Planting in Palanpur game and the Fishermen's Dilemma in Figure 1.8?

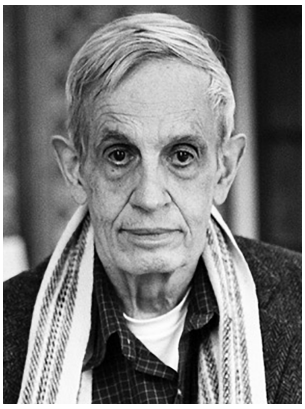


Figure 1.9 John F. Nash (1928–2015) was an American mathematician who contributed the to the theory of bargaining and to the concept of equilibrium that bears his name. He won the Nobel Prize in economics in 1994. His life was documented in the book and movie *A Beautiful Mind*.[9]

Image courtesy of Peter Badge/Typos1, CC-BY-SA-3.0.

1.7 PREDICTING ECONOMIC OUTCOMES: THE NASH EQUILIBRIUM

As you already know, to predict a game outcome—the strategy profile that will result—we need more than the description of the game alone. We need a solution concept—a statement about *how* players will behave in the game—that can be the basis of a *prediction* of the game's outcome. Predicting the outcome of a game—based on the rules of the game and the solution concept—is especially important if we are evaluating policies to improve the functioning of the economy by changing the rules of a game so as to change the outcome of a game.

Equilibrium and prediction

The key idea on which a solution concept is based is **equilibrium**. An equilibrium is a state in which there is nothing in the situation that will cause the state to change. A *predicted outcome* will be an equilibrium, that is, an outcome that is stationary (not changing). To understand why, imagine this were not the case. You make a prediction, but then the outcome changes. Your prediction would *no longer be true* because the outcome had changed.

Applying this reasoning to games, if we were to predict the outcome of a game to be a strategy profile under which one or more players would have reason to change their strategy, then the prediction would be falsified as soon as they carried out the change. So the status of stationarity—changelessness—is a property of a prediction; and this is why equilibrium is fundamental to making predictions about game outcomes.

Think of a concrete example. Suppose you want to predict where a marble will be if all that you know is that it is going to be somewhere in a round bottomed salad bowl sitting on a table. If I predicted that the marble would be somewhere halfway up the side of the bowl you would doubt my prediction. The reason is that any marble in that position would move downward in the bowl; that is, its position would not be stationary, so, if it ever were (for some reason) where I predicted it would be, it would not be there any longer. It is not that the prediction would necessarily be wrong. It could be true for a millisecond after I placed the marble in the bowl just above my predicted spot, for example.

The only predicted position in the salad bowl that would not immediately falsify itself in this sense is the bottom. So a reasonable prediction of the location of the marble would be "the bottom of the bowl."

There are some situations in which a prediction based on an equilibrium would be likely to be incorrect. Change the marble-in-bowl example by

EQUILIBRIUM An equilibrium is a situation that is stationary (unchanging) in the absence of a change external to the model.

filling the bowl with very thick honey. Then if you were asked to predict where the marble would be found, you would want to know how long it had been in the bowl, did it have time to reach the bottom? If the marble had been placed in the bowl just a millisecond ago, then you might be better off predicting that it would be where it had been placed, rather than the bottom of the bowl.

The marble-in-bowl-of-honey example is often a better illustration of how economic processes work than the initial example. Markets are often out of equilibrium. Predicting things in motion is a much more challenging task than predicting them when they are stationary. We provide an example in a model of residential segregation (section 1.15) where we follow the process of change step by step. But for the most part we study equilibria and how to change them so as to improve outcomes.

In the marble-in-bowl illustration (without the honey) what is the solution concept that lets us arrive at the "bottom of the bowl" prediction? It is gravity, which is our understanding about a reasonable way for the marble to "behave." In modeling an economic interaction, the game structure is analogous to the salad bowl. What is the analogy to gravity? The answer is the player's *best response.*

Best-response strategies

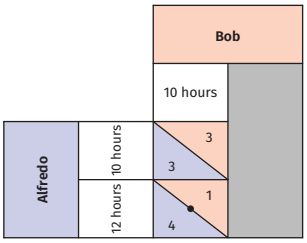
By far the most widely-used solution concept, the Nash equilibrium, is based on the idea that players choose **best-response** strategies; they do the best they can given the strategies adopted by everyone else.

To understand better what a best response is, think about Alfredo in the Fisherman's Dilemma and imagine each of the possible situations that might occur and what would be best for him in each of these hypothetical situations.

First, what strategy should Alfredo adopt in order to gain the highest payoff if Bob were hypothetically to play Fish 10 hours as shown in Figure 1.10. We do not ask why Bob would do this. We are mapping all of the possible situations that Alfredo might encounter.

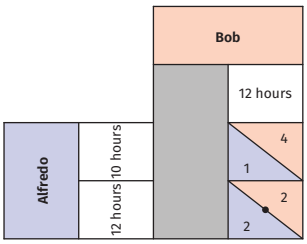
- Against Bob playing Fish 10 hours, Alfredo can get a payoff of 3 for fishing 10 hours or a payoff of 4 for fishing 12 hours.
- 4 > 3 therefore Fish 12 hours is Alfredo's *best response* to Bob playing Fish 10 hours.
- Place a solid dot in the cell (Alfredo plays Fish 12 hours, if Bob plays Fish 10 hours) to indicate that it is Alfredo's best response. We will use this "circle and dot" method to find the Nash equilibrium.

Figure 1.10 Imagine Bob playing Fish 10 hours to assess which of Alfredo's strategies is a best response.



BEST RESPONSE A strategy is a player's best response to the strategies adopted by others if no other strategy available would result in higher payoffs.

Figure 1.11 Imagine Bob playing Fish 12 hours to assess which of Alfredo's strategies is a best response.



Let's repeat the analysis and imagine Bob playing Fish 12 hours, as shown in Figure 1.11.

- Against Bob playing Fish 12 hours, Alfredo can get a payoff of 1 for playing Fish 10 hours or a payoff of 2 for playing Fish 12 hours.
- 2 > 1 therefore Fish 12 hours is Alfredo's *best response* to Bob playing Fish 10 hours.
- place a solid dot in the cell (Alfredo plays Fish 12 Hours, Bob plays Fish 12 hours) to indicate that it is Alfredo's best response.

CHECKPOINT 1.7 A best response for Bob Repeat the process we went through for Alfredo, but do it for Bob instead. Notice that when you do so, you will blank out a *row* for Alfredo to imagine him playing the strategy in the other row, whereas you blanked out a *column* for Bob. What are Bob's best responses? Show his best responses using a hollow circle.

M-CHECK Strong and weak best response. A best response may be either *strong* or *weak*. A strong (also called strict) best response yields higher payoffs than any other: it is strictly "better" than any other strategy. There can be no strategy that is better than a strong best response but a weak best response need not be better than any other; it may be "as good as" (the payoffs to the strategy and some alternative strategy being equal).

EXAMPLE The rock–paper–scissors game (also called ro–sham–bo and by many other names in other languages) originated in China about two thousand years ago. It does not have a Nash equilibrium.

Nash equilibrium and the outcome of a game

Some games do not have a Nash equilibrium and you will see shortly that some have more than one.

Using the best responses of the players we can now predict the outcome of a game using as our solution concept the **Nash equilibrium**. A Nash equilibrium is a profile of strategies—one for each player—each of which is a best response to the strategies of the other players. A Nash equilibrium is also called a *mutual best response*. Because at a Nash equilibrium all players are playing their best response to all of the others, it follows that no player has a reason to change his or her strategy as long as the other players do not change theirs. In Figure 1.12, Alfredo's best responses are shown by the solid black dot in the cell. Bob's best responses are shown by the hollow circle. Their best responses coincide at the Nash equilibrium (Fish 12 hours, Fish 12 hours) with payoffs (2, 2) shown in the cell where the solid dot is inside the hollow circle. You can use the "dot and circle" method to find one or more Nash equilibria (if they exist) for games that can be represented by a payoff matrix like Figure 1.12.

The outcome demonstrates how Nash equilibrium can initially seem counter-intuitive. Both would have had higher payoffs if they could have agreed to restrict their fishing to 10 hours (they could have had 3 each if they both fished 10 hours and 3 > 2). But suppose both were restricting their fishing to 10 hours; then each would an incentive to fish for 12 hours (because 4 > 3) and unless they had a binding agreement to continue fishing less, each would choose to fish more.

NASH EQUILIBRIUM A Nash equilibrium is a profile of strategies—one strategy for each player—each of which is a best response to the strategies of the other players.

Figure 1.12 Payoff matrix for the Fishermen's Dilemma. The solid dots indicate Alfredo's best responses. The hollow circles indicate Bob's best responses. A Nash equilibrium is a cell that contains both. In this case there is just one Nash equilibrium: both fishing 12 hours.

		Bob	
		10 hours	12 hours
Alfredo	10 hours	3, 3	1, 4
	12 hours	4, 1	2, 2

The Fishermen's Dilemma is therefore a coordination problem and it returns us to the classical institutional challenge. Without institutions to align the individual interest of the participants with their shared interest, they get an outcome that is worse for both of them than other possible outcomes. We will later show how a change in the institutions regulating how Alfredo and Bob interact—that is, changing the rules of the game—might address this coordination failure.

CHECKPOINT 1.8 Nash equilibrium

a. Explain why none of the other three outcomes (those that are not (Fish 12 hours, Fish 12 hours)) of the Fishermen's Dilemma satisfy the definition of Nash equilibrium.

b. At each of the other three outcomes, which player has an incentive to change strategy and in what way? Explain.

c. Explain why a game like rock–paper–scissors would not be much fun if there were a Nash equilibrium.

Dominant strategies

In the Fisherman's Dilemma (and all Prisoners' Dilemmas) there is a single strategy that yields the highest payoffs to a player independently of which of the strategies the other player adopts. A strategy is a player's **dominant strategy** if it is the player's best response to all possible strategy profiles of

DOMINANT STRATEGY A strategy is dominant if it yields the highest payoff for a player for any strategy chosen by the other players. Weak dominance refers to the case where there are one or more other strategies yielding the same payoff.

the other player or players. That is, a strategy is a dominant if by playing it the player's payoff is *greater than or equal to* the payoff they would get by playing any other strategy for every one of the other player's profiles of strategies.

Likewise we say that strategy A is *dominated* by another strategy B if the payoff to playing B is at least as great or greater than playing A for every strategy profile of the other players. If there is a strategy that dominates all of the other strategies that a player may choose, then it is a dominant strategy. If each player in a game has a dominant strategy, then the strategy profile in which all players adopt their dominant strategy is called a **dominant strategy equilibrium**.

We can apply the concept of *dominant strategy equilibrium* to the Fishermen's Dilemma. To do so, we need to understand whether each player has a dominant strategy.

- When Alfredo fishes 10 hours, his payoff is 3 if Bob fishes 10 hours and 1 if Bob fishes 12 hours.
- When Alfredo fishes 12 hours, his payoff is 4 when Bob fishes 10 hours and 2 when Bob fishes 12 hours.
- So, when Bob fishes 10 hours, fishing 12 hours gets Alfredo a higher payoff ($4 > 3$) and when Bob fishes 12 hours, fishing 12 hours gets Alfredo a higher payoff ($2 > 1$).
- Therefore, Alfredo gets a higher payoff from fishing 12 hours against each of Bob's strategies.
- Fish 12 hours is therefore Alfredo's *dominant strategy*.

Fishing 12 hours is also Bob's dominant strategy. Because each player has a dominant strategy to fish 12 hours, the *dominant strategy equilibrium* is (Fish 12 hours, Fish 12 hours) with payoffs (2, 2). The dominant strategy equilibrium of a game is always a Nash equilibrium.

The fact that the Fishermen's Dilemma has a dominant strategy equilibrium makes it a particularly simple problem for us, studying it. It also makes it simpler for Bob and Alfredo because what is best for each does not depend on what the other does. But this does not mean that they will be happy with the result.

CHECKPOINT 1.9 Dominance and Nash equilibrium

a. Repeat the analysis we did for Alfredo for Bob and confirm that Fish 12 hours is a dominant strategy for him too.

b. We said that a dominant strategy equilibrium is always a Nash equilibrium. But do you think that a Nash equilibrium is always a Dominant Strategy equilibrium? Why or why not?

DOMINANT STRATEGY EQUILIBRIUM A dominant strategy equilibrium is a strategy profile in which all players play a dominant strategy.

1.8 EVALUATING OUTCOMES: PARETO COMPARISONS AND PARETO EFFICIENCY

The Nash equilibrium can help us predict the result of an interaction. But it does not tell us anything about whether some outcome is good by any standard, or even better or worse than some other outcome. Economists, policymakers and others would like to evaluate whether some outcomes are better or worse. We do this so that we can try to work out which rules of the game would make the better outcomes Nash equilibria, and therefore more likely to be what we observe as the real outcomes in the economy in question.

The challenge in making these comparisons is that whether some outcome is better than another depends on what you value. There is no agreed upon standard of what makes one outcome better than another. Returning to our fishermen, here are some of the values that we could use to evaluate an outcome:

- *Fairness in the distribution of payoffs* among the players; is it fair that Alfredo receives four times what Bob gets when Alfredo does *not* limit his fishing hours and Bob does?
- *Are the rules of the game itself fair?* In the Fishermen's Dilemma the same rules applied to both players; but were the game a bit different, many would think it unfair if Alfredo could simply order Bob to fish 10 hours, or to hand over half of all the fish Bob caught.
- Setting aside fairness, is the outcome a *reasonable use of available resources* including the working time of the two fishermen and the sustainability of the lake itself and the living things that it supports.

There are many other standards that could be proposed.

A concept that is widely used to evaluate economic outcomes involving two or more people is called **Pareto efficiency**. The idea is simple: an objective of public policy and institutional design—the rules of the game—should be to avoid those outcomes—like traffic jams, planting late in Palanpur, and overfishing the lake—that are worse for everyone, compared to an alternative outcome that also would have been feasible.

Pareto comparisons

Pareto efficiency is based on Pareto comparisons of outcomes. Consider two outcomes, A and B, with resulting payoffs for two or more players.

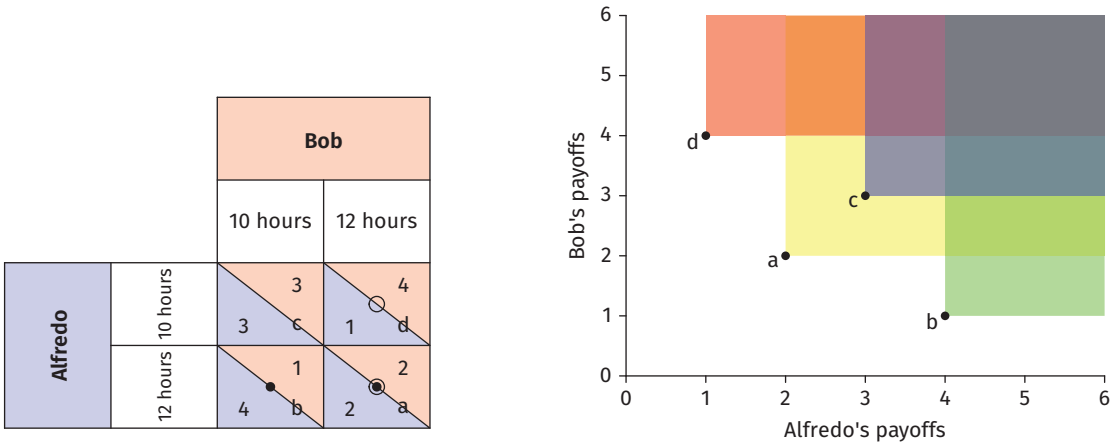
PARETO EFFICIENCY A Pareto-efficient allocation is an allocation with the property that there is no alternative technically feasible allocation in which at least one person would be better off and nobody would be worse off. If an allocation is Pareto efficient, then there is no alternative allocation that is Pareto superior to it.

Outcome A is **Pareto superior** to outcome B if in outcome A at least one player is better off than in outcome B without anyone being worse off. A change in the outcome from a Pareto-inferior situation like B to a Pareto-superior outcome like A is called a *Pareto improvement*. This is a Pareto comparison. An outcome is Pareto efficient if no other feasible outcome is Pareto superior to it.[10]

Figure 1.13 depicts the outcomes of the interaction between Alfredo and Bob. Figure 1.13 (a) a is the Fishermen's Dilemma payoff matrix with each outcome given a label **a**, **b**, **c**, or **d**. These payoffs are indicated by points in Figure 1.13 (b) where you can read on the vertical and horizontal axes the payoffs to the two players that you see in the payoff matrix.

The Pareto comparison is easy to see in this type of plot. An outcome A is Pareto superior to another outcome B if the point indicating the payoffs from A lies to the "northeast" of point indicating the payoffs from B. "Northeast" in this figure is "better for both." So looking at the colored areas whose lower-left corners are points **a**, **b**, **c**, and **d**, then a Pareto-efficient point is one that has no other point in its "colored shadow" extending

Figure 1.13 Three Pareto-efficient outcomes of the Fishermen's Dilemma. Panel (a) is the same as Figure 1.12 except that each of the four squares in the payoff matrix has been assigned a letter. In panel (b) we show the Fishermen's Dilemma indicated by the payoffs of the two at the four possible outcomes given by the same letters that appear in each of the cells of the payoff matrix. We use shaded colors indicating 90-degree angles to the northeast of the feasible outcomes (each of the lettered points).



(a) Fishermen's Dilemma with labeled points

(b) Analyzing points for Pareto efficiency

PARETO-SUPERIOR Outcome A is Pareto-superior to outcome B (it Pareto-dominates outcome B) if, in outcome A, at least one player is better off than in outcome B without anyone being worse off.

upward and to the right of the point. By this standard, three of the points–**b**, **c**, and **d**–are Pareto efficient, while **a** is not, because point **c** is in the yellow "color shadow" of point **a**.

We say that two outcomes can be Pareto compared, or Pareto ranked, if one of them is Pareto superior to the other. But as you can see from the figure Pareto comparisons (or rankings) are often not possible. Specifically, when two outcomes are both Pareto efficient, they cannot be Pareto compared or Pareto ranked. We could rank **c** above **a** because both players were better off, but with **b**, **c**, and **d** we cannot move from one outcome to another without worsening outcomes for at least one of the players.

HISTORY Pareto efficiency
The idea is attributed to the Italian economist and sociologist Vilfredo Pareto (Pa RAY to) but it was first introduced by the British economist, Francis Edgeworth. However, an important concept that we use extensively in later chapters, called the Edgeworth Box, was actually invented by Pareto.[11]

CHECKPOINT 1.10 Pareto improvements in the Fishermen's Dilemma
Referring to Figure 1.13 (b) do the following:

a. Is any point dominated by some other point? Say which, if any?
b. At which point is the total payoff of the two fishermen the greatest?
c. Would a change from any other point to that "total payoff maximum" point be a Pareto improvement?
d. Explain what you think is the meaning of the expression "there is no such thing as a free lunch" and say whether this saying is true in Figure 1.13 (at all of, some of, one of the, none of the points).

1.9 THE VALUE AND LIMITATIONS OF PARETO EFFICIENCY

Pareto efficiency gives us a way to identify "lose-lose" outcomes we should seek to avoid, namely those "that are worse for all than they could be." But except in special cases, Pareto efficiency does not provide a rule to select what we might call "the best" outcome.

To see why this is true, suppose we have a cake and we are dividing it among people, all of whom enjoy eating cake. An outcome in which one person gets the entire cake is surely Pareto efficient because in any other allocation that lucky person would get less. Likewise an allocation in which everyone got the same sized slice of the cake is Pareto efficient, for in any other allocation at least one person would have to get less.

Pareto efficiency is *not* about how something of value should be divided up. All it says is "make sure there's no cake left on the table!"

Most economic problems that we face are similar to the cake example in that there are a great many Pareto efficient outcomes. Think about the Fishermen's Dilemma game: all of the possible outcomes of the game *except one* are Pareto efficient. When there are many Pareto-efficient outcomes there is a *conflict of interest* among players over which Pareto-efficient outcome they would prefer. We cannot say that one is "more Pareto efficient" than the other.

It is also perfectly sensible to prefer an outcome that is not Pareto efficient but is more fair over an alternative Pareto-efficient outcome that is unfair. To continue the cake example, if there are two people between whom the cake will be divided many people would reject the (Pareto-efficient) outcome in which one person gets the entire cake in favor of a Pareto-inefficient alternative in which each gets one-third of the cake (the remaining one-third perhaps being thrown away or destroyed in the conflict over its distribution). So we would prefer a Pareto-inefficient outcome over a particular efficient outcome (one person gets the whole cake). But the Pareto comparison would remind us that each person getting half of the cake is preferable to each getting a third with the rest being wasted.

Pareto efficiency is a useful device for screening out those outcomes (like throwing away some of the cake in the above example, or planting late in Palanpur, or over-fishing the lake) that should not be among the list of candidate feasible outcomes among which the choice of better or best should be made on grounds of fairness or other bases.

CHECKPOINT 1.11 **Pareto efficiency** Consider these questions about Pareto efficiency.

a. True or false (and explain): "The fact that an outcome is Pareto efficient does *not* imply that it is preferred by all the actors to all the other outcomes."

b. Can two Pareto-efficient outcomes be Pareto compared? Why or why not? Explain.

c. Imagine you are an impartial observer evaluating the possible outcomes that might occur for Bob and Alfredo. Are there any reasons why you might judge the outcome **a** in the figure to be better than the Pareto-efficient outcomes **b** and **d**, despite the fact that **a** is Pareto inefficient?

1.10 CONFLICT AND COMMON INTEREST IN A PRISONERS' DILEMMA

You know that the game the fishermen are playing is a particular case of the **Prisoners' Dilemma**. We now point out some of the general characteristics of this particular kind of coordination problem.

PRISONERS' DILEMMA A Prisoners' Dilemma is a 2-by-2 social interaction in which there is a unique Nash equilibrium (that is also a dominant strategy equilibrium), but there is another outcome that gives a higher payoff to both players (and a higher total sum of payoffs than any other outcome), so that the Nash equilibrium is not Pareto efficient.

A Prisoners' Dilemma is a two-person interaction in which there is a unique Nash equilibrium (that is also a dominant strategy equilibrium), but there is another outcome that gives a higher payoff to both players, so that the Nash equilibrium is not Pareto efficient. This means that in the Prisoners' Dilemma both players get their second-worst payoffs in the game by playing their strictly dominant best-response strategies.

In Figure 1.14 we show the familiar payoff matrix for the Fishermen's Dilemma, but instead of the numbers indicating the payoffs of the players now we label the payoffs w, x, y, and z. We label the Fish 10 hours strategy "Cooperate" because it is the mutually beneficial action the two fishermen could take if they could coordinate their actions. The strategy Fish 12 hours is labeled "Defect" because choosing to fish 12 hours instead of 10 is deviating from a mutual cooperate outcome in which the two fishermen might be able to coordinate.

The interaction is a Prisoners' Dilemma if two conditions hold:

- $y > w$ and $z > x$ means that playing Defect is a strict dominant strategy
- $w > z$ means that mutual cooperation is Pareto superior to mutual defection.

For Alfredo, 12 hours is a best response to Bob playing 10 hours because $y > w$; 12 hours is also a best response to 12 hours because $z > x$ (both best responses are shown in Figure 1.14 by the solid dot). Similarly, for Bob, 12 hours is a best response to Alfredo playing 10 hours because $y > w$; 12 hours is also a best response to 12 hours because $z > x$ (both best responses are shown by the hollow circle). The dot inside circle in Figure 1.14 confirms that if the game is played noncooperatively the Nash equilibrium is

M-CHECK A third condition is sometimes added, namely $x + y < 2w$ which means that the sum of payoffs when both players cooperate is greater than the sum of payoffs when one cooperates and the other defects. This condition makes (Cooperate, Cooperate) preferable to the outcome in which one defects and the other cooperates.

Figure 1.14 A general Prisoners' Dilemma. For the game to be a Prisoners' Dilemma, we require $y > w > z > x$ and $2w > y + x$ (this is $4 > 3 > 2 > 1$ and $2 \times 3 > 4 + 1$ in the numerical example).

		Bob	
		10 hours (Cooperate)	12 hours (Defect)
Alfredo	10 hours (Cooperate)	w, w	x, y
	12 hours (Defect)	y, x	z, z

(12 hours, 12 hours) with payoffs (z,z) when by coordinating their choices they could have each received (w,w) (where $w > z$).

CHECKPOINT 1.12 When is cooperation *not* the best they could do? Show that if the condition $x+y<2w$ is violated (i.e. if $x+y>2w$), then the two players could do better by

a. one defecting on the other and then sharing their total payoffs equally or

b. if the game is played many times, alternating who cooperates and who defects.

Economic rent: The incentive to coordinate

Both players have a good reason to try to change the rules of the game so that they can agree on both cooperating. How much more they would get if they were to mutually cooperate than if they mutually defected—in this case $w-z$—is called an **economic rent**, meaning the difference between the payoff that they would get if they cooperated and their next best alternative. Their next best alternative to cooperating, in this case, is mutual defection, also known as their **fallback position**.

Economic rents and the fallback position play a central role in microeconomic theory, so it is a good idea to master them. The meaning of the term fallback position is intuitive: it is what you *fall back* to if some particular outcome is not possible, in this case if the mutual cooperation should not work out. A player's **fallback position** or **fallback option** is the *payoff* they receive in their *next best alternative*.

The term "economic rent" may at first seem surprising, because the word "rent" also means a payment for the temporary use of something, for example, to a landlord or a car rental agency. The term economic rent means something entirely different. A participant's *economic rent* is the payoff they receive in excess of what they would get in their fallback position.

We shall use the idea of a fallback often, from social interactions like the Prisoners' Dilemma, to worker-employer relationships where a worker's next best alternative may be unemployment so the rent she receives as an employee is the difference between her wage and the government transfer she would receive were she to lose her job. The next best alternative for a person applying to a bank for a loan is trying to get money from friends

FALLBACK POSITION A player's fallback position (or reservation option) is the payoff they receive in their next best alternative.

ECONOMIC RENT A participant's economic rent is the payoff they receive in excess of what they would get in their fallback position.

or family along with future obligations. As these examples indicate, the fallback position will differ depending on the details of the situation and what our modelling assumptions are designed to illuminate.

Impediments to coordination: Limited information and conflicts of interest

If $w - z$ is substantial—meaning substantial rents associated with cooperation for each player—then it might seem a simple matter for the players to agree to cooperate. But people often fail to reach or enforce such an agreement, for two main reasons:

- *Limited information:* The participants may lack the information needed to monitor and enforce an agreement. How can a participant know or verify what other participants do?
- *Conflict over distribution of the economic rents from cooperation:* Disagreement about who gets what—for example who gets to fish more—may make it impossible for the two to agree.

Concerning the information problem, the fishermen, for example, may have no way of enforcing an agreement, or even knowing if the agreement has been violated. While each may know how many hours the other has fished on a day with clear and sunny weather, on a foggy day it may be impossible to know. Even if one fisherman knows how much the other fished, that knowledge may be insufficient to enforce an agreement through a third party such as a court of law.

This is the problem of **asymmetric information** or non-**verifiable information**. Information is asymmetric if people know different things, or if what one person knows (for example how many hours he fished), the other person does not know. Information is not verifiable if people cannot use it to enforce an agreement or a contract. Asymmetric and non-verifiable information will play a central role in our analysis of how the labor market, the credit market, and other markets work. For example most courts will not accept "hearsay" (meaning "second-hand") information, so if one of the fishermen had heard from someone else that the other had fished 12 hours, this would be non-verifiable information.

Concerning the second problem for coordination, conflicts over the distribution of the economic rents from cooperation, in the Fishermen's

EXAMPLE The term economic rent is what is known in the study of language as a "false friend," a term that you think you know the meaning of but means something entirely different in the new language you are now learning. "Sensible" in English means "reasonable" but in Italian it means "sensitive."

ASYMMETRIC INFORMATION Information is asymmetric if something that is relevant to the parties in an economic interaction is known by one actor and is not known by another.

VERIFIABLE INFORMATION Information that can be used in legal proceedings to enforce a contract or other agreement is termed verifiable information.

Dilemma, the agreement to restrict fishing to 10 hours a day divides the benefits of restricting fishing in a particular way, namely equally. But the fishermen need not agree on 10 hours each. Alfredo might insist that he will fish 12 hours and Bob only 10 hours. Or Bob might insist on the opposite.

Or Bob might insist that both fish 10 hours, but that Alfredo give him most of Alfredo's catch, leaving Alfredo with just enough of his catch to be no worse off than had they both fished 12 hours, namely with a payoff of z. Which of these we will observe depends on rules of the game we have not yet introduced, including differences in the bargaining power exercised by the two. Unless they can find a mutually acceptable solution to the distribution problem they may end up having no agreement at all, and then simply fish at 12 hours each, as their fallback position.

The fishermen's distribution conflict highlights a challenge that arises in any voluntary economic interaction. Consider their possible agreement to limit their fishing time:

- The agreement is voluntarily entered into. This means that neither player can force the other to accept terms worse than their fallback position.
- The agreement therefore must allow each participant to achieve a payoff greater than (or at least not worse than) what would have resulted had the individual not agreed to cooperate. In other words, there must be some economic rents made possible by a voluntary cooperative outcome.
- This being the case, the participants have to find a way that the total rents will be divided. If they are to agree to cooperate by restricting the total time they spend fishing, they must also agree on how these economic rents will be distributed.
- Conflict over the distribution of the economic rents (who gets what amount of economic rent) may prevent the fishermen from coming to an agreement.

✓ **FACT CHECK** In the next chapter we will see that across many cultures of the world, people would rather get nothing than get what they consider to be an unfair share of the economic rents, and as a result cooperation breaks down and nobody gets any rent at all.

We sometimes think of cooperation and conflict as opposites, as for example when members of a team cooperate in their efforts to win some conflict with another team. But the Prisoners' Dilemma is a scenario of *conflict and cooperation among the very same people*. They have common interests in getting some share of the economic rents by cooperating; but they have conflicting interests in how the total will be divided into the rents received by each.

A catalog of games: And their challenges to coordination

Some interactions present greater impediments to coordination than others; the Prisoners' Dilemma is in some respects the most challenging of all.

We can classify coordination problems and the challenges they present by the relation between Nash equilibria and Pareto-efficient outcomes of the games that represent them.

- In the Prisoners' Dilemma, as you know, there is a unique Nash equilibrium that is Pareto inefficient. Because this outcome is also a dominant strategy equilibrium, coordination on mutual cooperation will require some change in the rules of the game (making it a cooperative game) or a change in the players' payoffs, for example, if they dislike harming the other player by defecting on them.
- In interactions like Planting in Palanpur, which are often called *Assurance Games*, there are two Nash equilibria, (both Plant Early and both Plant Late) one of which (Plant Early) is Pareto superior to the other (Plant Late). In these games if one of the players plays the strategy making up the Pareto-superior equilibrium (Plant Early), then the best response of the other will be to do the same. Finding institutions that will implement the preferred plant early outcome in a game like this will be a lot less challenging than in a Prisoners' Dilemma.
- Another important class of coordination problems arise in what we call *Disagreement Games* where there are two Nash equilibria each of which is Pareto efficient, so that they cannot be Pareto ranked, and players disagree about which Nash equilibrium they would like to occur. These are like the Planting in Palanpur game but with the additional challenge stemming from a conflict over which Nash equilibrium will be implemented.

We start with an even less challenging game in which players' self-interests lead them to a Pareto-efficient Nash equilibrium.

CHECKPOINT 1.13 **Guilty prisoners** Referring to Figure 1.13 (a), imagine that both Bob and Alfredo have become ethical and now would feel guilt if one defected when the other cooperated. Should they do this, their guilt results in a subtraction from the payoff points shown in the figure.

a. What is the smallest value of this guilt that each feels that would make Defect no longer be the dominant strategy?

b. If Alfredo but not Bob acquired this "defection guilt" so that for him defecting on a cooperator was no longer a best response, but Bob continued with the values (and payoffs) in the figure, is there a Nash equilibrium of the game, and if so, what is it?

1.11 COORDINATION SUCCESSES: AN INVISIBLE HAND GAME

The characteristic of what we call an Invisible Hand Game is that it has a single Nash equilibrium that is Pareto efficient. Apologies to Adam Smith: our game is much simpler than Smith's reasoning and Smith did not use

Figure 1.15 The Corn-Tomatoes Game: An Invisible Hand Game. The players' best responses are indicated by dots (Arkady) and hollow circles (Barbara). Arkady's payoffs are listed first in the bottom-left corner. Barbara's are listed second in the top-right corner. The game captures Adam Smith's ideas of *specialization* and *gains from trade* (that is, the opportunity to obtain economic rents from trade).

		Barbara	
		Corn	Tomatoes
Arkady	Corn	1 / 2.5	2 / 2
	Tomatoes	4 / 4	2.5 / 1

ideas like Pareto efficiency. But our **Invisible Hand Game** illustrates Adam Smith's core insight that through the competitive buying and selling of privately owned goods on competitive markets, self-interested people can implement outcomes that are by some standards socially desirable (we spell out what this means and the conditions under which it might come about in Chapter 14).

Consider a 2-by-2 game with two players, Arkady and Barbara, both farmers. Each player can choose one of two strategies: planting corn or planting tomatoes. The payoffs that they assign to the various outcomes of the game are provided in the matrix in Figure 1.15, which we call the Corn-Tomatoes game.

The payoff matrix reflects two facts about the problem that the two farmers face.

- Either because of their skills or the nature of the land they own, Arkady is better at growing tomatoes; Barbara is better at growing corn.
- They both do poorly when they produce the same crop because the increased supply of whichever good it is that they both produce drives down the price.

INVISIBLE HAND GAME An Invisible Hand Game has a single Nash equilibrium that is Pareto efficient.

The Nash equilibrium of the Corn-Tomatoes game is (Tomatoes, Corn), that is, Arkady plants tomatoes, and Barbara plants corn, at which the players receive payoffs (4,4). The equilibrium (Tomatoes, Corn) is Pareto efficient as there is no alternative outcome which is Pareto superior to it. This is the best they could do. There was no need for them to explicitly agree on how to coordinate to achieve this result.

Just as in Adam Smith's reasoning about his invisible hand, Arkady and Barbara, are in a situation in which by simply following their self-interest they coordinate to their mutual benefit.

CHECKPOINT 1.14 **Invisible Hand Game** Which entries in the payoff matrix would you have to compare in order to show the following:

a. They each do better when Arkady specializes in tomatoes and Barbara specializes in corn then vice versa.

b. They each do worse when both produce the same crop.

c. Growing corn is Barbara's dominant strategy.

d. Arkady growing tomatoes and Barbara growing corn is the dominant strategy equilibrium.

e. Explain why the Nash equilibrium of the game is Pareto efficient.

1.12 ASSURANCE GAMES: WIN-WIN AND LOSE-LOSE EQUILIBRIA

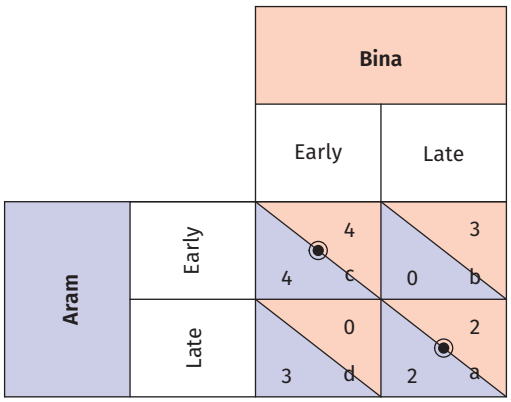
Return to the farmers in Palanpur. There are two Nash equilibria in this game, one in which both participants Plant Early and one in which both Plant Late. The best response to the other farmer's planting early is also to plant early, while the best response to the other farmer's planting late is also to plant late. The outcome where both farmers plant early is Pareto superior to the outcome when both farmers plant late.

The players do not have any conflict of interest: both would share equally in the gains from cooperation, should they find a way to coordinate on planting early. The problem for the real-life farmers of that village is that they are stuck in the Pareto-inefficient Nash equilibrium of what is called an Assurance Game. Their challenge is how to move to the Pareto-superior Nash equilibrium.

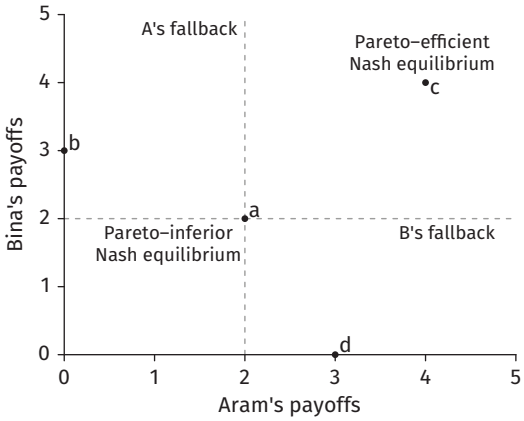
This could happen if all the participants had confidence (were *assured*) that the other participants also play the strategy yielding superior outcome. This is why it is an "Assurance Game."

Figure 1.16 is the payoff matrix for two players, Aram and Bina, choosing when to plant their millet in the village of Palanpur, India. (It is the same as the earlier figure about the two farmers, except that we now have numbers representing the farmers' payoffs.) Coordination failures arise in the Assurance Game because of *positive feedbacks*: if one plants late

Figure 1.16 Planting in Palanpur: an Assurance Game. In panel (a), Aram's payoffs are listed first in the bottom-left corner. Bina's payoffs are listed second in the top-right corner. Aram's best responses are shown by the solid point and Bina's are shown by the hollow circle. The Nash equilibria of the game are (Plant Early, Plant Early) and (Plant Late, Plant Late), with payoffs (4,4) and (2,2). The Plant Early Nash equilibrium is Pareto efficient. The Plant Late equilibrium is not. In panel (b), the payoffs are plotted against each other. Aram's payoffs are plotted on the horizontal axis, increasing as you move rightward. Bina's payoffs are plotted on the vertical axis, increasing as you move upward.



(a) Assurance Gam with labeled payoffs



(b) Payoffs plotted against each other

the more is the incentive for the other to plant late, and vice versa. The strategies are **strategic complements**.

CHECKPOINT 1.15 Graphing Palanpur

a. Using the graphical method for identifying Pareto-efficient outcomes as shown in Figure 1.16, show which outcomes in the Palanpur game are Pareto efficient.

b. Can you explain why **a** and **c** are Nash equilibria?

STRATEGIC COMPLEMENTARITY Strategic complementarity exists when (a) a strategy is a strategic complement to itself: the payoff to playing a particular strategy increases as more people adopt that strategy as a result of some form of positive feedbacks, or (b) one strategy and another are strategic complements to each other. In this case, for two activities A and B, the more that A is performed, the greater the benefits of performing B, and the more that B is performed, the greater the benefits of performing A.

Assurance game and strategic complementarity

Social media, dating platforms, and other matching services are examples of strategic complementarities. They are more valuable to everyone if many people participate.

Strategic complementarity exists when either of two conditions hold.

1. *A strategy is a strategic complement to itself*: The payoff to playing a particular strategy increases as more people adopt that strategy as a result of some form of *positive feedbacks*. Plant Early in Palanpur is an example. Dating platforms are another. The strategy could be "Open a dating app account." The positive feedback arises because the more that other people are using the dating app the more people you will "meet" (which is better for you and better for them). Tinder, Bumble, OkCupid, Hinge, Grindr, and other dating apps are social platforms illustrating what are called **network externalities** or *network external effects* which occur when the benefits to members of a social or physical network increase when more people join the network. In this case the strategy of joining the network is a strategic complement to *itself*.
2. *One strategy and another are strategic complements to each other*: The payoff to playing one strategy (say, A) is greater the more people adopt the other (B). In this case we say that strategies A and B are *strategic complements*. An example is the Invisible Hand Game shown in Figure 1.15. The payoff to Arkady from planting tomatoes is greater if Barbara plants corn (instead of tomatoes), and the payoff to Barbara from planting corn is greater if Arkady plants tomatoes (instead of corn). Growing corn and growing tomatoes are strategic complements.

We predict and evaluate the possible outcomes of the Planting in Palanpur Game using the concept of best response (using the dot and circle method introduced earlier). We see that the game has two Nash equilibria (Early, Early) with payoffs (4, 4) and (Late, Late) with payoffs (2, 2). The outcome (Early, Early) is Pareto superior to (Late, Late) and it is Pareto efficient because no alternative outcome is Pareto superior to (Early, Early).

Even though there is a Pareto efficient Nash equilibrium, that does not guarantee players will actually play it. A population—like the people of Palanpur—may get stuck in the Pareto-inferior Nash equilibrium. From the Assurance Game we have learned two things applicable across many kinds of social interaction:

ASSURANCE GAME An Assurance Game is a two-person, symmetric, strategic interaction with two strict Nash equilibria, one of which is Pareto superior to the other.

- the fact that a Pareto-efficient outcome is a Nash equilibrium does not mean that it will be the one we observe; getting there is not assured; and
- in cases where there is more than one Nash equilibrium, we need more information than is provided by the Nash equilibrium and Pareto-efficiency concepts to make a prediction about the strategy profiles we will see in practice.

The second requirement to make a prediction—called equilibrium selection—becomes a serious challenge in cases where, unlike the Assurance Game, the players *disagree* about which equilibrium they would like to occur. This is the case in the next game in our catalog.

CHECKPOINT 1.16 **Assurance Game** Which payoff table entries would you have to compare in order to show that:

a. Planting early is Pareto efficient.

b. Planting late is a Nash equilibrium.

c. The best response to the other planting early is to plant early.

1.13 DISAGREEMENT GAMES: CONFLICT ABOUT HOW TO COORDINATE

An example of what we call a **Disagreement Game** is illustrated in Figure 1.17. In a Disagreement Game there are two Pareto-efficient Nash equilibria and

Figure 1.17 **The Language Game: A Disagreement Game.** The players need to coordinate on an equilibrium, but each player prefers one equilibrium to the other, so there is a *conflict of interest*. If they fail to coordinate on one of the Nash equilibria because of the conflict of interest, the outcome will be a *coordination failure*.

		Ben	
		Improve Swahili	Stick to English
Aisha	Stick to Swahili	4, 2	0, 0
	Improve English	1, 1	2, 4

DISAGREEMENT GAME A game in which there are two or more Pareto-efficient Nash equilibria that are ranked differently by the players so that in the two-person case the players prefer different equilibria is a Disagreement Game (which also goes by other names).

the players are in conflict over which Nash equilibrium each prefers. But both of the players prefer both of the Nash equilibria to the *other* outcomes. The players' problem is to find a way to *coordinate* on one of the Nash equilibria to ensure that *no* coordination failure results. They do *not* want to end up with an outcome where both of them do worse than at one of the Nash equilibria.

Consider two players, a home-language Swahili speaker (Aisha) and a home-language English speaker (Ben) who have recently met. Each person can speak the other language, but prefers to speak their home language. They share many common interests but do not communicate as well as they would like. Each has two strategies: Stick to your home language or Improve the other language.

Among the possible outcomes are that he could learn better Swahili and they could routinely converse in that language; or she could learn better English and they could converse in English. They do not need to both be fluent in both languages.

So for Aisha, if Ben becomes fluent in Swahili, then her best response is not to take the time and trouble to improve her English. For Ben, similarly, if Aisha were to become fluent in English, then he would see little point in taking the Swahili courses.

The result is two Nash equilibria (Stick to Swahili, Improve Swahili) with payoffs (4, 2) and (Improve English, Stick to English) with payoffs (2, 4). The two Nash equilibria are both Pareto efficient because there are no alternative outcomes which are Pareto-superior to these strategy profiles.

But, as shown in the payoffs in Figure 1.17, Aisha would prefer the (Stick to Swahili, Improve Swahili) Nash equilibrium and Ben would prefer the (Improve English, Stick to English) Nash equilibrium.

The Disagreement Game is *similar* to the Assurance Game in that:

- There are two Nash equilibria.
- Both players do better if they coordinate (that is, speak the same language at one or the other of these equilibria).

The Disagreement Game *differs* from the Assurance Game because:

- each player in the Disagreement Game prefers one of the Nash equilibria while the second player prefers the other, while both prefer the Pareto-superior Nash equilibrium in the Assurance Game; so as a result
- the players in the Disagreement Game have a *conflict of interest* concerning which equilibrium gets selected.

HISTORY One of the first game theoretic studies of coordination problems—by David Lewis—was concerned with how we coordinate on a common language.

Disagreement Games highlight how there can be social interactions with multiple Nash equilibria, each of which is Pareto efficient, but there may be no 'middle ground' to coordinate on and as a result conflict over who gets to benefit the most is unavoidable. Both players in the Disagreement Game would both be worse off out of equilibrium than at one of the Nash equilibria in the game. They have a common interest in coordinating

somehow as opposed to *not* coordinating; but their interests conflict in *how* they coordinate.

CHECKPOINT 1.17 **Language Game** Label the outcomes of the Language Game (Figure 1.17) like we did for the Prisoners' Dilemma Game in Figure 1.13. Plot the outcomes using axes with the players' payoffs, and determine which outcomes are Nash equilibria and which are Pareto efficient.

1.14 WHY HISTORY (SOMETIMES) MATTERS

As we have seen from Disagreement Games and Assurance Games, strategic complementarities in games may give rise to more than one Nash equilibrium. When this is the case we cannot say which Nash equilibrium is our prediction of how the game will be played. The best the Nash equilibrium concept could do is to say that the outcome of the game is likely to be one of the (perhaps many) Nash equilibria.

We need more information to make a prediction. Think about the Palanpur game, and imagine that all you know is the payoff matrix (not how the farmers played the game in recent years). Though you would be on solid ground predicting that it is likely that you'd see both farmers planting either early or late, you would not have much confidence in which it would be.

But now suppose you were told that last year they planted late. Then, unless they had discovered some way to coordinate a switch to planting early, you would be correct when you predicted that they would both be planting late this year too.

When history matters in this sense, we say that outcomes are **path dependent**. When the outcome of a game is path dependent, knowing the recent history of a social interaction is valuable information to predict which equilibrium will occur. So, quite different equilibrium outcomes–poverty or affluence, for example–are possible for different groups of players with identical preferences, technologies, and resources but with different histories. This is how "history matters."

The Palanpur payoff matrix describes a **poverty trap**. A poverty trap occurs when identical people in identical settings may experience either

PATH DEPENDENCE A process is path dependent if the most likely state of something this period depends on its state in recent periods.

POVERTY TRAP A poverty trap occurs when identical people may experience either an adequate living standard or poverty, depending only on chance events of their histories. Poverty in this case is a result of a person's circumstances, not personal attributes.

an adequate living standard or poverty, depending only on chance events of their histories, for example were their parents rich or poor, or were they citizens of Norway or Nigeria. The possibility of poverty traps alerts us to the fact that people may be rich or poor not because of anything distinctive about their skills, hard work, or other personal attributes, but because of the situation they find themselves in. Poverty may be inherited as it is in Palanpur not by anything that parents pass on to children but instead by the inheritance of a common history.

The same is true about other aspects of how we interact in society, for example in the ways our lives may be highly segregated in interacting with people who differ in the groups with which they are identified, whether that be ethnicity, or religion, or even loyalty in sports teams. Segregation can be both a cause and effect of **group inequalities**, or economic differences between sets of people distinguished by some common attribute .

CHECKPOINT 1.18 Drain the meadow: name that game

a. Write down a payoff matrix for Hume's "drain the meadow" game, with the two actions open to farmers Adams and Brown being "drain" and "do not drain," and assuming that the value of the drained meadow (to each farmer) is 5, the value of the undrained meadow is 3, and if the two farmers jointly work on the draining it costs them 1 each, while if a single farmer does the draining alone it costs him 3.

b. What kind of game is this? Explain how it might be solved if there were just two farmers, and why with many farmers (as Hume wrote) it would be "difficult and indeed impossible" for them to agree on a common course of action to avoid a coordination failure.

1.15 APPLICATION: SEGREGATION AS A NASH EQUILIBRIUM AMONG PEOPLE WHO PREFER INTEGRATION

Segregated communities–whether on grounds of ethnicity, race, religion, or class–often cultivate intergroup prejudice and hostility and are the basis for systematic denial of equal dignity to all citizens. The correlates of segregation typically include systematic deprivation of adequate schooling, health facilities, personal security, and other necessities of life to a politically subordinate demographic group. Racial segregation in New York City is illustrated in Figure 1.18.

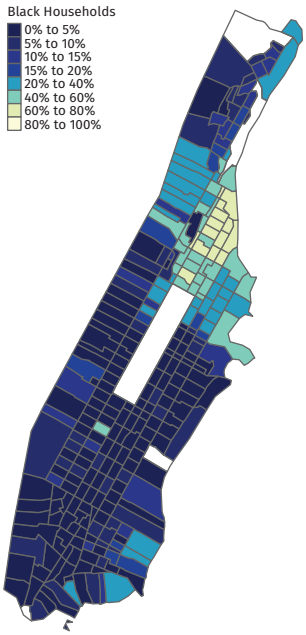


Figure 1.18 Segregation in Manhattan, the central borough of New York City in 2018. The rectangle without shading is Central Park

Source: Sethi and Somanathan (2009) using updated 2018 block data from the American Community Survey, 2018.

GROUP INEQUALITY Economic differences between sets of people distinguished by some common attribute—men and women and people of different nations, ethnic or racial groups—are called group inequalities.

✓ **FACT CHECK** Two decades into the twenty-first century in Seattle, Washington, unenforceable but still on the books "racially restrictive covenants" covering more than 20,000 homes prohibit sale or rental to particular groups. One stipulated that, "No person or persons of Asiatic, African or Negro blood, lineage, or extraction shall be permitted to occupy a portion of said property."[12]

Segregation often results from deliberate policies of discrimination by governments, banks, and homeowners. Examples are the apartheid system of enforced racial separation in South Africa that persisted until 1994 and legally mandated housing segregation in the US–the so-called racial covenants that were finally outlawed in 1968. Deliberate attempts to sustain segregated communities continue to the present; in the US for example in state laws–"single-family zoning"–that effectively make it impossible to build inexpensive housing in high-income neighborhoods.

But segregation can also result from the uncoordinated decisions of people who would actually prefer to live in integrated communities. This counterintuitive result illustrates the use of the Nash equilibrium concept.

It also underlines the lesson already learned from the interaction among the Palanpur farmers. The lesson is that there may be more than one Nash equilibrium–one Pareto superior to the other–and a society can find it difficult to escape the inferior equilibrium. The example of segregation is also a reminder–like the case of the overfishing Nash equilibrium–that the fact that an outcome is a Nash equilibrium does *not* mean that it is something that the players would choose, if they could coordinate and decide jointly on the outcome.

✓ **FACT CHECK** Between-nation group inequalities are much greater than within-nations inequalities as you can see from this interactive data set (https://jackblun.github.io/Globalinc/), which is a link to CORE global income distribution data. Another example of group inequality: the most educated member of upper-caste Brahmin households in India has on average twice the years of schooling as the most educated in lower-caste households. And in all castes women have much less schooling then men. Membership in all of these groups—castes, genders, nations—are accidents of birth.[13]

The setup of the model

Here is a model. There are two groups of people, Greens and Blues, and they live in homes arrayed around a circle representing a neighborhood. The homes are identical except that they may differ in the group identities of the immediate neighbors. The neighborhood is the circle as a whole. A household's immediate neighborhood is made up of the two households on either side of it.

If a citizen would like to live at some other location around the circle, they can switch with some other person currently occupying that position, as long as the other person is willing. The homes just change occupants with no money changing hands. We would like to know what the neighborhood will look like when all the switching that people can do has been carried out, so that the neighborhood's composition *stops changing*.

When *no citizen is able to benefit* by switching home the distribution of homes among the groups around the circle is a Nash equilibrium.

Greens and Blues have identical preferences and they care only about the group identity of their two immediate neighbors. All people in the neighborhood would prefer to have one neighbor of each group, as is shown in Figure 1.19. But they are "satisfied" as long as they either have an immediate neighbor of each group or if both are of their own group. People are "dissatisfied" if both immediate neighbors are of the *other* group An ideal neighborhood, then, is shown above in Figure 1.20 (c): Each person has one neighbor of each group.

Figure 1.19 The preferences of a household depending on the kinds of neighbors that surrounds it. Household B will either be satisfied or dissatisfied depending on the groups of neighbors they have. B's choice to play Signal Dissatisfaction or Do Nothing therefore depends on the composition of their immediate neighborhood.

(a) Household B has one neighbor of each group, their *ideal* situation, is *not dis-satisfied*, and will play Do Nothing

(b) Household B has two neighbors of their own group, is *not dis-satisfied*, and will play Do Nothing

(c) Household B has two neighbors of the other group, is *dissatisfied*, and will play Signal Dissatisfied

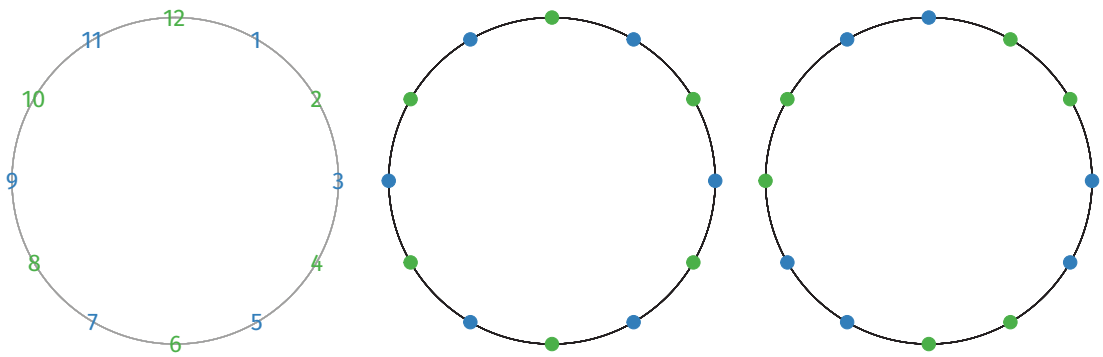
People have two strategies: "Do Nothing" or "Signal Dissatisfaction." Signaling dissatisfaction means being willing to switch positions with another person–anywhere in the neighborhood–who has also "signaled dissatisfaction." People are willing to switch only if they prefer the new location to their old location. For this reason people of either group will *never* switch with a person of the *same* group. This is because, for example, if a Green is dissatisfied with her current location and would like to move, all other Greens would be equally dissatisfied were they to take her position, so no other Green would agree to a switch. So all of the switches will be with different groups: a Green will switch with a Blue, but a Blue will never switch with a Blue and a Green will never switch with a Green. This means that switches will change three things:

- *both switchers' own immediate neighbors*: for the two who switched, they now have immediate neighbors that differ in their group identity from before the switch (this is the reason for the switch);
- *the switcher's new immediate neighborhood*: those on either side now experience having a neighbor of a new group identity given the switcher's arrival and the previously dissatisfied person's departure; and
- *the switcher's old immediate neighborhood*: those who were previously on either side of the switcher have a neighbor of a new group given the arrival of the person with whom the previously dissatisfied person switched.

A segregated Nash equilibrium

We begin with 6 Greens and 6 Blues occupying alternating positions in the 12 "houses" at the locations on the circle that are numbered as if from time on a clock (so, 12 is the top). The twelve homes on the circle are "the neighborhood." We call the assignment of different groups to the homes around the circle in Figures 1.20(b) and (c): an *allocation*. An allocation in this game is an assignment of homes to the groups at a given stage of the

Figure 1.20 The neighborhood and the citizen's ideal integrated outcome. Panel (a) is the "geography" of the neighborhood, showing that, for example, the citizen at position 2 on the circle has two immediate neighbors, the people at positions 1 and 3. Panel (b) shows that the person at position 2 is a Blue and her two immediate neighbors are both Greens is just a starting point at which the neighborhood is as integrated as possible in the sense that the two immediate neighbors of each citizen are of the *other* type. Panel (c) shows the distribution of types across locations that the citizens prefer: each citizen has one immediate neighbor of each type.



(a) The circle as a clock (b) The baseline (c) The citizens' ideal integrated neighborhood

game. The allocation before the game starts is the *initial* allocation. The allocation after the game ends is the *final* allocation.

The game proceeds as follows.

At each step, each of the 12 people plays either Do Nothing, or Signal Dissatisfaction. Their choice of a strategy is known to all other players.

Then, one of the 12 citizens is randomly selected and given the opportunity to make a switch if she wishes and can find another person willing who has also signaled dissatisfaction and is willing to make a switch.

At step 1, for example, we might ask the Green at position 10 o'clock in Figure 1.22 (a) below if she would like to switch. She would, because both of her neighbors are Blues (she signals "D" as in the figure). Whether she is able to make a switch depends on whether there are others who have chosen the strategy Signal Dissatisfaction. Because everyone else is also dissatisfied, she has many choices.

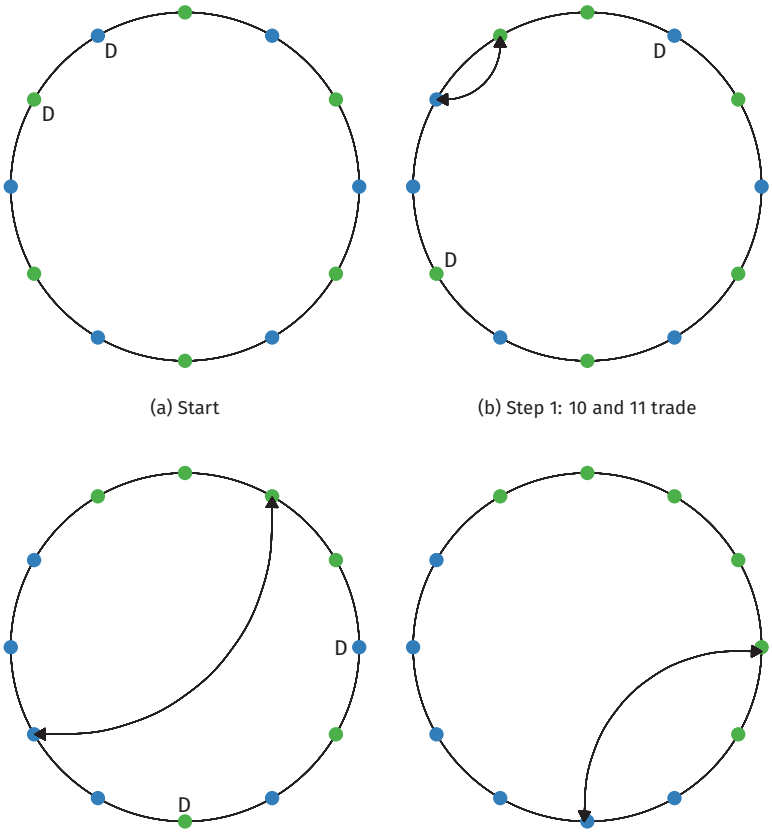
Suppose she switches with her friend and immediate neighbor, the Blue at position 11 (who is also signaling "D"), shown by the colors of position 10 and 11 changing from Panel a Start to Panel b Step 1. The two people are still friends and neighbors, but each now also has a same-group neighbor on the other side.

Suppose, next, that it is the Blue at position 1 who is picked to stay or switch. If he plays "Signal Dissatisfaction" (D), he could switch with

his friend at position 8. We continue this process until either no one is dissatisfied, or if someone is dissatisfied, there are no others playing the strategy Signal Dissatisfaction with whom a switch is possible. This process could continue as shown in the figure, resulting at the end of three steps in the completely segregated neighborhood shown in Figure 1.22 (d).

At step 4 (not shown), each of the 12 would choose the strategy Do Nothing, because 8 of them have members of their own group as neighbors

Figure 1.22 From integration to a segregated Nash equilibrium. The figure shows one of many possible progressions from an integrated non-equilibrium situation to an entirely segregated Nash equilibrium. Panel (a) shows the starting point from the previous figure. In step 1 the Green at position 10 and the Blue at position 11 switch positions, shown by the double headed arrow, and resulting in the neighborhood shown in panel (b). The remaining panels show the successive steps to the final fully segregated Nash equilibrium. A "D" next to a household's position indicates that that household is signaling dissatisfaction and will switch places with the other household signaling dissatisfaction in the next stage.



(a) Start

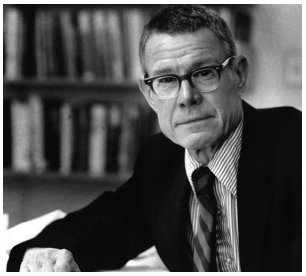
(b) Step 1: 10 and 11 trade

(c) Step 2: 1 and 8 trade

(d) Step 3: 3 and 6 trade; Complete Segregation

Figure 1.21 Thomas Schelling (1921–2016) was an American economist who won the Nobel Prize in economics in 2005 for his contributions to our understanding of conflict and cooperation in what are now called "non-market social interactions" that go beyond simple exchanges of the type typically taught in economics courses. He sought to establish "an inter-disciplinary . . . theory of bargaining . . . that could be useful to people concerned with practical problems." The model of segregation here is based on his work.[14]

Photo by Alvaro conde/wikimedia.org



EXAMPLE As a practical matter, they would not have to implement their ideal integrated solution. Taking account of some of the policies promoting segregated communities (that are not in the model), for example, they could vote to repeal "single-family zoning" regulations that, in the US, promote more homogeneous neighborhoods by limiting the kinds of housing (including lower-cost housing) that can be constructed in the neighborhood.

only and the other four have one neighbor of each group. So no one is dissatisfied. As a result, we observe no further moves: the allocation is *stationary* (meaning unchanging). It is a Nash equilibrium.

Avoiding outcomes that nobody prefers

The conclusion is *not* that complete segregation will necessarily be the result. This is true for two reasons.

- *There is also a Nash equilibrium that is integrated rather than segregated.* In Figure 1.20 (c), the allocation has each person's immediate neighborhood composed of both groups. You can confirm that, like the completely segregated allocation, this integrated allocation is also a Nash equilibrium: every citizen has their ideal immediate neighborhood, so no citizen is dissatisfied and each are best responding with Do Nothing. This allocation could have come about by the same rules of the game that resulted in complete segregation. It was just a matter of chance whether the ideal or fully segregated neighborhoods occurred. This is an example of implementing a desirable allocation within the given set of rules of the game.
- *The citizens could play the game cooperatively rather than noncooperatively.* If the citizens had realized that playing the game noncooperatively could lead them to a complete segregation outcome that nobody wanted, they could have acted cooperatively—that is jointly agreed—to implement their ideal allocation. This is an example of implementing a desirable allocation by changing the rules of the game: agreeing to act jointly was not an available strategy in the noncooperative variant of the game above.

The outcome in the segregation model shares three features with a game representing what would appear to be a very different situation: Planting in Palanpur.

- A *Pareto-inferior Nash equilibrium*: There is a Nash equilibrium—planting late and a segregated community—in which everyone is worse off than they could be at some other allocation.
- A *path-dependent outcome*: History matters because an outcome that is preferred by all participants is also a Nash equilibrium, so if the preferred outcome were to occur, it could persist.
- A *change in the rules of the game can avoid the inferior outcome*: By coordinating their actions—changing the interaction to a cooperative game—they could escape the Pareto-inferior outcome

In these three respects the two interactions—when to plant and where to live—are *not unique* or even unusual in these three respects.

CHECKPOINT 1.19 Segregation as a Nash equilibrium

a. Show that the segregated neighborhood in Figure 1.22 is a Nash equilibrium.

b. Explain why the 'ideal neighborhood' in Figure 1.20 is also a Nash equilibrium.

1.16 HOW INSTITUTIONS CAN ADDRESS COORDINATION PROBLEMS

Game theory has given us a catalogue of coordination problems: Prisoners' Dilemmas, Invisible Hand Games, Assurance Games, and Disagreement Games (there are many more!). Knowing how the structure of these games differ will help to diagnose the nature of a coordination problem and to devise policies and constitutions—changes in the rules of the game—to avoid a coordination failure.

Changing the rules of the game

This is an example of using the concept of equilibrium to understand how to change an undesirable outcome. The idea is simple: a change in the rules of the game can eliminate an undesirable Nash equilibrium, so that it is no longer our prediction of how the game will be played. Instead it may be possible to make some preferable strategy profile a Nash equilibrium which then could be the predicted outcome of the game.

A common approach to averting coordination failures in a Prisoners' Dilemma is to devise policies or institutions that transform the payoff matrix so that the game is no longer Prisoners' Dilemma. There are two possibilities to consider:

- change the Prisoners' Dilemma to an Assurance Game; or
- change the Prisoners' Dilemma to an Invisible Hand Game.

Changing the Prisoners' Dilemma game into an Assurance Game means making the mutual cooperate outcome a Nash equilibrium (it was not in the Prisoners' Dilemma) even if mutual defect remains a Nash equilibrium. In section 5.12 we show that one way this can be accomplished is to let the same players interact many times in what is called a repeated game. In this setup cooperating to restrict fishing can be sustained as a Nash equilibrium because those who defect—overfishing and exploiting the cooperation of others who fish less—can be punished in future interactions.

The second option, the one we will explore here, is more ambitious: converting the game from a Prisoners' Dilemma to an Invisible Hand Game. To see how this might work, remember that the coordination failure that results in the Prisoners' Dilemma is a consequence of the fact that players

take actions that inflict costs on others—negative external effects—that are not part of their thinking when they decide what to do.

To see that internalizing these external effects can address the coordination failure, we examine the implementation of a liability rule in the Fishermen's Dilemma. Tort is a branch of law dealing with damages inflicted by one person on another (or another's property). Among other things, tort law establishes the responsibility—called the *liability*—of the person inflicting the damages to compensate the harmed individual. The requirement to compensate the harmed individual internalizes the external effect.

Compensating for external effects by liability law

How would a liability system work in the Fishermen's Dilemma? Look again at Figure 1.14. We will let the payoff numbers measure something—like kilos of fish caught—that can be transferred from one fisherman to another.

Suppose Alfredo and Bob decided to jointly adopt "Cooperate" (Fish 10 hours) as an agreement. In their agreement, they also choose to adopt a *liability rule* requiring compensation be paid to the other party if one's actions result in lower payoffs than would have occurred had the agreement to cooperate (fish only 10 hours) been observed.

With the liability rule the following will happen:

- if both Cooperate as they have agreed, then they both get 3 as before;
- if Alfredo Defects on Bob (plays Fish 12 hours), Alfredo initially gets 4;
- but because of the liability rule, then Alfredo must compensate Bob for the costs that his defection inflicted on Bob, who got a payoff of 1 rather than 3 (the payoff Bob would have obtained had Alfredo not violated the agreement);
- so, Alfredo pays Bob 2 who ends up with 3; Alfredo ends up with 2.

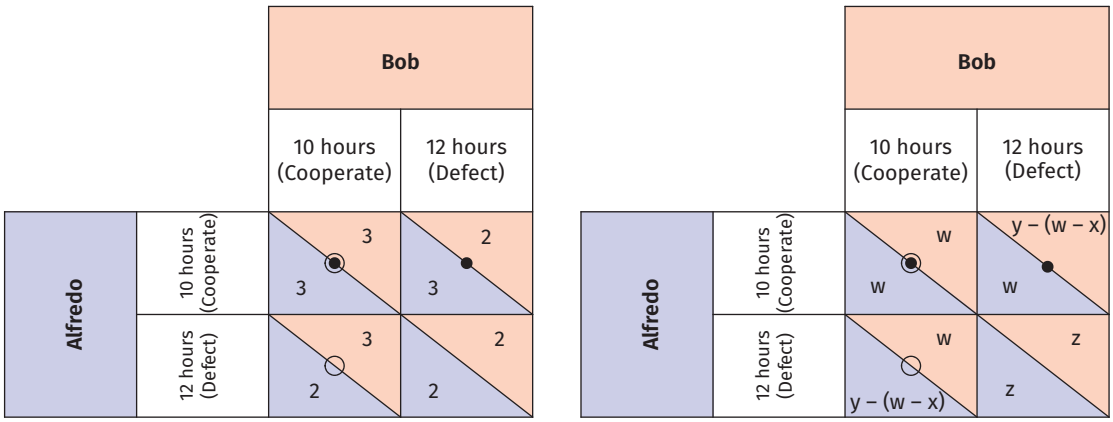
M-CHECK Because $y - w + x < w$, Cooperate is now a best response to Cooperate and (Cooperate, Cooperate) is a Nash equilibrium. Cooperate is also a best response to Defect (because $w > z$), so Cooperate is the dominant strategy with the liability rule in force, and (Cooperate, Cooperate) is the dominant strategy equilibrium. Note too that $x + y < 2w$, which is important for the liability rule to work.

We can use these changes to the payoffs to construct a transformed payoff matrix. The transformed payoff matrix for Alfredo's and Bob's payoffs is given by the entries in Figure 1.23.

Did the change in the rules of the game work? Put yourself in Alfredo's position, contemplating defecting on Bob. If he honors the agreement and fishes 10 hours, like Bob he gets 3. If he defects and fishes 12 hours he ends up with 2 after having paid Bob the compensation required by the liability rule. So Defect is no longer a best response to Bob playing Cooperate; Bob will honor the agreement. If Bob were to consider defecting on Alfredo, he would reach a similar conclusion.

From Figure 1.23 using the circle-and-dot method you can see that (Cooperate, Cooperate) is now the only Nash equilibrium. Redefining property rights—to take account of liability for damages—can implement a Pareto-efficient outcome by inducing each player to account for how his actions affect the other player. By redefining property rights to include the

Figure 1.23 Fishermen's Dilemma with a liability rule. Players can implement a desired outcome by transforming property rights using a liability rule (the harm a player does to another player is deducted from their payoff). This payoff matrix is based on Figure 1.14 modified by the liability rule. Alfredo's payoffs are listed first in the bottom-left corners and shaded blue. Alfredo's best responses are shown by the solid point. Bob's are listed second in the top-right corners and shaded red; his best responses are shown by the hollow circle.



(a) Liability rule with numbers (b) General liability

liability of the damages (external effects) that one inflicts on others, we have transformed the game to an Invisible Hand game.

CHECKPOINT 1.20 Limited liability by the numbers Use the model of the liability rule in Figure 1.23 to complete the following tasks.

a. Redraw the payoff table, but substitute in the values for x, y, w and z from Figure 1.12. (Hint: the payoffs should only be 2s and 3s).

b. Solve your new game using best response analysis (the circles-and-dots method) to find the Nash equilibrium of the game. What is it? Explain.

c. Does either player have a strictly dominant strategy? Is there a dominant strategy equilibrium? Explain.

1.17 GAME THEORY AND NASH EQUILIBRIUM: IMPORTANCE AND CAVEATS

We have started off this introduction to modern microeconomics with game theory. The reasons are that

- Many important economic relationships–in labor markets, families, credit and financial markets, between citizens and governments, among

neighbors, between nations seeking to address climate change, and many more—are strategic, and require the tools of game theory.

- Focusing on people as actors often in conflict with each other, but also sharing common interests, is essential to economics as a social science, and game theory allows us to do this.
- How these interactions work out depends on the institutions that regulate them, and game theory allows us (even requires us) to be very specific about the varieties of possible rules of the game under which we now interact, and how we might change these rules for the better.

For game theory the Nash equilibrium is a key economic idea and it provides a way to answer the question: What will be the outcome if each of the actors adopts a strategy that will not lead any other actors to change what they do?

In many situations the Nash equilibrium among players who independently pursue their individual interests provides a good prediction of what we observe in the real world. But not always. We will consider three caveats.

- *Individualism*: Overlooked opportunities for collective, not just individual best responses to the strategies of others.
- *Equilibrium selection*: The need for a method to predict outcomes when there is more than a single equilibrium.
- *Dynamics*: We are interested in what happens out of equilibrium, in part because we need to know an equilibrium will come about.

The predicament of the Palanpur farmers illustrates the first caveat. If the farmers could have *all* agreed to plant early—as would be the case in a cooperative game—then they could have easily solved their planting late problem. However, following what one of the farmers said (we quoted him in the introduction), we assumed in modeling their situation that an agreement among the entire set of players was not possible.

But in using the Nash equilibrium concept we went to the other extreme: we assumed that an outcome would be an equilibrium (meaning undisturbed by any players changing their strategy) as long as *no single individual (acting alone)* could do better by altering strategy. But perhaps two or three jointly deciding to plant early could have done better.

A more adequate equilibrium concept would take account of collective best responses where there is reason to think that small groups might be able to decide to act together even when the entire population could not jointly coordinate.

The second caveat—about equilibrium selection—is not a criticism of the Nash equilibrium concept itself. Instead, it is a reminder that the Nash equilibrium concept by itself is *insufficient* to make predictions in cases

where there are two or more Nash equilibria, as in the Planting in Palanpur Game, the Language Game, and the model of residential segregation. In these cases, knowledge of the recent past play of the game would be an important part of making predictions based on the Nash equilibrium.

We will return to the third caveat–about understanding what happens out of equilibrium–when we consider such questions as firms, decisions whether or not to enter an industry and how a group of buyers and sellers could get to a competitive equilibrium (Chapters 9 and 14).

CHECKPOINT 1.21 **Equilibrium selection** How does the Disagreement Game shown in Figure 1.17 illustrate the need for a method of equilibrium selection in order to predict the outcome of the game? Why does a similar problem not arise in the Prisoners' Dilemma?

1.18 APPLICATION: COOPERATION AND CONFLICT IN PRACTICE

If all that is needed to address a coordination failure is to require that people pay the costs that their actions impose on others then why are coordination failures so common?

Overexploitation of fisheries is an international problem that humans as a world community have failed to solve. Many overexploited fisheries will not recover for a long time. But local communities and groups of fishermen have found ways to combat overfishing, and we can learn from what they do. Many groups–from farmers to fishermen–face equivalent problems worldwide. These outcomes provide a concrete motivation to study the Prisoners' Dilemma Game and other coordination problems.

What we learn from these games is that an effective liability rule requires two things:

- the injured party or the courts have to have verifiable information (that is information sufficient to enforce the liability aspect of the property right); and
- there has to be a court or some other body willing to and capable of enforcing the contract.

When we turn from game theory to the study of real fishing communities we find that both conditions are unlikely to be met, which is why the over-exploitation of fisheries continues in many cases.

- *Limited information.* The lack of verifiable information is common in social interactions and this limits the policies that governments or private actors can design in response to the persistence of Pareto-inefficient Nash equilibria.

- *Conflicts of interest*. Governments may not have the capacity or the will to enforce the necessary policies especially in cases where doing this would impose costs on a powerful group. An example is the failure of many countries to address the problem of climate change, which is in part the result of the fossil fuel companies' opposition to putting a sufficiently high price on carbon emissions.

Fishing communities, of course, are not acting out a tragic script, as were the herders in Hardin's tale about the tragedy of the commons. They are not prisoners of the dilemma they face. Real fishermen are resourceful and seek solutions to the problem of overfishing.

- Lobstermen in the US state of Maine limit how many lobster they catch using highly local restrictions on who can set traps where (the state government provides the legal framework for this).[15]
- Turkish fishermen allocate fishing spots by lottery and then rotate the spots so the distribution is fair.[16]
- The fishing community of Kayar in Senegal adopted the rule that only one trip to the fishing grounds per day is permitted (a bit like Alfredo and Bob limiting their hours of fishing) and appointed a committee to check that the rule was being observed. They also limited the number of boxes of fish that could be offloaded by a single canoe.[17]
- Shrimp fishermen in Toyama Bay, Japan have a rule that they offload their daily catch at the same time and place, so that the size of each boat's catch would no longer be asymmetric information.[18]

These rules and practices based on small local fishing communities are often disrupted by the entry of other groups whose members are not bound by the local rules. Conflicts of interest within the local community also sometimes limit the effectiveness of attempts to limit the catch. One reason is that restrictions on fishing are often supported as a way to raise the wholesale price of fish and hence the incomes of fishing families. But fish sellers—who buy the fish wholesale at the port and then sell to local consumers—would profit if they could pay less.

The rules regulating access to fishing that we see in existence are a small selection from a much larger set of rules that people have tried out at some point. What we see are the institutions that have succeeded well enough to allow the communities using them to persist and not to abandon their rules. The persistence of such rules does not require the rule to implement a Pareto-efficient outcome; it only requires that the rule be reproduced over time by people adhering to it. By this reasoning, even if the rules of the game do not implement Pareto-efficient outcomes, we might expect a fishing community that has hit on a way of sustaining cooperation in the long run to do better in competition with groups that overfish, and that successful groups may be copied by other groups.

CHECKPOINT 1.22 Institutions and Palanpur: why history matters Supposing that the only voters involved in approving the Palanpur village council's decision to require planting early were themselves farmers. Explain why they would unanimously support the measure. What would happen if after implementing the law requiring early planting one year, the next year the law was revoked?

1.19 CONCLUSION

Early in 2021 the people of the world faced urgent questions: What are the rules that should govern the distribution of the newly available COVID-19 vaccines? Should the most vulnerable be vaccinated first, irrespective of their income or nationality? Should those deemed to be essential workers also have priority? Should governments determine the order of priority or should those willing and able to pay substantial sums have access to vaccinations first? If schools will be open for a limited number of students, whose children should have priority?

These are all questions about the right rules of the game, ones that avoid coordination failures, and also are consistent with other criteria possibly including, for example, the values stressed by the classical philosopher-economists such as liberty, dignity, and fairness.

The classical institutional challenge which we stated was "How to design institutions so that people could be left free to make their own decisions and at the same time avoid outcomes that were inferior for everyone?"

With the terms you have learned this can now be rephrased "How can social interactions be structured so to avoid Pareto-inefficient Nash equilibria resulting from people's choice of their own actions?" The Fishermen's Dilemma is an example of a challenging coordination problem because an inefficient outcome is the unique Nash equilibrium. The negative external effect of overfishing in our model is intended as an analogy for coordination problems going far beyond the lake they share. The analogy includes the external effects of burning carbon-based fuels and the resulting change in global climate or the external effects associated with the spread of a pandemic.

To study a game and its likely outcomes and also how to improve these outcomes we have proceeded in three steps:

- first, use the Nash equilibrium concept to identify one or more likely outcomes of the game;
- second, use Pareto comparisons to identify outcomes that are "worse for everyone"; and
- third, devise changes to the relevant institutions—the rules of the game—or that would shift the population to a superior Nash equilibrium either pre-existing (as in the case of Planting in Palanpur, or the segregation case) or novel (as with the transformed Prisoners' Dilemma Game).

We have illustrated the third step by a legal remedy: the introduction of tort liabilities for damages in the Prisoners' Dilemma Game so as to internalize the external effects accounting for the coordination failure.

This approach—using game theory to understand how the workings of the economy and society might be improved—draws on three foundational concepts:

- *The rules of the game* as a description of the situation in which individuals take their actions;
- *Best response* or what we call "doing the best that you can" under the circumstances defined by the rules of the game and the actions of others (elaborated further as constrained optimization in Chapter 3); and
- *Nash equilibrium* as a way of understanding, given the rules of the game, how the best responses that people take will implement some societal outcome.

These three ideas will be deployed together in each of the chapters that follow. Whether the question at hand is how best to organize a work team or business or to develop government policies to address the climate emergency and to achieve a more just economy, these ideas can be used to devise new rules of the game that improve how the economy works. Taken as a whole this toolkit conveys a simple message: better outcomes require better rules of the game.

MAKING CONNECTIONS

Institutions and the rules of the game: To predict or explain the outcome of a social interaction, it is essential to know the "rules of the game" that determine who *knows* what and when, who *gets to do* what and when and as a result who *gets* what.

Equilibrium: Equilibrium describes an outcome that will persist until some aspect of the situation is altered as a result of externally caused changes. A Nash equilibrium is a special kind of equilibrium widely used in economics.

External effects: People often take actions without considering the effects of these actions on others. The resulting external effects—positive and negative—pervade social interactions.

Pareto inefficiency of Nash equilibria: A common result of these external effects is that the outcomes of social interactions (the Nash equilibria of the games) are Pareto inefficient, meaning that opportunities for mutual gains remain unrealized.

Rents: When players interact they face opportunities for mutual benefit, or *common interest*. But this creates opportunity for rents and for a *conflict* over how the benefits resulting from the interaction will be distributed.

Policy and changes in the rules of the game: Improving property rights (such as making people legally responsible for the harms they inflict on others) can lead people to internalize external effects. Other institutions that would facilitate people making decisions to act jointly can also provide solutions to coordination problems. Policy may result in a shift to a Pareto-efficient equilibrium.

Positive feedbacks, increasing returns, and strategic complementarity: Often players' strategies are *strategic complements* due to *positive feedback* and *increasing returns*. As a result, in some social interactions there may may be many equilibria as in the Assurance Game and the Disagreement Game.

IMPORTANT IDEAS

institutional challenge	fallback	Pareto superior/inferior
coordination problem	next best alternative	Pareto efficient
player	best response (weak/strong)	(economic) rent
strategy	dominance (weak/strict)	payoff
dominant strategy equilibrium	institution	Nash equilibrium
interdependence	positive external effect	negative external effect
Prisoners' Dilemma	(non)cooperative games	tragedy of the commons
Invisible Hand Game	Assurance Game	Disagreement (Language) Game
increasing returns	strategic complement	path dependence
poverty trap	liability rule	positive feedback

CHAPTER

2

PEOPLE

PREFERENCES, BELIEFS, AND CONSTRAINTS

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it.

Adam Smith,
The Theory of Moral Sentiments (1759), Chapter 1

DOING ECONOMICS

This chapter will enable you to:

- Understand that people make decisions based on the actions open to them (*constraints*), which of these possible actions they believe they must take to bring about the outcomes (*beliefs*) they most prefer (*preferences*).
- Use this approach to analyze economic situations involving risky outcomes, bargaining, and contributing to the public good.
- Analyze sequential games and games with multiple Nash equilibria, showing how being the first mover in these games may confer advantages on a player.
- Explain the institutional challenges arising in the case of public goods and common property resources.
- Show that experiments based on this "preferences, beliefs, and constraints approach" provide evidence that people's preferences go beyond self-interest and include generosity, reciprocity, fairness, and hostility toward others.
- See how changes in the rules of the game can result in better outcomes for all.
- Understand that these other-regarding preferences are as much part of what we consider to be rational action as is self-interest.
- Give examples of the importance of social norms and culture for decision-making and economic policy-making.

2.1 INTRODUCTION: "THE CUSTOM OF THE COUNTRY"

Chicago, a city famous for its pizza, sports, jazz, and its skyline, built its fortune on the farming of the state of Illinois. Today Illinois farmers use high tech machinery and advanced business plans, some cultivating a thousand acres of land or more. But many of the farmers don't *own* the land they cultivate; they rent land and work it as a **sharecropper**. Sharecroppers are farmers—"tenants"—who pay the owners of land a share of the total harvest that they cultivate.

Figure 2.1 Farming in Illinois is big business.

Photo courtesy of the Illinois Farm Bureau.

In the mid-1990s, over half of the contracts between farmers and owners were sharecropping agreements, and in northern Illinois 95 percent of these contracts stipulated a fifty–fifty division of the crop between the owner and the sharecropper. An equal split of the crop means that a tenant on fertile land will have higher income for the same amount of effort and other inputs.[1] Because a tenant on fertile, or high-quality, land will reap a larger harvest than a tenant on low-quality land, the fifty-fifty sharecropping contract presents us with a puzzle.

Here's the puzzle: if half of the crop on poor-quality land is sufficient to attract tenants, why should the owners of high-quality land give up half of the crop to their tenants? Those tenants must be earning more than what was necessary to get them to work the owner's land. So, why don't the owners of the better land propose a tenant's share less than half, giving the tenants just enough so that they are willing to farm the land?

We would expect owners to insist on lower tenant's shares to sharecroppers on higher quality land and offer higher shares to sharecroppers on low-quality land. Because land varies in quality by small gradations, this would result in a pattern of sharecropping contracts with a range of shares depending on the land quality. But this is not what we see. Almost all of the contracts are fifty-fifty.

Illinois sharecropping contracts allow the sharecroppers on good land to receive income attributable to the superior land quality, income the owners would otherwise have received if the owners had insisted on a lower tenant's share on the high-quality land. The fifty-fifty split effectively transfers millions of dollars annually from owners to sharecroppers simply because of the fifty-fifty division. This is not a peculiarity of Illinois. Fifty-fifty is the norm in sharecropping around the world.

HISTORY In 1848 the British philosopher-economist John Stuart Mill noted the striking global pattern of equal division in cropsharing, calling it "the custom of the country" and "the universal rule."[2]

Rice cultivation in West Bengal, India during the 1970s provides another example. There, poor illiterate farmers in villages isolated by impassable roads for much of the year and lacking electronic communication eked out a

SHARECROPPER A sharecropper is a farmer who cultivates land owned by another person with whom he or she contracts to give a share (often one half) of the crop produced.

bare living on plots that average just 2 acres rather than the 1,000-acre plots farmed in Illinois. The Indian farmers shared one similarity with farmers in Illinois: the division between sharecroppers and owners was fifty-fifty in over two-thirds of the contracts.[3]

Why was the contract the same in these distant parts of the world? The short answer is that where most contracts are fifty-fifty, that particular division is a social norm, something people feel they are morally obliged to respect. The fact that around the world land owners respect a social norm that overrides their material self-interest tells us that many people are committed to acting fairly, being treated fairly, and conforming to ethical standards of appropriate behavior.

But the sharecroppers in Illinois and West Bengal, like farmers everywhere, are also trying to make a decent living, or even to become affluent. They are not simply following social norms. They carefully weigh alternative methods of cultivating their crops at the least possible cost and marketing their harvest at the highest possible price.

2.2 PREFERENCES, BELIEFS, AND CONSTRAINTS

Understanding economic behavior requires a model that takes account of what people care about (for example, the farmers' incomes, and also their desire to uphold social norms) and how, from the set of actions they are able to undertake, they adopt those actions that they think will bring about desired results. We will develop a model of economic behavior based on:

- **constraints:** the feasible set of actions, meaning actions that are open to us,
- **beliefs:** our understanding of the outcomes that will result from the actions that are open to us, and
- **preferences:** our evaluation of the outcomes that we believe will result from the actions we take.

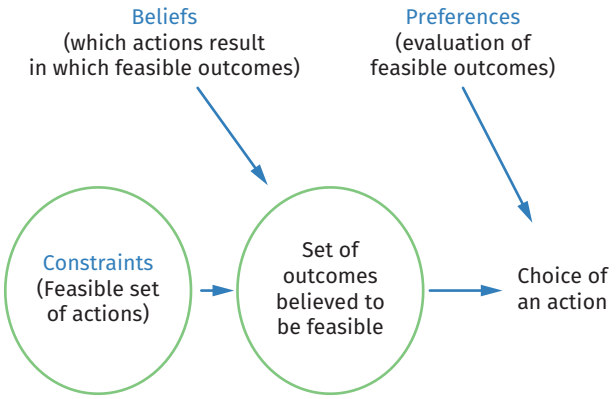
This is called the **preferences, beliefs, and constraints approach**.

The relationship between these three elements of the preferences, beliefs, and constraints approach is described below and is shown in Figure 2.2. Game theory, which you have already studied, is an important example of the preferences, beliefs, and constraints approach.

EXAMPLE The preferences, beliefs, and constraints approach is sometimes called rational choice theory or the rational actor model, but we prefer the more specific label that we use here as it identifies the three important elements making it up.

PREFERENCES, BELIEFS, AND CONSTRAINTS APPROACH According to this approach, from the feasible set (which includes all of the actions open to the person given by the economic, physical, or other constraints she faces), a person chooses the action that she believes will bring about the outcome that she values most as given by her preferences.

Figure 2.2 Preferences, beliefs, and constraints. The actor may choose from a set of feasible actions (the constraint set on the left). Combining that set with her beliefs about the outcome produced by each of the actions in the constraint set, she then has a set of outcomes that she believes are feasible, depending on her choice of an action. From all of these outcomes in the set believed to be feasible, she identifies the one that is ranked highest according to her preferences and then takes the action that she believes will bring about this outcome.



Constraints: Limits on action

From a long list of things a person might consider doing, constraints define a *more limited possible set of actions*, namely the shorter list of all of those so-called feasible actions she can carry out. In the game theory introduced in Chapter 1, the constraint was the set of possible actions, that is, a list such as "Fish 10 Hours," "Fish 12 Hours" or "Plant Early," "Plant Late."

Constraints may be imposed by personal limitations, by laws of nature, or by the force of law. A constraint can also reflect a decision by the actor to eliminate some action from the feasible set of actions on moral grounds, irrespective of the payoffs. Reasons for eliminating some actions from the feasible set include keeping promises, not harming a friend, or obeying the law.

In Table 2.1 we give examples of how the preferences, beliefs, and constraints approach can be applied. The list of feasible actions given by constraints need not be just a list of particular actions, like drive or take the bus. When marketing their output (first row of the table), the owners of a firm, for example, can set any price they like (anywhere from 0.00 by penny increments up to some very high number).

Wealth, the availability of credit, and the prices of goods impose constraints on an actor's consumption. The institution of private property also imposes limits: it means that theft is not an option for increasing your consumption. Given private property and in the absence of gifts or transfers from a government, the total amount of goods and services you

Table 2.1 Applications of the preferences, beliefs, and constraints framework. Real choice situations are typically not as simple as shown below. The urban resident, for example, may care both about travel time to work and his carbon footprint.

Actor	Constraints (feasible set of actions)	Beliefs (information about which actions will result in the preferred state)	Feasible outcomes (states that could result from the actions)	Preferences (ranking of all outcomes)
Firm owner	High or low prices	The demand curve (how quantity depends on price)	Various levels of profits	Maximize profits
Urban resident	Drive or take the bus	How many others will drive	Travel time	Minimize travel time
Ordering a meal	The menu; your budget	Simple: just order the best you can pay for	Meal quality, money left over	Maximize payoffs (choose the meal you rank highest, taking account of the cost)

can consume is limited by your wealth and how much you can borrow. So when we study someone's consumption, their *budget constraint* is a critical factor as people have a certain budget determined by wealth, access to credit, and prices all limiting how much they can buy.

Beliefs: Translating actions into outcomes

EXAMPLE The word "belief" is often used to refer to spiritual matters ("religious beliefs"); but in game theory a belief is a statement about how the world works, namely what action is required to bring about some particular outcome.

Beliefs are a person's understanding of the outcomes that her actions will bring about.

In many cases what I must do to get the outcome that I prefer depends on what other people do. I would like to spend the evening with friends, but where I should go to make it happen depends on where I think my friends will go. Given that I cannot communicate with my friends (the batteries in their phones have run out), my action (where I will go) will therefore depend on my belief about where I will find my friends.

In Table 2.1 the owners of firms are not constrained to set any particular price, but if they want to translate their choice of a price into what they care about—profits—they must form an opinion about the number of units they will be able to sell at each price. This is the demand curve, and it expresses

BELIEFS Beliefs are an individual's understandings of the relationship between an action she may take and the outcome of the action.

the owner's beliefs about the relationship between their action (the price) and an outcome (how many goods they will sell).

Preferences: Reasons for preferring one outcome over another

Preferences are evaluations of outcomes that provide motives for actions. A person's preferences are the reason why she takes the action that she believes will bring about the outcome that is better than or at least as good as the others. In Chapter 1, preferences were represented by the payoffs in games that people played. For each player, a strategy profile was associated with a number—her payoff—and players chose actions that they believed would result in the strategy profile with their most preferred (highest payoff) outcome.

In many games, preferences are represented by money payoffs. But, more broadly, preferences represent the favorable (positive) or unfavorable (negative) feelings a person has about an outcome that leads them to try to make an outcome happen (high payoff) or that leads them to try to avoid an outcome (low payoff). Preferences include:

- tastes (food likes and dislikes, for example);
- habits (even addictions);
- emotions (such as disgust and anger) often associated with visceral reactions (such as nausea or an elevated heart rate);
- social norms (for example, preferences that induce people to prefer to be honest or fair); and
- psychological tendencies (for aggression, extroversion, and the like).

The difference between preferences and beliefs is simple. A preference says: I like the outcome X more than the outcome Y. A belief says: I believe I can get X to happen if I do some action Q.

Self-regarding and other-regarding preferences

A feature of the preferences, beliefs, and constraints approach is that it allows us to model choices based on the entire range of preferences whether they be entirely self-regarding, caring for others (wishing them well or wishing to harm them), reflecting religious commitments, or any of the other reasons we may have to value some outcome more than another.

EXAMPLE While most widely used in economics, the preferences, beliefs, and constraints approach is also used in political science, for example, to understand the strategies followed by elected officials seeking to maximize their chances of re-election, in law to design criminal or civil penalties to effectively deter illegal activity, and even in biology to study the evolution of genes, modeled as if they are "trying to" increase their numbers.

PREFERENCES Preferences are evaluations of outcomes of one's actions that provide motives for taking one course of action over another.

HISTORY In 1977 Amartya Sen wrote "Rational Fools" in which he pointed out that the preferences beliefs and constraint approach ignores the importance of promises, what he called commitments. The reason is that the approach seeks to explain behavior entirely on the basis of the actor's anticipation of what her actions will bring about in the future. Honoring a past commitment—not because she would otherwise feel guilty in the future, but because it is the right thing to do—cannot be modeled in the preferences, beliefs, and constraints approach.

A key distinction about our preferences is whether in evaluating the results that we believe our actions will bring about (the right-hand part of Figure 2.2) we think about the results that we ourselves experience only, or whether we also consider the results that are experienced by others. This gives us two categories of preferences:

- If we think only about the results experienced by ourselves, we have **self-regarding preferences**.
- If we *also* think about the results experienced by others, then we have **other-regarding preferences**.

An example will clarify why we introduce these two terms.

Abraham Lincoln is said to have remarked: "When I do good, I feel good. When I do bad, I feel bad. That is my religion." Does this mean that Lincoln's "good" acts were in fact self-regarding because they made him feel "good?"[4] That does not follow. He had other-regarding preferences leading him to act differently than if he cared only about the outcomes that he personally experienced.

In the preferences, beliefs, and constraints model *all* actions are motivated by preferences, so doing a preferred thing cannot be termed "selfish" without making all behavior selfish by definition. That is why we use the term self-regarding rather than "self-interested" or "selfish."

For example, if you (like Lincoln) enjoy helping others, and you act on these preferences, does this mean you are selfish (because, for example, that's what gives you a sense of leading a good life)? No, it does not. You are acting on your preferences, but they are other-regarding because you enjoy trying to make the results that others experience be what they would want. Do not think that 'other-regarding' means "good" or "admirable." Other-regarding preferences include feelings of altruism toward others, but they also include negative feelings about others, such as envy, spite, racism, and homophobia.

In sections 2.10, 2.13, and 2.11 we provide some evidence from experiments about other-regarding preferences and how common they are across the world.

OTHER-REGARDING PREFERENCES A person with other-regarding preferences, when evaluating the outcomes of her actions, takes into account their effects on the outcomes experienced by others as well as the outcomes she will experience.

SELF-REGARDING PREFERENCES When choosing an action, a self-regarding actor considers only the effect of her actions on the outcomes experienced by the actor, not outcomes experienced by others.

"Rationality"

The term **rational** in economics means acting on the basis of:

- **Complete preferences** This means that for any pair of possible outcomes that a person's actions may bring about, A and B, it is the case that the person prefers A to B, or B to A, or is **indifferent** between the two (A and B are equally preferred).
- **Consistent preferences** If an individual with consistent (also called *transitive*) preferences prefers a bundle of goods A to another bundle B, and bundle B to a third bundle, C, they cannot prefer C to A.

A person with complete preferences, which requires only that she can rank all pairs of outcomes, might nonetheless violate the consistency assumption. So she could prefer A to B, B to C, and C to A. All that matters for completeness is that she can rank each pair. But rationality requires both completeness and consistency.

In the heading at the start of this section, we put quotation marks around rationality to underline the difference between how economists use the term and how it is generally used, that is, to mean "based on reason." But in economics, as you can see from the above definition, it means something entirely different.

- *Rationality does not say anything about what it is that the person values*: A completely generous and ethical person is rational as long as her preferences are consistent and complete.
- *Rationality does not mean being intelligent or well informed*: The beliefs that (along with preferences) determine the choices a person makes need not be true.

Moreover, people with incomplete preferences would hardly be called "irrational" in the ordinary meaning of that term, meaning "not logical" or "unreasonable." Ask yourself if your preferences are complete for the following outcomes: express preference or indifference over which of your two dearest friends will be tortured to death. If you were to say "I cannot

EXAMPLE Preferences are not complete if there is some other pair, say *A* and *D* for which none of the above three comparisons can be made. For example, if you ask someone to choose one of the three statements "I prefer A to D," "I prefer D to A," and "I am indifferent between A and D" the person responds "none of the above."

RATIONAL A rational person has complete and consistent (transitive) preferences and can therefore rank all of the outcomes that their actions may bring about in a consistent fashion.

COMPLETE PREFERENCES Complete preferences specify for any pair of possible outcomes that a person's actions may bring about, A and B, that A is preferred to B, B is preferred to A, or they are equally preferred.

INDIFFERENCE When a person is indifferent between two outcomes, they do not prefer one over the other.

rank those two outcomes, nor am I indifferent between them" you would not be "rational" by the economic definition, but nobody would think your behavior was unreasonable either. We might be more inclined to worry about the person who *would* be able to make such a ranking.

CHECKPOINT 2.1 **Why beliefs matter** Consider the coordination problems studied in Chapter 1.

a. Explain why in the Assurance Game representing planting in Palanpur, the action taken by each farmer depends on what they believe the other farmer will do.

b. In the same game explain why the farmer who believes most other farmers will plant late, will also plant late.

c. Explain why Ben's belief about what Aisha will do matters for how he will play in the Disagreement Game.

d. Are there any games you have learned so far in which beliefs about what the other does did not affect the outcome of the game?

Figure 2.3 **Amartya Sen**, born in 1933 in the then British colony of India and a Nobel Laureate in economics, has combined economic and philosophical reasoning to propose a new view of human well-being based not on what we *have* but what we *can do*—our capabilities. His book *Development as Freedom* makes the case that the success of an economy should be judged by the scope of the real choices that are open to all. His essay "More than 100 million women are missing" documented the differential mortality of males and females—especially as infants—associated with limited rights and power of women.

Photo © Norman McBeath/National Portrait Gallery, London.

2.3 TAKING RISKS: PAYOFFS AND PROBABILITIES

Beliefs become especially important in cases where we have to take some action without knowing for sure what the outcome will be. You make many of this kind of choice every day, from the important choices of what to study at university, to more trivial choices like whether to take an umbrella to class. The theory of decision-making in these cases rests on the idea that the evaluation of how good a course of action is depends on:

- how much the decision maker values each of the possible but uncertain outcomes of the action, and
- the decision maker's beliefs about how likely each outcome is.

Here we introduce a basic concept for decision-making with **risk**—expected payoffs—that will be used throughout the book. In Chapter 13 we return to the topic of risk including preferences about taking risks and the value of insurance.

CONSISTENT PREFERENCES Preferences are consistent if whenever an individual prefers a bundle of goods A to another bundle B, and bundle B to a third bundle, C, they cannot prefer C to A. Consistent preferences are also known as transitive preferences.

The value of uncertain outcomes: Expected payoffs

There are two possible but uncertain outcomes of the action "take an umbrella to class," namely, "keep dry walking home in the rain" and "carry the umbrella to and from class without even opening it, because it does not rain." The feasible actions of the decision-maker are just: take the umbrella or not.

According to the preferences, beliefs, and constraints approach, the decision maker assigns numbers indicating how much she values each of the possible four outcomes shown in Table 2.2. These numbers give the ranking of the four possible outcomes: (Don't take the umbrella, No rain) is better than (Take the umbrella, Rain) and so on. But if they are to provide a framework for making a decision when you do not know for sure if it is going to rain or not, the numbers have to be more than a ranking. They have to indicate *how much* the actor values each of the possible four outcomes. So for example taking the umbrella when it rains is five times better than *not* taking the umbrella when it rains.

We call these numbers the *payoffs* to each of the four possible outcomes.

Table 2.2 Two contingencies (Rain or No rain) and two actions (Take the umbrella, or Don't). The payoffs correspond to the coincidence of an action and a contingency, so Anoushka receives 15 if she plays Take the umbrella when the contingency is *Rain*, and she receives 8 if she plays Take the umbrella and the contingency is *No rain*.

		Uncertain event (contingency)	
		Rain	No rain
Action	Take	15	8
	Don't take	3	20

The likelihood of uncertain outcomes: Beliefs

Only one of these two uncertain events will occur. Whether, at the end of the day, it turned out to have been a good idea to have brought the umbrella is said to be *contingent* on (meaning: depends on) whether it rains or not. The payoff to the two actions in this case is said to depend on a **contingency**. The contingency in this case is whether or not it rains, and the payoff to taking the umbrella is contingent on (depends on) its occurrence.

When you decide what to study at university before knowing what kind of work you'll do after, you're making choices about contingencies too: do you go risky and study drama, or do you go safe and do accounting? In this case, the contingencies include the uncertainty about how good you will be at the field you choose and your chance of getting a job in your field. We return to risky choices about education in Chapter 13.

CONTINGENCY A contingency is a state of the world that may or may not happen and that affects the payoff to some action.

PROBABILITY DISTRIBUTION A probability distribution for n contingent outcomes of a decision is a list of non-negative numbers $\{P_1, P_2, \ldots, P_n\}$ that add up to 1. These probabilities express the decision maker's belief about the likelihood that each of the n contingent outcomes will occur.

RISK The term risk is conventionally used in economics to describe situations in which payoffs depend on contingencies, and the probabilities of each contingency occurring are known.

HISTORY In 1947 John von Neumann and Oskar Morgenstern showed a way to think about how people can evaluate risky choices. When the outcome of an action is subject to a risky contingency, how much we value an action that we can take can be expressed as a weighted sum of how much we value the alternative outcomes of our actions (depending on the realization of the contingency). The weights in the sum are the probability of each outcome occurring *if* we take the action.

The theory of decision-making about risky outcomes concerns a decision maker, call her Anoushka, who has beliefs about the probabilities (P_i) that each of the contingencies $i = 1, \ldots, n$ will occur. Her beliefs can be based on observation, on empirical studies, guessing, experience, or superstition. They need not be correct.

For simplicity we consider contingencies with just two outcomes (like "it rains" or "it does not rain" above). The basic principles of decision-making are the same no matter how many contingent outcomes there are. In this case, we use the symbol P for the probability the contingency occurs, understanding that $1 - P$ is the probability the contingency does not occur.

The decision rule: Maximize expected payoffs

Often we must take an action prior to the realization of the contingency, that is, *before* the contingency happens. But you have to make a choice anyway.

To take account of the "action now, contingency later" aspect of the decision problem we distinguish between:

- *Expected payoff*: how much the actor values taking the action given her beliefs about the probability that the various contingencies will occur and
- *Realized payoff*: how much she values the outcome that actually happens, that is, after a contingency has been realized ("realized" here means really happened, or actually occurring).

EXAMPLE If Anoushka's payoffs for the four possible outcomes of her actions are as in Table 2.2, and the probability of rain in the afternoon as 0.6, her system of contingent payoffs for taking the umbrella is $((15, 0.6), (8, 0.4))$. These numbers can be interpreted as follows: since there is a 60% chance of rain, Anoushka has a 60% chance of receiving a payoff of 15 if she takes the umbrella. Further, this means that there is a 40% chance of no rain; therefore, if Anoushka takes the umbrella she has a 40% chance of having a payoff of 8.

The **expected payoff** of an action is the basis for her choosing one course of action over another: Anoushka chooses the action with the highest expected payoff. Here is how she can calculate expected payoffs.

For each contingency, i, and each action she can take, x, Anoushka knows the payoff of taking action x conditional on i happening, which we write as $\pi(x|i)$. For example, if i is the contingency of rain in the afternoon, and x is the action of taking her umbrella with her in the morning, then her realized payoff is $\pi(x|i)$ associated with her having the umbrella when it rains. The vertical line | is read "conditional on," or "given," so $\pi(\text{umbrella}|\text{rain})$ is Anoushka's payoff to having the umbrella (x) conditional on (|), or given, rain (i) in the afternoon. For a contingency with two outcomes—numbered 1 and 2—we have to consider only two payoffs and the corresponding probabilities of each, $(\pi(x|1), P), (\pi(x|2), 1 - P)$. The *expected payoff* to an action x given a list of contingent payoffs is the weighted average of the payoffs for each contingency where the weights are the decision maker's belief about the probability of each contingency being realized. We abbreviate the expected

EXPECTED PAYOFF In a situation of risk, the expected payoff to an action is the sum of payoffs occurring under each contingency multiplied by the probabilities that each contingency occurs.

payoff to choosing x given the probabilities (P) of contingencies 1 and 2 being realized as $E(\pi_x, P) = E((\pi(x|1), P), (\pi(x|2), 1-P))$:

$$\text{Expected Payoff} \qquad E(\pi_x, P) = P\pi(x|1) + (1-P)\pi(x|2) \tag{2.1}$$

Equation 2.1 expresses the fact that the greater the probability of an outcome, the greater its weight in the weighted average calculated by the expected payoff. For example, using the values from Table 2.2, Anoushka's expected payoff to taking the umbrella, assuming the probability of rain, $P = 0.6$, would be $0.6 \cdot (15) + 0.4 \cdot (8) = 9 + 3.2 = 12.2$, that is, closer to 15 than to 8 because the probability of rain is greater than one half.

Calculating expected payoffs with probabilities is essential to understanding strategic interactions, such as the games we introduced in Chapter 1. But in games—that is strategic interactions with other people—the contingencies include the strategies chosen by the other player, not just things like whether it rains.

CHECKPOINT 2.2 Basis of probability assessments

a. Imagine that you are rolling two six-sided dice with sides corresponding to one of each of the numbers 1, 2, 3, 4, 5, and 6. You calculate the sum each time you roll the two dice simultaneously, for example, 1 + 2 = 3. Explain why the probability of getting a total of 7 from rolling the two dice is 1/6.

b. What is the expected payoff if you get paid $5 for rolling a sum of 6 or 8 on a roll of the two dice and $0 otherwise?

c. Go back to Table 2.2: What would Anoushka's expected payoff to *not* taking the umbrella be given the probability of rain being $P = 0.6$?

2.4 EXPECTED PAYOFFS AND THE PERSISTENCE OF POVERTY

In games like the Prisoners' Dilemma which have a dominant strategy equilibrium, the action that will maximize your payoffs does not depend on what the other player does, so it does not matter that you do not know what the other will do.

! REMINDER A player's *dominant strategy* is one that is a best response for all of the other players' strategies; a *dominant strategy equilibrium* is a strategy profile in which all players play a dominant strategy.

But if—like in most games—you are a player who does not have a dominant strategy, your best response will depend on what the others do. We need to take account of this in our decision-making rule. We can use expected payoffs to understand the choice of which strategy to play in an Assurance Game, like a farmer's choice between Planting Early or Planting Late in the Planting in Palanpur Game.

The game is shown in Figure 2.4 to remind you of the game's structure. The payoffs in each cell indicate how much the farmer values the outcome resulting from the strategy profile given by the particular row and column.

Figure 2.4 Planting in Palanpur: an Assurance Game. Aram's payoffs are listed in the blue bottom-left corner. Bina's payoffs are listed in the red top-right corner. Aram's best response to Bina's choice of strategy is indicated by a black dot in the relevant cell, while Bina's best responses are indicated by hollow circles. The Plant Early Nash equilibrium is Pareto efficient. The Plant Late equilibrium is not.

		Bina	
		Early	Late
Aram	Early	4 / 4	3 / 0
	Late	0 / 3	2 / 2

As you can see from the circles and dots, the game has two Nash equilibria: (Early, Early) and (Late, Late). Comparing the payoffs at the two Nash equilibria you also see that (Early, Early) is Pareto superior to (Late, Late) because (4,4) is better for both than (2,2). But recall that the actual Palanpur farmers plant late.

To see why this occurs, place yourself in the situation of one of the farmers: you will consider what some other farmer will do as a contingency, with P the probability that she will plant early. A farmer believing with probability P that the *other* farmer will plant early and probability $(1-P)$ that the other farmer will plant late is an example of decision-making under risk, since the farmer assigns probabilities to a contingency. In this case, the contingency is the other farmer's behavior.

We do not explore where these beliefs about probabilities come from, but we can imagine that the farmer will form beliefs based on what other farmers tell him or on the basis of their behavior in past planting seasons. We will include just Aram and Bina in the game, but remember we use only two players to simplify our analysis of what is really a much larger population of many people like Aram and Bina.

If Aram believes that the probability of Bina planting early is P we can construct his expected payoffs to each of his strategies, each part of which is shown by Figure 2.5. In the equations below we use a "hat" on a variable to mean "expected," so $\hat{\pi}$ reads "π hat." Using these probabilities, Aram's expected payoff ($E(\pi)$ or $\hat{\pi}$) to playing Plant Early is:

$$\hat{\pi}_{Early} = \hat{\pi}(\text{Plant Early}) = P\pi(\text{Plant Early}|\text{Bina plays Plant Early}) + (1-P)\pi(\text{Plant Early}|\text{Bina plays Plant Early})$$

Figure 2.5 Aram's view of Planting in Palanpur. The figure shows Aram's payoffs only and his belief about the probability that Bina will play her two strategies: Plant Early with probability P and Plant Late with probability $1-P$. In any given row Aram's payoffs in a cell are multiplied by the probability that Bina plays the strategy given by the column that the cell is in. We can calculate Aram's expected payoffs if he plants early by adding the payoff in each cell multiplied by the probability that he will receive that payoff if he plants early. So we have: Plant Early: $\hat{\pi}_{Early} = 4 \cdot P + 0 \cdot (1-P)$. And similarly for the other strategy: Plant Late: $\hat{\pi}_{Late} = 3 \cdot P + 2 \cdot (1-P)$.

		Bina	
		Early (P)	Late (1–P)
Aram	Early	4P	0(1–P)
	Late	3P	2(1–P)

Aram's expected payoff to planting *late* is:

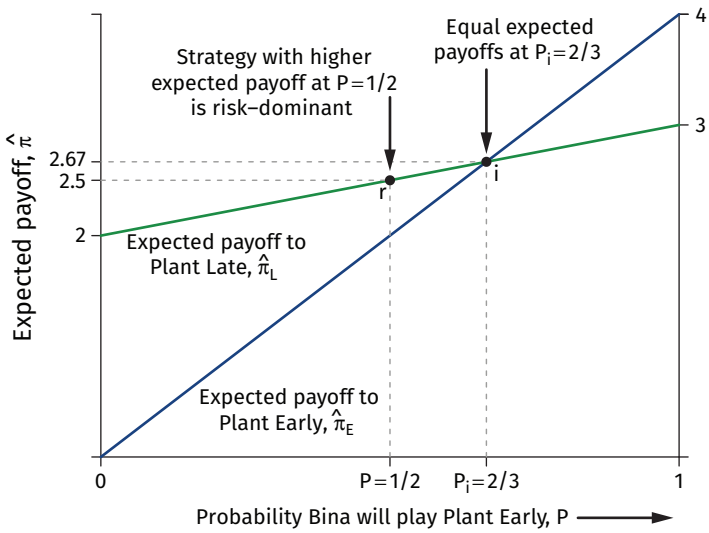
$$\hat{\pi}_{Late} = \hat{\pi}(\text{Plant Late}) = P\pi(\text{Plant Late}|\text{Bina plays Plant Early}) + (1-P)\pi(\text{Plant Late}|\text{Bina plays Plant Late})$$

An expected payoff-maximizing farmer will choose to plant early or late depending on which expected payoff is higher. As Figure 2.6 shows, for Aram, which action this will be depends on the probability that he thinks Bina will play Plant Early. The vertical axis is the expected payoff to each strategy: Plant Early or Plant Late. The horizontal axis is the probability, P, that Aram attributes to the contingency that Bina plants early: from left to right P goes from $P = 0$ (the Bina plants late with certainty) to $P = 1$ (the Bina plants early with certainty).

The two upward-sloping lines plot Aram's expected payoffs to the two strategies, Plant Early and Plant Late, showing how these depend on his belief about the probability that Bina will play Plant Early (that is, for each value of P).

The blue line graphs the equation for the expected payoff to Aram playing the strategy Plant Early which (repeating it from above) is $\hat{\pi}_{Early}(P) = P \cdot 4 + (1-P) \cdot 0 = 4P$. When the probability the other farmer plants early is zero, i.e. $P = 0$, the payoff to Plant Early is zero. When the probability the other farmer will Plant Early is 1, i.e. $P = 1$, the payoff to Plant Early is 4.

Figure 2.6 Aram's expected payoffs to playing Plant Early or Plant Late depend on his belief about the probability that Bina will Plant Early. Aram evaluated the expected payoffs to his strategies based on the probability that Bina will play Plant Early. The indifference probability where the two strategies have the same expected payoff is $P_i = \frac{2}{3}$, and the payoff to playing Plant Late is greater than the payoff to playing Plant Early for $P = \frac{1}{2}$. The intercepts of the vertical axes are the payoffs in the payoff matrix for the planting game in Chapter 1 (Figure 2.4).



We can draw the expected payoff line for Plant Late in the same way, where the expected payoff to Plant Late is $\hat{\pi}_{Late}(P) = P \cdot 3 + (1-P) \cdot 2 = 2 + P$ depicted in green, and where $\pi_{Late}(P=0) = 2$ and $\pi_{Late}(P=1) = 3$. We can then interpret the expected payoffs as follows:

- Plant Late provides a higher expected payoff when $P < \frac{2}{3}$.
- Plant Early provides a higher expected payoff when $P > \frac{2}{3}$.
- The expected payoffs to the strategies are equal at the indifference probability $P_i = \frac{2}{3}$ (where a farmer is indifferent between Plant Early and Plant Late).

The result is that Aram will choose Plant Late as long as he believes that the probability that Bina will Plant Early is less than two-thirds. Bina, facing the identical situation, has the same decision rule: Plant Late unless you think that Aram is going to Plant Early with a probability of at least two-thirds.

They will remain poor even though, had they somehow started off both planting early, they would both have had twice the payoff (4 rather than 2). The poverty trap in which they find themselves is not the result of rudimentary technology or infertile soil. What they lack is the "social

technology" that would allow them to coordinate on the Pareto-superior strategy profile, planting early. Their poverty is due to the rules of the game, which make coordination difficult.

CHECKPOINT 2.3 **A change in payoffs** Redraw Figure 2.6 to represent a new situation in which the payoff to playing Plant Late when the other farmer plays Plant Late is 1 and the payoff to Plant Late when the other plays Plant Early is 2. Explain why in this new situation if a player believes that the other is equally likely to playing Plant Late or Plant Early, then this person's expected payoffs to playing Plant Late or Plant Early are equal.

2.5 DECISION-MAKING UNDER UNCERTAINTY: RISK-DOMINANCE

So far we have assumed that Aram and Bina have some idea (maybe a guess) of the likelihood that the other would Plant Early. They faced risk (they had some information on the probability of the contingent event), but not **uncertainty** (no information at all). Decision-making under uncertainty is especially important in the field of climate change, where there are some contingencies for which there is no way to assign probabilities of their occurrence. We can explore uncertainty by continuing with the Palanpur farmers, but under slightly altered assumptions.

What is the farmer facing uncertainty to do? Economics does not have a very good answer.

HISTORY The "principle of insufficient reason" due to the Swiss mathematician Jakob Bernoulli (1655–1705) states that if we have no information on which to estimate the probability that one of two contingencies will occur, we should consider them to be equally likely. Not everyone finds this satisfactory. John Maynard Keynes found it "paradoxical and even contradictory."[5]

A two-person risk-dominant equilibrium

Economists often use what is called the "principle of insufficient reason" when a player has no information on which to place a probability on some contingency. This principle holds that the farmer who has no information on likely strategy choice of his neighbor will assign equal probability to the two events and hence use the probability $P = \frac{1}{2}$ that the other will Plant Early. What is termed the **risk-dominant strategy** is that strategy which yields the highest expected payoff when a player attributes equal probability to the two actions of the other player.

UNCERTAINTY The term uncertainty describes situations where the decision maker does not know and cannot learn the probabilities of the contingencies affecting their payoffs.

RISK-DOMINANT STRATEGY The strategy in a 2×2 game that yields the highest expected payoff when the player attributes equal probability to the two actions of the other player.

In the Planting in Palanpur Game, you can see from Figure 2.6 that a farmer who assigns the probability $P = \frac{1}{2}$ to the contingency that the other farmer will Plant Early will himself Plant Late (the green Plant Late expected profits line is above the blue Plant Early expected profits line). His expected payoffs are $2 = \frac{1}{2} \cdot 4$ for Plant Early, and $2.5 = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 2$ for Plant Late.

Plant Late is therefore the risk-dominant strategy, that is, the strategy that maximizes the farmer's expected payoffs when $P = \frac{1}{2}$. You can confirm this by going back to Figure 2.6: at $P = \frac{1}{2}$ the green line (expected payoff to playing Plant Late) is above the blue line (expected payoff to playing Plant Early). Because this is true for the other farmer as well, both farmers playing Plant Late is the *risk-dominant equilibrium*.

Plant Late in the Planting in Palanpur Game is risk dominant because planting early when the other plants late is much worse (you get zero rather than the payoff of two you would have received had you also planted late) than planting late when the other plants early (you get three rather than the four you would have received had you also planted early).

CHECKPOINT 2.4 Risk dominance and the worst-case outcome

a. Redraw the expected payoff line for planting early with the payoff to playing Plant Early when the other farmer plays Plant Late to be even worse than shown in Figure 2.6, e.g. −2 instead of 0.

b. In this case what is the indifference probability?

c. What is the least payoff to planting early when the other farmer plays Plant Late that would make playing Plant Late no longer risk dominant?

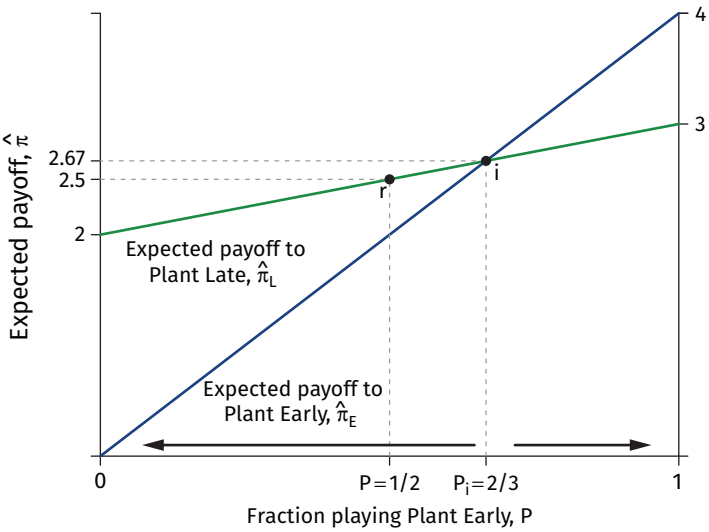
A risk-dominant equilibrium in a large population

Instead of thinking about only two farmers, we can interpret the model as portraying a population of farmers in a village like Palanpur itself. Like Aram and Bina, the farmers face a multiplayer coordination problem: doing well if they all Plant Early and doing poorly if they all Plant Late. They are currently stuck in the poor equilibrium

We can repurpose Figure 2.6 such that the horizontal axis is the fraction of the population going from 0 to 1 who choose Plant Early, P (reading left to right) as shown in Figure 2.7. The payoff lines in the figure have the same interpretation as before: They are the expected payoffs for any one of the large number of identical farmers in the village. The probabilities translate to population fractions too:

- $P < \frac{2}{3}$: When less than two-thirds of the population choose Plant Early (i.e. more than one-third play Plant Late), the Plant Late strategy has a higher expected payoff. So if last year $P < \frac{2}{3}$, then *all* farmers whether they planted early or late, will reason that they should play Plant Late this year. If they do that, then all of the farmers will end up with a payoff of 2.

Figure 2.7 Fraction of farmers planting early. P is the fraction of farmers playing Plant Early. $1-P$ is the fraction of farmers playing Plant Late. In the case of the population as a whole, the indifference probability (or in the case of a population of players, the indifference fraction) shown at point **i** with fraction P_i corresponds to the fraction of the population at which the players are indifferent between the strategies (Plant Early or Plant Late). In the case of the whole population, point **i** is also the *tipping point*: when a fraction of the population less than P_i plays Plant Early all farmers will want to play Plant Late; when a fraction of the population greater than P_i plays Plant Early all of the farmers will want to play Plant Early. The arrows indicate this movement to the extremes of $P=0$ or $P=1$.



- $P > \frac{2}{3}$: When more than two-thirds of the population select Plant Early (i.e. less than one-third select Plant Late), the Plant Early strategy has a higher expected payoff. At any fraction $P > \frac{2}{3}$, *all* farmers will do better by choosing Plant Early. If they do so, they will end up with a payoff of 4.
- $P = \frac{2}{3}$: At two-thirds playing Plant Early and one-third playing Plant Late, the expected payoffs are equal. The point at which the expected payoffs are equal is a **tipping point** as a small change in the fraction planting early will drive all players to adopt one or the other strategy: Plant Early or Plant Late.

Now imagine that, as in the village of Palanpur, virtually all of the farmers have been planting late year after year (maybe even generation after gen-

TIPPING POINT An unstable equilibrium at the boundary between two regions characterized by distinct movements in some variable.

eration). There would not be much uncertainty about what fraction of the population would Plant Late the next planting season. Each of the farmers would hold the belief that P is close to zero and as a result they all would Plant Late, confirming their beliefs. The belief that almost nobody would Plant Early sustains *both* the low income of the farmers, and the belief itself, which year after year turns out to be correct.

Why does this occur? In the Fishermen's Dilemma the best outcome for one of the players is the worst for the other, so there is a conflict of interest between the two. And this contributes to the difficulty of finding some way of coordinating so as to avoid over-exploitation of the fishing stock.

This is not the problem in the Assurance Game. There is no conflict of interest: All of the Palanpur farmers prefer the outcome when they all Plant Early to any other outcome. Their failure to implement the mutually desired outcome is the result of their inability to coordinate on planting early rather than late.

Is there a way things could have turned out better for the farmers? What may seem to be a minor tweak to the rules of the game under which the farmers are interacting can help them escape their poverty trap.

2.6 LEADERSHIP IN SEQUENTIAL GAMES: WHEN ORDER OF PLAY MATTERS

The game we introduced to model the coordination problem facing Aram and Bina was unlike many real-world social interactions, they were total strangers who had no way of coordinating their actions, and they acted simultaneously (or, at least, without knowledge of what the other had done).

But it might be that rather than playing simultaneously, they play sequentially. Playing sequentially is a change in the rules of the game; it represents a change in the institutions governing their interaction. We will see that this seemingly small change makes it an entirely different kind of game possibly even allowing Pareto-efficient outcomes.

To see how this could work, suppose the Planting in Palanpur Game (Assurance Game) is now sequential. Aram moves first (he is called the **first mover**) and Bina moves second. How will Aram reason?

He has to think about what Bina will do in response to his planting early or late. He knows that:

- Bina's best response to his playing Plant Late is to Plant Late and the best response to his playing Plant Early is to Plant Early; and
- his payoff is greater if they both Plant Early.

FIRST MOVER A player who can commit to a strategy in a game before other players have acted is a first mover.

So he will announce that he will Plant Early, and Bina will respond with Plant Early. Rather than being stuck planting late with a small harvest, they have now solved their coordination failure. How did they manage it?

Game trees and extensive-form games

The answer is that the sequential nature of the game gave them a way of acting together even if they had no way of coming to some kind of enforceable agreement. By looking ahead to how Bina would respond to his move, he got the highest possible payoff (4). And while this was not his intention, he acted so that they would together implement the single Pareto-efficient outcome.

What Aram did is called **backward induction**, a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen in response by other players in subsequent steps.

Sequential games provide opportunities for coordination among players; and they also require a new way of modeling a game. So far we have studied games in which we could represent the payoffs of each player in a matrix in which each cell is a particular strategy profile. This is called a **normal-form** (or strategic form) representation of a game. To study sequential games we need to keep track of *before* and *after*, so we use what is called a game tree (we will show one for the Palanpur farmers).

Game Trees have the same basic information as a normal-form representation by a payoff matrix–they show the strategy set and the payoffs associated with each strategy profile–except that the tree-like structure tells us who moves when; and a strategy profile is now a *path* through the branches of the tree. Including the time dimension–who knows what when and the sequence of moves–in addition to the strategies with payoffs is called an **extensive form** representation of a game.

› EXAMPLE The timing of a sequential game does not depend on actions being taken in that sequence, as long as *commitments* to those actions can be taken in sequence. A professor commits to a grading policy in her syllabus even if her students haven't written a midterm exam or solved a problem set yet. To design the syllabus, the professor, using backward induction, thought through what a student would most likely do in response to her commitments in the syllabus.

BACKWARD INDUCTION Backward induction is a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen by other players in subsequent steps in response to her choice.

NORMAL-FORM REPRESENTATION OF A GAME The description of a game by a matrix of strategies with payoffs associated with each strategy profile is the normal-form (or strategic) representation of a game.

EXTENSIVE-FORM REPRESENTATION OF A GAME An extensive-form representation of a game includes, in addition to the strategies with payoffs associated with each strategy profile, the time dimension—who knows what, when they know it, and the sequence of moves as described by a game tree.

A game tree for the sequential version of the Planting in Palanpur Game is shown in Figure 2.8. The player at the top of the tree moves first, with subsequent players moving in sequence after the first player. (The "top" of the tree is the trunk, and the branches extending from it are shown below the trunk. The passage of time is shown as a movement down in the figure from the trunk to the branches).

A strategy in a sequential game is a statement of the action a player will take at any point in the tree at which it is her turn to act (whether or not that point will ever be reached). This differs from the strategies we have considered so far, which were simply actions like Fish 10 hours or Plant Late. Strategies are now contingent on what has happened so far in the game. So in the example above Bina's strategy was Plant Early *if Aram Plants Early*.

Aram is the first mover, so he is at the top of the game tree. Bina is the second mover, so she is shown farther down the tree, acting knowing what strategy Aram has chosen. Each player's action—Plant Early or Plant Late—is shown alongside a *branch* of the tree to indicate which action the player chooses as they move along that branch.

Players best respond based on their payoffs, shown at the *end* of a branch of the game tree that indicates a specific path to that end point, Aram planting early, then Bina planting early; Aram planting early, then Bina planting late; and so on. The payoffs (first-mover's payoff, second-mover's payoff) are those shown in the normal-form representation of the simultaneous game in Figure 2.4 (which we saw in Chapter 1). Because of the branching tree-like structure of the figure there is only one path from the start of the game to each of the end points.

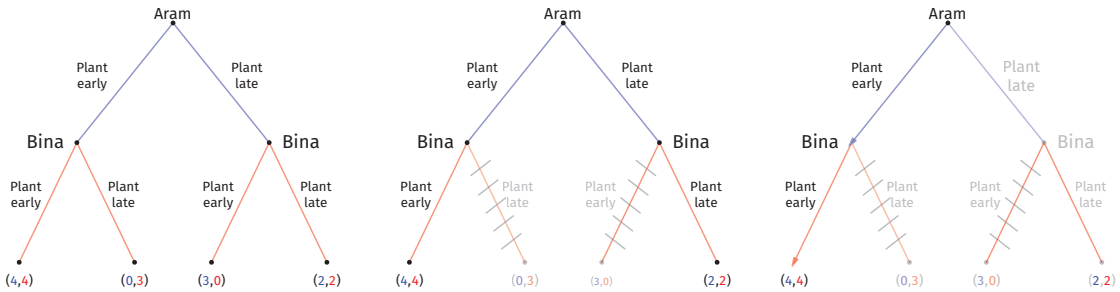
In Figure 2.8 on the left-hand side we have the full game tree, showing all the potential payoffs for the game. Bina is the second mover and she needs to decide what to do at each point where she could move. If Aram plants early, Bina can get a payoff of 4 for planting early, or a payoff of 3 for planting late. So, if Bina is self-interested, then she will plant early when Aram plants early ($4 > 3$).

EXAMPLE Think of a strategy in a sequential game as a complete list of instructions covering any possible situation that could come up that the player could leave with an assistant, if the player could not be present for the actual play of the game.

Bina also has to make a choice between her actions if Aram plants late. Bina can get a payoff of 0 if she plants early given Aram planting late or a payoff of 2 if she plants late given that Aram plants late ($2 > 0$). So Bina will plant late when Aram plants late.

We now know what Bina will do, but what will Aram choose to do knowing this? Using backward induction Aram will have a choice between a smaller set of payoffs, shown in the Figure 2.8 (b): either 4 if he plants early or 2 if he plants late. So he will choose to plant early. As a result, the only Nash equilibrium of the game is (Plant Early, Plant Early) with payoffs (4,4).

Figure 2.8 Game tree of the sequential Planting in Palanpur (Assurance) Game. Panel (a) presents the full game tree for both players. The color on the branches representing actions taken by the player (blue for Aram, red for Bina) match the payoff numbers at the end of the branches (first-mover's payoff first, second-mover's payoff second). In panel (b) we illustrate how Aram uses backward induction by crossing out the branches that he knows Bina will not take if that point in the game is reached. If he has planted late she will definitely not plant early. Considering the remaining branches Aram's possible payoffs are now reduced to 4 if he plants early and 2 if he plants late. So backward induction leads to the solved game in panel (c) with the arrows indicating the path to the Nash equilibrium (Plant Early, Plant Early).



(a) Full game tree (b) Aram's reduced choice of Actions (c) Fully solved Game

CHECKPOINT 2.5 Back in Palanpur Making the game sequential solved the problem for Aram and Bina. But would that work for the couple of hundred families in Palanpur? Suppose some order of play was determined and that the first family had announced that they would Plant Early. Would the second family then follow? And the third? What would you do if you were first mover in this game? If you were 27th mover and the first 26 had all chosen Plant Early?

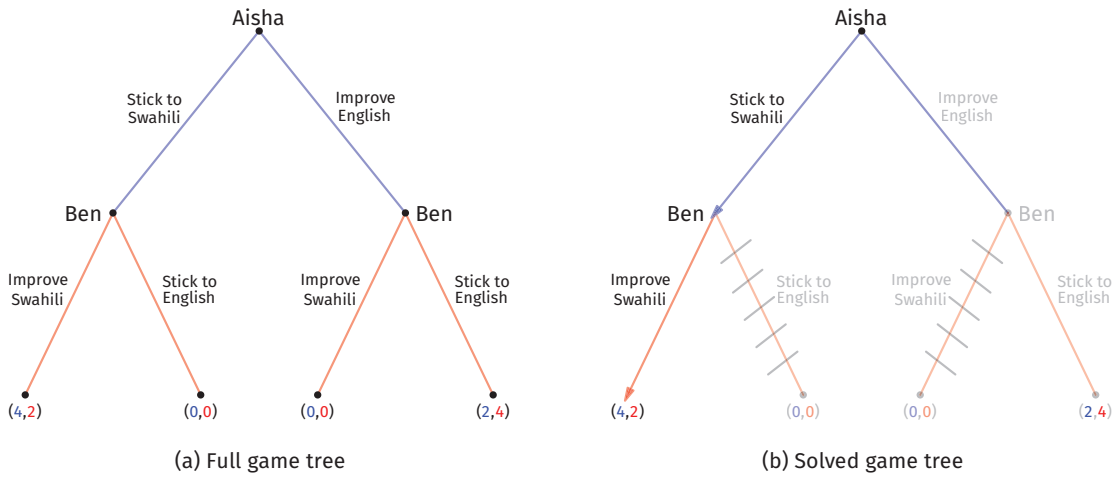
2.7 EQUILIBRIUM SELECTION: FIRST-MOVER ADVANTAGE IN A SEQUENTIAL GAME

Being first mover did not give Aram any particular advantage over Bina in the Planting in Palanpur Game, it just allowed him and Bina jointly to coordinate on the Pareto-efficient Nash equilibrium. The result would have been the same had Bina been first mover.

But sometimes it is advantageous for a player to move first; this person then has what is called first-mover advantage.

Think about the Disagreement Game from Chapter 1. Recall that two players, Aisha and Ben, have a disagreement over which (or perhaps both)

Figure 2.9 Game tree of the Language (Disagreement) game. Panel (a) presents the full game tree for both players. Panel (b) shows the solved game tree with the arrows indicating the path to the Nash equilibrium (Stick to Swahili, Improve Swahili). Aisha's actions are shown by the blue branches and Ben's by the red branches. Aisha's actions are reduced because she has projected forward in time and used backward induction to work out what Ben will do. This reduces Aisha's choices to a payoff of 4 if she plays Stick to Swahili and a payoff of 2 if she plays Improve English. So backward induction leads to the Nash equilibrium of the game being (Stick to Swahili, Improve Swahili) with payoffs (4,2). The outcome favors Aisha over Ben because she has a first-mover advantage.



of them should study to improve the language spoken by the other. Both prefer that they both be good at speaking *some* common language. But, Aisha prefers that it be Swahili and Ben prefers that it be English. What happens in this game when Aisha is the first mover rather than when they both move simultaneously?

Considering the game tree in Figure 2.9, we can solve the game by backward induction and see that the Nash equilibrium of the game is (Stick to Swahili (for Aisha), Improve Swahili (for Ben)) with payoffs (4,2). The outcome (Improve English, Stick to English) which was one of two Nash equilibria in the simultaneous version of the game is no longer a solution in the sequential version of the game if Aisha is first mover. Aisha does better as a first mover because she obtains her preferred outcome. Ben would have benefited in the same way had he been first mover.

The reason why being first mover gave Aisha an advantage is that the simultaneous game has two Nash equilibria–one preferred by Aisha and the other by Ben. In the sequential game the first mover determines which of the two Nash equilibria will occur. Once Aisha has moved and has established that she will Stick to Swahili (and *not* try to Improve

English), Ben needs to take Aisha's move as *given*. He must therefore choose his best response to Aisha choosing Stick to Swahili. Given that he would like to communicate with Aisha, his best response is to Improve his Swahili.

Remember that Aram's first-mover status in the Planting in Palanpur game did not allow him to benefit at Bina's expense; but this was not the case with Aisha, her first-mover status gave her an advantage over Ben.

First movers in a modern economy are more like Aisha:

- *Employers*: they commit to the wage, job requirements and working conditions; workers—actual and prospective—best respond to that.
- *Banks and other lenders*: they set the interest rate, repayment schedule, and other aspects of a loan contract. Borrowers and would-be borrowers best respond to that.
- *Owners of major companies*: in the US Walmart, Amazon, Apple—commit to prices and delivery schedules. Consumers best respond.

The fact that people occupy different positions in our economy—employers and workers, lender and borrowers—interacting under rules of the game that give some first-mover status and other special advantages is an important part of the explanation of inequality of wealth and income, as we will see in Chapters 11, 12, 13, and 15.

CHECKPOINT 2.6 Ben has the first mover advantage

a. Consider the sequential Disagreement Game shown in Figure 2.9. Redraw the game tree, with Ben as the first mover rather than Aisha. Show that (Improve English, Stick to English) is the Nash equilibrium of the game.

b. Assuming that the payoffs in the Disagreement Game are in hundreds of dollars and that you are Ben, how much would you pay for the privilege of being first mover (i) if otherwise Aisha would be first mover, and (ii) if the game were to be played simultaneously (so that there is no first mover)?

2.8 INSTITUTIONAL CHALLENGES: COMMON PROPERTY RESOURCES, PUBLIC GOODS, AND CLUB GOODS

In the games you have studied—Prisoners' Dilemma, Assurance, Disagreement, and others—coordination problems arise because when we interact with others we affect their well-being—positively or negatively—and these external effects are not taken into account when we decide on a course of action. The nature of these external effects and how changes in the rules of the game can avoid or lessen the resulting coordination failures depend on the nature of the goods in question.

A taxonomy of goods

To better understand the kind of coordination problems that we face and how we might design effective remedies, we classify goods according to the kinds of external effects associated with them and the reason why these are a problem. To do this we ask two questions, introducing two new terms:

- Is the good **rival** or non-rival?
- Is the good **excludable** or non-excludable?

When a good is rival, the benefits of its use are limited: more people using the good reduces the benefit available to others. Your phone is a rival good: your using it prevents me from using it at the same time. But information typically is non-rival: the fact that I know what time it is and give this information to you does not deprive me of the same information, as would be the case if I gave you my phone.

So, to remember the distinction between rival and non-rival goods think how different the reaction would be if you met someone in the street and politely asked:

- "Excuse me, could you give me the time of day?" or
- "Excuse me, could you give me your phone?"

When a good is excludable a potential user may be denied access to the good (or excluded from its usage) at low or zero cost. Your home is an excludable good: all you have to do is lock the door. The music from an outdoor concert in a park is non-excludable.

We make use of these distinctions to provide the taxonomy shown in Table 2.3. The four categories shown there are "pure cases" introduced to clarify distinctions. In reality many goods or resources have some aspects of a public good (they may be a little bit rival and a little bit excludable). The same is true of the other three categories.

RIVAL A good is rival when more people using the good reduces the benefits available to other users.

EXCLUDABLE A good is excludable when a potential user may be denied access to the good at a low or zero cost.

Table 2.3 Public, private, common property, and club goods. In parentheses are examples of the kinds of goods.

	Excludable	Non-excludable
Rival	Private good (clothing, food)	Common property (pool) resource (fishing stocks, potential buyers)
Non-rival	Club good (streaming music, online movies)	Public good (global climate, rules of calculus)

Non-excludability and external effects

If we just think about the pure cases for now, we have the following classification. **Common property resources** are rival and non-excludable, like in the Fishermen's Dilemma in Chapter 1. The more one fished, the less others caught; but in the absence of an enforceable agreement between the two, no fishermen could be stopped from fishing, so the common property or pool (the lake) was non-excludable.

Examples of common property resources and their associated coordination problems include congestion in transportation and communications networks, overuse of open-access forests, fisheries, water resources. Status is another common property resource; not everyone can be high status (there is a limited amount to go around) so it is rival. But nobody can be excluded from acquiring status symbols and engaging in other social-climbing activities. Using common property resources imposes external costs on others. As a result, common pool resources will be overexploited.

A **public good** is both non-rival and non-excludable. A **private good** is neither: it is both rival and excludable. A slice of pizza is a private good: it is rival because if you eat it nobody else can enjoy it. It is excludable because the pizza seller can exclude you from eating it if you do not pay for it. By contrast, weather forecasts (on your phone, website, or the radio) are a public good. As more people use the weather forecasts the benefits that those already using the forecasts receive do *not* decrease, the benefits of the weather forecasts are non-rival. No person can be excluded

COMMON PROPERTY RESOURCE A common property resource is rival and non-excludable. See also *public good, private good, club good*

PUBLIC GOOD A public good is non-rival and non-excludable. See also *common property resource, private good, club good.*

PRIVATE GOOD A private good is rival and excludable. See also *common property resource, public good, club good.*

EXAMPLE Club goods. The physical book you are reading is a private good—though you can access a free pdf online so its content is a public good—but another recent economics textbook *The Economy* by the CORE team (including one of your current authors) is a public good entirely available in open access digital form on any device.

from access to the information about the weather; therefore the benefits are non-excludable.

When a person contributes to a public good–for example by producing some new information of value to everyone–she is contributing benefits to others, so she confers external benefits on others. The problem here is that the person does not benefit from the positive external effects that her actions convey on others. So unless the actor values the well-being of others as much as her own (very unlikely) the public good will be under-provided.

In contrast with public goods and common property resources, there are **club goods**. Club goods are non-rival, but people can be excluded from their consumption. Common examples include collecting a toll on a little used highway, charging admission to an uncrowded museum, or requiring people to pay for streaming video and music.

Intellectual property rights such as patents and copyrights are club goods. These legal devices give a person–the patent or copyright holder–a monopoly over a piece of information or a design. This monopoly allows the owner to exclude people from the use of information, which in the absence of the intellectual property rights would be a public good.

This makes it clear that how some good or resource is classified in our two-by-two taxonomy depends not only on the nature of the good itself, but also on the rules of the game that determine whether it is excludable or not. Information is typically a public good, but it can be made a club good if a person is granted a copyright or patent making some piece of information excludable.

We analyze the coordination problem that occurs when our economic activities are overexploiting a common property resource in Chapter 5. Here we study the problem of public goods.

CHECKPOINT 2.7 A taxonomy of types of goods

a. Look again at Table 2.3 and think of at least two further examples for each of the four categories of goods.

b. Why are the rules of calculus a public good, but the formula for making Coca-Cola not?

c. What kind of good is the formula for making Coca-Cola?

CLUB GOOD A club good is non-rival and excludable.

2.9 THE PUBLIC GOODS GAME

To better understand public goods we start with an example. Think about a group project in which, say, five students collaborate on a report and a presentation and all receive the same grade based on the quality of their joint work. A project like this is a public good.

A more pressing example is global climate: it is experienced by everyone. Efforts to address the problem of climate change contribute to a public good: that is, a more sustainable environment. Another example is the rules of calculus: if you learn how to differentiate, that does not deprive others of the knowledge of the same rules of differentiation.

This sounds like a good thing. But there is a problem. Why do people produce or contribute to the provision of a public good? If nobody can be excluded from enjoying the good, it's hard to see how it would be possible to make money by providing it. (Imagine trying to make a living by selling or renting the rules of calculus!)

We can describe the problem of provision of a public good by a game. It shares with the Prisoners' Dilemma Game the feature that everyone could do better if they agreed on a common course of action (i.e. they all contribute) but the dominant strategy for a self-regarding player is not to contribute. A player who does this is called a **free rider**. Because it has the same incentive structure, the Public Goods Game is sometimes called an n-person Prisoners' Dilemma. The Public Goods Game has been played as an experiment around the world.

Rules of the Public Goods Game

Here are the rules of the experimental game:

- n players are each given some amount of money z called an endowment.
- Each player simultaneously selects an amount e^i, $0 \leq e^i \leq z$ to contribute to the public good (think of e^i as player i's "effort" in contributing to the public good).
- The amount of the public good produced depends on the level of contributions. For example it could be half of the sum of all of the contributions. In this case the productivity of contributions would be one-half.
- Each player, regardless of whether they contribute or not, obtains the entire benefit of the total amount of the public good produced.

As a result of the rules, each player's payoff can be read as follows:

$$\text{Own payoff} = \text{Endowment} - \text{Contribution} + \text{Productivity} \times \text{Total contributions}$$

FREE RIDER A free rider is a person who benefits from the cooperation or generosity of others, while not reciprocating in a cooperative or generous way, for example, not contributing in a Public Goods Game.

Figure 2.10 A four-player Public Goods Game with choices to contribute or not. Each player can play either Contribute or Don't Contribute, and, as there are four players, this means that the number of others contributing can be any of the numbers 0, 1, 2, or 3 players playing either of the strategies. Playing Don't Contribute yields a higher payoff for the player regardless of how many players play Contribute or Don't. Therefore, Don't Contribute is a strictly dominant strategy. The payoffs are consistent with M-Note 2.1.

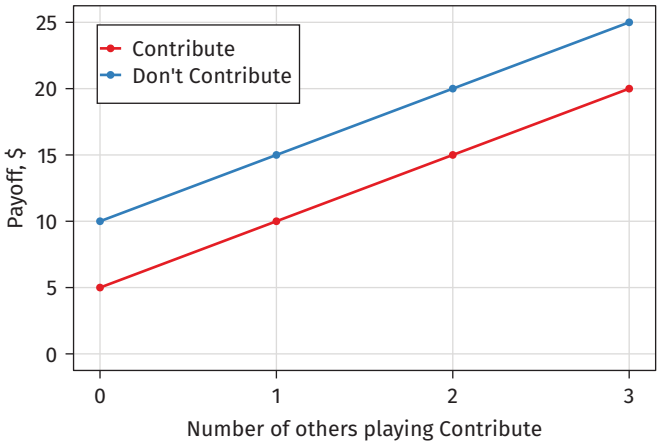


Figure 2.10 illustrates the benefits of the public good minus the costs of contributing to a public good in a four-person Public Goods Game. In the version of the game we depict, they can each contribute $10 or $0: which we call "Contribute" or "Don't." Now compare how a player does if they Contribute (red line) or Don't (blue line) if they are the only one who contributes, or there are 1, 2, or all 3 others contributing. You can see that in every case she will earn higher payoffs by not contributing. Therefore, if all players are self-regarding, the dominant strategy equilibrium is Pareto inefficient; but an alternative outcome, full contribution by all, which is not a Nash equilibrium, is Pareto efficient.

As a result, economists expected that when this game is played for real money that no player would contribute. They were in for a surprise. But first we will explain the logic of the experiments that provided the surprise.

M-NOTE 2.1 The Public Goods Game: another coordination problem

In Figure 2.10, there are four players and we limited their actions to either contributing $10 (Contribute) or contributing $0 (Don't Contribute), but in a standard Public Goods Game players can contribute any amount up to and including their entire endowment (such that for player i, $e^i = z$). Then player i's payoff is given by :

$$\pi^i = z - e^i + M\sum_j e^j \text{ for } j = 1, \ldots, i, \ldots, n \tag{2.2}$$

continued

where:

- z is the endowment of money the player receives from the experimenters.
- e^i is the contribution player i makes.
- M is the multiplier, or the productivity of contributions < 1.
- n is the number of players with $nM > 1$.
- $\sum_j e^j$ is the total amount contributed by *all* players including player i.

Why the public good is good. The requirement that $nM > 1$ ensures the total benefits of contributing to the public good exceed the costs. Because contributing one more unit produces some fraction M of a unit of the public good that is enjoyed by all n players, the total benefit is $nM > 1$. The total benefit (to all members of the population) exceeds the cost of a single member's contribution.

Why not contributing is a dominant strategy. You can also see from Equation 2.2 if you differentiate π with respect to e^i that contributing, say, one unit more changes person i's payoff by $-1 + M$. This is the cost of contributing minus the public good that the contributor herself enjoys as the result of her own contribution. But $M < 1$ so contributing anything reduces the contributor's payoffs. And this is *independent* of the amounts contributed by others. This is why *not* contributing is the dominant strategy for a self regarding player.

CHECKPOINT 2.8 Two-action Public Goods Game

a. Draw a payoff table with two players, A and B, playing the Public Goods Game. Limit their actions to contributions (e) of $e = 10$ and $e = 0$ with $M = 0.5$. Check which is the dominant strategy and explain why. What happens if $M = 0.75$?

b. Revise your payoff table and check what would happen if the strategies were $e = 1$ and $e = 0$ with $M = 0.5$? Would anything change? What happens if $M = 0.75$?

c. Think about the condition $M < 1 < Mn$. Why must this be true for the game to be an n-person Prisoners' Dilemma game? (Hint: Think about what would happen if it were *not* true. What would happen if $M > 1$? What would happen if $Mn < 1$?)

2.10 APPLICATION: EXPERIMENTS ON ECONOMIC BEHAVIOR

Suppose you wanted to know if someone has altruistic preferences, that she is willing to help others at a cost to herself. How would you find out? Would you ask her? Well, that could provide some information, but merely asking might not be entirely convincing, because many people would like others to think they are altruistic even when they are not, so they might lie.

What about observing her behavior—for example the help that she actually offers to others—and comparing her behavior to how others behave? This would be informative, but how much she helped others would be influenced not only by her preferences, but also by how much free time she has, how wealthy she is, and many other difficult to observe influences.

Economists use experiments to study preferences because at least ideally this allows us to control for (hold constant) other influences on a person's behavior—the constraints they face and their beliefs—to focus on their preferences. Experiments allow economists to implement the **ceteris paribus**—other things equal—assumption that we think is so important when we are trying to identify causes and consequences of some change or difference.

✓ **FACT CHECK** Behavioral experiments are a recent addition to economists' tool kits; but they have been used in psychology for almost a century and a half. The main innovations that economists have made to experimental social science are the use of game theory to clarify the role of beliefs, preferences, and the nature of incentives, as well as the common use of monetary payoffs.

To understand how common different types of preferences are, and how they affect our behavior, economists use laboratory experiments in which subjects, the people participating in the experiments, interact in games like the ones you have already studied, designed to elicit the nature of their motivations.

Experiments play a central role in science: they allow predictions made from theories to be tested empirically. This has been done, for example, with the prediction that players in a Prisoners' Dilemma experiment choose the dominant strategy equilibrium, that is, Defect.

But in Prisoners' Dilemma experiments, in which payoffs took the form of money that a player could win, the proportion of subjects who cooperate rather than defect is commonly between 40 and 60 percent.[6]

This means the *prediction* based on the assumption that people are entirely self-regarding was borne out for some but far from all of the subjects. The finding therefore provoked some rethinking of the assumption that people are entirely self-regarding.

Many subjects prefer the mutual cooperation outcome and are willing to take a chance on the other player also not defecting, rather than the higher material payoff they can obtain by defecting when the other cooperates. When subjects defect, experimental evidence suggests it is because they dislike being taken advantage of, not because defection is the payoff-maximizing strategy independently of the other participant's actions.

We use a specific vocabulary when we talk about behavioral experiments in economics. The following terms will come up often:

- *Subject/participant*: A *subject* or *participant* is a person who participates in an experiment.
- *Endowment*: The *endowment* is an initial amount of money or tokens later converted to money that subjects receive at the beginning of the experiment, and later make decisions about in the experiment.

CETERIS PARIBUS A Latin term that means 'other things equal.' In an economic model it means an analysis that 'holds other things constant.'

- *Incentives*: The fact that players stand to win material rewards in varying degrees depending on how they play the experimental game means that the experiment mimics many real economic interactions.
- *Payoffs*: In Chapter 1 we introduced the term payoff (for example in a payoff matrix of a game) as a number indicating the player's evaluation of a particular strategy profile, so that a player will best respond by choosing a strategy with the highest possible payoff. But we have already seen (just above) that people do not always choose strategies that maximize the money they receive from an experiment. In experimental economics a "payoff" is money that a player gets from the game. In the next section we will see more evidence that payoffs are not the only thing people care about.
- *One shot vs. repeated*: A *one-shot* experiment occurs once and subjects make one decision in the experiment as a whole and are paid for that one decision. A *repeated* experiment involves subjects making repeated decisions often with information about the play of others on previous rounds, sometimes with the same subjects in a group or sometimes with different subjects.
- *Replication*: Experimental evidence carries little weight unless the experiment can be *replicated*, different independent researchers reaching the same results.

CHECKPOINT 2.9 **Where experiments do not work** Think of important questions that cannot be answered by experiments.

2.11 APPLICATION: CHANGING THE RULES MATTERS—EXPERIMENTAL EVIDENCE

The prediction of the model based on self-regarding preferences that all players in a Public Goods Game will contribute nothing is consistently contradicted by the experimental evidence. The evidence we have comes from people playing one-shot games and from people playing repeated games with as few as five rounds and as many as 50 rounds.[7] In one-shot games, contributions average about half of the endowment. In repeated games, as you can see in the first ten periods of play in Figure 2.11, contributions start at a substantial level but then decline so that a majority of players contribute nothing in the final round of a ten-round game.

Researchers have interpreted the decline in the first half as a reflection of people getting disappointed about the expectations they had that other people would contribute more, along with the desire people have to punish low contributors (or at least not to be taken advantage of) in a situation in which one person can punish a low contributor only by reducing their own contributions.

In this interpretation it is the higher contributing subjects disappointed or angry about their free-riding fellow subjects that explains why cooperation unravels. So the decline in contributions becomes a vicious circle: only by reducing how much they contribute can people punish others, but in so doing other people might want to punish them for their low contributions by contributing yet less again.

The idea that the decline in contributions is due to the fact that in the standard game contributing less is the only way to punish low contributors is supported by an ingenious experiment. This has the same public goods structure but with what turned out to be a major difference: after subjects contributed, the contributions of each—by a code, not the player's name—were then made public to all the group members. Members then had the opportunity to punish others in the group, reducing their target's total payoff. In order to impose this cost, however, the Punisher also had to pay a cost themselves.

The change in the rules of the game—adding the punishment option—represents a change in the institutions governing contributions to the public good. In the language of experiments the new rules are termed a new treatment. So the standard game is one treatment and the game with punishment is a second treatment.

In the experiment, subjects engaged in extensive punishment of low contributors. At the start of the game people contributed over half of the endowment and then, apparently in response to punishment of low contributors, they contributed more over the course of the game. The change in institutions modeled by adding the punishment option altered the result dramatically as you can see from Figure 2.11.

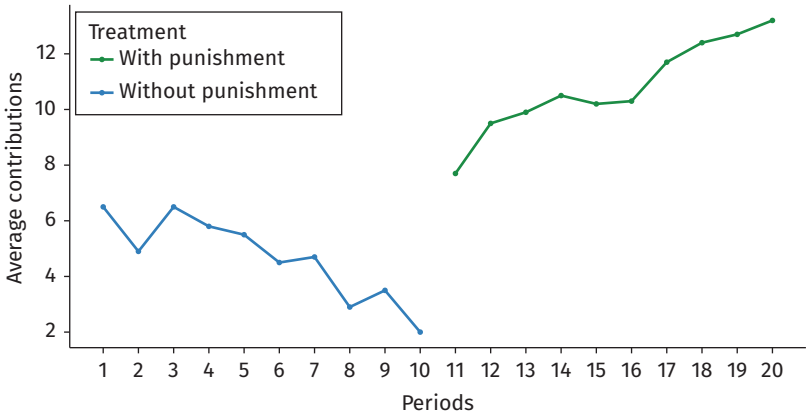
To see if subjects' willingness to punish could be based on the expectation that they would benefit in subsequent rounds of the game, a slightly different experiment was tried. The researchers adopted what they called a "perfect strangers" treatment: after each round of the ten-round experiment the groups were reshuffled, so that no player ever encountered any other player more than once. The "perfect strangers treatment" turned the experiment into a series of one-shot games.

Since every player would encounter every other player only once, if low contributors responded to punishment by contributing more in subsequent rounds, they would raise the payoffs of others but not the punisher (who would never again be in the same group with the target of her punishment).

In this experiment, *punishment itself* is a public good. This is because a punisher incurs a cost, except that the punisher is not a beneficiary of the good. In the perfect stranger treatment, for a self-regarding player not punishing, like not contributing to the public good is the dominant strategy. Even in the perfect strangers treatment subjects avidly punished low contributors.

Figure 2.11 Public Goods Game with punishment. Average contributions over periods 1 to 10 decrease without punishment. Over periods 11 to 20, subjects can be punished by their peers and average contributions are higher on average than in the first 10 rounds. The vertical axis is the average contribution each round. The horizontal axis is the period. At period 11 the subjects are given the opportunity to punish each other. There are three treatments in this Public Goods Game experiment. This figure portrays the behavior in the "Strangers" treatment where players are randomly re-matched each round. The two other treatments, which show similar results, are "Partners" where players are in the same group for all the rounds; and "Perfect Strangers" where players are re-matched, but no player will encounter any other more than once during the experiment.

Source: Fehr and Gächter (2000a).



Further evidence comes from the fact that people punish low contributors even in the last round of the game when punishment cannot be motivated by the expectation that the punisher will benefit from their target's improved behavior in the future. There is no future (the game ends after they punish). So the pleasure of punishing someone who is violating a social norm is most likely involved.[8]

CHECKPOINT 2.10 Changing the rules of the game Explain why in the Public Goods with Punishment game, punishing a low contributor is itself a public good.

2.12 SOCIAL PREFERENCES: BLAME ECONOMIC MAN FOR COORDINATION FAILURES?

The fact that for self-regarding people, not contributing to the public good is the dominant strategy definitely constitutes an institutional challenge. But we will also see that although being concerned about how your actions

HISTORY The idea of basing economics on the assumption that people are entirely self-regarding—"solely as a being who desires to possess wealth"—goes back to the last of the great classical economists, John Stuart Mill author of *Principles of Political Economy* (1848), considered to be the first economics textbook in the English language. He considered this view of people to be "an arbitrary definition of man." In other words, like economists since then, he was making a simplifying assumption in order to model human behavior.[9]

affect others will help to address coordination failures, it will not be sufficient.

Homo economicus or "economic man" is the term economists have used to designate an entirely self-regarding and amoral actor, a person who is not motivated by either a concern for others, or a desire to conform to any ethical principles. The term is often put in italics to parallel the biological terminology for a species (like *Homo sapiens*). *Homo economicus*, however, is a *fictional* character representing one possible variety of human behavior.

Models based on *Homo economicus* have provided predictions about behavior that are borne out by empirical studies that range from how American windshield installers and Tunisian sharecroppers respond to different work incentives to the effect of taxes on cigarette consumption. But, as we shall see, *Homo economicus* is not an accurate depiction of how most people behave:

- People volunteer for firefighting, delivering food to the sick during a pandemic, and other dangerous but socially beneficial tasks, and contribute substantial sums to charity.
- People participate in joint activities such as strikes or protests even knowing that their individual participation is unlikely to affect the success of the event and that, if successful, the benefits would be widely shared, not confined just to those people participating in the protest.
- People donate blood for the health of strangers, and wear masks in public places during a pandemic, even knowing that the primary benefit of the mask is to prevent spreading the virus to strangers, not protection against being infected by others.
- In public opinion polls and in voting, people support taxes that transfer incomes to the poor even when they are sufficiently rich and unlikely ever to benefit directly from these policies.

! REMINDER Remember that in Chapter 1 we saw how 'internalizing the external effects' means getting people to pay for the external costs they imposed on others and this resulted in the fishermen choosing to cooperate and fish less in the Fishermen's Dilemma.

Motivated by these and similar observations and augmented by controlled experiments about human behavior (that we will review below), economists have revised our assumptions about *Homo economicus* to recognize that people are capable of ethical, generous, and other motivations as well as self-regarding motives.

This is important because as you learned in the first chapter, coordination failures occur because we fail to take adequate account of the effect that our actions have on others. Our concern for others can help to internalize these external effects whether it be our willingness to curb our carbon footprint or willingness to protest for causes whose benefits would be widely shared.

But coordination failures cannot be blamed entirely on people seeking to maximize their own payoffs. Think again about the real farmers in Palanpur, all planting late when they could all do better if they all switched to planting early. Suppose one of those farmers was deeply concerned about

the poverty of his entire village, and wished to improve living standards for everyone. He could not do this by individually planting early.

Now suppose that every villager shared his concerns for all members of their community. Each one would know that their own decision to plant early would change nothing (except that their seeds would be eaten by the birds). What has captured the people of Palanpur in a poverty trap is not that they care only about their own harvest (they surely care about others'), but their inability to come to a common agreement to plant early. Their poverty stems from a problem of institutions, not motivation.

To understand individual behavior and its social consequences we need an approach that allows for the full range of human motivation.

CHECKPOINT 2.11 ***Homo economicus* goes to the polls**

a. Given that it costs time to cast a vote (going to the voting station, standing in line, and the opportunity cost of your time), do you think a person with *Homo economicus* preferences would vote in most elections? Why or why not?

b. In what circumstances do you think someone with the preferences of *Homo economicus* would vote?

While answering these questions, think about the *beliefs* the person with *Homo economicus* preferences would have about the probability his vote will affect the outcome of the election.

Types of social preferences

While *Homo economicus* is among the kinds of actors this approach considers, there are other characters, representing other sides of human behavior such as generosity, fairness, reciprocity, and spite. What these four aspects of behavior have in common is that they are *other-regarding*: the outcomes that a person considers in choosing an action include things experienced by others, not just outcomes affecting the person herself. Here are some forms of other-regarding preferences that experiments of the type surveyed below have shown to be common:

- Those with *altruistic* preferences, such as basic generosity, are motivated to help others even at a cost to themselves; they place a positive value on the well-being or payoffs of others.
- **Inequality-averse** or fairness-based preferences motivate people to seek to reduce unjust or unfair economic differences even if the actor is herself a beneficiary of these differences.

INEQUALITY AVERSION A preference for more equal outcomes and a dislike for both disadvantageous inequality that occurs when others have more than the actor and advantageous inequality that occurs when the actor has more than others.

- A person with *reciprocal* preferences is motivated to help others who have themselves behaved generously or upheld other social norms, and also to punish those who have treated others badly.
- *Spite* and '*us versus them*' distinctions that place a negative value on outcomes experienced by others, often motivate hostility toward members of religious, racial, ethnic, and other groups. Therefore a *negative* outcome another person experiences, can result in a *positive* value for someone who feels *spiteful*.

The term "social preferences" is used to describe all types of other-regarding preferences.

CHECKPOINT 2.12 Social preferences and social norms

a. Give an example of a preference you have that is *not* self-regarding.
b. Can you think of any social norms that lead you to act in an other-regarding way?
c. Suppose that Aram and Bina (in the Planting in Palanpur Game) were of different religions between which there is hostility, so that each would gain some pleasure from the misfortunes of the other. Can you show how this could change the game so that instead of having the Pareto-efficient mutual early planting as one of its two Nash equilibria, it becomes a Prisoners' Dilemma with Plant Late as the dominant strategy equilibrium?

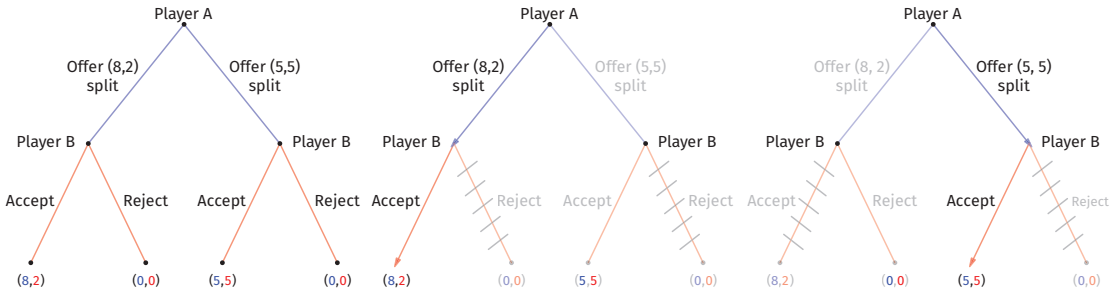
2.13 THE ULTIMATUM GAME: RECIPROCITY AND RETRIBUTION

Observing substantial levels of cooperation in the Prisoners' Dilemma game was a shock to the standard *Homo economicus* assumptions. But the experiment that has sparked perhaps the greatest reconsideration of the *Homo economicus* model is the *Ultimatum Game*.

Here is the game with its basic treatment:

- Subjects are anonymously paired for a one-shot interaction with another person.
- The role of "Proposer" who will be the first mover, is randomly assigned to one of the subjects; the other is then the "Responder."
- The Proposer is given an endowment, the "pie" (e.g. $10), by the experimenters and the Responder knows the size of the pie.
- The Proposer then proposes how to divide the endowment between Proposer and Responder, transferring to the Responder any amount

Figure 2.12 Game tree of the ultimatum (bargaining) game. Panel (a) presents the full game tree for both players. Player A is the Proposer and their actions are shown by the blue branches. Player B is the Responder and their actions are shown by the red branches. Panel (b) shows the backward induction thinking of the Proposer, A, if she believes that B is self-regarding, that is, cares only about her own payoff being as large as possible. Panel (c) shows the same process if A believes that B will reject the (8,2) offer as "unfair."



(a) Full game tree (b) Self-regarding players (c) Player B cares about being treated fairly

between nothing and the entire endowment, e.g. the Proposer chooses to keep \$8 and give \$2 to the Responder.

- If the Responder accepts the proposed division, the Responder gets the proposed portion, and the Proposer keeps the rest and the game ends.
- If the Responder rejects the offer, both get nothing and the game ends.

Figure 2.12 presents a *game tree* for a variant of the Ultimatum Game in which the Proposer, player A, selects one of two offers to make to the Responder: divide the pie equally and each person gets \$5 for an outcome (5,5) or keep \$8 and offer the Responder \$2 for an outcome of (8,2). The Responder, player B, then chooses whether to accept or reject the offer. You can see from the (0,0) labels at the end of the two Reject branches, that if B rejects the offer, both players get zero. The payoffs to each player are listed in the order of play (Player A, Player B), so (8,2) means Player A gets 8 and Player B gets 2. If the Proposer cares only about her monetary payoffs in the game and believes that the Respondent is similarly self-regarding, then the Proposer (Player A) will reason backwards as follows:

EXAMPLE In this video (tinyurl.com/y47onzue), Juan Camilo Cardenas talks about his innovative use of experimental economics in real-life situations (from the CORE project. www.core-econ.org).

- Responder (Player B) will accept the offer of \$2 because \$2 is greater than \$0 which is what he gets if he rejects the offer.
- So A will propose the (8, 2) split.
- And the Responder (B) will accept.

That is not how the experiment worked out.

The Ultimatum Game has been played anonymously, sometimes for substantial sums of money, in hundreds of experiments with university

student subjects and other populations—businessmen, fishermen, farmers, civil servants—in all parts of the world.[10]

✓ **FACT CHECK** Did the subjects not understand the game? It is not that complicated a game, and later experiments in which subjects played the game many times with different partners showed they did understand it. Their behavior remained consistent with the one-shot experiments and their results continued to be reproduced with many people making 50-50 splits (or nearly so) and rejecting low offers.

✓ **FACT CHECK** Some have suggested that the results were due to the relatively low stakes in the game, such as the $10 mentioned earlier. But subsequent experiments conducted among university students in Indonesia for a 'pie' equal to three months' average expenditures replicated the results as did experiments with US students with a 'pie' ranging in size up to $100. Evidence from France showed similar behavior by proposers with stakes ranging from 40 French francs ($7.20) to 2,000 French francs ($360) (this was prior to the adoption of the euro). A further study in India observed stakes that varied by a magnitude of over 1,000: From 20 rupees ($0.41) to 20,000 rupees ($410) as the stakes.[12]

The prediction based on the assumption that people are entirely self-regarding and believe that others are too invariably fails as a description of how people behave. For example:

- Modal offers—the most common offers in the experiments—are typically half of the pie, and average offers generally exceed 40 percent of the pie, and
- Offers of 20 percent of the pie or less are often rejected; people in the position of Responder choose to reject and get zero rather than accept and get a payoff of, say, $2 offered from the Proposers $10 pie.

As a possible explanation of these results Figure 2.12 c shows how the game might be played if Player B cares both about monetary payoffs and also about being treated fairly. In this case, Player B views an offer of (8,2) as unfair or demonstrating greed on A's part, and they would rather get a payoff of zero dollars than accept a deal in which they are treated poorly, so they would *reject*.

If, on the other hand, Player A offers (5,5), then Player B views that as fair or demonstrating good will and they would prefer a payoff of 5 in that context to a payoff of 0, so they would *accept*. Player A prefers a payoff of 5 to a payoff of 0 and so the Nash equilibrium of the game is (Offer a (5,5) Split, Accept) with payoffs (5,5).

These rejections of small but positive offers from the Proposer are interpreted as evidence for *reciprocity* motives on the part of the Responder. Why? Because the Responder is willing to pay a price (giving up a positive payoff) to punish the Proposer for making an unfair offer (an offer the Responder considers too low). Responders apparently consider a low offer to be a violation of a norm of fairness, and a person with reciprocal preferences responds by *depriving* the proposer of any payoffs at all.

Explaining the behavior of Proposers is more complicated. The outcomes of the experiments are not sufficient to say whether the large number of even splits (and other seemingly fair offers) is explained by adherence to fairness norms or altruism by the Proposer or to self-regarding preferences informed by fear that the Responder will reject an unfair offer. The evidence for reciprocity motives therefore comes from the Responders' behaviors, not the Proposers' behaviors.[11]

CHECKPOINT 2.13 Fairness in the Ultimatum Game Explain why it might make sense for an entirely self-regarding Proposer in the Ultimatum Game to offer half of the "pie" to the Responder. Would the Proposer do this if she knew that the Responder was also entirely self-regarding?

2.14 APPLICATION: A GLOBAL VIEW—COMMON PATTERNS AND CULTURAL DIFFERENCES

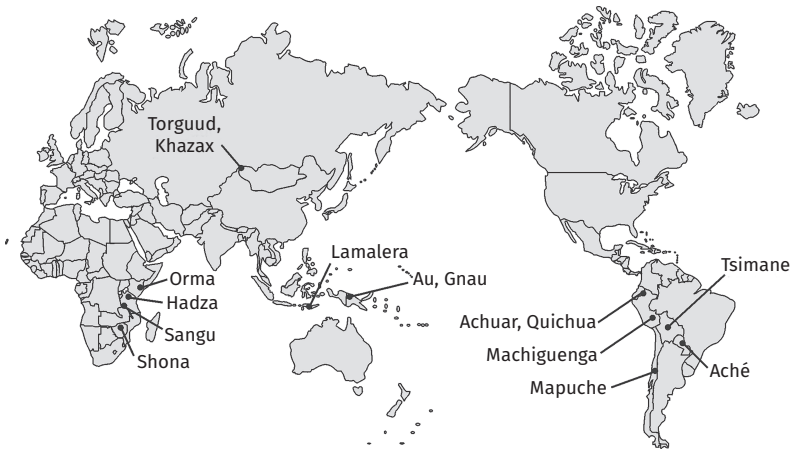
Anthropologists and others were surprised that the results of experiments with the Ultimatum Game have been so similar across the many countries in which they have been conducted. One observed that in virtually all of the early experiments the subjects were from WEIRD countries, meaning western, educated, industrialized, rich, and democratic.[13] A team of anthropologists and economists (including one of your current authors) designed a series of experiments to explore whether the results reported so far are replicable in societies with quite different cultures and social institutions and whether results differed across the different societies.[14] These societies included hunter-gatherers, herders, and farmers (some using modern methods, others not even having cattle, horses, or plows). In their Ultimatum Game experiments the pie was substantial, approximately a day's average wages or other income.

Figure 2.13 shows the location of the 15 small-scale societies around the globe. The team was wondering if they would find cultural differences, and they found them.

Among the Au and Gnau people in Papua New Guinea offers of more than half of the pie were common, and many of these high offers were rejected. In fact Responders among the Au and Gnau peoples were as likely to reject an offer of much more than half as an offer of much less than half.

Figure 2.13 Small-scale societies where the Ultimatum Game experiments were conducted. The researchers wanted to ensure cultural diversity in their sample. So they selected communities living in very different physical environments, making their living in diverse ways and very little influenced by the homogenizing influences of markets, governments, and other modern institutions.

Source: Figure 1 from: Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q. and Tracer, D. (2005), "'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and Brain Sciences*. Cambridge University Press, 28(6), pp. 795–815. doi: 10.1017/S0140525X05000142. Reproduced with permission.



Though this seemed odd to the economists on the team, it did not surprise the anthropologists who study New Guinea. They know that people in New Guinea compete with each other to see who can give more or better gifts. Gift-giving conveys status in their society and people use giving gifts as a way to obtain status over others. Refusing a gift suggests that you are *not* subordinate to the gift-giver, while accepting it means their status is higher than yours.

By contrast, among the highly individualistic Machiguenga slash-and-burn farmers in Amazonian Peru, almost three-quarters of the offers were a quarter of the pie or less and there was just a single rejection, a pattern strikingly different from other experiments. The Machiguenga came as close to acting like *Homo economicus* as any population yet studied. Even among the Machiguenga, however, the mean offer was still 27 percent of the pie, more than close to zero that we'd expect if they all were consistently self-interested.

The researchers who analyzed the experiments in the 15 small-scale societies made the following conclusions:

- Although behaviors vary greatly across societies, not a single society approximated the behaviors that would be observed if everyone cared only about their own payoffs and believed others were the same.
- Between-society differences in behavior seem to reflect differences in the kinds of social interaction people experience in everyday life.

Here is some evidence that the experimental game behavior reflected the lived experiences of the people.

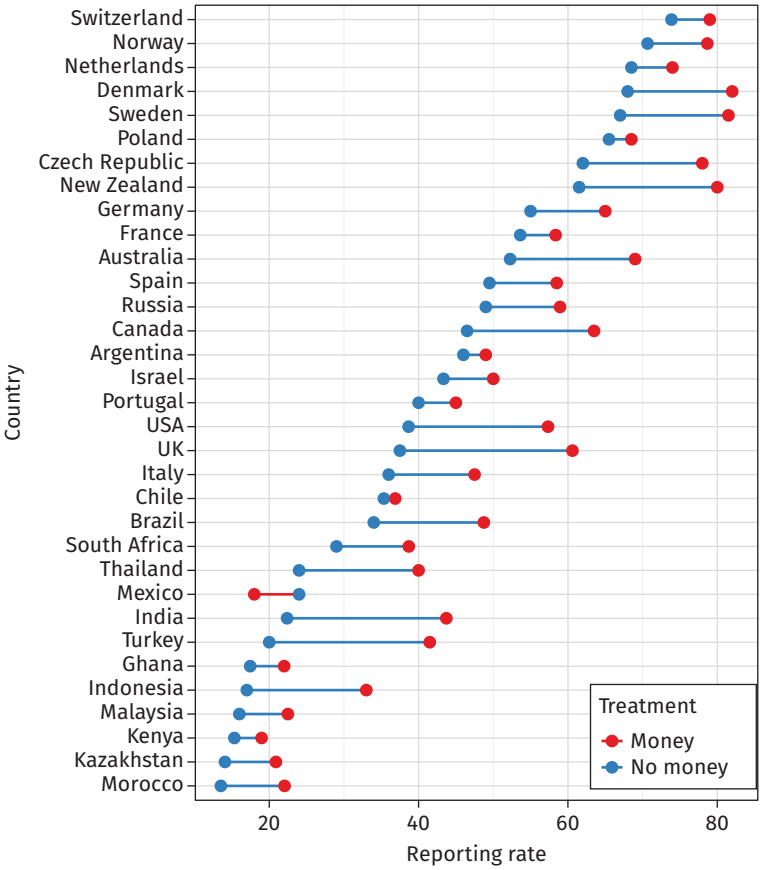
- The Ache (Ah-CHAY) hunter-gatherers in Paraguay share meat and honey equally among all group members. Ache Proposers contributed half of the 'pie' or more.
- Among the Lamalera whale hunters of Indonesia, who hunt in large crews and divide their prey among the entire community according to strict sharing rules, the average proposal was to give the Responder 58 percent of the pie.

Given the evidence from small-scale societies like the Lamalera and the Ache, we might ask whether we find other-regarding behavior in real-world situations elsewhere in the industrialized world. A different team of researchers were interested in exactly this question and designed an experiment that mirrors a real-life dilemma: what would you do if you found a wallet someone had lost: would you return it?

The team distributed a total of 17,303 "lost" wallets, some with money in them, some without, in 355 cities across 40 countries.[15]

Using transparent wallets with a business card, grocery list, key and cash, the researchers could check how many people contacted the "owner" of the wallet given in the email address listed on the business card to return the wallet.

Figure 2.14 Wallets were more likely to be returned to their owners when they contained money than when they did not. The "reporting rate" is the fraction of wallets that were "returned."



Before reading on, ask what you think would happen in your community: how many people would try to return the wallet? Would more people return the wallet if it had money in it, than if it did not?

The results of people's choices are shown in Figure 2.14. Though there are differences across countries, with just two exceptions among the 40 countries people were more likely to contact the "owner" if the wallet contained money ($13.45, the treatment) in it than if it did not ($0, the control). In a subset of cases—in the US, UK, and Poland—the researchers added a treatment with even more money in the wallet ($94.15). With a really substantial sum of money in the wallet, people were as likely, if not more so, to contact the listed email address on the business card in the wallet.

Keep in mind that the countries differ greatly in how much an additional $13.45 would make to a person's standard of living. Per capita income in the richest countries in the sample (Norway for example) is ten and even in

some cases 20 times the per capita income in others (Kenya for example), even when account is taken of the differing purchasing power of each national currency at domestic prices.

The evidence from both the Ultimatum Game and the wallet experiments suggests two important takeaways:

- *Culture matters*: people from different parts of the world live by different social norms and mutual expectations—what we call "culture." People from different cultures differ in what they consider fair offers and whether they think it's acceptable to make a self-regarding offer. They also differ substantially in whether they will return a lost wallet.
- *People are similar in many important respects*: people across the world have other-regarding motives including altruism, fairness, and reciprocity. In the "lost wallet" experiment in most countries a substantial fraction of people attempted to return the wallet.

CHECKPOINT 2.14 Not just for the money Why do you think that wallets with money in them (in some cases a substantial amount) were more likely to be returned to their owners than wallets without money?

2.15 SOCIAL PREFERENCES ARE NOT "IRRATIONAL"

People sometimes think of other-regarding and ethical preferences as something special—different from the taste for ice cream, for example—and requiring a model different from the preferences, beliefs, and constraints approach. But the desire to contribute, to punish those who do not free ride on others' contributions, and otherwise to act on the basis of social preferences, like the desire to consume conventional goods and services, can be represented by preferences that conform to standard definitions of rationality.

What we know from experiments is that whether it's ice cream or contributions to the public good, people respond to trade-offs, taking account of the costs and how much they value the activity in question: the higher the cost of helping others, the less its frequency. In other words, other-regarding preferences are consistent with rationality, namely consistency (transitivity) and completeness.

Researchers tested the rationality of seemingly altruistic choices by asking 176 subjects to play a version of what is called the Dictator Game.[16] One player (the Dictator), Alice, is given a sum of money by the experimenter, and asked to transfer whatever proportion of the money that she wishes to an other (anonymous) subject, Bob. Alice is told that that for every dollar that Bob receives from her, she will have to pay p dollars. So p is the price of altruism: how much she has to pay for every dollar that Bob gets. After Alice makes her decision, the money is transferred, and the game is over.

In this experiment, 75 percent of the Dictators gave away some money, demonstrating altruistic preferences. The average amount given away was a quarter of the endowment when the price $p = 1$ (a dollar-for-dollar transfer).[17] However, the higher the price of generosity, the less money was transferred. For instance, when each dollar transferred to Bob cost Alice two dollars ($p = 2$), only 14.1 percent of the endowment was given away on average, and when each dollar transferred cost four dollars, only 3.4 percent of the dictator's endowment was transferred. The higher the price of altruism, the less did Alice "purchase."

✓ **FACT CHECK** In a Public Goods Game with Punishment experiment researchers found that the level of punishment that subjects inflicted on others was less when each dollar subtracted from the payoffs of the target cost more in foregone payoffs to the punisher.[18]

It may be, as the old saying goes, that "virtue is its own reward." But that does not mean that people will act virtuously no matter what the price. This finding is perfectly consistent with the fact that people respond to the price of virtuous behavior just as the preferences, beliefs, and constraints model predicts.

CHECKPOINT 2.15 **Dictator Game?** Is the Dictator Game a game? Think about how we've defined games (check back to Chapter 1 if necessary).

2.16 APPLICATION: THE LAB AND THE STREET

Do people behave in the real world the way they do in experiments? The experimental evidence for reciprocity or related forms of other-regarding behavior would not be interesting if was not matched by similar behavior outside the lab. We therefore need to check whether laboratory evidence is **externally valid**, that is, consistent with behavior observed outside of the laboratory in similar circumstances to those found in the lab. External validity is particularly important for policy questions because policymakers and governments need to know whether a policy will work outside of the controlled conditions of the laboratory.

✓ **FACT CHECK** In an experimental game about trust and reciprocity played by groups of students and groups of chief executive officers of Costa Rican businesses, the businessmen were both more trusting of others and also reciprocated the generosity of their game partners to a far greater degree than did the students. Based on existing experimental evidence, students are not particularly other-regarding.[20]

Generalizing directly from experiments to behavior in other contexts is often unwarranted. For example, in the Dictator Game typically more than 60 percent of the Dictators allocate a positive sum to the recipient, and the average given is about a fifth of the endowment.[19] But we would be sadly mistaken if we predicted on the basis of this experimental result that 60 percent of people would spontaneously give money to an anonymous person passing them on the street, or that the same subjects would offer a fifth of the money in their wallet to a homeless person asking for help.

Many researchers have asked whether behavior in lab experiments predicts behavior outside the lab.

Along the coast of northeastern Brazil, for example, people catch shrimp in large plastic bucket-like contraptions. The shrimpers cut holes in the

EXTERNAL VALIDITY Results of experiments or other scientific research that can be generalized to circumstances outside (external to) the laboratory or other setting in which the research was produced are said to be externally valid.

bottoms of the traps to allow the baby shrimp to escape, thereby preserving the stock of shrimp for future catches.

The shrimpers face a real-world coordination problem: the expected income of each would be greatest if he were to cut smaller holes in his traps (increasing his own catch) while others cut larger holes in theirs (preserving future stocks). In Prisoners' Dilemma terms, small trap holes are a form of defection that maximizes the individual's material payoff irrespective of what others do (it is the dominant strategy if the shrimper is self-regarding). But a shrimper might resist the temptation to defect if he were both public spirited toward the other fishers and sufficiently patient to value the future opportunities that they all would lose were he to use traps with smaller holes.

Figure 2.15 A shrimping bucket with holes in it.

Photo from Fehr, E., and Leibbrandt, A. (2011), "A field study on cooperativeness and impatience in the Tragedy of the Commons," *Journal of Public Economics*, Vol. 95 (9–10), October, pp. 1144–55.DOI: https://doi.org/10.1016/j.jpubeco.2011.05.013. Reproduced with permission from Elsevier.

Economists Ernst Fehr and Andreas Leibbrandt implemented both a Public Goods game and an experimental measure of impatience with the shrimpers. They found that the shrimpers with both greater patience and greater cooperativeness in the experimental game punched significantly larger holes in their traps, thereby protecting future stocks for the entire community.[21]

Additional evidence of external validity comes from a set of experiments and field studies with 49 groups of herders of the Bale Oromo people in Ethiopia, who were engaged in forest-commons management. Economist Devesh Rustagi and his coauthors implemented public-goods experiments with a total of 679 herders, and also studied the success of the herders' cooperative forest projects.[22]

The most common behavioral type in their experiments, constituting just over a third of the subjects, were reciprocators who responded to higher contributions by others by contributing more to the public good themselves. The authors found that groups with a larger number of reciprocators were more successful—they planted many more trees—than those with fewer reciprocators.

CHECKPOINT 2.16 Masks in a pandemic: Not just a game During the COVID-19 pandemic that began in 2000, public health experts advised (and some governments required) people to wear face masks when in public places. The masks were more effective in preventing the mask wearer from infecting others than in protecting the wearer themselves. People found it somewhat uncomfortable to wear a mask.

a. Suppose there are just two people, and that both are entirely self-regarding (they care only about their own comfort and health). Write down a payoff matrix for the two strategies: Wear (the mask) and Don't.

b. What kind of game is this?

c. Write down a payoff matrix in which both playing Wear is a Nash equilibrium and so is both playing Don't.

d. What kind of game is that?

2.17 APPLICATION: A FINE IS A PRICE

How might a policymaker or CEO of a business make use of the fact that people care about what happens to others and they value behaving ethically?

Think about a set of rules for compensating employees. The rules typically specify pay and provision for time off, sick days, and the like. But problems arise with using purely material incentives to influence how people behave. Here is an example.

Having noticed a suspicious bunching of sick call-ins on Mondays and Fridays, the Boston Fire Commissioner on December 1, 2001 ended the Department's policy of unlimited paid sick days. Instead, the Commissioner imposed a 15-day sick day limit. The pay of firefighters exceeding that limit would be cut. The firefighters responded to the new incentives: those calling in sick on Christmas and New Year's Day increased ten times over the previous year's sick days.

The Fire Commissioner retaliated by canceling their holiday bonus checks. The firefighters were unimpressed: the next year they claimed 13,431 sick days; up from 6,432 the previous year.[23]

Many of the firefighters, apparently insulted by the new system, abused it, or abandoned their previous ethic of serving the public even when injured or not feeling well. In the language of the Ultimatum Game, they responded reciprocally to an offer they disliked by rejecting it. They were trying to *punish* the Commissioner at a cost to themselves.

The Commissioner's difficulties are far from exceptional.

Consider the following experiment in Haifa, Israel.[24] Parents everywhere are sometimes late in picking up their children at daycare centers. Uri Gneezy and Aldo Rustichini wanted to understand whether fining parents if they were late would result in parents arriving on time. So they implemented an experiment in a set of daycare centers.

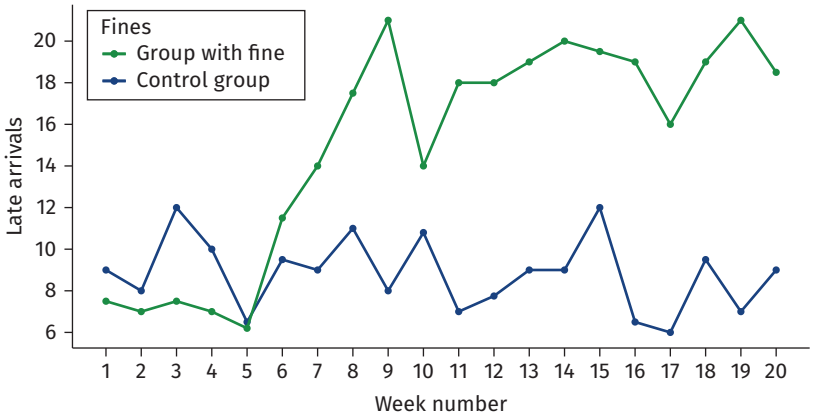
- *Treatment*: At six randomly chosen daycare centers, a fine was imposed for parents picking up their children late.
- *Control*: In a control group of daycare centers no fine was imposed.

Researchers expected parents to arrive on time because of the fine. But parents responded to the fine by arriving late more often: the fraction of parents picking up their kids late more than doubled. When the fine was taken away after 16 weeks, the parents continued to arrive late, showing no tendency to return to the status quo prior to the experiment. Over the entire 20 weeks of the experiment, there were no changes in the degree of lateness at the day-care centers in the control group.

The researchers reason that the fine was a contextual cue, unintentionally providing information about appropriate behavior. The effect was to convert lateness from the violation of a social norm or obligation that the parents were to respect, to a choice with a price that many were

Figure 2.16 The effect of a fine for lateness in Haifa's daycare centers.

Source: Gneezy and Rustichini (2000a). The fine was imposed in week 5 and retracted in week 17.



willing to pay. They titled their study "A Fine Is a Price" and concluded that imposing a fine labeled the interaction as a market-like situation, one in which parents were more than willing to buy lateness for money. Revoking the fine did not restore the initial context.

When monetary incentives undermine social preferences as they did among the Boston firefighters and Haifa parents, this is called **motivational crowding out**. These two cases are cautions that the use of monetary incentives may be inappropriate where the targets of the incentives are motivated by other-regarding preferences. But they are not reasons to think that incentives are ineffective, as we will see in many examples to follow. We have no doubt that had the fine for lateness in Haifa been 500 Israeli new shekels rather than just 10, the parents would have found a way to pick up their kids on time.

CHECKPOINT 2.17 Crowding out Why do you think the parents of children in the treatment group (with the fine) in Haifa continued arriving late to pick up their kids after the fine was discontinued?

2.18 COMPLEXITY: DIVERSE, VERSATILE, AND CHANGEABLE PEOPLE

The experimental and observational evidence suggests that an adequate understanding of preferences should recognize four aspects in human social behavior.

MOTIVATIONAL CROWDING OUT Motivational crowding out occurs when monetary or other material incentives or attempts to control someone diminish that person's other-regarding or ethical preferences or intrinsic motivation.

- *Diversity*: people differ in their preferences both within populations and across cultures.
- *Versatility*: a single person has a diverse set of preferences, and which of these is salient for making a decision depends on the situation, for example, when shopping as opposed to when spending time with friends.
- *Changeability*: people learn new preferences—often unwittingly—under the influence of their experiences.

These three aspects of our preferences contribute to a fourth attribute of how human beings interact:

- *Complexity or "the whole is not the sum of its parts"*: the outcome of an interaction of many people cannot be deduced in any simple way from the characteristics of the individual people involved.

Diversity

What motivates people differs, both locally and across different cultures and across time. Using data from a wide range of experiments, researchers estimate that between 40 and 65 percent of people exhibit other-regarding preferences of some kind. The same studies suggest that between 20 and 35 percent of the subjects exhibit conventional self-regarding preferences.[25] The authors of another study (in the US) termed 29 percent of their experimental subjects as "ruthless competitors" (presumably resembling *Homo economicus*) and 22 per cent as "saints."[26]

Versatility

A common observation about human behavior made by psychologists is that the same person can act differently depending on the situation. As a result, we say that people are *versatile*: we change how we act in response to what our situation seems to require of us, for example, being self-regarding while shopping and other-regarding with one's neighbors.

In the Ultimatum Game, people randomly assigned to the role of Proposer often offer amounts which maximize their expected payoffs given how likely low offers are to be rejected. But people randomly assigned to be a Responder rarely act in ways that maximize their payoffs. If they did, they would never reject a positive offer. The fact that in the role of Proposer people are more like "ruthless competitors" while people from the same culture in the role of Responder are more like "saints" is evidence of our versatility.

✓ **FACT CHECK** In experimental games about dishonesty, people who grew up in Communist Party–ruled East Germany are more likely to cheat than those who grew up in West Germany.[27]

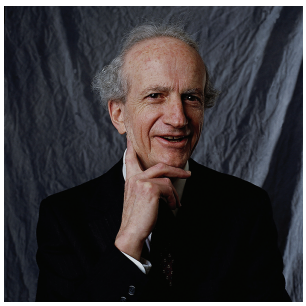
Changeability

Some preferences are part of our genetic makeup, having a taste for sweet and fatty foods, for example.

But most preferences are *learned* rather than *given* by our genetic inheritance. Durable changes in an individual's evaluations of outcomes

Figure 2.17 Gary Becker (1930-2014) was a professor of economics and sociology at the University of Chicago for four decades. In 1977 he coauthored an article "De Gustibus Non Est Disputandum" the Latin expression usually translated in English as "there's no accounting for tastes" in which he and his coauthor George Stigler analogized preferences to "the Rocky Mountains—both are there, will be there next year too, and are the same to all men." The book he published more than two decades later was *Accounting for Tastes*, in which he analyzed how preferences change. He was awarded the Nobel Prize in economics for contributions to our understanding of marriage, crime, politics, discrimination, and other aspects of social interactions.[28]

Photo by Business Picture/ullstein bild via Getty Images.



often take place as a result of experience. When this occurs we say that **preferences are endogenous**. This means that they change as a result of influences such as where a person lives, how they make their living or the rules of the game that govern how they interact with others. By contrast, when preferences do not change or change only as a result of changes occurring external to the interactions being studied, preferences are termed **exogenous.**

Over a lifetime or even generations, migrants to a new country, or those moving from a rural to an urban area often adopt new preferences (for example concerning food tastes). The fact that preferences are learned may account for the fact that, as we saw from the experiments in small-scale societies, people who hunt large animals tend be generous with the meat they acquire; and they seem to generalize these habits to other realms of life.

A consequence: Complexity

In everyday language the word "complexity" refers to the state of being intricate or complicated. The term is used in quite a different way in the study of interactions of a large number of independent entities–whether particles or people. A complex system is one for which the results of these interactions for the system as a whole cannot be predicted in any simple way from even the most detailed knowledge of the interacting entities. The economy is a complex system.

The best example of complexity in the social sciences is Adam Smith's invisible hand. What Smith suggested two and a half centuries ago, and modern economics has shown (as seen in Chapter 14) is that under some conditions uncoordinated interactions among entirely self-regarding total strangers through competition in markets among private property owners can (unwittingly) create an outcome that is better for all than many of the alternatives.

The idea of complexity is often expressed by the saying: "the whole is different from the sum of the parts." The key here is not that the "whole" may be greater or less than the sum; it is that *summing* the parts is not the right way to calculate the whole. Averaging the components of some interacting system will not give what their interactions will actually add

ENDOGENOUS PREFERENCES If one's experiences result in durable changes in preferences, then they are termed endogenous.

EXOGENOUS PREFERENCES Preferences are exogenous if they change in response only to influences external to the economy or at least outside of the economic subject matter under study.

up to. The results of the interaction for the whole—called their *emergent property*—may be surprising given the nature of the interacting entities.

Here are some examples of surprises (with which you are already familiar) in the properties that emerge from people with diverse and versatile preferences interacting.

- Small differences in the distribution of types of people—the presence in a population of people willing to punish those who do not contribute in a Public Goods Game, for example—can have large effects on how everyone behaves, getting self-regarding people to act as if they were cooperators. You have seen this in Figure 2.11.
- Seemingly small differences in institutions can make large and surprising differences in outcomes. Why did adding the punishment option so radically change the outcomes in the Public Goods Game? We know that cooperation—contributing to the public good—unravels in the absence of the punishment option. But the incentives to punish would seem identical to the incentives to contribute to the public good in the first place: everyone would like someone else to bear the cost of punishing the free riders. So not contributing and not punishing should be the dominant strategy in this game. But we now know that that is not what we observe.
- While imposing a fine or other cost on socially undesirable behaviors may create socially desirable outcomes in certain circumstances such as getting people to stop using plastic grocery bags, a fine on parents arriving late to pick up their kids backfired. We saw that the nominal fine decreased parents' willingness to pick up their children on time perhaps because they viewed the fine as a price: the fine changed what they viewed as socially acceptable behavior.
- Letting a self-regarding player be the first mover in a Prisoners' Dilemma Game when she knows that the other player has strong reciprocity motives can avert the coordination failure resulting in mutual cooperation. Letting the Reciprocator be the first mover would have the opposite result: both players would defect, resulting in the Pareto-inefficient outcome. You can confirm this by doing the checkpoint below.

CHECKPOINT 2.18 Sequential Prisoners' Dilemma For a sequential Prisoners' Dilemma Game where the first player is known to be self-regarding and the second player is known to be reciprocal draw a game tree in which the Nash equilibrium may be (Cooperate, Cooperate) and explain why it could occur.

2.19 CONCLUSION

Recognizing the complexity of social interactions makes it harder to reach simple conclusions about the economy. But this is a good thing, not a shortcoming of the approach we have outlined; a feature, not a bug.

We have introduced the preferences, beliefs, and constraints approach and showed how games can help us understand the coordination problems that communities of people face. Examples are poverty traps that occur in Assurance Games, and the under-provision of public goods such as a sustainable environment in the Public Goods Game.

We also showed how changing the rules of the game can sometimes avert or mitigate a coordination failure. Examples include introducing the possibility of leadership by letting the Assurance Game be played sequentially, and introducing the option of peer punishment of low contributors in the Public Goods Game.

Finally, the preferences, beliefs, and constraints approach and game theory are the basis of experiments that allow us to study preferences empirically with some surprising results. Included is the finding that in most populations studied many people are not entirely self-regarding but are also other-regarding, caring for better or worse about how their actions affect other people. Among the preferences the experiments have identified are: altruism, fairness, reciprocity, and spite (or "us versus them").

A key concept introduced in this and the previous chapter is the Nash equilibrium based on the idea of a best response. The choices we have posited for our actors have been overly simplified: Contribute to the public good or Don't Contribute, Accept or Reject the Proposer's offer in the Ultimatum Game.

The preferences, beliefs, and constraints approach is capable of a far more realistic view of the strategy sets open to us allowing us to contribute some or a lot, for example. But to benefit from this we need to develop the mathematical tools of constrained optimization. We take up this task in the next chapter.

MAKING CONNECTIONS

Preferences, beliefs, and constraints: This framework for analyzing decisions will be used throughout the rest of the book.

Risk and uncertainty: Many, maybe most, of the important decisions that people make are risky because the resulting outcome depends on something occurring in the future that is not known.

The rules of the game and coordination problems: Sequential rather than simultaneous play may result in a better outcome in an Assurance Game (or even a Prisoners' Dilemma). The reason is that the leadership exercised by the first mover can help to coordinate play in the game. Another example: allowing other players to punish low contributors in a Public Goods Game dramatically changes the outcome. Leadership (first-mover advantage) may also benefit the leader at the expense of the follower.

External effects and Pareto-inefficient Nash equilibria: The Public Goods Game illustrates an extreme form of positive external effects (each person's contribution benefits everyone equally).

Evidence: Economists have recruited novel experimental evidence—from the laboratory and the field—to examine our theories about how people behave. Economists have used the evidence to modify and improve existing models and to develop entirely new models of how people behave.

Diversity/heterogeneity: People differ in their preferences (self-regarding, other-regarding) and in the advantages associated with their positions (first mover, second mover).

Inequality: In part as a result of the way that the rules of the game confer differential advantages on people, the mutual gains made possible by an interaction are often unequally shared (e.g. being first mover in the Language Game).

IMPORTANT IDEAS

preferences	beliefs	constraints
rationality	self-interest	social preferences
fairness	altruism	reciprocity
spite	endogenous preference	exogenous preference
institutions	first-mover advantage	external validity
laboratory experiment	field experiment	endowment
poverty trap	crowding out	versatility
Ultimatum Game	Public Goods Game	learning
complexity	inequality aversion	game tree
diversity/heterogeneity	changeability	Dictator Game
normal form	extensive form	

MATHEMATICAL NOTATION

Notation	Interpretation
i	a contingency
P	probability that a contingency will occur
$\pi(\ )$	a player's payoff
$E(\ ), \hat{\pi}$	a player's expected payoff
$E(\pi(x\|i))$	Expected payoff to an action (x) conditional (|) on a contingency (i)
z	individual endowment in Public Goods Game
e^i	individual contribution in Public Goods Game
M	return factor (productivity of contribution) in Public Goods Game
n	number of participants in Public Goods Game
p	price of altruism in Dictator Game

Note on superscripts and subscripts: i: an individual.

CHAPTER 3

DOING THE BEST YOU CAN

CONSTRAINED OPTIMIZATION

"What a useful thing a pocket-map is!" I remarked.

"That's another thing we've learned from your Nation," replied Mein Herr, "map-making. But we've carried it much further than you."

"What do you consider the largest map that would be really useful?"

"About six inches to the mile."

"Only six inches!" he exclaimed.

"We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of a mile to the mile!"

"Have you used it much?" I enquired.

"It has never been spread out, yet,"

"The farmers objected: they said it would cover the whole country and shut out the sunlight!"

Lewis Carroll,
Sylvie and Bruno Concluded (1893)

DOING ECONOMICS

This chapter will enable you to:

- See how the preferences, beliefs, and constraints framework from Chapter 2 forms the basis for mathematical models of economic behavior.
- Recognize how preferences—whether entirely self-regarding or altruistic—can be represented both in mathematical form (a utility function) and graphical form (an indifference map).
- Understand that constrained optimization is a method that economists use to explain the actions that people take; it is not a description of the thoughts or feelings making up individuals' decision-making processes (e.g. studied by a psychologist).
- Explain how people are constrained—for example by limited time—and how these constraints give rise to opportunity costs and, along with our preferences, to trade-offs.
- Use the preferences, beliefs, and constraints framework to analyze difficult policy-making choices, including how much of society's resources should be devoted to the abatement of environmental damages.
- Understand ordinal and cardinal utility, explain how they differ, and how cardinal utility provides a way to represent the societal cost of economic inequality.
- Understand the shortcomings and limits as well as the insights of these models.

3.1 INTRODUCTION: THE MAP AND THE TERRITORY

Lewis Carroll, the author of this dialogue (not to mention *Alice in Wonderland*) was also a mathematician and a philosopher. The point Carroll made about maps also goes for economic models. Maps are useful because they convey the necessary information, not because they are an *exact* representation of the territory, as the people from Mein Herr's country discovered. Carroll's point? The map is not the territory.

A good model is *not* reality, but it's a helpful guide.

What qualifies a map or a model as useful depends on what we need it for: six inches to the mile might be adequate for a map of hiking trails, but such a hiking map would not be much use to an airplane pilot. The same is true of economic models.

Think of a model as a lens. A good economic model is a way of focusing on what is important given the question that one wants to address without complicating the picture with things that do not matter for the question at hand.

A key component of many economic models—those using the preferences, beliefs, and constraints approach—is that we can understand the actions people take by assuming that they are doing the best they can under the circumstances that they are in.

When implemented using mathematical reasoning, this is called **constrained optimization**, a mathematical method by which we can determine a course of action that accomplishes a goal (reflecting a person's preferences), given the information that the person has (beliefs), and the actions they may feasibly take (a constraint.)

We illustrate both a model and the method of constrained optimization by something that matters to all of us: time, and how we use it.

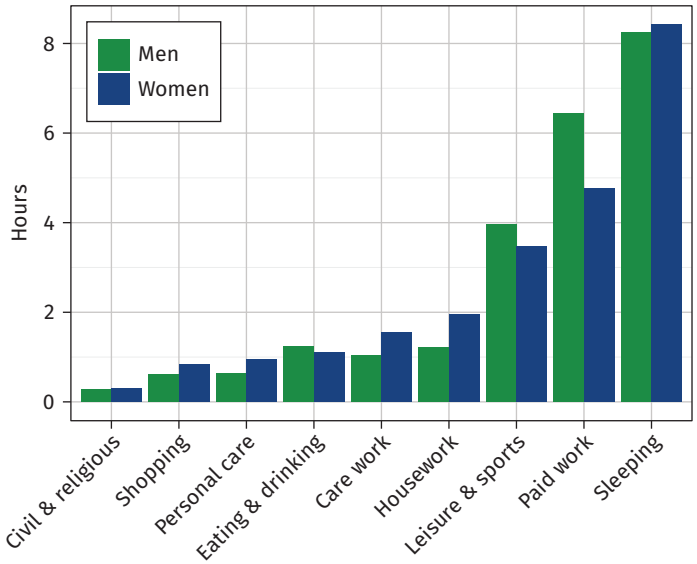
3.2 TIME: A SCARCE RESOURCE

Benjamin Franklin (1706-1790)—the American politician and inventor—once said, "Time is money." Franklin was referring to the presence of trade-offs in how people choose to spend their limited time. His three-word sentence is therefore a model with a simple message: it is illuminating to think about how people choose their daily actions to achieve their goals under the constraint of limited time in the same way that we think about how they spend their limited budgets.

CONSTRAINED OPTIMIZATION A constrained optimization problem is one in which a decision maker chooses the values of one or more variables to achieve an objective which is subject to a constraint that determines the feasible set of actions or outcomes.

Figure 3.1 Daily time use of American men and women. These data—for hours in each activity measured on the horizontal axis for US adults in 2013—differ from data restricted to those with small children, or retired people, or students.

Source: Hofferth et al. (2013).



Spending an hour or minute on an activity provides us value of some kind: we enjoy the activity itself (e.g. eating) or the results of the activity (e.g. being paid a wage with which we can buy our food). But, since time is limited, choosing one activity also means we *give up* that time to do something else. We incur a cost of doing an activity because we forfeit the value of the next best thing we could have spent our time on instead: this is the **opportunity cost** of our time.

Unless we have time to spare and are wondering how we will fill up our day, there is an opportunity cost to our use of time. As a result, we can model how we could best use our time by evaluating the benefits and costs (including opportunity costs) of pursuing one set of activities rather than another. To do this we use constrained optimization.

Before developing the concepts on which constrained optimization is based, let's look at the kinds of facts that a model of time use should be able to explain.

Figure 3.1 shows how men and women from the USA used their time each day during the year 2013. The largest time use is for the categories sleep, work (meaning for pay), leisure, and housework. Men and women differ

OPPORTUNITY COST Where x and y are both valued positively, the opportunity cost of x in terms of y is how much y a person must give up to get a unit more of x.

typically in the hours they devote to paid work and housework and care work, often reflecting differing social norms that they hold about the kinds of activities that it is "appropriate" or "natural" for men and women to do.

But these social norms also change, sometimes in ways that show that the differences in the distribution of work time between men and women are far from determined by "nature" but instead reflect changed economic conditions. During the second half of the twentieth century in the rich countries the fraction of women doing paid work outside the home dramatically increased. While we do not have detailed information like that shown in Figure 3.1 for the mid-twentieth century on how men and women spent their time, there almost certainly has been a decline in the total amount of time spent doing housework. Part of the change in the distribution of women's time between housework and work for pay is due to the availability at affordable prices of new technologies—household appliances—that reduced the amount of time required to clean the house, wash clothes, and to carry out other housework tasks. These appliances include washers, refrigerators, and vacuum cleaners, which in the US became common from the late 1940s onward, and dryers, dishwashers, and microwaves somewhat later.

EXAMPLE In this video (tinyurl.com/yyop8rzh) Juliet Schor addresses the question of why we work so hard (from the CORE project. www.core-econ.org).

FACT CHECK For a long historical view of why the washing machine was a "miracle" have a look at this video by Hans Rosling:

Evidence that these new technologies contributed to the change in the distribution of women's work time comes from a comparison across countries of increases in the fraction of women working outside the home—called the labor force participation rate—and decreases in the price of these labor-saving household appliances (compared to other prices).[1] The results are in Figure 3.2, which shows that in countries such as the US where the prices of these appliances fell the most, women's labor force participation rate rose substantially. By contrast, in Germany where prices of household appliances fell the least, the increase in labor force participation was half as great as in the US.

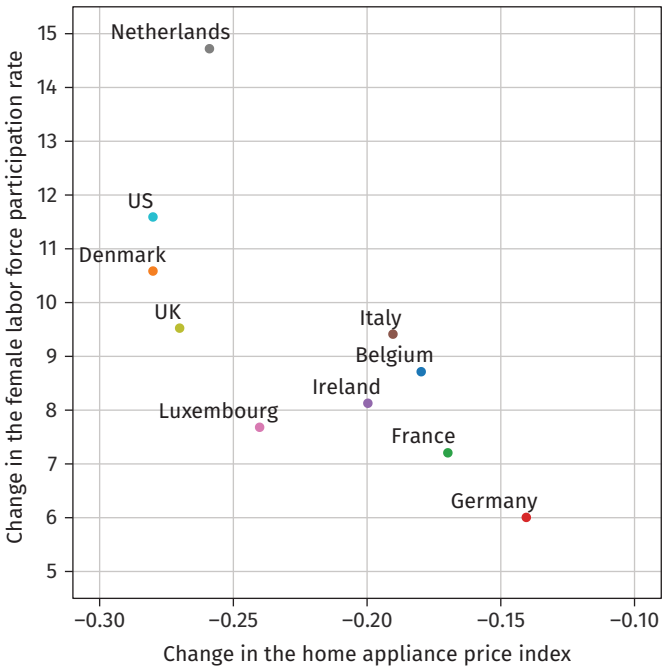
Other factors contributed, of course, most importantly the reduction in the number of children born per woman. But the fall in the prices of appliances, the study reporting these data concluded, was of approximately equal importance. It appears that economic changes—the new household appliances and their falling prices—changed how women spent their time—more working outside the home. This in turn may have been both a result and a cause for the changing social norms about "women's work" and the decreased adherence to the ideal of a family with a husband income earner and a wife raising (many) children and taking care of the home. This is an example of endogenous preferences, that is, preferences—for example, social norms about "women's work"—changing as economic conditions—the prices of home appliances—change.

REMINDER Preferences are endogenous if one's experiences result in durable changes in preferences.

We begin with these examples because methods of constrained optimization—the preferences, beliefs, and constraints framework—provide a way of posing and in some cases answering questions like: Why do men

Figure 3.2 The relative price of home appliances and the female labor force participation rate. The vertical axis represents the change in the percent of adult women working outside the home, termed the female labor force participation (FLFP) rate. The change in the home appliance price index is on the horizontal index. Notice that a *bigger* price decrease would be shown by a larger *negative change* (further to the left on the x-axis) so the US, Denmark, and the Netherlands had big *decreases* in the prices of home appliances and big *increases* in the FLFP rate. Household appliances—like TVs—that did not reduce the amount of time necessary to perform housework tasks are *excluded*. The data are for the period 1975 to 1999 and include those OECD countries for which data during the 1970s exist.

Source: de V. Cavalcanti et al. (2008).



and women spend the time they do on the various activities shown? Or why did work hours fall so dramatically in some countries over the twentieth century? (In Chapter 7 we use a constrained optimization model to provide one answer to that question.)

We begin with preferences, before turning to constraints later in the chapter. Because we are not considering strategic interactions or other situations in which the relevant facts are not known, we do not introduce beliefs into our modeling of constrained optimization until the next chapter.

! REMINDER Preferences represent the favorable (positive) or unfavorable (negative) feelings that could lead a person to choose one outcome over another. Included are tastes (food likes and dislikes, for example), habits (or even addictions), emotions (such as anger and disgust) often associated with visceral reactions (such as nausea or an elevated heart rate), social norms (for example, those that induce people to prefer to be honest or fair), and psychological tendencies (for aggression, extroversion, and the like).

CHECKPOINT 3.1 Labor-saving household appliances and women's labor force participation Think about a form of the family that was common in the countries shown in Figure 3.2 in the 1970s: in those days described as a male "breadwinner" (working for pay) and a female "housewife."

a. In that setting what was the *opportunity cost* of the time that the woman in the family spent working for pay outside the home?

b. Explain how the availability of lower priced and more effective household appliances changed the opportunity cost of women working for pay outside the home.

3.3 UTILITY FUNCTIONS AND PREFERENCES

In Chapter 1, we represented preferences—our evaluations of the outcomes our actions may bring about—as *payoffs*, that is numbers indicating how much the decision maker values each of the possible outcomes. We discussed, as an illustration, the choice of whether to take an umbrella or not, with a decision (Don't take the umbrella, It rains) resulting in a payoff of 3. The payoff to (Take the umbrella, It rains) was 15, meaning that if it rains the person valued having the umbrella by five times as much as not having it.

In that example we simplified things by limiting the actions and the outcomes to just a few, for example, it either rained or it did not. The simplification allowed us to focus on two-by-two payoff matrices with just four possible outcomes, or equally simple game trees.

But most of the economic interactions that we study are not that simple: we can contribute any amount to the public good (not just $10 or nothing), the farmers in Palanpur have the choice to plant a little bit earlier, or much earlier, and so on. Or, to return to the question of time: how we divide up our day among the activities in Figure 3.1 could be measured in variations of minutes devoted to each of the nine activities, giving us trillions of "outcomes" to choose from.

We need a way to represent preferences when there are a great many outcomes, without expanding our payoff matrices to the unusable size of the 1:1 maps in the Lewis Carroll fable at the beginning of the chapter.

Why we use utility functions to represent preferences

To do this we use a **utility function**, a mathematical expression that translates the full range of possible outcomes into a person's valuation of the outcome—her payoffs.

UTILITY FUNCTION A utility function is an assignment of a number $u(x,y)$, to every bundle (x,y) representing a person's valuation of that bundle. This means that if given the choice between two bundles (x,y) and (x',y'), the individual will choose the first if $u(x,y) > u(x',y')$.

The word "utility" (in ordinary language, "usefulness") is used to mean the same thing as "payoff." It is a number assigned to a particular bundle that has the property that when choosing between alternative bundles, a person will select the one with the highest (utility) number. Both "utility" and "payoff" sound like some monetary or other amount of something you take home as the outcome of a game. But in economics utilities, like payoffs, are not something you get or even experience. You don't take them home; they are nothing more than numbers that indicate the course of action you will take.

For simplicity, we call this number "how much the person values the outcome" but the utility function tells us nothing about *why* the bundle has a higher number. It could be any of the reasons for the collection of pro or con evaluations that make up our preferences for some bundle, ranging from food tastes and phobias, to addictions and ethical norms.

What the function allows us to do is to take account of more complex outcomes than "Don't take the umbrella" and "It rains." The decision maker, as before, will choose the actions the she believes will result in the highest utility outcome.

Suppose that our decision maker, Anmei ("Ahn-may," an Uber driver), is deciding how much time to work, x, and what fraction of the resulting income to spend on food, y. The utility function then assigns a number–the level of utility–to each possible combination of x and y, one of which, say, is work for 4 hours and 15 minutes and spend 35 percent of the resulting pay on food. Any other combination, say, work four hours and spend 40 percent of the resulting income on food, will also be assigned a number, representing Anmei's valuation of that particular outcome. The number assigned to the second bundle can be greater than, less than, or the same as the previous bundle depending on whether the first is preferred, or the second, or she is indifferent among them.

This assignment of numbers is a *utility function*, $u(x,y)$: for every outcome (x,y) the value of the utility function is the number representing a person's valuation of the outcome. If we know what combinations of x and y are available to Anmei based on the relevant constraints, then we can predict the choice Anmei will make, namely the combination with the highest utility (or if there were more than one bundle tied for highest, then one of these tied top bundles).

What do the utility numbers measure?

We measure how much a person values various outcomes in two ways, either:

- by indicating how valuable each is on some absolute scale; or
- by simply ranking them in order.

If Anmei compares two bundles (or outcomes), namely (x,y) and (x',y') with $u(x,y) = 3$ and $u(x',y') = 9$ there are two different statements we could make about Anmei, one much more informative than the other:

! REMINDER "Payoffs" in experiments Economists refer to the amount of money that a player in an experiment receives as her payoffs, as we did in Chapter 2. But as we saw in the Ultimatum Game and the Public Goods Game, many people do not select the strategy with the highest possible monetary gain. We often think of payoffs as some kind of material gain—like the quantity of fish caught or the amount of grain harvested—but remember the word payoff like utility is just a number indicating what the actor will choose, and this is often not adequately measured by material gain.

M-CHECK We read x' as "x prime" and x'' as "x double prime." We usually denote a bundle other than (x,y) as (x',y') to indicate a different composition of x and y.

! REMINDER Consistency (or transitivity) requires that when considering three bundles (x,y), (x',y'), and (x'',y''), if (x,y) is preferred to (x',y') and (x',y') is preferred to (x'',y''), then (x'',y'') *cannot* be preferred to (x,y).

! REMINDER Completeness requires that all possible outcomes can be ranked. For any two bundles (x,y) and (x',y') either the person prefers (x,y) to (x',y'), or the person prefers (x',y') to (x,y), or the person is *indifferent* between (x',y') and (x,y).

- Anmei values (x', y') three times as much as (x, y) and
- Anmei values (x', y') more than (x, y)

In the first case above, utility is a number indicating *by how much* Anmei prefers (x', y') to (x, y). Utility is therefore called a **cardinal** measure (cardinality in mathematics refers to the *size* of something). In Chapter 2 we represented people's preferences by the payoffs associated with particular bundle of games like (x', y') or (x, y). When we defined the expected payoffs to some course of action we added up the payoffs of each possible outcome (weighting them by the probability of each outcome occurring). Doing this required that utility is a measure of size. The numbers representing payoffs and expected payoffs in Chapter 2 are cardinal utilities.

In the second case the utility function gives us an ordering of better-worse for the pair of outcomes. When the utility function is measured in this way, we say that Anmei has **ordinal** preferences or that utility is ordinally measured. Ordinal utility says nothing about how much better the preferred outcome is.

Instead of assigning numbers to the outcomes, in the case of ordinal utility, it would be clearer if we just assigned ranks, like instead of 1, 2, 3, 4, and so on, we used 1st, 2nd, 3rd, 4th (and in cases of indifference: for example, tied for 7th). In the cartoon figure about the Planting in Palanpur game (Figure 1.3), we listed the four possible outcomes as "Best, Good, Bad" and "Worst": this is an example of ordinal utilities.

If utility is just an ordering, there is no way that we can say that the top-ranked bundle is twice as good as the second-ranked bundle or ten times as good as the tenth-ranked bundle. Nor could we add up the ranks, saying, for example, that getting your second-ranked bundle and your third-ranked bundle with equal probability is as good as getting your first- and fourth-ranked bundle with equal probability. None of these statements make any sense. This is why when dealing with decisions involving risk, we used a cardinal measure.

CARDINAL UTILITY A cardinal utility function assigns a number to each bundle, such that, with a cardinal utility function, $u(x, y) = 10u(x', y')$ means that (x, y) is preferred ten times as much as (x', y').

ORDINAL UTILITY Let $a \succ b$ mean "a is preferred to b." An ordinal utility function ranks bundles, e.g. $(x, y) \succ (x', y') \succ (x'', y'')$, without specifying how much (x, y) is preferred to (x', y') or (x', y') is preferred to (x'', y''). The assignment of numerical utilities representing ordinal preferences is meaningful only to express an ordering: $u(x, y) > u(x', y')$ implies only that the first bundle is preferred to the second but not by how much.

So, for example, when we introduced the Palanpur farmers' uncertainty about when the other farmer or farmers would plant their crops, calculating their expected payoffs required adding up the values that each farmer attaches to an outcome. Because you cannot add up ordinal measures, we gave the payoffs numeric values (the numbers in the payoff matrix) representing cardinal utility.

For some questions in economics the ordinal–better or worse–meaning of utility is all we need to understand and predict the actions that people will take. But in many situations, those involving risk and uncertainty, as we have just seen, or in evaluating the effects of differing rules of the game–policies to ensure competition in markets or concerning fairness, for example–addressed in section 3.13, the cardinal measure is required.

M-CHECK For simplicity, we generally restrict our analysis to outcomes that can be described in terms of two variables x and y, though it is straightforward to generalize this model to outcomes described by more than two variables. The actor therefore makes choices among "bundles" that combine different amounts of x and y.

CHECKPOINT 3.2 **Utility and payoffs** Give examples of preferences that might lead people to act in ways that they would regret.

3.4 INDIFFERENCE CURVES: GRAPHING PREFERENCES

Indifference curves are a useful way to visualize a person's preferences. We will illustrate the concept of an indifference curve by Anmei, who is choosing among differing amounts of kilograms of coffee (x) and gigabytes of data on her cell phone (y).

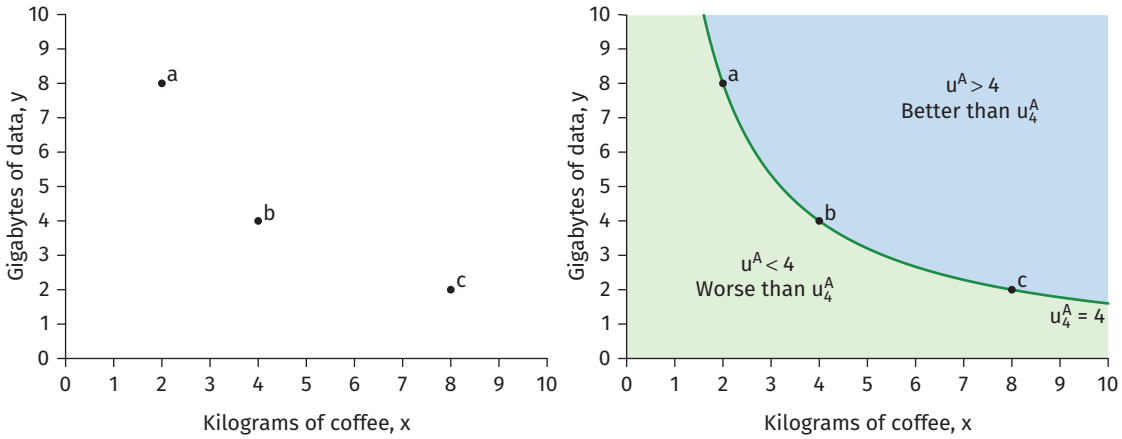
Every point given by the coordinates (x,y) in Figure 3.3 (a) is a pair of the quantities of the two goods, called a **bundle**. Points **a, b**, and **c** therefore represent three bundles of differing amounts of coffee and data. Suppose that Anmei ranks the points **a**, **b**, and **c** equally–she is indifferent among the three bundles–then these three points lie on the same indifference curve, as shown in Figure 3.3 (b). Her indifference curve represents the combinations of bundles among which she is indifferent. This means that for either bundle **a**–8 gb of data and 2 kg of coffee–or bundle **b**–4 gb of data and 4 kg of coffee–or bundle **c**–2 gb of data and 8 kg of coffee–$u(2,8) = u(4,4) = u(8,2) = 4$.

Figure 3.3 (b) shows the indifference curve made up of all bundles for which Anmei's utility is equal to 4. Her indifference curve is labeled by a u with a subscript which represents the level of *utility* that is the same for all points on that indifference curve. Anmei prefers to consume *more* of both

INDIFFERENCE CURVE The points making up an individual's indifference curve are bundles—indicated by (x,y), (x',y'), and so on—among which the person is indifferent, so that $u(x,y) = u(x',y')$ and so on.

BUNDLE A bundle is a list of an individual's goods (or bads).

Figure 3.3 One of Anmei's indifference curves: coffee and data. The dark-green indifference curve u_4^A represents all the combinations of x and y that provide Anmei (A) with the same level of utility, 4. The blue area above and to the right of Anmei's indifference curve shows combinations of the amounts of coffee and data that provide her with utility greater than 4. The light-green area beneath her indifference curve shows the bundles of x and y that she values at less than 4. She would therefore rather choose a combination of x and y on the indifference curve shown than any point to the left or below it.



(a) Consumption bundles

(b) An indifference curve

data and coffee, so she would like to be anywhere in the blue-shaded area where her utility would be *greater* than 4. She would rather not consume *less* of both data and coffee, so she would not like to be down in the area shaded in green where her utility would be *less* than 4.

The single indifference curve shown in Figure 3.3 (b) divides the space of all possible bundles of x and y into three categories: bundles that are respectively better or worse than any of the bundles making up u_4 and bundles that are equally valued with a utility of 4.

To predict the action that will be taken by a person in some given situation, we proceed in four steps:

- *Step 1*: In this and the next section we use many such indifference curves–her indifference map–to evaluate all of the bundles that she could consider; we can do this because her utility function assigns a utility number to each bundle.
- *Step 2*: In section 3.7 we then limit the decision maker's choices to those that are feasible for her (that is, choices that are actually open for the decision maker to take).

- *Step* 3: Putting steps 1 and 2 together, we use the evaluations in Step 1 to rank all of the feasible outcomes, showing us the one the decision maker ranks the highest.
- *Step* 4: We conclude that she will select the bundle identified in the previous step.

Figure 3.4 shows three indifference curves, u_3, u_4, and u_5, part of Anmei's indifference map. Anmei prefers more of both goods–that's why they are called "goods." Therefore, indifference curves to the upper right, like u_5, are higher (corresponding to the blue-shaded area in Figure 3.3). Indifference curves representing less preferred combinations, like u_3^A are to the lower left (corresponding to the green-shaded area in Figure 3.3). Of the three indifference curves plotted on the indifference map of Figure 3.4, u_3^A provides Anmei with her lowest utility, whereas u_5^A provides Anmei with her highest utility. A different person, one who valued coffee more than Anmei would have a different indifference map.

If you think of her indifference curves as a kind of contour map, Anmei can be pictured standing somewhere on a mountain wanting to get to the top. She might, for example, be in the lower-left corner of the contour map of a hill shown in Figure 3.5 wanting to reach the 800-meter-plus top of the hill.

Figure 3.4 An indifference map for kilograms of coffee, x, and gigabytes of data, y. The quantity of good x is on the horizontal axis and the quantity of good y is on the vertical axis. Three indifference curves are shown: u_3^A, u_4^A, and u_5^A, where the rank of the utilities is $u_5^A > u_4^A > u_3^A$. The constant level of utility for $u_4^A = 4$. Points **a**, **b**, and **c** all lie on u_4^A and give Anmei the same utility of 4. Point **d** would give Anmei lower utility and point **e** would give Anmei higher utility (because every bundle is associated with some utility number, we could draw indifference curves through those points, and through any point in the figure).

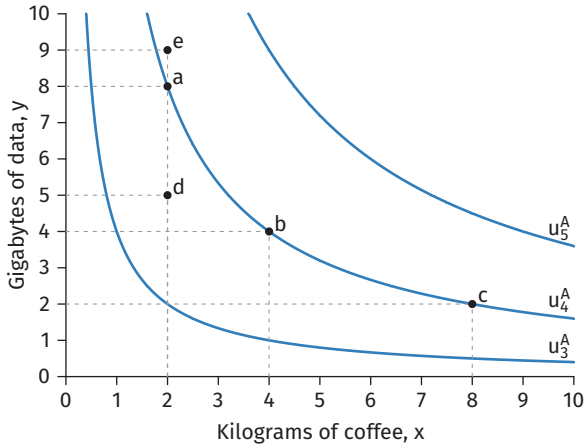
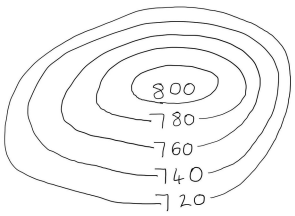


Figure 3.5 A contour map of a hill showing altitudes. Indifference curves are similar to contour lines, which are composed of all the points in the landscape which are at the same altitude. The lower-left quarter of the contour map resembles the indifference map in Figure 3.4.



Her utility is the altitude where she is standing, say, at a point on the 720 meters above sea-level contour. Her indifference curves are the numbered contour lines on a map of the mountain she is climbing, each indicating locations on the mountain the same height above sea level.

A map, as the quotation at the beginning of this chapter reminds us, is a representation of territory. The territory represented by Anmei's indifference map is her evaluation of all possible outcomes she might experience. An indifference curve runs through every point in the (x,y) plane, but just like maps that could not possibly show *every* contour line, we can plot only a selected number of them in any case.

Anmei wants to climb as high as she possibly can up the utility mountain, given whatever limitations she faces, including her own physical capacities and possibly impassible cliffs blocking her way. As Anmei advances up the mountain, she crosses contour lines, moving from lower-to-higher indifference curves. She is engaging in a constrained optimization process.

CHECKPOINT 3.3 Maps, points, and bundles Sketch your own version of the indifference map in Figure 3.4. Add two new points to your graph:

a. A bundle, labeled **f**, where Anmei holds the same amount of y as she does at point **b,** but Anmei prefers bundle **b** to **f**.

b. A bundle, labeled **g**, where Anmei holds the same amount of y as she does at bundle **b**, but which Anmei prefers to bundle **b.**

c. Explain why the following is true: consistency of preferences implies that indifference curves cannot cross. Draw two *intersecting* indifference curves and label points on them that enable you to show that these points violate the consistency assumption. Hint: where $(x,y),(x',y'),(x'',y'')$ are bundles you will need to show something like $(x,y) \succ (x',y') \succ (x'',y'')$ but $(x,y) \not\succ (x'',y'')$ where $\not\succ$ means "is not preferred to."

3.5 MARGINAL UTILITY AND THE MARGINAL RATE OF SUBSTITUTION

Indifference maps are used to summarize the values that an individual places on differing bundles of goods. But "goods" go beyond things like Anmei's coffee or data. Goods can be anything a person values, such as free time. (Indifference curves, as we will show later in the chapter, can also summarize the preferences people have about "bads" such as environmental degradation, that, unlike goods, are things that people would prefer to avoid.)

To see this, we will move from the choice about coffee and data, and think instead about a new person, Keiko (KAY-i-ko), who is a student making a choice about the use of her time. One decision she has already made is that she will sleep eight hours every night, so she has 16 remaining hours of the day that she will use in some way. As Keiko progresses through her studies (no doubt fueled by coffee and using data), she has two important priorities, which she thinks of as "Living" and "Learning."

- *Learning* comprises all the aspects of her life as a student that contribute to her goals of becoming an educated person and becoming qualified for an interesting career.
- *Living* comprises everything else, including keeping up with friends, meeting new people, and taking care of herself.

As there are only so many hours in a day, and because Learning takes time, Keiko faces a **trade-off** between Learning and Living, the more she has of one the less she will have of the other. So she is facing a constrained optimization problem.

We explain in section 3.15 that constrained maximization is not a description of the mental and emotional processes by which a decision maker adopts one course of action over another. It is a research strategy that economists use to understand what people do, not how they come to do it. To illustrate the method we will suppose that Keiko consciously maximizes her utility function subject to her only-24-hours-in-the-day constraint, by comparing the utility associated with each of the combinations of Learning and Living that are open to her. (OK: only a student in economics would actually do this!)

Keiko is a systematic and quantitatively oriented person, and decides to measure her Learning quantitatively with a number. In calculating her Learning, she takes account of feedback from her teachers, such as grades (marks), but also evaluates this feedback in terms of her own estimation of how much she has learned, such as how much her study improves her writing skills and general understanding.

Keiko measures the amount of Living by the hours she can spend *not studying*, x, and the amount of her Learning by her personal rating, y.

Key to how the preferences, beliefs, and constraints approach works is the fact that for most of the things that we may value, if we have little of it, we highly value having more of it, but the more of the thing we have,

TRADE-OFF A trade-off is a situation in which having more of something desired (a "good") requires having less of some other "good" or more of something that the actor would like to have less of (a "bad").

✓ **FACT CHECK** Diminishing marginal utility in economics is often based on the psychological principle of satiation of wants, which states that satisfying our wants is pleasurable, that our wants (for example hunger) are limited, when the resources allowing satisfaction of wants are limited we satisfy our most urgent wants first, and that the more satisfied is the want (by eating) the less pleasure do we derive from further satisfying the want.

the less valuable will be the next additional unit that we *could* have. This is called **diminishing marginal utility**, where the new idea here is "marginal."

M-NOTE 3.1 The meaning of marginal

The change in the value of a function—like utility, $u(x,y)$—when just one argument of the function x or y changes is a basic concept in calculus. The partial derivative of the function with respect to an argument—that is either $u_x(x,y)$ or $u_y(x,y)$— is approximated by the effect of a small change in the argument on the value of the function, holding constant the other argument. If the decision maker increases her consumption of x by a small amount Δx, then her utility becomes $u(x+\Delta x,y) \approx u(x,y) + u_x(x,y)\Delta x$, so $\Delta u \approx u_x(x,y)\Delta x$.

So the marginal utility of x is $u_x(x,y) = \frac{\Delta u}{\Delta x}$ where Δx is small. Conventionally this is expressed as the effect on u of a one unit change in x.

If the marginal utility of any thing that we value positively is less, the more of it that we have—diminishing marginal utility—then this means that:

- the *first* partial derivative of the utility function with respect to good x is positive, $u_x > 0$ (because more x is better than less); and
- the *second* partial derivative of the utility function with respect to good x is negative, $u_{xx} < 0$ (because as x increases, utility is increasing ($u_x > 0$), but at a diminishing rate, therefore giving us diminishing marginal utility).

M-CHECK We also use the symbol for partial differentiation $\frac{\partial u(x,y)}{\partial x}$ to mean the marginal utility of x. When it is not necessary to be reminded of the other variables (held constant) that the marginal utility depends on, we eliminate the (x,y) and just use $\frac{\partial u}{\partial x}$ or u_x.

Diminishing marginal utility

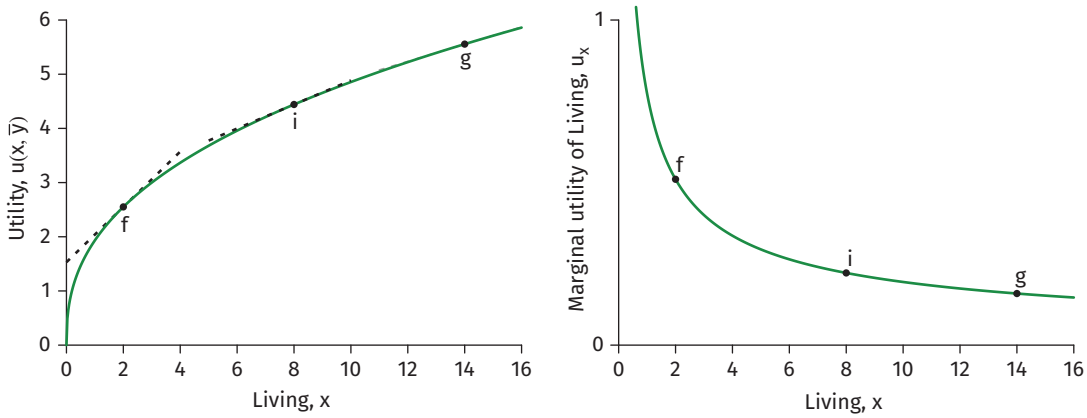
A change of one variable–like Keiko's Living–by one very small unit while holding constant everything else, including her Learning, is a *marginal change*, meaning the change is very small and in only one variable. The change in utility corresponding to a marginal change in x or y is called the *marginal utility* of x or y. Keiko's marginal utility of Living, which we denote as u_x, like her utility itself, depends on how much Living and Learning she is currently experiencing. So we write u_x as a function of x and y: $u_x(x,y)$.

Similarly, Keiko's marginal utility of Learning, $u_y(x,y)$ or using the alternative notation $\frac{\Delta u(x,y)}{\Delta y}$, is how much her utility changes as she changes her Learning (y) by one unit, holding constant the amount of Living she does (x).

Figure 3.6 shows just a slice of Keiko's preferences, namely how they vary with the level of Living she experiences, when the level of Learning she experiences is fixed at $y = 3$. We can study the full range of her preferences

DIMINISHING MARGINAL UTILITY A property of some utility functions according to which each additional unit of a given variable results in a smaller increment to total utility than did the previous additional unit.

Figure 3.6 Diminishing marginal utility. In panel (a), utility is an increasing and concave function of Living, meaning that the curve is positively sloped, but with a decreasing slope for higher levels of Living. The slope of the curve is the marginal utility and this is shown in panel (b). The points in panel (a) correspond to the same points in panel (b). For example, the height of point **f** in panel (a) shows the level of utility when Keiko experiences just two hours of Living and the slope of a tangent to the curve at that point is the marginal utility of Living. The height of point **f** in panel (b) shows the value of that slope, that is the marginal utility of increased Living when Keiko experiences two hours of Living.



(a) Utility of Living holding Learning constant (b) Marginal utility of Living holding Learning constant at 3

when the value of both goods varies by looking at her entire indifference map.

The marginal rate of substitution

The **marginal rate of substitution** is the maximum amount of y that Keiko would be willing to give up to get a unit more of x. The marginal rate of substitution is also the least amount of y that Keiko would view as an adequate *substitute* for losing a unit of x. The marginal rate of substitution should be read as "units of good y per unit of good x."

We show in M-Note 3.2 that the marginal rate of substitution is equal to the ratio of the marginal utilities of the two goods:

$$mrs(x,y) = \frac{u_x(x,y)}{u_y(x,y)} = \frac{\text{marginal utility of x}}{\text{marginal utility of y}} \tag{3.1}$$

MARGINAL RATE OF SUBSTITUTION The marginal rate of substitution is the negative of the slope of an indifference curve. It is also the maximum willingness to pay for a small increase in the amount x expressed as how much of y the person would be willing to give up for this. In a model with y as money, this is called the offer price.

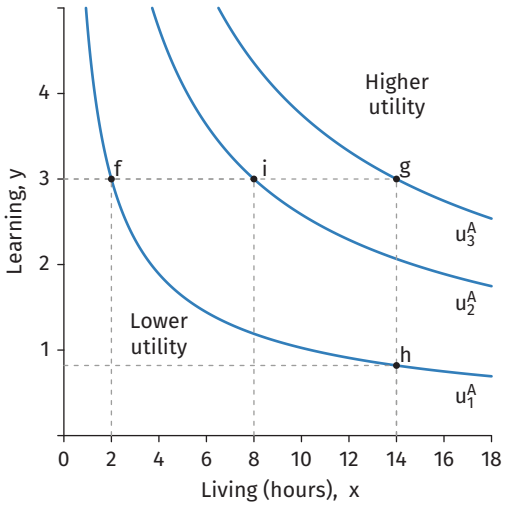
This is true because the amount of y that compensates Keiko for a small loss of x is the ratio of her marginal utility of x, which tells us how much she misses the x she has lost, to the marginal utility of y, which tells us how much she appreciates the compensating gain in y.

The marginal rate of substitution provides us with an essential piece of information. Imagine that Keiko had some bundle (x, y) and she were offered the following exchange–trade away some of her y in order to get more x. You already know that the *mrs* tells us the greatest amount of y that she would be willing to give up to get one more unit of x in such a trade. This is why we call the *mrs* the willingness to pay y to get more x.

Why does the *mrs* tell us her *maximum* willingness to pay? She would happily pay less than the *mrs* to get one more unit of x because this would increase her utility (put her on a higher indifference curve). But she would *not* pay more. This is why we call the *mrs* the maximum *willingness to pay*.

This is shown in Figure 3.7. At point **f**, Keiko spends 14 hours studying and attending classes and has two hours left over for Living, so, a lot of

Figure 3.7 An indifference map showing Keiko's evaluation of bundles of Living (x) and Learning (y). The negative of the slope of the indifference curve is the *marginal rate of substitution* of Learning (y) for Living (x), $mrs(x, y)$, capturing the trade-offs of Keiko's preferences for the two goods. The lettered points correspond to the same letters in Figure 3.6. At **f**, Keiko has a high level of Learning (3) and little Living (two hours) and she is willing to give up a lot of Learning to get more Living (the slope is *steep* at point **f**; therefore her marginal rate of substitution is large). At **h**; Keiko has a low level of Learning (0.82) and a lot of Living (14 hours) and she is willing to give up very little Learning to get more Living (her slope is *relatively flat* at point **h**; therefore her marginal rate of substitution is small). Keiko has a Cobb-Douglas utility function (introduced in the next section) with $u(x, y) = x^{0.3}y^{0.7}$.



Learning and not so much Living. As a result her indifference curve at point **f** is steep. The largest amount of Learning that Keiko would be willing to give up in order to get one more unit of Living is the negative of the slope of her indifference curve at that point (0.64 at point **f**), which is the *marginal rate of substitution* at that point or $mrs(x,y)$.

M-CHECK When considering two *goods*—things that people value positively, like data and coffee, or living and learning—the indifference curves are downward-sloping. That is, they have a *negative* slope. The negative of the slope of an indifference curve is just its slope with the sign changed.

You can see from Equation 3.1 that Keiko's indifference curve is steep (large mrs) because the marginal utility of additional Living ($u_x(x,y)$) is large (she has little Living) and the marginal utility of Learning ($u_y(x,y)$) is small (she has a lot of Learning).

Comparing points **f** and **g**, we can see that if she were to have the same level of Learning (3) but much more Living (14 hours) we already know from the same points in Figure 3.6 (b) that her marginal utility of Living would be lower. So her indifference curve would be flatter, meaning that she would be unwilling to give up as much Learning to get another hour of Living than she was at point **f** when she had just two hours of Living.

The same reasoning shows (and Figure 3.7 confirms) that Keiko's willingness to pay (in units of good y) for another unit of good x increases as she has more of good y (compare the slopes at points **h** and **g**).

Before going on to the constraints facing Keiko we will now show how what you have learned so far can be used with an explicit mathematical function.

M-NOTE 3.2 **The *mrs* is the ratio of marginal utilities**

To derive the marginal rate of substitution using calculus, we use the method of *total differentiation* (covered in the Mathematics appendix). First of all, along an indifference curve the amount of utility is a constant, $u(x,y) = \bar{u}$.

To find the slope of the indifference curve we ask what changes in the quantities of x and y (one increasing, the other decreasing) are consistent with $u(x,y)$ not changing. This is what total differentiation tells us. When we totally differentiate the utility function with respect to its arguments we get the change in Keiko's utility as the sum of the changes due to changes in her consumption of each good. The indifference curve is defined as the changes in x and y such that the change in Keiko's utility is zero.

So we set the total derivative of her utility function equal to zero:

$$du = u_x(x,y)dx + u_y(x,y)dy = d\bar{u} = 0 \tag{3.2}$$

This equation says that since her utility is constant on an indifference curve by definition, the change in her utility is zero.

Recall, too, that the derivative of a constant like $\bar{u}$ is 0.

We can now rearrange equation 3.2 to find the $mrs(x,y)$:

Subtract $u_y(x,y)dy$ from both sides $\quad -u_y(x,y)dy = u_x(x,y)dx$

Divide by $u_y(x,y)$ and dx

$$mrs(x,y) = -\frac{dy}{dx} = \frac{u_x(x,y)}{u_y(x,y)} \tag{3.3}$$

As a result, the negative of the slope of the indifference curve $-\frac{dy}{dx}$ is equal to the ratio of the marginal utilities of the goods, $\frac{u_x(x,y)}{u_y(x,y)}$. But the negative of

continued

the slope of the indifference curve is the marginal rate of substitution of y for x, so we have shown that the marginal rate of substitution is the ratio of the marginal utilities, and correspondingly that the

$$\text{slope of an indifference curve} \quad = -mrs(x,y) = -\frac{u_x(x,y)}{u_y(x,y)} \tag{3.4}$$

The mrs has the dimensions of an amount of good y per unit of good x because the marginal utility of y has the dimensions utility per unit y, and the marginal utility of x has the dimensions utility per unit x.

CHECKPOINT 3.4 **Diminishing marginal utility** Explain the relationship between the slopes of the curves at points **f**, **i**, and **g**, in Figures 3.6 panels a and b and Figure 3.7.

3.6 APPLICATION: *HOMO ECONOMICUS* WITH COBB-DOUGLAS UTILITY

HISTORY The Cobb-Douglas function is named after the economist and later US Senator Paul Douglas and his then Amherst College colleague, mathematician and economist Charles Cobb, who jointly came up with the function in 1928 for an econometric study of the contributions of labor and capital goods to output in the US economy.

In Chapter 2, we saw that people may have some combination of preferences including self-regarding and other-regarding in its many forms: altruistic, fair-minded, reciprocal, and spiteful. Representing these preferences mathematically requires knowledge of what Keiko values including:

- How important to her are Learning and Living?
- Is her own Living and Learning all she cares about, or does she value other people's Living and Learning?

In this section we study the preferences of a self-regarding Keiko: she does not care about the Living and Learning of others. We use what is called a Cobb-Douglas utility function to illustrate how we can model the difference it makes what value she places on the two elements in her choice bundle.

Here is a Cobb-Douglas utility function:

$$u(x,y) = x^{\alpha}y^{(1-\alpha)} \tag{3.5}$$

The size of α, which is a positive number less than 1, is a kind of baseline measure of how much the individual values x independently of how much x and y she has. For example, we show in Chapter 7 that the fraction of a utility-maximizing person's budget that will be spent on good x is α. The fraction spent on y will be $1-\alpha$. So if Keiko's α is greater than her friend's, then we would expect to see Keiko consuming more of the x good.

When a person's preferences are described by a Cobb-Douglas utility function, then as long as the Keiko has some of each good, $x > 0$ and $y > 0$, the following will be true:

- her utility $u(x,y) > 0$ is positive; and
- her utility increases as she consumes more of either good x or y, meaning that the marginal utility of both goods is positive.

Because the marginal utilities for both goods is positive, Keiko will select a bundle with more of each over a bundle with less of either if both bundles are available to her.

Here is an example of a Cobb-Douglas utility function where a consumer, Anmei from earlier, has a stronger preference for y than for x because $\alpha = 0.4$ and $(1-\alpha) = 0.6$.

$$u(x,y) = x^{0.4}y^{0.6} \tag{3.6}$$

Let's assume that x is kilograms of coffee and y is gigabytes of data as we did earlier. The values of α and $(1-\alpha)$ show that Anmei has a stronger preference for data than for coffee because $\alpha = 0.4 < 0.6 = (1-\alpha)$.

M-CHECK In M-Note 3.4 we show that if Anmei is consuming the same amount of x and y, then the maximum number of units of y that she would be willing to pay for one unit of x is $\frac{\alpha}{1-\alpha}$. So if $\alpha = 0.4$ her willingness to pay for a unit of x would be $\frac{0.4}{0.6}$ or two-thirds of a unit of y.

M-NOTE 3.3 **Cobb-Douglas diminishing marginal utility**

How do we check that marginal utility is diminishing? Let us examine the marginal utility of Living in the Cobb-Douglas utility function, assuming $x > 0$, $y > 0$.

$$\text{Utility function} \qquad u(x,y) = x^{\alpha}y^{1-\alpha} \tag{3.7}$$

To find the marginal utility of x we differentiate Equation 3.7 with respect to x:

$$\text{Marginal utility of x} \qquad u_x = \frac{\partial u}{\partial x} = \alpha x^{\alpha-1}y^{1-\alpha} \tag{3.8}$$

Because $0 < \alpha < 1$, the marginal utility of x is positive, that is $u_x > 0$. Why? x and y are both positive, as is the parameter α, as is the exponent $1-\alpha$. The exponent $\alpha - 1 < 0$, but this simply means that x can be read as being in the denominator of the marginal utility (because $\alpha - 1 = -(1-\alpha)$). For example, for $\alpha = 0.6$, the marginal utility of x is:

$$u_x = 0.6\frac{y^{0.4}}{x^{0.4}} = \alpha\frac{y^{1-\alpha}}{x^{1-\alpha}} \tag{3.9}$$

You can see from Equation 3.9 that the larger is x the smaller will be the marginal utility of x. You already saw this in Figure 3.6 (in fact, we used a Cobb-Douglas utility function to make the indifference curves in that figure). To confirm that the marginal utility of x is diminishing, we need to differentiate the marginal utility of x with respect to x. That is, we need to find the second derivative of the utility function with respect to x, $\frac{\partial^2 u}{\partial x^2}$. So we partially differentiate Equation 3.8 with respect to x.

$$\text{Change in } u_x \qquad \frac{\partial^2 u}{\partial x^2} = (\alpha)(\alpha-1)x^{(\alpha-2)}y^{(1-\alpha)} < 0$$

Because $0 < \alpha < 1$, $\alpha - 1 < 0$. Therefore, $\alpha(\alpha-1) < 0$. Therefore, the rate of change of the marginal utility with respect to x is negative (marginal utility is diminishing), or what is the same thing: utility increases at a decreasing rate as x increases.

M-NOTE 3.4 Cobb-Douglas/Coffee-Data

We derive the marginal rate of substitution for the general Cobb-Douglas utility function.

$$u(x,y) = x^{\alpha}y^{(1-\alpha)} = \bar{u}$$

To find the marginal rate of substitution, we need the marginal utilities of x and y. Consequently, we differentiate the utility function with respect to x to find u_x, the marginal utility of coffee, and with respect to y to find u_y, the marginal utility of data.

$$u_x = x^{(\alpha-1)}y^{(1-\alpha)} \tag{3.10}$$

$$u_y = (1-\alpha)x^{\alpha}y^{(1-\alpha)-1} \tag{3.11}$$

We substitute the marginal utilities (Equations 3.10 and 3.11) into the definition of marginal rate of substitution, $mrs(x,y)$ (Equation 3.4) to find the formula for the marginal rate of substitution.

$$\begin{aligned} mrs(x,y) = -\frac{dy}{dx} &= \frac{u_x(x,y)}{u_y(x,y)} \\ &= \frac{x^{(\alpha-1)}y^{(1-\alpha)}}{(1-\alpha)x^{\alpha}y^{(1-\alpha)-1}} \\ \text{Factor out } x^{-1} \text{ and } y^{-1} &= \frac{\alpha x^{\alpha}x^{-1}y^{(1-\alpha)}}{(1-\alpha)x^{\alpha}y^{(1-\alpha)}y^{-1}} \end{aligned}$$

Remember that $x^{-1} = \frac{1}{x}$ and $\frac{1}{y^{-1}} = y$ and cancel the terms $\frac{x^{\alpha}}{x^{\alpha}}$ and $\frac{y^{(1-\alpha)}}{y^{(1-\alpha)}}$:

$$mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x} \tag{3.12}$$

Equation 3.12 shows that if Anmei is consuming the same number of gigabytes of data and kilograms of coffee (say, five of each) she will evaluate them at ratio $\frac{\alpha}{(1-\alpha)}$. The preferences for each good (α and $(1-\alpha)$) determines the ratio at which Anmei is willing to trade data for coffee, together with the amount of coffee and data she is actually consuming. You can see that if Anmei had a different level of current consumption of the two goods, say, more x and less y her mrs would be lower.

CHECKPOINT 3.5 Diminishing *mrs* as $\frac{y}{x}$ falls Go back to Figure 3.4 and explain why the mrs is lower at point **c** than at point **b**, and lower at point **d** than at point **a**.

3.7 THE FEASIBLE SET OF ACTIONS: OPPORTUNITY COSTS AND THE *MRT*

Keiko's preferences and the resulting utility numbers she assigns to each bundle are a reflection of what she wants to achieve, what her goals are. But her preferences do not tell us what she can *feasibly* obtain. To understand the bundles that are feasible for her, we need to know how she obtains

Learning from spending her time studying. Remember, Keiko sleeps eight hours every night and she is not considering changing that. Her choice is what she will do with the 16 hours in the rest of the day.

A production function: How studying produces Learning

The relationship between the time Keiko spends studying and the amount of learning she achieves is given by an equation that shows for the time (in hours) spent studying (h), how much Learning (y) results, $y = f(h)$. This is a **production function**–a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows. Production functions are more often used to study things other than success in coursework, that is outputs such as meals served, lines of code written, or bushels of corn harvested.

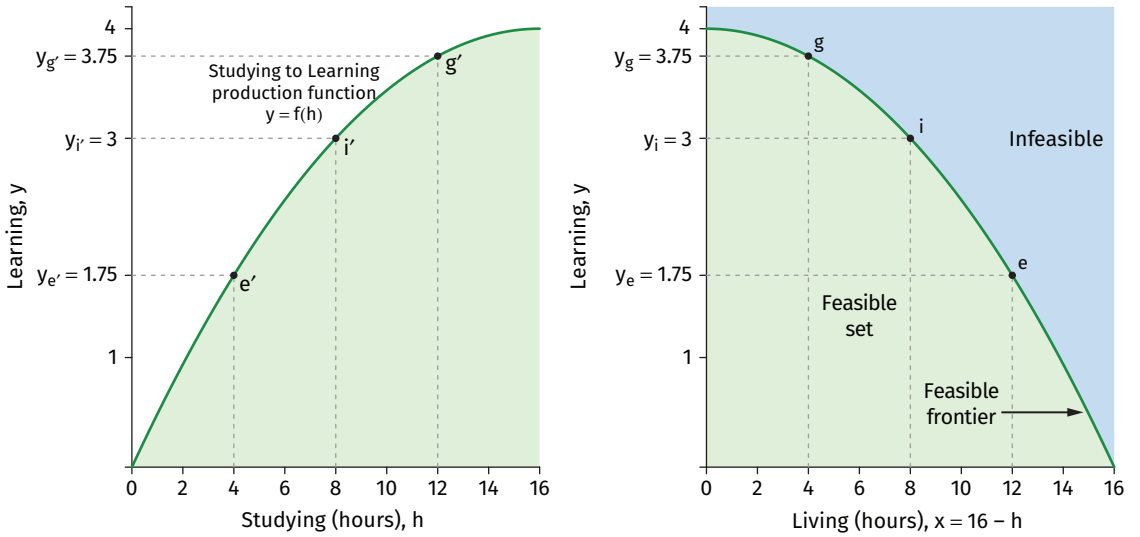
Keiko's production function is depicted in Figure 3.8 (a). From the figure you can see that for Keiko to obtain Learning she must spend hours (h) studying. Up to a maximum of 16 hours, she can increase her learning by studying more. But starting from studying just a few hours, doubling the amount of studying she does does not double her Learning. We can see this by comparing points $\mathbf{e}'$, $\mathbf{i}'$, and $\mathbf{g}'$. Four hours of study ($h = 4$), gets Keiko $y = 1.75$ points of Learning, as shown by point $\mathbf{e}'$. But doubling her studying to eight hours gets her just three units of Learning, far from a proportional increase. This is because if she has just four hours, then she focuses on the really important key points. While if she has eight hours, then she gets into the details, which add to her Learning, but not as much as the key ideas do.

Keiko's learning production function illustrates an important common economic phenomenon: diminishing marginal productivity. The marginal productivity of hours studying is the effect of a small increase in studying time on the resulting Learning. As you can see from the fact that the production function in Figure 3.8 is flatter for more hours of study, marginal productivity of studying hours is therefore diminishing.

This is similar to diminishing marginal utility. Just as the person satisfies her most pressing needs if she has very limited expenditures, but can turn to frills if she has more to spend, Keiko focuses on the essential points if her study time is limited but can turn to the examples and further illustrations if she has more time to spend.

PRODUCTION FUNCTION A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows.

Figure 3.8 Production of Learning by studying and the feasible frontier of Living and Learning. Points **e′**, **i′**, and **g′** on the production function show combinations of hours of study and the maximum amount of Learning she could accomplish in that time. Point **i′**, for example, shows that if she studies eight hours she could attain learning equal to 3 (she could also attain less if she spent the "studying" time texting with friends). The amount of Living that she can have is her 16 hours minus the time she spends learning, i.e. $x = 16 - h$, as shown in panel (b). Panel (b) shows the feasible frontier (dark-green curve), which is the border of the feasible set (shaded in green). The feasible frontier is just a flipped version of her production function. Points beyond the feasible set (shaded in blue) are *infeasible* given the number of hours in the day and her Learning production function. In this figure the equation for the feasible frontier is given by: $y = 4 - \frac{1}{64}x^2$. The equation for the production function is given by: $y = 4 - \left(\frac{1}{64}\right)(16-h)^2$, or simplifying: $y = h\left(\frac{1}{2} - \frac{h}{64}\right)$. (Remember $x = 16 - h$.)



(a) The production of Learning by studying

(b) The feasible frontier of Living and Learning

M-NOTE 3.5 An education production function

The education function relating hours of studying to the resulting learning that we used to draw Figure 3.8 (a) is:

$$\text{production function} \qquad y = f(h) = h\left(\frac{1}{2} - \frac{h}{64}\right) \tag{3.13}$$

The marginal product of time spent studying is the derivative of y with respect to h:

$$\text{marginal product of studying} \qquad y_h = \frac{dy}{dh} = \frac{1}{2} - \frac{h}{32} \tag{3.14}$$

So you can see that when studying time is limited, the marginal product is almost one-half; and with more study time the marginal product is less. When

continued

$h = 12$, for example, you can see from Equation 3.14 that the marginal product of studying time is only one-eighth.

The marginal rate of transformation and opportunity cost

Because her waking hours are just 16 and Keiko defines Living as her time *not* studying, she has two ways to use her time—studying or not studying, so:

$$\text{Living} = 16 \text{ hours} - \text{hours of studying}$$

This makes it clear that

- She has just one decision to make not two: if she chooses hours of studying, that also determines her hours of Living.
- Because more time living means less time studying this means that the opportunity cost of living more is some amount of learning less.

To understand what this opportunity cost is, see Figure 3.8 (b), showing the feasible set of outcomes that Keiko might experience. The feasible frontier shown there is the mirror image of the production function in the panel (a). The horizontal axis is no longer studying hours but instead 16 minus studying hours, which is the amount of Living she can have for each level of studying she chooses.

At **e**, Keiko studies for four hours, which means she is Living for 12 hours and her Learning is 1.75. Or (point **g**) she could study for 12 hours and have learning of 3.75. All of the points like **e**, **g**, **i**, and the rest of the feasible frontier are choices that she could make.

The **feasible set** is the area bounded by the **feasible frontier** and the x and y axes composed of all combinations of Living and Learning that she could experience.

Turning to Figure 3.9 we can also contrast two points on the feasible frontier, such as points **a** and **b**. At point **a**, Keiko spends 13 hours studying and has three hours left over for Living, with the result of a lot of Learning and not so much Living. At point **b**, Keiko spends eight hours studying and attending classes and has eight hours left over for Living, but her Learning is lower than at point **a**. The difference between the two points on the feasible

FEASIBLE SET All of the combinations of the actions or outcomes that a decision maker could choose, given the economic, physical, or other constraints.

FEASIBLE FRONTIER The boundary of a feasible set. In the case of two goods, it is the curve made of points that defines the maximum feasible quantity of one good for a given quantity of the other.

HISTORY The 1975 collection of essays by Nobel Laureate Milton Friedman titled *There's No Such Thing as a Free Lunch: Essays on Public Policy* popularized the idea that there is an opportunity cost to having more of anything that we value.[2]

M-CHECK When x is a bad (like studying time) rather than a good (like Living) the mrt is still the negative of the slope of the feasible frontier but the opportunity cost is now the amount of the y-good (Learning in this case) that will be sacrificed by studying *less*. In this case we have $dy/dx > 0$ so the mrt as the negative of this is $-dy/dx < 0$, and the opportunity cost (a positive amount) is the effect on y of having less x or $-(-dy/dx) > 0$.

frontier illustrates another trade-off that is central to Keiko's choice: more living means less learning. And vice versa.

If we apply the same reasoning to very small differences of the two goods, we can see that the opportunity cost in less Learning that is required to get more Living is the negative of the slope of her feasible frontier at that point, namely $-\frac{\Delta y}{\Delta x}$. This is called the **marginal rate of transformation** or $mrt(x,y)$.

The marginal rate of transformation is the smallest amount of y that Keiko has to give up to get a small unit more of x.

The mrt is therefore Keiko's *opportunity cost* of x in terms of y or the minimum amount of y she has to sacrifice in order to get a small unit of x. The interpretation of the mrt as the opportunity cost of the x-good plays a major role in the reasoning in this book.

$$\text{Opp. cost of x} \quad = -\text{Slope of feasible frontier} = -\frac{\Delta y}{\Delta x} = mrt(x,y)$$

The marginal rate of transformation should be read as "units of good y per unit of good x." What is being transformed into what? Free time is being given up and devoted to studying that is transformed into Learning. The marginal rate of transformation is determined by how productive are the hours of schooling she spends. You already know that the marginal product of study time declines the more she studies.

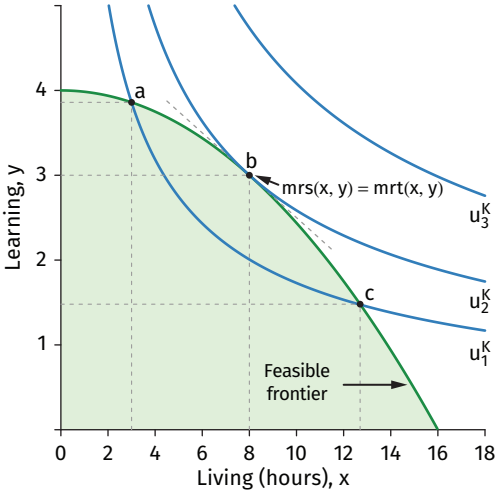
As a result the feasible set is steep when she is devoting little time to studying, but as she studies more—moving upward and to the left along the feasible set—the additional learning associated with more studying (and more Living given up) declines.

This is also why the opportunity cost of studying less increases as you move downward and to the right along the curve. At point **a** where Keiko is studying most of the time—13 hours—she is reading every page of the assigned readings twice, doing all of the practice problems, and even reading the footnotes. Cutting back a bit on her studying is not going to cost her much in terms of Learning.

In this case, as shown in Figure 3.9, her feasible frontier exhibits an *increasing marginal rate of transformation*. Her feasible frontier starts off with low opportunity costs at point **a**, but her marginal rate of transformation increases as she moves along the feasible frontier—studying less—toward having more x and less y. This steepening of the frontier reflects the increasing opportunity cost of free time. In this case, Keiko has to sacrifice more y for x the more x and the less y she has.

MARGINAL RATE OF TRANSFORMATION The marginal rate of transformation is the quantity of some good that must be sacrificed to acquire more of another good. It is equal to the negative of the slope of the feasible frontier (constraint).

Figure 3.9 The *mrs* = *mrt* rule for two goods: Living and Learning. Keiko's feasible frontier for Living and Learning is shown in green. Three of her indifference curves are shown by u_1^K, u_2^K, and u_3^K in blue ($u_3^K > u_2^K > u_1^K$). At point **b**, her utility is maximized: the marginal rate of substitution equals the marginal rate of transformation by choosing to spend eight hours Living which gives her a subjective Learning score of 3.



M-NOTE 3.6 A Living-Learning feasible frontier

An equation for the feasible frontier has the form:

$$\text{Feasible frontier} \qquad y = \bar{y} - c(x) \tag{3.15}$$

The parameter $\bar{y}$ is the maximum amount of y when $x = 0$, the y-intercept of the feasible frontier. The term $c(x)$ is the cost of x, that is, how many units of y (Learning) one must give up to get the value of x (Living) that she chooses.

We get the equation for the feasible frontier in Figure 3.8 (b) for Keiko's Living, x, and Learning, y, by starting with the education production function, Equation 3.13:

$$\text{Production function} \qquad y = f(h)$$

$$y = h\left(\frac{1}{2} - \frac{h}{64}\right)$$

And then placing h in terms of x using $h = 16 - x$ and simplifying yields:

$$y = 4 - \frac{x^2}{64} \tag{3.16}$$

The negative of the slope of the feasible frontier, $-\frac{dy}{dx} = \frac{1}{32}x$. This is the amount of x that she has to give up in order to get one more unit of y. The opportunity cost of additional Living is greater the more living she is doing because the more she "lives" the less she studies, and when she is studying just a few hours the marginal product of studying is high, and so studying less reduces Learning a lot.

CHECKPOINT 3.6 **Diminishing marginal productivity of studying** What does diminishing marginal productivity of studying mean, and why does it occur in the learning process?

REMINDER A bundle (x, y) is constrained utility-maximizing if there is no point in the feasible set that is on a higher indifference curve.

3.8 CONSTRAINED UTILITY MAXIMIZATION: THE *MRS* = *MRT* RULE

By combining the insights of feasible frontiers and indifference curves—as in Figure 3.9—we can understand what a constrained utility-maximizing choice by Keiko would be.

- *Constraints*: She can choose some point on or within her feasible frontier given by her production function and the limits of her time.
- *Preferences*: From among the points in her feasible set, she will prefer the bundle with the highest utility, meaning on the highest indifference curve.

To understand Keiko's **constrained utility-maximizing** problem, we contrast points **a**, **b**, and **c** in Figure 3.9. A bundle (x, y) is *constrained utility-maximizing* if there is no other point in the feasible set with a higher utility.

Point **a** is on Keiko's feasible frontier and lies on indifference curve u_1^K. But, **a** is not constrained utility-maximizing because Keiko could increase her utility by increasing her Living time and decreasing her Learning, by moving along the feasible frontier to the southeast. By similar reasoning point **c** cannot be the highest indifference curve she can reach.

Keiko's constrained utility-maximizing point is **b** in Figure 3.9, the point on the feasible frontier that is on the highest indifference curve. We label it **b** because it is the point where Keiko does the **b**est she can.

Figure 3.9 suggests a useful way to think about Keiko's constrained utility-maximization problem. In the figure, we see that the constrained utility-maximizing bundle is the point where Keiko's indifference curve is tangent to her feasible frontier. This means the indifference curve and the feasible frontier have the same slope at the constrained utility-maximizing point.

The slopes of the indifference curve and the feasible frontier express *trade-offs* between the two goods. This is the basis of what we call the $mrs = mrt$ rule.

M-NOTE 3.7 Equating *mrs* to *mrt* to find the constrained maximum

Suppose Keiko's utility for Living (x) and Learning (y) is described by a Cobb-Douglas utility function with parameter $\alpha = 0.4$ and $(1 - \alpha) = 0.6$:

$$u(x, y) = x^{0.4} y^{0.6}$$

continued

We find her marginal rate of substitution by using the marginal utilities (see M-Note 3.4 if you have difficulty here) and substituting them into the equation $mrs(x,y) = \frac{u_x^K}{u_y^K}$

$$\begin{aligned} u_x^K &= 0.4(x^A)^{-0.6}(y^A)^{0.6} \\ u_y^K &= 0.6(x^A)^{0.4}(y^A)^{-0.4} \\ mrs(x,y) = \frac{u_x^K}{u_y^K} &= \frac{2}{3}\frac{y}{x} \end{aligned} \tag{3.17}$$

Suppose her feasible frontier is the one we used in making the figures:

$$y = 4 - \frac{1}{64}x^2$$

Keiko's constrained maximum must be on her feasible frontier. We find her marginal rate of transformation by differentiating y with respect to x:

$$-\frac{dy}{dx} = mrt(x,y) = \frac{1}{32}x$$

To find Keiko's constrained maximum, we use the two expressions above for mrs and mrt, equating them to find a point on the feasible frontier consistent with the $mrs = mrt$ rule:

$$mrs(x,y) = \frac{2}{3}\frac{y}{x} = \frac{1}{32}x = mrt(x,y) \tag{3.18}$$

Then multiplying through by $\frac{3}{2}x$:

$$\begin{aligned} y &= \frac{3}{64}x^2 = 4 - \frac{1}{64}x^2 \\ x^2 &= 64 \\ x &= 8 \\ y &= 3 \end{aligned}$$

Keiko chooses a consumption bundle with 8 hours of Living, which means she studies for 8 hours, and obtains a Learning level of 3.

Doing the best you can: The *mrs* = *mrt* rule

Summarizing the results so far, in Figure 3.9:

1. The negative of the slope of the feasible frontier is the *opportunity cost of getting a unit more of the x-good*, in terms of the amount of the y-good foregone.
2. The negative of the slope of an indifference curve is a measure of the person's *willingness to pay for a little more of the x-good* in terms of how much of the y-good she would be willing to give up to get an additional unit of the x-good.

Using these two statements we can see why point **a** in Figure 3.9 could not be the utility-maximizing bundle. The indifference curve is steeper than the feasible frontier, so the value of getting more Living (by studying less) exceeds the associated opportunity cost in foregone Learning (point 2

! REMINDER *mrs* and *mrt*

- *mrs*, the *marginal rate of substitution*, is the negative of the slope of an indifference curve.
- *mrt*, the *marginal rate of transformation*, is the negative of the slope of the feasible frontier.

above is greater than point 1). So she could do better by giving up some Learning in favor of more Living.

The opposite is true at point **c**: the feasible frontier is steeper than the indifference curve. By giving up a unit of Living (studying more) she would get a substantial increase in Learning (that is what the steep feasible frontier means). Giving up a unit of Living could be compensated by a *modest* increase in Learning (that is what the flatter indifference curve means). So the *benefits* of giving up some Living in return for more Learning *outweigh the cost*. So any point like **a** and **c** where the feasible frontier and the indifference curve intersect cannot be the constrained utility maximizing output bundle. This gives us the $mrs = mrt$ rule: the utility-maximizing output bundle is a point where

$$\textit{Slope of feasible frontier} = \textit{Slope of indifference curve}$$

which requires that:

$$\textit{Marginal rate of transformation (mrt)} = \textit{Marginal rate of substitution (mrs)}$$

Or, what is the same thing,

$$\textit{Opportunity cost of } x = \textit{Willingness to pay for } x$$

The rule expresses a simple and true idea: if the opportunity cost of something is less than your willingness to pay, you should choose more of it (if you can) and if the opportunity cost is greater than your willingness to pay, you should choose less of it (if you can).

But there are cases in which the utility-maximizing bundle is not a tangency of the feasible frontier and an indifference curve:

- It may be that an indifference curve is steeper than the feasible frontier, but there is no way to get more of the x-good. In this case the slope of feasible frontier does *not* measure the opportunity cost of getting more of the x-good; that is impossible (its cost is infinite). You will find an example of this situation below: the utility-maximizing bundle at point **b** Figure 3.13—called a corner solution—there the $mrs = mrt$ rule does not work.
- We show in M-Note 3.8 that there are conditions under which a bundle such that $mrs = mrt$ can also be a *minimum* not a maximum. We provide an example of this in Chapter 6.

M-NOTE 3.8 When the *mrs* = *mrt* rule fails

The rule can fail to identify the constrained utility maximum under two conditions: when the maximum is a corner solution (so the rule is not satisfied) and when the rule is satisfied at a minimum rather than a maximum. Positing a case with diminishing opportunity cost of obtaining one good in terms of the other good foregone will illustrate both cases.

Setup: Assume that a person's utility varies with the amount of goods x and y:

continued

$$u(x,y) = x + y$$

and the maximum feasible amount of good y is a function of good x:

$$y(x) = (1-x)^2 \tag{3.19}$$

The rule may select a minimum, not a maximum. The marginal rate of substitution and marginal rate of transformation are:

$$mrs(x,y) = \frac{u_x}{u_y} = 1$$

$$mrt(x,y) = -\frac{dy}{dx} = 2(1-x)$$

Equating the mrs and mrt $\quad 2(1-x) = 1$

So, $x = \frac{1}{2}$

Using Equation 3.19 $\quad y = \frac{1}{4}$

Note that using these values, the utility is $u = \frac{3}{4}$. Alternatively, we could set $(x,y) = (1,0)$, or $(x,y) = (0,1)$: both allocations are in the feasible set. In both cases, $u = 1$, which is higher than the one that we have reached using the condition $mrs = mrt$.

The condition $mrs = mrt$ will not give the utility maximum if the second order condition is violated: the second derivative of the utility function with respect to the variables must be negative. Let's calculate it, replacing Equation 3.19 in the utility function:

$$u = x + (1-x)^2$$

$$\frac{du}{dx} = u_x = 1 - 2(1-x)$$

$$\frac{d}{dx}\left(\frac{du}{dx}\right) = u_{xx} = 2 > 0$$

The utility maximum may be a corner solution. In the example the utility maximums at both $x = 1$ and $y = 1$ are corner solutions (only one of the goods is consumed).

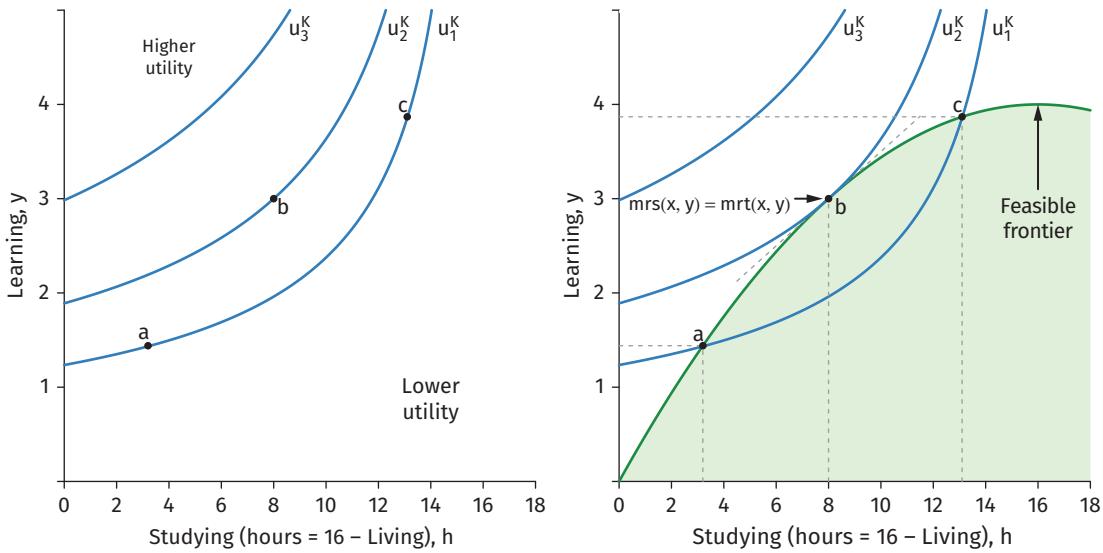
There are also cases in which the rule cannot be applied. Where either the indifference curves or the feasible frontier are not smooth but instead are kinked (are not differentiable), the derivatives on which the *mrs* and *mrt* are based will not exist at the kinks.

Trade-offs between goods and bads

In many situations it is easier to understand decisions in terms of a trade-off between a good and a bad rather than a trade-off between two goods. Recall that a bad is something that you would prefer to have less of, such as working harder than is comfortable or safe.

For example, Keiko might think of her decision in terms of a trade-off between her time studying that she does not enjoy, $h = 16 - x$, and her Learning, y. The more time Living the better for Keiko; therefore x is a good. The more time studying the worse for Keiko; therefore h is a bad. But as before since $x = 16 - h$, choosing (h,y) to maximize utility, $u(16-h,y)$ is the same thing as choosing x to maximize utility $u(x,y)$. These are just different ways of posing the same problem.

Figure 3.10 The *mrs* = *mrt* rule: Keiko's problem of choosing (h, y) when $h = 16 - x$ = time studying, is a bad. Studying time, h, is plotted on the horizontal axis, and Keiko's Learning, y, is plotted on the vertical axis. Keiko's feasible frontier is shown in green in the right-hand panel. Three of her indifference curves are shown by u_1^K, u_2^K, and u_3^K in blue in both panels. The points **a**, **b**, and **c** are the same as in Figure 3.9. Keiko maximizes her utility at the point on her feasible frontier on the highest indifference curve, that is, at point **b**, choosing to spend eight hours on Living and eight hours on Learning. The production function, that is, the feasible frontier is $y = h(\frac{1}{2} - \frac{h}{64})$.



(a) Indifference curves of Learning and studying

(b) The utility-maximizing choices with bads

M-CHECK When x is a bad (like studying time) rather than a good (like Living) the *mrt* is still the negative of the slope of the feasible frontier but the opportunity cost is now the amount of the y-good (learning in this case) that will be sacrificed by studying less. In this case we have $dy/dx > 0$ so the *mrt* as the negative of this is $-dy/dx < 0$, and the opportunity cost (which is a positive number) is the effect on y of having less x or $-(-dy/dx) > 0$.

Figure 3.10 shows Keiko's indifference curves and feasible frontier plotted in terms of study time, h and Learning y. Her indifference curves slope upward because hypothetically holding constant the level of Learning, an increase in studying, h, *lowers* Keiko's utility. So an increase in her study time requires an *increase* in Learning, y to compensate, in order to stay at the same level of utility. Utility increases as we move to the northwest and decreases as we move to the southeast in this plot.

Similarly, Keiko's feasible frontier slopes *upward*, because an increase in study time, h, leads to *more* Learning, y. This is her "learning production function" introduced earlier. So the slope of the feasible frontier is the marginal productivity of studying time or $\frac{\Delta y}{\Delta h}$ and this is also the marginal rate of transformation of study time into Learning. (In this case "transformation" actually describes the process underlying the feasible frontier.)

As was the case for trade-offs between two goods, a bundle in the feasible set is the utility maximizing output bundle if there is no other feasible bundle with greater utility. And this is the bundle for which the $mrs = mrt$ rule holds, namely the point on the feasible frontier where the

marginal rate of substitution equals the marginal rate of transformation ($mrs(x,y) = mrt(x,y)$).

M-NOTE 3.9 The marginal utility of the bad

The utility function for studying (h) and Learning (y) is given by:

$$u^A(h,y) = (16-h)^{0.4}y^{0.6} \tag{3.20}$$

To find the marginal utility of the "bad," studying, we need to partially differentiate Equation 3.20 with respect to h. Remember that when we partially differentiate we treat the other variable as a constant, so the term $y^{0.6}$ will simply remain where it is. We only have to think about the h term.

$$\frac{\partial u^A}{\partial h} = u_h^A = (0.4)(-1)(16-h)^{0.4-1}y^{0.6}$$

$$u_h^A = \underbrace{-0.4}_{<0}\underbrace{(16-h)^{-0.6}}_{>0}\underbrace{y^{0.6}}_{>0} < 0$$

The first term is negative whereas the second and third terms are positive. So the marginal utility of hours of study is negative. We call such a utility a *disutility* and will often talk about the *disutility of work* or the *disutility of effort*. Remember: the marginal utility of a bad is a negative quantity, the marginal *dis*utility of a bad is the same number but with the sign reversed, so it is a positive number.

CHECKPOINT 3.7 Understanding goods and bads Using the production function that is the feasible frontier, namely $y = h\left(\frac{1}{2} - \frac{h}{64}\right)$ confirm that:

a. If Keiko studies eight hours she will attain a value of 3 for Learning.

b. If she decided to sleep only seven hours and to study the entire time she is awake (17 hours), her Learning would be less than if she studied 16 hours.

3.9 THE PRICE-OFFER CURVE, WILLINGNESS TO PAY, AND DEMAND

We often want to know how people respond to different options for exchange in the form of prices. We may be interested in knowing, for each price at which she can purchase any amount of the good she pleases, how much Harriet, someone deciding on how much fish to buy, will purchase, namely the utility-maximizing amount. This is Harriet's **individual demand curve**.

DEMAND CURVE (INDIVIDUAL) A demand curve provides the answer to the hypothetical question: what is the maximum amount of a good that can be sold at each price. The individual demand curve refers to the purchase of a good by a person given the prices of the other goods and the individual's budget.

Remember that in explaining Keiko's indifference curves we asked what is the maximum amount of Learning she would be willing to give up in exchange for more Living. The answer is given by her maximum *willingness to pay*, or what is the same thing her marginal rate of substitution of Learning for Living.

We now ask almost the same question except that rather than giving up Learning to get more Living, Harriet is now giving up *money*—that is, paying for a good according to its *price*. Instead of a "time budget" (16 hours) she now has a money budget.

For each offered price she faces another constrained utility-optimization problem. The demand curve is constructed by a series of hypothetical constrained optimization problems, one for each possible offered price. Each price defines a *feasible set*; its boundary, the feasible frontier, defines the bundles of goods Harriet has access to, given her budget. For each of these feasible sets there is a bundle that maximizes her utility. This is a single point on her demand curve.

Indifference curves tell us the utility number that Harriet assigns to each possible consumption bundle. Using this logic, her utility will be greatest with the bundle in the feasible set that is on the highest indifference curve. This is a standard constrained utility maximization problem in which we can apply the $mrs = mrt$ rule.

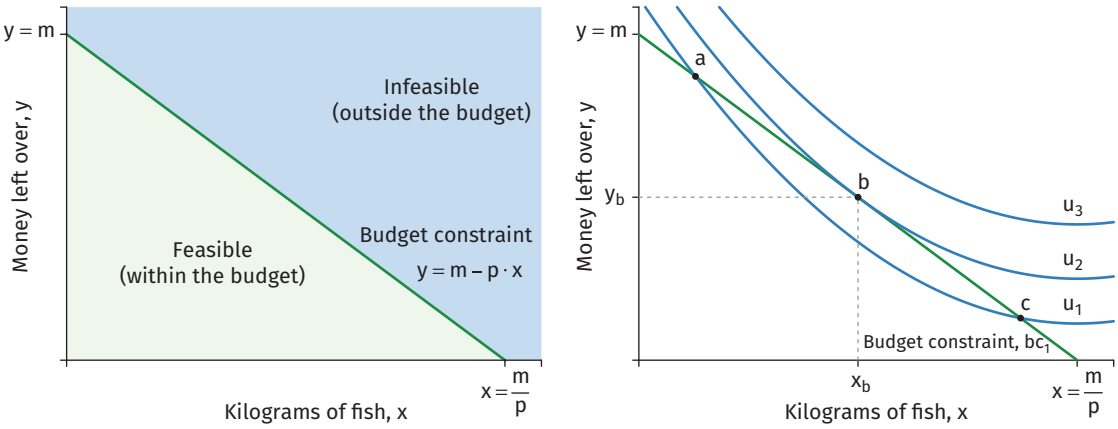
The budget constraint and feasible utility-maximizing choices

We shall use one particular kind of feasible frontier to think this through: the **budget constraint**. The budget constraint defines an amount of money m that a person has or has access to, through wealth and credit markets, which constitutes their budget to spend on goods and services. People can use their budget to spend on goods at prices that are given to them. Imagine that you want to buy fish at a fish market. The price (p) is measured in dollars per kilogram.

Figure 3.11 (a) shows the budget constraint for Harriet. The budget set is shaded in green and the budget constraint (feasible frontier) is the dark-green line on the border of the budget set (feasible set). Consumption bundles (x, y) in the budget set and on the budget constraint can feasibly be obtained with the current budget (m) at the price, p, for kilograms of fish, x. Outside the feasible set, in the shaded green area, the bundles of x and y cannot feasibly be obtained with the existing budget.

BUDGET CONSTRAINT An equation that represents all combinations of goods and services that one could purchase that exactly exhaust one's budgetary resources.

Figure 3.11 Budget constraint and utility-maximizing choice for fish and money for other goods. The budget set is shaded in green and the budget constraint (feasible frontier) is the dark-green line on the border of the budget set (feasible set). Harriet maximizes her utility subject to her budget constraint bc_1. She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x,y) = \frac{u_x}{u_y}$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x,y) = p$.



(a) The budget constraint (b) Utility-maximizing choice

We know how to find the utility-maximizing bundle for a given feasible frontier—or the budget constraint—with given indifference curves: we apply the $mrs = mrt$ rule finding the bundle where the marginal rate of substitution equals the marginal rate of transformation. We can combine these insights and calculate what the consumer's utility-maximizing bundle will be for every potential price of the good given a fixed budget and when the other good, y, is money for other goods.

Figure 3.11 (b) shows Harriet maximizing her utility subject to her budget constraint bc_1. To find her utility-maximizing choice, we must apply the $mrs = mrt$ rule to find where her marginal rate of substitution (her willingness to pay in money for kilograms of fish) equals her marginal rate of transformation, here the price for a kilogram of fish.

At point **a** she consumes too little of x and too much of y (her marginal utility of money for other goods (y) is much lower than her marginal utility of kilograms of fish (x), or her $mrs(x,y)$ is too high, and she would be better off if she consumed less y and more x. Conversely, at **c**, she consumes too little of y and too much of x (her marginal utility of x is much lower than her marginal utility of y, or her $mrs(x,y)$ is too low, and she would be better off if she consumed less x and more y). She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x,y) = \frac{u_x}{u_y}$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x,y) = p$.

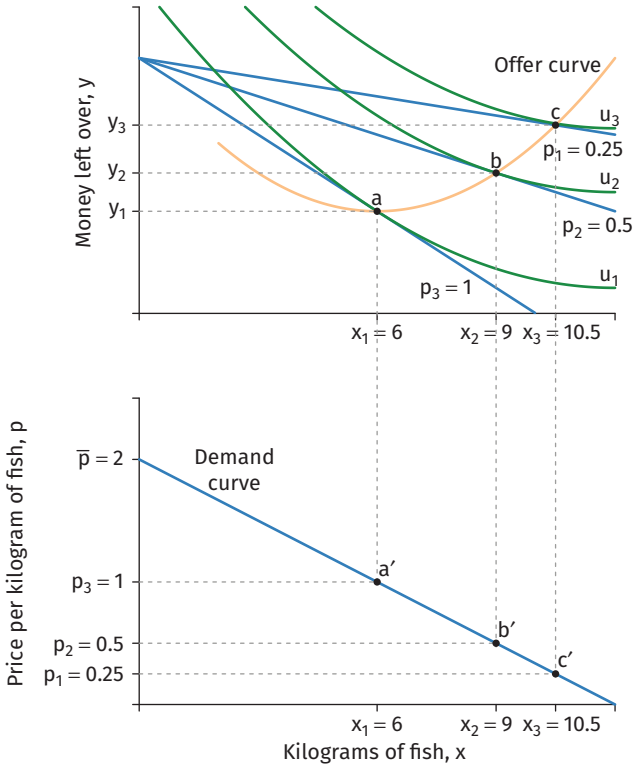
The demand curve: Utility-maximizing choices at difference prices

With every change in price, the consumer's budget constraint will pivot as shown in the top panel of Figure 3.12.

The budget constraint will pivot upward as a good's price *decreases*, because a consumer can buy *more* of the good with the same budget. The opposite is true for price increases. As the price of a good *increases*, the same budget buys less of the good, pivoting the budget constraint *inward*.

With every pivot of the budget constraint, at the utility-maximizing point, the new budget constraint will be tangent to a new indifference curve which will be higher if the price of the good decreases and lower if the price of the good increases.

Figure 3.12 Offer curve and demand curve for fish. The price of x in the top panel is in terms of the money Harriet sacrifices to get more fish. Similarly, in the lower panel the amount of money Harriet must sacrifice to get more fish—the price per unit of fish—determines Harriet's quantity of fish demanded along the demand curve. Points **a**, **b**, and **c** in the top panel correspond to points **a'**, **b'**, and **c'** in the lower panel.



Because we can calculate the utility-maximizing consumption bundle for each possible price, we can find a curve that records every utility-maximizing consumption bundle for each price, called the **price-offer curve**. Sometimes, for individual consumers, it is called the *price-consumption curve* because it indicates what the consumer will *consume* at different *prices*.

M-CHECK We can find the equation for the price-offer curve by using the equation for the budget line and combining it with the equation for the marginal rate of substitution.

Figure 3.12 maps three different utility-maximizing consumption bundles at three prices of x. With each price decrease, the budget constraint pivots outward from p_3 to p_2 to p_1. With each change in the price of x, the utility-maximizing bundle—the point at which the marginal rate of substitution is equal to the opportunity cost—changes. At $p_3 = 1$, the bundle includes $x = 6$, at $p_2 = 0.5$, the bundle includes $x = 9$, and at $p_1 = 0.25$, the bundle includes $x = 10.5$. With each price change, there is a new bundle for both x and y.

› EXAMPLE We demonstrate in Chapter 7 how to find the equation for the demand curve and how to see that a reduction in the price of fish has two effects. First, the lower price leads the person to buy less of some other foods and more fish; this is the *substitution effect*. Second, the lower price also allows the person to buy more of *everything* if she chooses (fish, rice, or whatever); this is called the *income effect*.

The different bundles suggest a price-quantity relationship between the quantity demanded of x and different prices of x. As the price of x decreases, the quantity demanded increases.

In the lower panel of Figure 3.12, we have taken each utility-maximizing consumption bundle from the different consumption bundles at each price and identified their coordinates on price-quantity axes. The price-quantity combinations provide a downward-sloping demand curve where quantity demanded, x, decreases as its price, p, increases.

Measured horizontally from the vertical axis, it tells us the amount that can be sold to the consumer at each particular price. Measured vertically from the horizontal axis, it also tells us what is the consumer's maximum willingness to pay for each amount on the horizontal axis.

CHECKPOINT 3.8 Willingness to pay In the lower panel of Figure 3.12, you can see that if the price is 2 the buyer will purchase no fish. Draw a figure like the top panel by first showing the budget constraint for $p = 2$, and then an indifference curve such that the buyer would purchase no fish if that were the budget constraint.

3.10 SOCIAL PREFERENCES AND UTILITY MAXIMIZATION

The preferences we have looked at so far have been entirely self-regarding, depicting a person who is concerned with their choices among bundles that they alone will experience. But people often make choices where they

PRICE-OFFER CURVE The price-offer curve shows every utility-maximizing consumption bundle for each price of the goods under consideration.

are not the only person affected, where what they choose can benefit or harm someone else. Think of someone shopping for an entire family. And we know from the Ultimatum Game in Chapter 2 that in these situations of interdependence people often care about the effects of their actions on others.

REMINDER A game is a mathematical representation of a *strategic* interaction, which means one in which players recognize that their payoffs depend on the actions taken by other players. So the so-called Dictator Game is not really a game at all, because the Dictator's payoffs do not depend at all on anything that the other player does.

Consider the Dictator Game that we mentioned in Chapter 2. In that game, a person, the Dictator, has an endowment of money, z, to split between themselves and another person in any way they choose, including giving nothing or giving the entire quantity.

Anmei is the Dictator and her endowment and the split she will make are $z = \pi^A + \pi^B = \$10$, where π^A is the amount in dollars that Anmei keeps for herself and π^B is the amount that she gives to Ben. We can rearrange the equation to find the feasible frontier for:

$$\text{Feasible dictator allocations} \qquad \pi^B = 10 - \pi^A \tag{3.21}$$

Looking at Equation 3.21, we can see that the feasible frontier is a line with a slope of –1: the feasible frontier slopes downward with a constant slope. Remember that the negative of the slope of the feasible frontier is the marginal rate of transformation: so a player in the Dictator Game who wishes to give \$1 to someone else has an opportunity cost of \$1 for doing so. If Anmei is like *Homo economicus*, she is purely self-regarding. She sets $\pi^B = 0$ and keeps everything for herself and therefore $z = \pi^A = 10$.

What happens when the Dictator is an *altruist* who values making a generous offer to the other?

To see what happens in these cases, let us contrast two pairs of people:

- Anmei (A) is paired with Ben (B). Anmei makes choices about how much money she gets and how much money Ben gets.
- Chen (C) is paired with Diane (D). Chen makes choices about how much money he gets and how much money Diane gets.

To think about the choices that Anmei and Chen make, let us consider two different kinds of Cobb-Douglas utility functions that Anmei and Chen might have.

$$\text{Anmei's utility function} \qquad u^A(\pi^A, \pi^B) = (\pi^A)^1(\pi^B)^0 \tag{3.22}$$

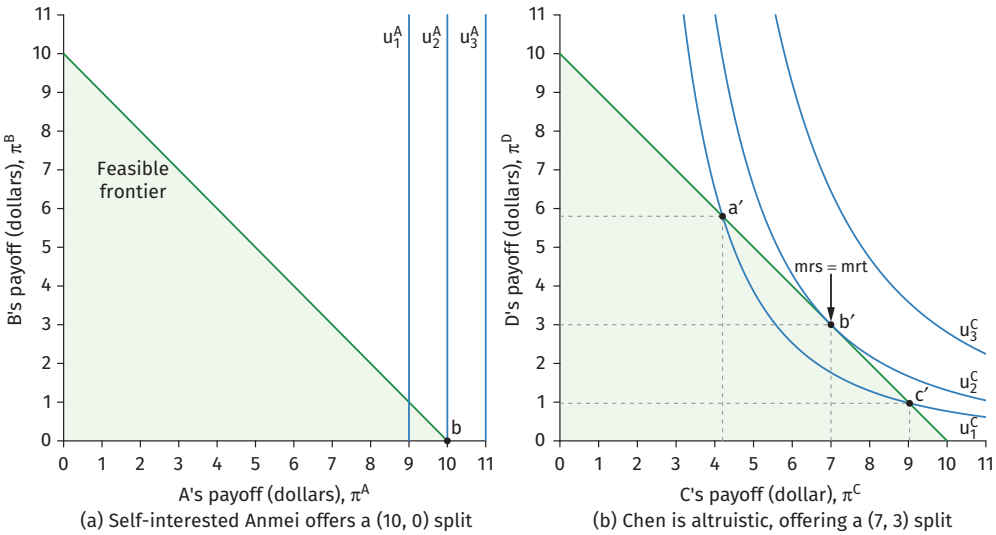
$$\text{Chen's utility function} \qquad u^C(\pi^C, \pi^D) = (\pi^C)^{0.7}(\pi^D)^{0.3} \tag{3.23}$$

M-CHECK Two things to remember when thinking about Equations 3.22 and 3.23.

- The exponents in Chen's Cobb-Douglas utility function Equation 3.23 mean that if they both had the same payoff, then Chen would value increasing his own payoffs more than he would value increasing Diane's.
- About Anmei's utility function, Equation 3.22, any number raised to a zero exponent is equal to 1, so because Anmei does not value Ben's payoffs at all (the exponent is zero), her utility is unaffected by the amount that he gets (her utility is simply how much she keeps for herself).

Chen is other-regarding, he cares about Diane's payoff as is indicated by the positive exponent on her payoff in his utility function, though not as much as he cares about his own (compare the two exponents). Anmei is entirely self-regarding, placing a zero weight on Ben's payoff; therefore all she cares about is her own payoff.

Figure 3.13 Utility maximization: Self-interested offer vs. altruistic offer. Anmei offers a split to Ben of (10, 0), whereas Chen offers Diane a split of (7, 3). Anmei's indifference curves are vertical because she gives no weight in her utility function to Ben getting any money ($(1-\alpha)=0$); therefore she keeps \$10 and Ben gets \$0. Between Chen and Diane, Chen gives some weight to Diane getting money ($(1-\alpha)=0.3$); therefore his indifference curves are shaped like indifference curves we've looked at previously and at his constrained utility maximum Chen gets \$7 and Diane gets \$3. Notice that if Chen gives any less or any more to Diane, then he would be on a lower indifference curve, such as at points **b'** and **a'** on u_1^C.



We display indifference curves for Anmei and Chen in Figure 3.13. The indifference curves in panel (a) are unusual: they are vertical because the only thing that Anmei values is what is on the horizontal axis, namely, her payoff. Using the $mrs = mrt$ rule, we find the constrained utility-maximizing point for each person where their highest indifference curve touches the feasible frontier.

In this case, though, the feasible frontier is given by a straight line because it represents a split of money. The maximum amount of money that Anmei or Chen can keep is \$10 and they can offer splits in 1 cent increments between themselves and their partners. The vertical intercept corresponds to the instance in which they give all \$10 to their partners. The horizontal intercept corresponds to the instance in which they keep all \$10 to themselves. Chen has preferences such that he would like a 70%-30% split of the \$10 (his $\alpha = 0.7$) and his highest indifference curve is tangent to the feasible frontier at a split of split of (7, 3) shown by point **b'** in Figure 3.13 (b).

Anmei has preferences such that she would like a 100%-0% split of the \$10 (her $\alpha = 1$, she places zero weight on Ben's payoff) and her highest indifference curve touches the feasible frontier at **b** in Figure 3.13 (a) at a

! REMINDER Payoffs and utility The term payoffs means money winnings in experimental games; utility is how much a person values a particular strategy profile. Here we have a case—Chen but not Anmei—in which payoffs are not the same thing as utility. Chen's utility function places a value on Diane's payoffs, not just his own.

split of (10,0) (she keeps all the money). We can interpret the slope of her indifference curves as her maximum willingness to pay in order to give Ben a small positive payoff, and ask: How much of her own payoffs would she be willing to give up to transfer a penny to Ben? The answer is that there is no amount, however small, that would motivate her to do this.

What allocation does she choose? Her highest utility is where her vertical indifference curve u_2^A touches her highest feasible allocation to herself of \$10. She keeps all the money. Her keeping all the money shouldn't surprise us because she gives no weight to Ben's payoff. In mathematics, a solution like this is called a *corner solution*. Notice that we couldn't use our standard requirement for finding the constrained utility maximum of $mrs = mrt$. But following the rules of constrained utility maximization, that Anmei would find the point in the feasible set with the highest utility, still applied to our problem and we found the solution.

M-NOTE 3.10 The *mrs* for a self-regarding dictator

Why are Anmei's indifference curves vertical in Figure 3.13? To answer this question, we need to find her marginal rate of substitution. To find her *mrs*, we need the marginal utilities of the two arguments of her utility function: π^A and π^B the money payoffs that Anmei and Ben respectively get.

Marginal utility to Anmei of Anmei's payoff:

$$u^A_{\pi^A} = \frac{\partial u^A}{\partial \pi^A} = 1 \cdot (\pi^A)^{(1-1)}(\pi^B)^0 = 1$$

Marginal utility to Anmei of Ben's payoff:

$$u^A_{\pi^B} = \frac{\partial u^A}{\partial \pi^B} = 0 \cdot (\pi^A)^1(\pi^B)^{(0-1)} = 0$$

Therefore Anmei's marginal rates of substitution is:

$$\begin{aligned} mrs(\pi^A, \pi^B) &= \frac{u_{\pi^A}}{u_{\pi^B}} \\ &= \frac{1}{0} = \text{undefined} \end{aligned} \tag{3.24}$$

Now, the result of Equation 3.24 should not surprise us because the slope of a vertical line is undefined. Anmei's indifference curves endlessly *rise* and have *no run*, so the negative of an undefined number (the slope) remains an undefined number (the *mrs*). Her indifference map therefore represents a range of vertical lines where the horizontal intercepts correspond to the amount of money she keeps which is also the utility number associated with the particular indifference curve.

Now, we might ask ourselves, what is Anmei's utility at her constrained utility maximum? Let's substitute in the values we have for $\pi^A = 10$ and $\pi^B = 0$.

$$\begin{aligned} u^A(\pi^A, \pi^B) &= (\pi^A)^1(\pi^B)^0 \\ &= (10)^1(0)^0 \\ &= (10)^1 = 10 \end{aligned} \tag{3.25}$$

Anmei has a utility that is equal to the amount of money she keeps for herself.

M-NOTE 3.11 An altruistic person splitting the pie

We will derive Chen's decision about splitting the pie between him and Diane. Using his utility function, Equation 3.23 and Equation 3.12, his marginal rate of substitution is:

$$mrs(\pi^C, \pi^D) = \frac{7\pi^D}{3\pi^C}$$

Now, let's assume that the size of the pie is $z = 10$, therefore, the feasible allocations are represented by $\pi^D = 10 - \pi^C$, so his mrt (the negative of the slope of the feasible frontier) is

$$mrt = -\frac{d\pi^D}{d\pi^C} = 1$$

Equating the mrs with the mrt, we can obtain how much Chen allocates to himself and to Diane:

$$\begin{aligned} \frac{7\pi^D}{3\pi^C} &= 1 \\ \pi^D &= \frac{3}{7}\pi^C \\ \text{Using the feasible set} \quad \frac{3}{7}\pi^C &= 10 - \pi^C \\ \frac{10}{7}\pi^C &= 10 \\ \therefore \quad \pi^C &= 7 \\ \text{and} \quad \pi^D &= 3 \end{aligned}$$

That is why Chen offers Diane \$3 of the total of \$10 that she is able to allocate.

Notice that the shares of the pie that Chen allocates to himself and to Diane are equal to the exponents of Chen's Cobb-Douglas utility function, Equation 3.23.

CHECKPOINT 3.9 Chen's choice and the *mrs* = *mrt* rule.

a. What is the marginal rate of transformation in the game described in Figure 3.13?

b. Why is the utility-maximizing bundle at point **b** in Figure 3.13 (a) an example of a case where the $mrt = mrs$ rule does not work? How does this case differ from the case shown in panel (b), where the rule does work?

c. Use the value of Chen's mrs at point **c'** in Figure 3.13 panel (b) along with the value of the mrt to explain why for Chen the opportunity cost of giving more money to Diane is less than his willingness to pay (give up his own payoffs) so that Diane can have more.

3.11 APPLICATION: ENVIRONMENTAL TRADE-OFFS

We think of environmental damage as something to be avoided, but stopping or slowing the damage—or "abating" the damage in the language of environmental science—is costly. Less damage means some combination of

HISTORY In the middle of the twentieth century, long before we worried about climate change and its unfolding calamities, Aldo Leopold, the American environmentalist raised an economic question: "Like winds and sunsets, wild things were taken for granted until progress began to do away with them. Now we face the question whether a still higher 'standard of living' is worth its cost in things natural, wild and free." [3] Leopold was articulating a trade-off between, on the one hand, consuming goods and services—Leopold's higher "standard of living"—and on the other, the costs of environmental damage—the "cost in things natural, wild and free."

less consumption, changing our consumption patterns to be less damaging to the environment, or diverting our productive potential from producing goods that we can now consume to discovering and installing new technologies. We therefore face a trade-off between consuming goods and maintaining the quality of the environment. How much of these opportunity costs of improved environmental quality are we willing to pay?

The constrained utility maximization method we have developed provides a way of posing and answering these questions using the preferences, beliefs, and constraints approach.

Feasible combinations of conventional goods and environmental quality

The opportunity cost of environmental quality is consumption of other (conventional) goods such as food, clothing, shelter, and transportation, which we must give up to secure a higher-quality environment. There is a feasible frontier showing the combinations of environmental quality, x, and conventional goods, y, that are possible for a society. The feasible frontier in the case of environmental quality depends on the *abatement technology*, which represents how much consumption of conventional goods we have to give up to achieve a given level of environmental quality.

Figure 3.14 shows a feasible frontier between conventional goods (y) and environmental quality (x). We measure environmental quality on a numeric scale from 0 (the environment that we would have if no abatement were done) to 20 (the environment resulting if we were to divert to abatement uses all of society's resources above some minimum level of consumption). We measure conventional consumption as billions of dollars.

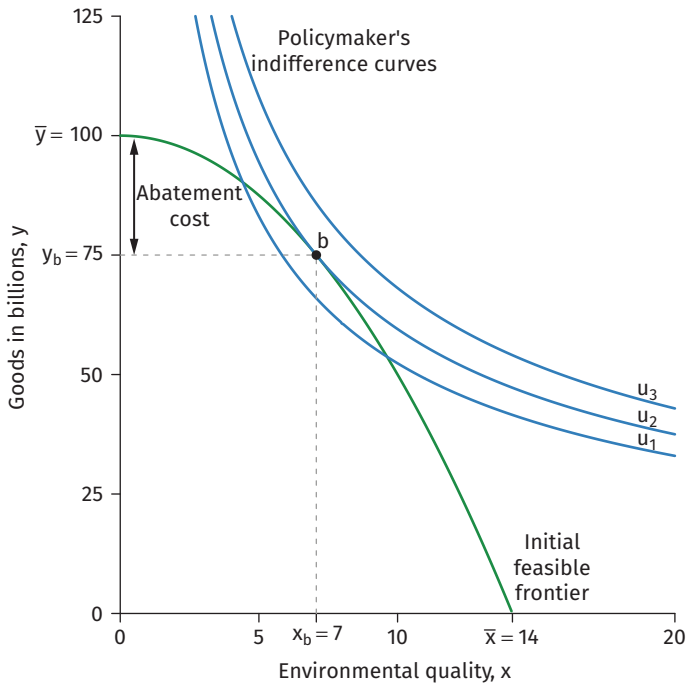
The negative slope of the feasible frontier at any point is the marginal rate of transformation of reduced environmental quality into increased conventional consumption, or $-\frac{\Delta y}{\Delta x}$. The steeper the frontier, the greater is the increase in feasible consumption allowed by a given small reduction in environmental quality.

This is also the *opportunity cost* of improved environmental quality. So a flatter frontier means a lower opportunity cost of abatement.

To see this, starting at no abatement expenditures ($y = \bar{y}$), the opportunity cost of improved environment is initially small (the frontier is nearly flat) and as the society implements more abatement, the cost of more abatement increases as the environmental quality increases. The shape of the feasible frontier reflects an increasing marginal rate of transformation, or an increasing marginal opportunity cost of environmental quality.

Put another way, if environmental quality is at its maximum at the intercept of the feasible frontier with the horizontal axis, people could consume a lot more conventional goods if it were willing to tolerate a

Figure 3.14 Trade-off between consumption of conventional goods and environmental quality. The constrained utility maximum is the point on the feasible frontier on the highest indifference curve u_2, shown as point **b** where the $mrs = mrt$ rule holds. The constrained maximum is at the point where the feasible frontier is tangent to the highest attainable indifference curve.



small deterioration of environmental quality (the frontier is steep where it intercepts the horizontal axis). But the feasible increase in consumption of conventional goods allowed by a reduction in environmental quality falls as the level of environmental quality declines.

CHECKPOINT 3.10 Increasing opportunity cost of environmental quality
Explain how the shape of the feasible frontier in Figure 3.14 illustrates the increasing opportunity cost of environmental quality.

3.12 APPLICATION: OPTIMAL ABATEMENT OF ENVIRONMENTAL DAMAGES

How much abatement is the right level, taking account of both preferences for conventional goods (consumption) and the quality of the environment along with the opportunity costs in lost consumption?

M-CHECK An example of the function representing the feasible frontier is shown below, where y is the goods available for consumption, $\bar{y} = 100$ is the level of y that is feasible when environmental quality is at its minimum, and x is environmental quality.

$$y = 100 - \frac{1}{2}x^2 \quad (3.26)$$

This is the equation for the feasible frontier that is graphed in Figure 3.14.

A citizen chooses a level of abatement of environmental damages

To begin with the simplest case, think of just one citizen, Anmei, who might be representative of the attitudes of the whole society, trying to decide on the level of abatement that she would like to see implemented. She cares about both the quality of the environment, x, and the amount of conventional goods that will be available for people to consume.

Anmei's utility function has the following form: $u = u(x, y)$. Anmei considers what she would like to see her society do about the environment (x), taking account of the effects on everyone. In other words, she is thinking from an other-regarding perspective, like an ideal policymaker.

Anmei's indifference curves between environmental quality and conventional goods are downward-sloping because she regards both environmental quality and conventional consumption as goods for which more is better. This means the marginal utility of both y and x are positive (e.g. $u_y > 0$ and $u_x > 0$). The negative slope of the indifference curves shown in Figure 3.14 at any point is Anmei's marginal rate of substitution between more consumption of goods and a better environment. Her marginal rate of substitution shows the amount of goods she would be willing to give up for a small improvement in the environment. As before, Anmei's indifference curves exhibit diminishing marginal utility of both environmental quality and consumption.

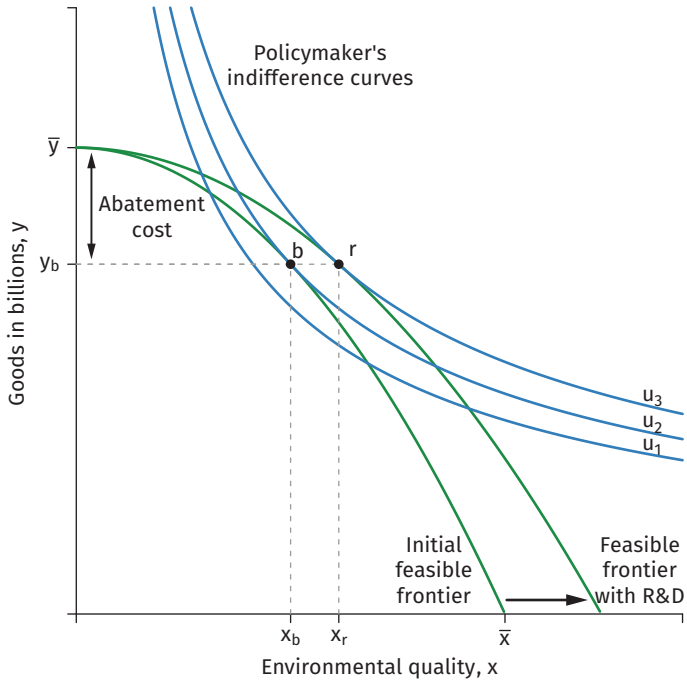
An example of a utility function that Anmei might have is the Cobb-Douglas utility function:

$$u(x, y) = x^{\alpha} y^{(1-\alpha)} = x^{0.4} y^{0.6} \tag{3.27}$$

Figure 3.14 shows three indifference curves defined by Equation 3.27: u_3 is unattainable given that the feasible frontier, u_1 intersects the feasible frontier twice, and u_2 is tangent to the frontier at point (x_b, y_b). Anmei's constrained maximum allows her and her fellow citizens to consume 75 million units of conventional goods and enjoy environmental quality of about 7 (see M-Note 3.12 for the worked solution). If she were able to implement relevant environmental and fiscal policies, this point is the best society can do in Anmei's opinion.

What is the total opportunity cost in foregone conventional consumption of a level of environmental quality of 7? The maximum feasible level of conventional consumption with no abatement is \$100 billion. The difference between the maximum feasible consumption of \$100 billion and Anmei's preferred choice of conventional consumption of \$75 billion is the opportunity cost of an environmental quality of 7. In our example, the abatement costs are equal to \$100 billion – \$75 billion = \$25 billion in conventional goods. A citizen with Anmei's preferences thinks that the sacrifice of \$25 billion consumption goods is worth paying to have an environmental quality of 7 instead of zero.

Figure 3.15 Trade-off between consumption of conventional goods and environmental quality with R&D. The choice between consumption of conventional goods and environmental quality with R&D pivoting the feasible frontier outwards leading to a new point of tangency on a higher indifference curve u_3. Given the functions we used to make the figure y_b does not change, but y_b could also increase or decrease.



New technologies, and conflicts of interest

If, with a mind to the future, some of the abatement costs are devoted to research to improve abatement technologies, this would pivot the feasible frontier outward, as shown in Figure 3.15.

As is shown in Figure 3.15, the shift of the feasible frontier would permit higher environmental quality of $x \approx 9.8$ at the same level of consumption of $75 billion. But there would still be a trade-off: more conventional goods would require less environmental quality, or more environmental quality would require fewer conventional goods to stay on the feasible frontier.

We can also use Figure 3.14 to see why people often disagree about environmental policy.

- Preferences: people's *preferences* for conventional goods and the environment may differ.
- Beliefs: people may disagree about the opportunity costs or the benefits of environmental quality.
- Conflicts of interest: the costs and benefits of abatement fall on different people; those whose jobs or profits depend on carbon-based energy or

✓ **FACT CHECK** The pace of environment-friendly innovation is astounding. Have a look at the reduction in costs of the photovoltaic cells used in solar panels dropping to one-onehundreth of their costs in 1975 in Figure 8.3: they cost about $105 in 1975 and the cost decreases to $0.72 by 2014.

who do a lot of airline travel, for example, stand to bear more of the costs of addressing climate change, while regions likely to be particularly hard hit by climate change in the absence of abatement like Africa bear a larger share of the benefits.

M-NOTE 3.12 The trade-offs and opportunity costs of the environment

Let us work through the process that Anmei the policymaker would go through to identify the combination of goods in billions of dollars with environmental quality.

First, let us calculate her marginal rate of substitution from her utility function, $u^A(x,y) = (x^A)^{0.4}(y^A)^{0.6}$. From earlier in the chapter, we know that the $mrs(x,y)$ is the ratio of marginal utilities and we have already calculated this for $\alpha = 0.4$ and $(1-\alpha) = 0.6$ in Equation 3.17 in M-Note 3.7.

$$mrs(x,y) = \frac{2y}{3x} \tag{3.28}$$

Anmei's feasible frontier, based on her beliefs and understanding of the existing science, is given by the equation $y = 100 - \frac{1}{2}x^2$, for which we can find her $mrt(x,y) = -\frac{dy}{dx}$:

$$\frac{dy}{dx} = -x$$

$$\therefore -\frac{dy}{dx} = x \tag{3.29}$$

We now set the $mrt(x,y)$ given by Equation 3.29 equal to the $mrs(x,y)$ given by Equation 3.28 and we isolate one of the variables, y:

$$x = \frac{2y}{3x}$$

Multiply through by $3x$ $\quad 3x^2 = 2y$

Divide through by 2 $\quad y = \frac{3}{2}x^2 \tag{3.30}$

We can now substitute Equation 3.30 into the feasible frontier to find x_b and y_b, the values that solve Anmei's constrained optimization problem, the best she can do:

$$\frac{3}{2}x^2 = 100 - \frac{1}{2}x^2$$

$$2x^2 = 100$$

$$x^2 = 50$$

$$\therefore x_b = \sqrt{50} = 7.07$$

Having found x_b, we can substitute it back into Equation 3.30 to find y^b:

$$y = 100 - \frac{1}{2}(\sqrt{50})^2$$

$$= 100 - \frac{1}{2}(50)$$

$$y_b = 75$$

So, as a result of Anmei's policymaking utility function and feasible frontier, she would choose a combination of environmental quality, x, of value 7.07 with consumption of good and services of \$75 billion; \$75 billion is \$25 billion less than the maximum consumption of goods and services, $\bar{y} = 100$ billion, so the cost of abatement is \$25 billion.

CHECKPOINT 3.11 **Differences and conflicts** Draw two versions of Figure 3.15 showing two people, one favoring a substantial amount of abatement and the other favoring little. Illustrate this in your two figures by differences in their indifference curves showing their preferences, and differences in the feasible frontier showing what they believe to be the opportunity costs of abatement.

3.13 CARDINAL INTERPERSONALLY COMPARABLE UTILITY: EVALUATING POLICIES TO REDUCE INEQUALITY

Most policy choices involve conflicts of interest like these concerning the abatement of environmental harms. Few policy choices are entirely win-win. Most policies—whether they concern taxation, immigration, health insurance, or the rate of inflation—result in benefits for some and losses for others.

Ordinal and cardinal utility in policy evaluation

How do we then evaluate competing policies? Don't think about this as a question about what would be a good outcome for *you* if you were a participant in the society. Instead, try to take the position of what Adam Smith called the Impartial Spectator who did not himself stand to gain or lose, but wanted instead to consider the gains and losses impartially.

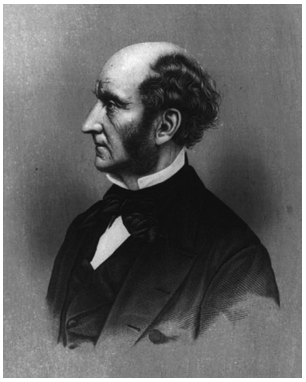
One answer you might give is just to count those who prefer each policy and select the most popular policy. All this requires is that people be able to rank the policies in question as better, worse, or indifferent. We could in this case treat utility as ordinal (that is, simply a ranking (or ordering) of outcomes).

Something like this might occur in a majority rule democratic political system, especially if citizens could vote on policies as they do in Switzerland and other countries in referendums asking citizens to vote for or against a particular policy.

But this way of evaluating policies might result in evaluating positively those policies that confer minor gains to those in favor, and substantial losses to those preferring another policy. This does not seem like an ideal rule. An alternative is to weigh the amount of the gains to the beneficiaries of each policy against the size of the costs incurred by those who would have done better under some other policy. This kind of comparison requires that we know not only *which* policies people prefer, but *how much* they prefer them. To do this we treat utility as a cardinal measure for which utility is not just an ordinal ranking, but instead a number indicating how well-off the person is under the option in question. As we explained in

Figure 3.16 **John Stuart Mill (1806–1873)**. Philosopher-economist John Stuart Mill referred to what we would now call the sum of the total utilities of a population as "a good" that should be promoted: "the general happiness is desirable . . . each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons."[4]

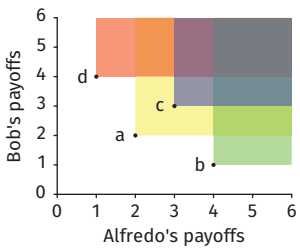
Photo credit: LOC.



HISTORY Adam Smith in *The Theory of Moral Sentiments* conceived of the Impartial Spectator as follows, "We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it. If, upon placing ourselves in his situation, we thoroughly enter into all the passions and motives which influenced it, we approve of it, by sympathy with the approbation of this supposed equitable judge. If otherwise, we enter into his disapprobation, and condemn it."[5]

HISTORY Lionel Robbins (1898–1984) was a leader in the "ordinal revolution" in economics. Economics, he wrote, does not need "to compare the satisfaction which I get from the spending of 6 pence on bread with the satisfaction which the Baker gets by receiving it. That comparison . . . is never needed in the theory . . ." Moreover, "There is no way of comparing the satisfactions of different people."[6]

Figure 3.17 Pareto comparisons in the Fisherman's Dilemma Game. The Pareto criterion selects **c** over **a** (which it dominates) but cannot rank points **b**, **c**, and **d** because all of them are Pareto efficient.



section 3.3, treating utility as cardinal allows us to say two very different things:

1. for Anmei, the outcome (x', y') is twice as good as (x, y) because for example $u^A(x', y') = 2u^A(x, y)$;
2. the sum of the Anmei's and Brenda's utility is greater with outcome (x', y') than with outcome (x, y) because $u^A(x', y') + u^B(x', y') > u^A(x, y) + u^B(x, y)$.

Both statements involve cardinal utilities, but they differ. The first statement compares how much *Anmei* values two different states that she will experience. It does not compare her evaluation of a state that she will experience with someone else's evaluation of the state they will experience. The first statement is an example of the cardinal utility that we introduced in Chapter 2 as the basis of expected payoffs (or expected utility) and the analysis of decision-making in risky situations.

The second statement compares Anmei's utility *with Brenda's utility*. When utility is represented in this way it is called *interpersonally comparable cardinal utility* (or sometimes "cardinal full comparable utility"). If we consider utility to be cardinal in this interpersonally comparable sense, then we can compare how well-off two or more people are, and how much better off or worse off a policy would make each of them. This provides a way to evaluate which policies should be implemented by asking whether the gains of those who benefit from a policy exceed the losses of those who do not.

Why do these two methods of comparing utility matter? Remember that one of the problems with Pareto efficiency as a criterion for policy outcomes is that *many* outcomes can be Pareto efficient. So Pareto efficiency does not provide an adequate basis for an Impartial Spectator preferring one outcome over the other. Using the second–stronger–conception of cardinal utility along with the judgment that one outcome is better than another if total utility is greater provides a rule for evaluating *which* Pareto-efficient outcome we might prefer as a society.

In the payoffs for the Fishermen's Dilemma in Figure 1.13 (shown here in the margin for easy reference) three of the four outcomes of the game are Pareto efficient. The Pareto criterion provides no way to choose among them. By contrast the rule–maximize total utility–selects the mutual cooperate outcome (point **c**) with total utility of 6.

Cardinal utility and the distribution of wealth

To see what adopting the "maximize total utility" rule would mean, imagine that the Impartial Spectator is given the task of dividing a given amount of wealth between Anmei and Brenda. The amount of wealth she has to divide is equal to 1, so each person can get a fraction of that wealth and, as long as the fractions sum to 1, then the outcome will be Pareto efficient. Let the fraction that Anmei gets be a and Brenda's fraction be $b = 1 - a$. Anmei and

Brenda have identical preferences for wealth given by the identical cardinal utility functions: *own utility* = *u*(*own wealth*) so:

- Anmei's utility: $u(a)$.
- And, because $b = 1 - a$, Brenda's utility: $u(b) = u(1 - a)$.

The horizontal axis in Figure 3.18 shows all possible distributions of wealth between Anmei and Brenda.

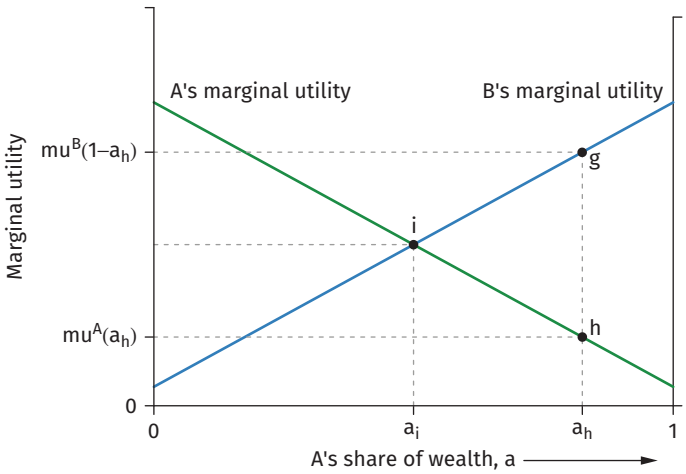
- Anmei's share of wealth, a, varies from 0 to 1.
- At $a = 0$, Anmei gets nothing and Brenda gets everything.
- At $a = 1$, Brenda gets nothing and Anmei gets everything.

Figure 3.18 also shows the two marginal utility of wealth functions. For each of them the marginal utility of wealth declines the more wealth they have. So, in the figure Anmei's marginal utility curve slopes *downward* as she gets more wealth (moving from left to right). Remembering that for Brenda farther to the right means less wealth (the opposite of Anmei) her curve slopes *upward* because as she gets less wealth (moving from left to right) her marginal utility of wealth is higher.

Suppose the status quo is a_h, a situation in which Anmei is wealthy and Brenda is poor (Anmei's has share of wealth a_h and Brenda's share of wealth

EXAMPLE When you say "I'll do the shopping; it'll be less trouble for me than for you" you are representing utility (the trouble of shopping) as cardinal and making an interpersonal comparison of utility (less trouble for me than for you). In fact, much of our everyday ethical reasoning and even simply being a good friend require making interpersonal comparisons of the benefits and costs that people experience.

Figure 3.18 Distribution of wealth, marginal utility, and total utility. In the figure a is the proportion of wealth belonging to Anmei (A). $(1 - a)$ is the proportion of wealth belonging to Brenda (B). As a person's wealth increases, the marginal utility of wealth decreases. A's total wealth *increases* as you move along the bottom line from left to right, and as a result her marginal utility *decreases*. Because B's wealth increases as the division moves toward the left, B's marginal utility decreases from right to left. The utility functions used to create this figure are shown in M-note 3.13.



is $1-a_h$). Anmei's marginal utility at a_h, $u_a^A(a_h)$ is lower than Brenda's utility at the same point $u_b^B(b_h)$. The vertical difference between points **g** and **h** shows the magnitude of the difference in their marginal utilities.

Because **g** > **h**, we know that a policy that takes a small amount of wealth from Anmei and transfers it to Brenda (moving to the left from a_h) reduces Anmei's utility by less than it increases Brenda's. Redistributing wealth from Anmei to Brenda therefore increases *total* utility (the sum of Anmei's and Brenda's utilities).

Applying this reasoning to other points in the diagram, we find that the distribution of wealth that maximizes total utility is a_i, where Anmei's marginal utility of wealth equals Brenda's marginal utility of wealth. Because Anmei's and Brenda's utility functions are identical, the total utility-maximizing point distributes wealth equally, $a_i = \frac{1}{2}$.

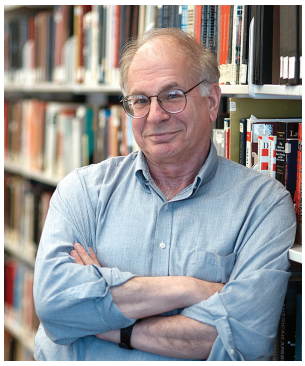


Figure 3.19 Daniel Kahneman (1934–), a psychologist and Nobel Laureate in economics, has advocated a *hedonistic* (meaning concerning pleasure and pain) theory of utility. Kahneman titled one of his papers "Back to Bentham?" to pay homage to the early nineteenth-century philosopher economist Jeremy Bentham's utilitarian theory.[7]

Photo credit: U.S. National Institutes of Health.

M-NOTE 3.13 Maximizing total utility

Adam Smith's Impartial Spectator would like to find the distribution of wealth, a^W, such that the sum of the utility of the two will be as large as possible. Here are their utility functions (the same as those used to create Figure 3.18).

$$\text{A's utility:} \quad u^A(a) = a\left(1-\frac{a}{2}\right) \tag{3.31}$$

$$\text{B's utility:} \quad u^B(b) = u^B(1-a) = (1-a)\left(1-\frac{1-a}{2}\right) \tag{3.32}$$

$$\text{The Impartial Spectator maximizes} \quad W = u(a)+u(1-a) \tag{3.33}$$

By differentiating W with respect to a, we can find a_i, the Impartial Spectator's welfare-maximizing choice of a:

$$W_a = \frac{du(a)}{da} + \frac{du(1-a)}{da}$$

Differentiating the two utility functions (Equations 3.31 and 3.32) with respect to a to find the two marginal utilities and setting the result equal to zero:

$$= (1-a)-a = 0$$

$$\therefore \quad a_i = \frac{1}{2}$$

CHECKPOINT 3.12 Redistribution of wealth Explain how Figure 3.18 shows that A owning the fraction a_h of the wealth does not maximize the sum of the utilities of A and B, and why redistributing some of A's wealth to B would increase total utility.

3.14 APPLICATION: CARDINAL UTILITY AND SUBJECTIVE WELL-BEING

A century ago economists thought that while ordinal comparisons like better or worse are possible, empirical interpersonal comparisons expressed by a number indicating the degree of preferences were impossible to

make. But today researchers are actively engaged in measuring individual happiness and life satisfaction, using techniques ranging from surveys and natural observation to the methods of experimental neuroscience. They are asking such questions as: "How important is income for happiness?" "Is being without a job a bigger source of unhappiness than being without a spouse?" These researchers refer to happiness or life satisfaction as *subjective well-being*.

To measure "pleasures and pains" in the lab, volunteers are exposed to an electrical shock and asked to report on their experience of that on a numerical scale. Others are asked to plunge their hands into extremely cold water for as long as they can stand it and immediately report their level of unhappiness having done so. Respondents in surveys are asked their "life satisfaction."

This research has sought to understand the activities that make people most happy. Almost all people surveyed seem to like sex quite a lot, ranking "intimate relations" as having a high subjective well-being value. Ranked after sex, people like socializing, relaxing, sharing meals with friends, praying, and exercising. People don't like housework, childcare, commuting, or working. People also report major changes in subjective well-being from painful events, like sudden loss of a job, a death in the family, or divorce, or from positive events like marriage, or the birth of a child.[8] But when you ask someone about their happiness over time the measures are surprisingly consistent: people are likely to report similar activities or outcomes as providing them with happiness when you ask them at different intervals.

What are the take-home messages about subjective well-being?

First, we can measure happiness and the degree to which people are satisfied with their lives and also identify the things that contribute to a person's subjective well-being.

Second, people who report greater subjective well-being are also better off by physical and biological measures. For example, they are less likely to be ill. Subjective well-being also manifests in hormone levels, brain patterns, and palm temperature.[10]

Third, while income matters for happiness (especially for people without much income) people value social relationships—marriage, a job, friendships—more than they value income.[11] Making the transition from unemployed to employed boosts a person's subjective well-being by much more than would be predicted simply by the increase in income. This is because having a job is a source of respect and dignity, especially as it provides a way for people to express autonomy, competence in their expression of their abilities, and relatedness to other co-workers and people around their work.

✓ **FACT CHECK** The Satisfaction with Life survey is based on five questions each of which is rated on a 7-point scale from Strongly Disagree (1) to Strongly Agree (7). Here are the questions: In most ways my life is close to my ideal; The conditions of my life are excellent; I am satisfied with my life; So far I have gotten the important things I want in life; and If I could live my life over, I would change almost nothing.

✓ **FACT CHECK** Non-laboratory measures of subjective well-being suggest that people with higher subjective well-being tend to be less likely to contract a cold virus and to recover more quickly when they do contract the cold. Similar evidence exists for people who have recovered from wounds and had baseline and subsequent subjective well-being measured: those who are happier recover more quickly.[9]

› **EXAMPLE** The substantial subjective cost that people experience when they are out of work is one reason why employers (who have the power to terminate a person's job) have power over their employees. We shall return to this when we study the firm and the labor market.

CHECKPOINT 3.13 **Joy or misery?** Think about the kinds of activities discussed above that provided people with joy (that they ranked highest in terms of providing them with subjective well-being).

a. Compare them with their opposites: those that result in disutility or even misery.

b. Come up with a list of activities that you engage in that provide you with joy which you try to prioritize.

c. Why do you spend the time that you do on these activities? Why do you not spend more?

d. Do you engage in activities that in the moment are unpleasurable but which you believe provide you with benefit nonetheless?

e. Do you think such activities appear in the models we've developed?

3.15 PREFERENCES, BELIEFS, AND CONSTRAINTS: AN ASSESSMENT

Many scholarly disciplines in addition to economics are devoted to understanding human behavior including psychology, sociology, anthropology, archaeology, and history, but also more distant endeavors including literature, philosophy, neuroscience, computer science, and biology. The preferences, beliefs, and constraints approach, while a standard set of tools in economics that is widely used in other fields, is just one of many approaches.

People newly familiar with the approach often raise the following questions about it.

- *Are people really all that selfish*? This concern is based on a misunderstanding of the model, which says nothing about whether people are seeking to help others, aggrandize themselves, or a little of both. Our treatment of altruism, reciprocity, and fair-mindedness shows that the model—using indifference curves and feasible sets, for example—can apply to a variety of motives.
- *Do people consciously optimize*, for example, applying the $mrs = mrt$ rule when they shop? The model is not a description of how people actually think or their emotional states when they take a break from studying, or support a particular environmental policy. We model instead what people would do if they did the best that they could. The fact that the model often yields predictions similar to what we observe empirically (including by experiments, econometric, and other quantitative methods) does not require that the model is an accurate representation of the process by which people come to take one course of action over another.

HISTORY In his 1953 work, *Essays in Positive Economics*, Nobel Laureate Milton Friedman (1912–2006) observed that "predicting the shots made by an expert billiard player" could be done on the basis of "the complicated mathematical formulas that would give the optimum directions" of the shots. But this prediction would not be "based on the belief that billiard players, even expert ones, can or do" actually make these calculations.[12]

In some cases, people consciously optimize, going though mental calculations similar to the model. For example, a person buying a house or choosing between two job offers will weigh the pros and cons of the

alternatives. But in other cases, the actions may not even appear to us as a decision, for example, what to eat for breakfast, what to wear today, or what our personal values should be. Without consciously trying to do so, people may arrive at something like the solution to these optimization problems by trial and error, or by observing others who seem to be successful or happy with their choices, or by following habits that will remain in place unless changed by some dramatically adverse consequences of following them.

Other concerns about the model are more serious.

- *What about emotions and visceral reactions, aren't they important*? This question points to a shortcoming of the approach; but it is not that the approach excludes emotions like fear, shame, and attraction. The shortcoming is that the preferences, beliefs, and constraints approach says almost nothing at all about motives; that is, it says nothing about the reasons *why* people rank some outcome as superior to another. Knowing more about motives like this would help us understand economic and other behavior.
- *Commitments and consequences*. The framework is based on the idea that our behavior is based on our beliefs about the consequences our actions will bring about in the future. Don't we sometimes act to fulfill promises or other commitments made in the past, or just to "do the right thing" without regard to future consequences? Yes, we do, and a shortcoming of the approach is that it does not address that kind of behavior.
- *Predicting behavior and evaluating outcomes*. Economists use the same concept "utility" both in models designed to predict the actions that people will take and also to provide the basis for evaluating economic outcomes and public policies to improve them. The idea is that whatever it is that motivates people to make the choices they do should also be the objective of public policy and form the basis for our preferring one societal outcome over another. But treating actual behavior as if it were the pursuit of a concept of well-being that should be the basis of our judgment of societal outcomes is a mistake. The reasons for our actions (that is, our preferences) include addictions, weakness of will, shortsightedness, and other well-documented socially dysfunctional aspects of human behavior that in retrospect are often deeply regretted by those acting on them.

A sensible conclusion from reviewing these concerns about the preferences, beliefs, and constraints approach might be that the approach is better for answering some questions than others, and learning to distinguish which is which is an important learning objective. As we said at the beginning of the chapter: the map is not the territory. Good maps don't have all the information about the territory they depict and good economic models require us to leave some things out.

HISTORY Positive and normative economics The distinction between the economics of "what is" called positive economics and "what ought to be" called normative economics was made by John Maynard Keynes in his 1893 *Scope and Method of Political Economy* and by Milton Friedman in his 1953 *The Methodology of Positive Economics*. The distinction is controversial in part due to differences about the appropriate role for "what ought to be" statements in economics.[13]

HISTORY Jeremy Bentham (1746–1832) is considered the founder of the philosophical tradition called utilitarianism, which forms the basis of much economic thinking both positive and normative. His most famous book begins: "Nature has placed mankind under the governance of two sovereign masters, *pain and pleasure*. It is for them alone to point out what we *ought to do*, as well as to determine what *we shall do*. On the one hand the standard of right and wrong, on the other the chain of causes and effects are fastened to their throne." [14]

CHECKPOINT 3.14 Utility in the evaluation of policy Use the case of smoking to illustrate the difference between using the concept of utility to predict behavior and also to evaluate societal outcomes. Economists use the addictive nature of tobacco to better understand preferences for smoking. Should our evaluation of anti-smoking policies include, as a cost of the policy, the frustrated craving for smoking experienced by the targets of such a policy?

3.16 CONCLUSION

In this chapter we have studied the constrained optimization problems shown in Table 3.1. Though the problems concerned are quite different, the models and analytical tools we used to analyze them are very similar. In each case the analysis of the decision involves two kinds of trade-offs:

- The first trade-off that appeared in each of these situations is the actor's relative valuation of the things she cares about, measured by the negative of the slope of an indifference curve, that is, the marginal rate of substitution.
- The second trade-off is that at any point on the feasible frontier, the opportunity cost of having more of one good that the actor values is that she must have less of another good that she values. This opportunity cost trade-off is measured by the negative of the slope of the feasible frontier, that is, the marginal rate of transformation.

The result—the action taken doing the best she can under the constraints she faces—is determined in the same way in all the cases: by finding the point on the feasible frontier that is on the highest indifference curve. This will often be the bundle where the $mrs = mrt$ rule holds. Table 3.1 demonstrates that many seemingly different kinds of action can be studied with a common model, one that we will use often.

In this chapter, we have focused on single actors and for the most part excluded from the model something important: other people. With the

Table 3.1 The constrained optimization problems used in this chapter.

Actor	Utility depends on	Action	*Constraints*
Keiko	Learning, Living time	Time allocation	Learning-Living feasible frontier
Keiko	Learning, study effort	Study effort	Study-Learning production function
Anmei/Chen	Payoffs to two players	Distribution of endowment	The total endowment
Anmei/Brenda	Conventional goods Environmental quality	Choice of amount of conventional goods	Consuming conventional goods degrades the environment
Impartial Spectator	Wealth of A and B	Redistribute wealth	Limited amount of wealth to distribute

exception of the farmers of Palanpur, we have modeled the person facing a given situation defined by a feasible frontier and preferences represented by indifference curves. (We already explained that the second person in the Dictator Game is not really a player at all.)

We now turn to a world populated by people interacting strategically, and we ask how economic institutions—the rules of the game—affect the outcomes of these interactions and in particular:

- Is the resulting allocation one in which all of the potential gains from exchange have been realized?
- Have the rules of the game advantaged some players at the expense of others, resulting in unequal outcomes?

We shall continue to employ the tools of constrained utility maximization, understanding people's trade-offs through their marginal rates of substitution and of transformation. We will need these tools—and a few new ones—to understand the mutual gains made possible by people interacting and the conflicts that necessarily arise over how these gains will be distributed.

MAKING CONNECTIONS

Strategic and nonstrategic social interactions: In the previous chapters we considered strategic social interactions—like the fishermen and the farmers from Palanpur. Here we look at simpler aspects of behavior when a person is attempting to do the best they can in situations that are not strategic because the choice of how hard to study, or how much fish to buy is not greatly affected by others' choices.

Self-regarding and social preferences: In Chapter 2 we provided evidence that people can be self-regarding, altruistic, reciprocal, spiteful, and fair-minded. These diverse behaviors can be modeled using the preferences, beliefs, and constraints approach by means of indifference curves and feasible frontiers, as we showed for the case of an altruist.

Opportunity costs and trade-offs: Regardless of whether a person's preferences are entirely self-interested or not, people face trade-offs among the ends they wish to pursue and they face opportunity costs when trying to choose a course of action.

Public policy: Economics engaged: The idea of constrained utility maximization illustrated the trade-off between consuming more goods on the one hand or either consuming less and using some of the economy's resources to abate environmental damages, obtaining greater environmental quality. We also modeled the choices an altruistic person might make in sharing something of value thereby providing a model capable of analyzing the kinds of result observed in the experiments reviewed in the previous chapter.

Evaluating outcomes: Treating utility as cardinal and inter-personally comparable rather than ordinal allows us to compare the benefits and burdens

that a policy will impose on different people. This provides a basis (one of a number of alternatives) for saying that one policy or outcome might be preferred to another, as illustrated by the case of the distribution of wealth.

IMPORTANT IDEAS

preference	constraints	beliefs
Homo economicus	altruist	reciprocator
ordinal utility	cardinal utility	utility function
Cobb-Douglas utility	total utility	marginal utility
diminishing marginal utility	*mrs* = *mrt* rule	slope
indifference curve	marginal rate of substitution	diminishing marginal
trade-off	willingness to pay	rate of substitution
marginal product	feasible frontier	production function
opportunity cost	marginal rate of transformation	increasing marginal rate of transformation
feasible set	increasing opportunity costs	subjective utility
interpersonally comparable utility	abatement (of environmental costs)	offer curve
distribution of wealth	feasible frontier	corner solution
price line	pleasures and pains	

MATHEMATICAL NOTATION

Notation	Interpretation
$u()$	utility function
x	a good (or a "bad")
y	a good (or a "bad")
h	hours of studying
α	Cobb-Douglas exponent of good x
$\bar{y}$	vertical intercept of the feasible frontier
a	A's share of wealth
p	price of a good
$\bar{u}$	constant utility along an indifference curve
z	endowment in the Dictator Game
π	payoff in the Dictator Game

Note on superscripts and subscripts: A, B, C, D: different people; subscript b indicates where someone does the best they can; RD: feasible frontier with R&D.

CHAPTER

4

PROPERTY, POWER AND EXCHANGE

MUTUAL GAINS AND CONFLICTS

[T]he efforts of men are utilized in two different ways: they are directed to the production or transformation of economic goods, or else to the appropriation of goods produced by others.

Vilfredo Pareto,
Manual of Political Economy (1906)[1]

DOING ECONOMICS

This chapter will enable you to:

- Explain why, when people exchange goods, there are both mutual gains and conflict over the distribution of these gains.
- Understand how an allocation of goods can be evaluated on grounds of Pareto efficiency and fairness.
- Show how preferences affect the outcome, and how other-regarding social preferences may reduce the scope of conflicts over the distribution of the gains from exchange.
- Understand how property rights, the exercise of power, and other aspects of the rules of the game will affect the extent of mutual gains and the inequality of their distribution.
- Use mathematical tools (equations and graphs) to illustrate the above points.

Figure 4.1 Burning Man is an annual gathering of tens of thousands in the Nevada desert in the U.S. What you see is the city (tents and other dwellings) constructed by the "burners." Nothing will remain after they leave. Burners cannot use money to purchase goods. Instead, people give goods they bring or make or services they provide to others. See here how Nobel Prize–winning economist Paul Romer spent his time at Burning Man (https://www.nytimes.com/2019/09/05/upshot/paul-romer-burning-man-nobel-economist.html).

Photo: DigitalGlobe via Getty Images.

4.1 INTRODUCTION: "STRANGE AND HARD TO BELIEVE"

Ibn Battuta, the fourteenth-century Moroccan scholar, reported that along the Volga River in what is now Russia, long-distance trade took the following form:

> Each traveler . . . leaves the goods he has brought . . . and they retire to their camping ground. Next day they go back to . . . their goods and find opposite them skins of sable, miniver, and ermine. If the merchant is satisfied with the exchange he takes them, but if not he leaves them. The inhabitants then add more skins, but sometimes they take away their goods and leave the merchant's. This is their method of commerce. Those who go there do not know whom they are trading with or whether they be jinn [spirits] or men.[2]

The Greek historian Herodotus describes similar exchanges between Carthaginian and Libyan groups in the 5th century BCE. After having left their goods, Herodotus reports, the Carthaginians withdraw and the Libyans "put some gold on the ground for the goods, and then pull back away from the goods. At that point the Carthaginians . . . have a look, and if they think there is enough gold to pay for the cargo they take it and leave."

Herodotus describes how the process continues until an acceptable price is hit upon, remarking with surprise that "neither side cheats the other . . . [the Carthaginians] do not touch the gold until it is equal in value to the cargo, and [the Libyans] do not touch the goods until the Carthaginians have taken the gold."

Alvise da Ca da Mosto, a fifteen-century Venetian working for the Portuguese crown, reported a similar practice in the African kingdom of Mali, regarding it as "an ancient custom which seems strange and hard to believe."

But is the so-called silent trade really so odd? Transfers of goods among strangers can be dangerous. What one expected to be an exchange at mutually agreeable prices may end up as theft or an "offer you cannot refuse." But trade among strangers can also be highly profitable. The potential gains from trade are often greater, the more distant geographically or socially the parties are to the exchange: the salt brought by Tuaregs from the Atlas Mountains in North Africa across the Sahara by camel to the Kingdom of Ghana was not available at any price in West Africa.[3] The gold and tropical nuts Tuaregs gained in silent trade with Ghanaians was not available north of the Sahara.

The silent trade—with its unusual etiquette in which parties interacted only at a distance—allowed both Tuaregs and Ghanaians to get some of what

they lacked and wanted in return for giving up some of what they had in abundance and could readily part with.

They were exploiting the mutual gains that differences in geography, tastes, technologies, and skills allow. And the rules of the game for governing their exchange process—the institutions that we call "the silent trade"—were a way of doing this and dividing the mutual gains without risking violent conflicts.

Other than these mutually advantageous exchanges, there are many other ways that goods change hands: from the use of violent coercion by private parties (i.e. theft), or by the use of one nation's military force to acquire the resources of another people. People have also been violently coerced into work through enslavement by private actors and states alike.

A key characteristic of these coerced transfers is that they are *not* motivated by mutual gain, but instead by the gain of one party facilitated by superior force and institutional power. These transfers of resources and lives have shaped the course of history and have had important economic consequences and enduring legacies.

But here we set aside the use of physical coercion and ask how societies organize the process of exchange motivated not by fear of physical harm but instead by the prospect of mutual gain. We also provide terms that allow us to evaluate some of these outcomes as better or worse than others. And we will see that depending on the rules of the game the distribution of the mutual gains made possible by exchanges may be highly unequal.

4.2 MUTUAL GAINS FROM TRADE: CONFLICT AND COORDINATION

In a modern economy we engage in indirect monetary exchange: selling some of our goods or some of our working time for money and using the money we have acquired to purchase the goods we need. We typically do not barter directly as did the Libyans and Carthaginians. The principles of *barter exchange*, where goods are directly transferred between two parties without the use of money, however illustrate the fundamental considerations behind all types of exchange, including indirect monetary exchange.

! REMINDER As in Chapter 3 a bundle is just a list of the quantity of the goods (or other thing of value) that a person has. We refer to the bundles held by all of those involved in an exchange as an *allocation*.

We will simplify by thinking about just two people who exchange goods directly with each other, thereby modifying the goods that they hold. To do this we will introduce three terms describing the *bundles* that each has before and after exchange:

- *The endowment bundle*: or *endowment*, the quantities of goods a person has before exchanging goods.

- *The post-exchange bundle*: the bundle a person has after exchanging goods with another person.
- *The* **allocation:** bundles held by each of the people (either before or after exchange).

Voluntary exchange: mutual gains and conflict over their distribution

! REMINDER Recall from Chapter 1 that a change is a Pareto improvement if it makes at least one person better off and none worse off.

An exchange is **voluntary** if all parties to the exchange have the option to not engage in it but instead choose to engage in the exchange. So each party must expect to be better off, or at least expect to be no worse off, as a result of the exchange. This implies that each party prefers (at least weakly) their post-exchange bundle to their endowment bundle.

Recalling the meaning of a Pareto comparison, we can see that if an exchange is voluntary for both parties, the post-exchange allocation must be a Pareto improvement over the endowment; otherwise one or both of the parties would have refused to participate in the exchange.

To make the idea of voluntary exchange concrete we often let the fallback position of the players be a bundle of goods that is their **private property** which they are free to dispose of in exchange or by gift to others, or to retain for themselves, excluding others.

Let's review some of the terminology from earlier chapters and explain how they are used to study the process of exchange.

- A person's *fallback position* is what they experience in the absence of the exchange under consideration, and also the utility number they assign to that bundle (that is, the utility of their endowment bundle, which is considered to be her *next best opportunity*).
- The improvement in utility enjoyed by a party to an exchange is their *rent* resulting from the exchange, namely, the difference between the utilities associated with their post-exchange bundle and their fallback position.

ALLOCATION In a game, an allocation is a particular distribution of goods or other things of value to the players.

VOLUNTARY EXCHANGE An exchange is voluntary if neither party can coerce the other to participate nor require them to renew the exchange beyond an agreed-upon duration.

PRIVATE PROPERTY The right and expectation that one can enjoy one's possessions in ways of one's own choosing, exclude others from their use, and dispose of them by gift or sale to others who then become their owners.

- The total rents received by parties to an exchange, also termed the *gains from trade* are the utilities of the exchanging parties at the outcome of the exchange minus the utilities at their fallback positions.

The fact that an exchange is voluntary does not mean that it is fair. Some exchanges take place under conditions such that one party gains virtually all of the available rents. How the economic rents are divided between participants is the **distributional outcome** of the exchange. The rents may be captured by one party, leaving the other with a different set of goods than her endowment but no better off.

Or the rents may be split among the parties in a way that appears fair, or at least acceptable to both, as in the silent trade between the Carthaginians and the Libyans described by Herodotus. The division of the gains from exchange in the form of economic rents is parallel to the division of the pie in the Ultimatum Game of Chapter 2.

Exchange therefore has two aspects: *mutual benefit* and *conflict of interest*:

- *Mutual benefit* is possible because participants move from their endowment bundle to the post-exchange allocation where they share the gains from exchange and obtain an *economic rent*.
- A *conflict of interest* is present because the gains from exchange can be divided in many ways among the parties who find themselves in conflict over who gets the larger share.

Institutions and social norms govern the process of exchange that leads both to the reallocation of goods, and to the distribution of the gains from trade. We will see that institutions and social norms have effects on:

- *Pareto efficiency*, facilitating or obstructing the realization of every opportunity for mutual gain among the parties to an exchange; and
- *The degree of inequality of the distributional outcome*, favoring one party or the other in the conflict of interest in the distribution of the economic rents.

A major institutional challenge today is to find *rules of the game* that will have as a Nash equilibrium an allocation that is *both* Pareto efficient and fair.

CHECKPOINT 4.1 **Conflict of interest** Make sure that you understand the terms fallback position, voluntary exchange, mutual benefit, and conflict over the distribution of the mutual benefits made possible by exchange. (A good way to check is to use each term in a sentence of your own.)

REMINDER An economic rent is the difference between a player's *fallback* payoff and the payoff (profit or utility) they obtain from participating in an interaction. The total gains from exchange from an interaction is the sum of the economic rents of all participants.

DISTRIBUTIONAL OUTCOME How the gains from exchange—the economic rents—are distributed among the people in an exchange.

HISTORY The Edgeworth box is named after the British economist Francis Ysidro Edgeworth (1845–1926) who is credited with having invented this clever way to represent exchange and bargaining. It was actually invented by Pareto![4]

4.3 FEASIBLE ALLOCATIONS: THE EDGEWORTH BOX

Let's think about a concrete setting in which two people might consider alternative possible distributions of two goods between them. Let's say that Ayanda and Biko have to divide a total of 10 kilograms of coffee and 15 gigabytes of data between them. (For concreteness, they found the coffee and a burner phone with the data left behind by the students who moved out of the apartment they just rented.) At the start, nobody owns the goods, the two quantities are simply amounts available to the two of them. Ayanda and Biko might now ask each other: What allocation of the coffee and data between the two of us would be the best?

We use the notation $\bar{x} = 10$ and $\bar{y} = 15$ to stand for the *total* amount of coffee (x) and data (y) available. We define x^A and y^A as the quantities of goods x (coffee) and y (data) in Ayanda's bundle, and similarly x^B and y^B are the quantities in Biko's.

The amount of the two goods in their respective bundles can be anywhere from zero to the entire amount available, namely, $\bar{x}$ and $\bar{y}$. Then, an allocation is a particular assignment of coffee and data to the two people that we can write as $(x^A, x^B; y^A, y^B)$. An allocation is *feasible* if the amounts of coffee and data it gives to Ayanda and Biko is no greater than the amount available:

$$x^A + x^B \leq \bar{x}$$
$$y^A + y^B \leq \bar{y}$$

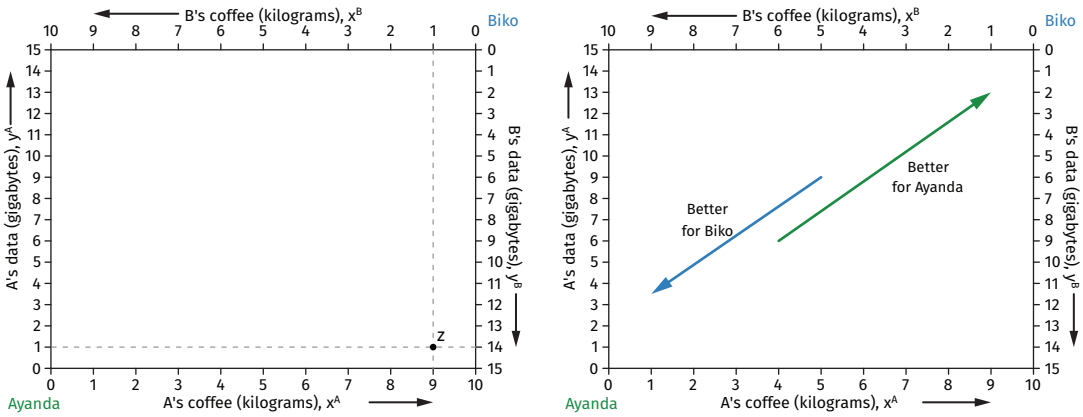
Figure 4.2 (a) represents the *total supply* of the goods, with width and height equal to the total amount of coffee (x) and data (y) available. The box's *width* is the total amount of x, $\bar{x}$ (kilograms (kg) of coffee) and its *height* is the total amount of y, $\bar{y}$ (gigabytes (gb) of data). We measure A's allocation, (x^A, y^A) from the lower-left-hand corner of the box, and B's allocation, (x^B, y^B) from the upper-right-hand corner.

Any point in the box (or on its edges) is a bundle representing a *feasible* allocation of the two goods between the two parties, with the property that it fully *exhausts the total supply* of the two goods.

Allocation **z**, for example, gives Ayanda 9 kilograms of coffee and 1 gigabyte of data and Biko 1 kilogram of coffee and 14 gigabytes of data (exhausting the 10 units of x and the 15 units of y).

There are also many feasible allocations of the two goods that are *not* shown in the box. For example, if Ayanda and Biko each got 1 kilogram of coffee and one gigabyte of data, that would be feasible given the total amounts, but it could not be shown in the Edgeworth box because the Edgeworth box shows only allocations where the two people divide up *all* of the goods so that they sum to $\bar{x}$ and $\bar{y}$.

Figure 4.2 Feasible allocations that exhaust the supply of both goods. Figure 4.2 (a) shows an example of a feasible allocation at point **z**. Figure 4.2 (b) shows the direction in which each person prefers to move to increase their utility. When indifference curves are plotted in this rectangle the graph is called an Edgeworth box.



(a) An allocation, **z** (b) Preferred bundles

A caution: it is natural to think that Ayanda might be at one point in the box and Biko at another. This is not possible. A point in the box is *not* just a single bundle for one or the other of them. It is an allocation of goods *between* the two. They differ in where they would like that allocation to be.

As we move to the northeast (up and to the right) in the box, Ayanda gets more of both goods, and as we move to the southwest (down and to the left) in the box, Biko gets more of both goods. Because we assume for now that both are self-regarding we show this on the figure with the arrows labeled: "Better for Ayanda" and "Better for Biko."

How can we evaluate whether some allocations are better than others? To do this we can represent the preferences of the two parties by plotting their indifference curves in the box. This allows us to say for both Ayanda and Biko that for any two allocations (points in the box) the first is preferred to the second, the second is preferred to the first, or the person is indifferent between the two. To do this we need to know the utility functions of the two.

Both Ayanda and Biko enjoy consuming both coffee and data. Their utility functions are:

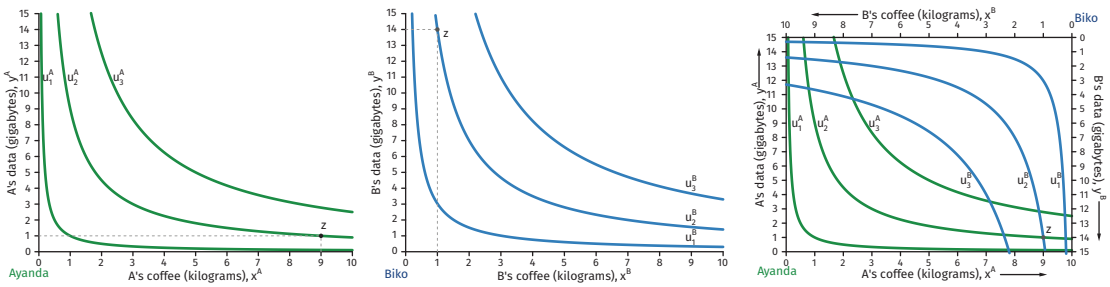
Ayanda's utility function $u^A(x^A, y^A)$

Biko's utility function $u^B(x^B, y^B)$

We assume that the indifference curves for both parties exhibit decreasing marginal utility for both goods. To provide a concrete example, we will

! REMINDER In Chapter 1, we used **z** to indicate the fallback position of people playing games in the general form of the Fishermen's game with ranked outcomes. At **z**, the people experience their utilities, u_z^A and u_z^B, as their utility at the fallback position, that is, their endowments if they do *not* trade.

Figure 4.3 Indifference curves in an Edgeworth box: Identical utility functions. In panels (a) and (b) we show three of Ayanda's and Biko's indifference curves respectively. In panel (c), Biko's indifference curves have been flipped 180° so that the origin in the lower left of panel (b) has become the origin of the Edgeworth box at the upper right.



(a) Ayanda's indifference curves (b) Biko's indifference curves (c) Indifference curves in an Edgeworth box

assume that both Ayanda's and Biko's utility functions are Cobb-Douglas, but in some cases that follow, with different preferences for coffee and data:

Ayanda's utility function $$u^A(x^A, y^A) = (x^A)^{\alpha^A}(y^A)^{(1-\alpha^A)}$$

Biko's utility function $$u^B(x^B, y^B) = (x^B)^{\alpha^B}(y^B)^{(1-\alpha^B)}$$

In numerical examples we will often contrast two cases:

- *Identical*: The two people have identical preferences for the two goods, such as $\alpha^A = \frac{1}{2}, \alpha^B = \frac{1}{2}$.
- *Different*: The two people have *different* preferences, for example, such that A's $\alpha^A = \frac{2}{3}$, whereas for B $\alpha^B = \frac{1}{3}$. So Ayanda has a stronger preference for coffee than Biko does.

An Edgeworth box allows us to see both people's indifference curves in the same space to identify mutually beneficial trades. Ayanda and Biko's indifference curves are shown separately in Figure 4.3 panels (a) and (b). In panel (c) we plot the same indifference curves together in the Edgeworth box. Ayanda evaluates the allocations from the point of view of the lower-left-hand corner, and her indifference curves represent higher utility as we move to the northeast in the box.

Ayanda's indifference map looks exactly the same in the Edgeworth box as it does in the separate plot, because in both cases the origin from which we measure her allocation is in the lower-left-hand corner. But Biko's map has been flipped so that the origin is in the upper-right corner and his indifference curves represent higher utility as we move to the southwest (down and to the left) in the box.

It may help you understand how we superimposed Biko's preferences on Ayanda's if you think about what we called their "point of view." In panels (a) and (c), imagine Ayanda standing at the lower-left origin and looking up her indifference map, as if the curves were contours of a mountain, the curves farther away being at higher altitudes. Now do the same with Biko, but for him when he looks to the northeast in panel (b), he is looking up his "utility mountain." But in panel (c) he is standing at the upper-right origin and the way up his utility map is to the southwest.

In the figures, at allocation **z** Ayanda and Biko have allocations $(x_z^A, y_z^A) = (9,1)$ and $(x_z^B, y_z^B) = (1,14)$. The indifference curves that go through allocation **z** provide Ayanda and Biko with utilities $u_z^A = u_2^A$ and $u_z^B = u_2^B$.

In panels (a) and (c), $u_2^A = u_z^A$ is Ayanda's indifference curve through **z**. In panels (b) and (c), $u_2^B = u_z^B$ is Biko's indifference curve through **z**. The indifference maps for both Ayanda and Biko have indifference curves through every point in the box, but (following "the map is not the territory" principle) we show only three in the figure.

M-NOTE 4.1 Evaluating utilities at an allocation

Consider the case in which both utility functions are Cobb-Douglas with Ayanda's $\alpha^A = \frac{2}{3}$ and Biko's $\alpha^B = \frac{1}{3}$. We can calculate their utilities at the allocation **z**. In this example, Ayanda likes coffee more than Biko does.

Ayanda has a Cobb-Douglas utility function $u^A(x^A, y^A) = (x^A)^{\frac{2}{3}}(y^A)^{\frac{1}{3}}$:

- She has 9 kilograms (kgs) of coffee and 1 gigabyte (gb) of data.
- So her allocation at point **z** is $(x_z^A, y_z^A) = (9,1)$.
- At her allocation **z** her utility is $u^A(x_z^A, y_z^A)$.
- So for 9 kgs of coffee and 1 gb of data: $u^A(9,1) = (9)^{\frac{2}{3}}(1)^{\frac{1}{3}} = 4.33$.

CHECKPOINT 4.2 Biko's utility at allocation z Using the two utility functions shown in the text with $\alpha = \frac{1}{2}$, calculate the utility of the two at the allocation **z**, and at an alternative allocation in which Ayanda has exchanged one kg of her coffee for one gb of Biko's data.

4.4 THE PARETO-EFFICIENT SET OF FEASIBLE ALLOCATIONS

Which allocations in the Edgeworth box are Pareto efficient?

It's easy to see that simply throwing away some of x or y cannot be efficient because allocating those portions to Ayanda and/or Biko instead would have made at least one of them better off without making the other worse off. So Pareto efficiency also requires that $x^A + x^B = \bar{x}$ and $y^A + y^B = \bar{y}$. By construction, any of the great many allocations in the Edgeworth box

! REMINDER In Chapter 3 we defined the Cobb-Douglas (CD) family of utility functions as:

$$u(x,y) = x^{\alpha}y^{(1-\alpha)}$$

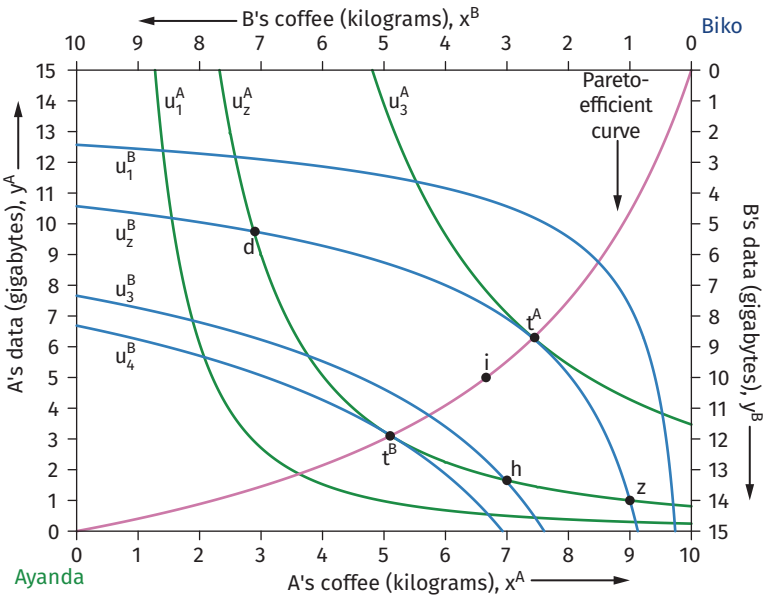
(with $0 \le \alpha \le 1$). The Cobb–Douglas utility function results in a marginal rate of substitution, $mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x}$.

M-CHECK Biko's indifference map would look exactly the same as in Figure 4.3 (b) if we rotated the Edgeworth box 180° to measure Biko's allocation from the lower-left-hand corner.

! REMINDER In games like the Ultimatum Game in Chapter 2 any allocation of the pie in which the entire endowment is allocated to one of the players or the other (in other words, "no money left on the table") is Pareto efficient. But the allocations resulting from the Ultimatum Game are frequently inefficient because when the Responder rejects the Proposer's offer both players get zero, and *all* of the money is left on the table.

M-CHECK Even if for some reason we were not to allow the allocation to involve fractional quantities of the goods and require that allocations be integers, there are 176 possible allocations (that's 11×16, in case you are wondering, because we would then have to include zeroes as possible allocations for the goods).

Figure 4.4 Pareto-efficient allocations: Different utility functions. To make this figure we set $u^A = (x^A)^{\frac{2}{3}}(y^A)^{\frac{1}{3}}$ and $u^B = (x^B)^{\frac{1}{3}}(y^B)^{\frac{2}{3}}$. So Ayanda has a stronger preference for coffee, and Biko has a stronger preference for data. Allocation **h** is Pareto superior to allocation **z**, but it is not Pareto efficient because an alternative point, e.g. allocation $\mathbf{t^B}$, is Pareto superior to point **h** (Biko is better off without Ayanda's being worse off). All points along the Pareto-efficient curve between **i** and $\mathbf{t^B}$ are both Pareto superior to **h** and **z** and Pareto-efficient.



! REMINDER The *marginal rate of substitution* is the negative of the slope of the indifference curve. It is also equal to the ratio of the marginal utilities of the two goods, x and y, i.e. $mrs^A(x,y) = u_x^A/u_y^A$. The marginal rate of substitution is also the willingness to pay for x in terms of y. The people's marginal rates of substitution have the dimensions data/coffee (data for coffee).

! REMINDER For an outcome to be *Pareto superior* to another, at least one participant must be made better off—get higher utility—and no participant can be made worse off—get lower utility.

allocates all of the coffee and data to one or the other participant, and meets this criterion.

To narrow things down, Ayanda and Biko could agree that the final allocation chosen must be Pareto-efficient. In Figure 4.4 we show Ayanda and Biko's indifference curves through some arbitrary endowment allocation **z** that they might consider as a way of dividing up the goods. The figure also shows two more indifference curves for Ayanda: one indifference curve higher and one indifference curve lower than for allocation **z**. The figure also shows three more indifference curves for Biko: two indifference curves higher and one indifference curve lower than for allocation **z**.

The endowment allocation is not Pareto efficient

Think about **z** as a hypothetical allocation, for example, if Biko said: "Ayanda, how about you have 9 kg of coffee and I get the 1 kg remaining, while I get 14 gb of the data, and you get the 1 gb remaining." We can see, however, that **z** in Figure 4.4 is not Pareto efficient. The reason is that at the allocations given by point **z**, Ayanda's and Biko's indifference curves:

- *intersect*, which means
- they have *different slopes*,
- indicating *different marginal rates of substitution*,
- which means their willingness to pay to acquire more of one or the other good differ,
- and this means that there is *a feasible Pareto-improving exchange that has not been realized*,
- so these allocations are *not Pareto efficient*.

The difference between the two people's marginal rates of substitution at the point **z** indicates that there must be a Pareto-improving alternative allocation—Ayanda having less coffee and more data and the opposite being the case for Biko. In other words there is some allocation to the "northwest" of allocation **z**, for example the allocation at point **i**, that is a Pareto improvement over the allocation at **z**. So we can eliminate point **z** in Figure 4.4 as a candidate for being a Pareto-efficient allocation.

M-NOTE 4.2 **The *mrs* in the Edgeworth box with different utility functions**

At allocation **z** $(9,1;1,14)$ in Figure 4.4, we can calculate each person's marginal rate of substitution and compare them. We computed what a person's $mrs(x,y)$ is when she has Cobb-Douglas utility in M-Note 3.4. We obtain Biko's from the same reasoning. We shall assume for this example that the two have *different* preferences as in Figure 4.4.

Let's start with Ayanda, given that $\alpha^A = \frac{2}{3}$ and recalling (from Equation 3.12) that with the Cobb-Douglas function $mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x}$:

- $mrs^A(x,y) = \frac{u_x^A}{u_y^A} = 2\frac{y^A}{x^A}$
- Substitute in A's allocation at **z**: $mrs_z^A(x_z^A, y_z^A) = 2 \times \frac{1}{9} = \frac{2}{9}$

Ayanda is willing to sell a kilogram of coffee for $\frac{2}{9}$ of a gigabyte of data.

Now for Biko, $\alpha^B = \frac{1}{3}$ and again recalling that $mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x}$:

- $mrs^B(x^B, y^B) = \frac{u_x^B}{u_y^B} = \frac{1}{2}\frac{y^B}{x^B}$
- Substitute in B's allocation at **z**: $mrs_e^B(x_e^B, y_e^B) = \frac{1}{2}\frac{14}{1} = 7$

Biko is willing to pay 7 gigabytes of data for a kilogram of coffee.

We can see that $mrs^A < mrs^B$ because $\frac{2}{9} < 7$. This means that Biko would pay up to 7 gigabytes of data for a kilogram of coffee, and Ayanda would give up 1 kilogram of coffee for as little as $\frac{2}{9}$ of a gigabyte of data. There is ample space between these two prices for a mutually beneficial trade. This shows up in Figure 4.4: the slope of Ayanda's indifference curve is steeper than the slope of Biko's indifference curve at allocation **z**.

Which allocations are Pareto efficient? The $mrs^A = mrs^B$ rule

The same reasoning allows us to eliminate most of the other points too. Remember the demonstration that showed point **z** to be Pareto inefficient started with "at the allocations given by these points Ayanda's and Biko's indifference curves intersect."

So *any* allocation at which the indifference curves intersect, like points **d**, **h**, and **z** in Figure 4.4 *cannot* be Pareto efficient.

To find the Pareto-efficient allocations, we need to determine which allocations remain after we have eliminated all of those at which the indifference curves cross. To do this we can run the above reasoning in reverse.

If the two indifference curves (one of Ayanda's, one of Biko's) share a common point (that is, that represent the utilities at a particular allocation) but do not *intersect*, then the two indifference curves must be *tangent*. This tells us (reversing the logic above about indifference curves that intersect) that if Ayanda's and Biko's indifference curves:

- are *tangent*, this means that
- they have the *same slopes*, indicating
- *identical marginal rates of substitution*,
- meaning that Ayanda and Biko have *the same willingness to pay for the two goods*.
- This is the same as saying that *their maximum willingness to pay to acquire more of the other's good is not greater than the least price at which the other would part with their good*
- and this means that there is *no feasible Pareto-improving alternative allocation that could be implemented by an exchange*,
- so the *status quo allocation is Pareto efficient*.

This gives us the following rule for an allocation between two players, A and B, being Pareto efficient:

$$\text{The } mrs^A = mrs^B \text{ rule:} \qquad mrs^A(x^A, y^A) = mrs^B(x^B, y^B) \tag{4.1}$$

This rule differs from the seemingly similar $mrs = mrt$ rule introduced in section 3.9. The $mrs = mrt$ rule applies to a single individual and it identifies a constrained optimum for that person. This new rule applies to interactions among two or more interdependent actors, of the kind that occur in markets for labor, credit, and many goods. It identifies a Pareto-efficient point for the people involved in the interaction. The superscripts A and B are there to remind you that two (or more) players are involved in this rule. The two tangency rules are compared in Table 4.1.

! REMINDER The $mrs = mrt$ rule We derived a similar rule for single-person interactions in Chapter 3. The $mrs = mrt$ rule (with a few exceptions) identifies the constrained utility-maximizing optimal allocation for a single individual as the bundle at which the marginal rate of substitution (the person's willingness to pay for more of the y-good) is equal to the marginal rate of transformation (the opportunity cost of getting more of the y-good).

M-CHECK Like the $mrs = mrt$ rule, $mrs^A = mrs^B$ does not work in every case. The Pareto-efficient point may be a corner solution (not a tangency) at which one of the goods is not consumed at all by one of the players. Moreover, just like the $mrs = mrt$ rule, a tangency identified by the $mrs^A = mrs^B$ rule may be a minimum not a maximum. The reasons are the same as were explained for the $mrs = mrt$ rule.

Table 4.1 Two rules: individual constrained optimization and societal Pareto efficiency.

Rules	Tangency of	Rule for what
$mrs = mrt$	An individual's feasible frontier and indifference curve	Individual constrained optimization
$mrs^A = mrs^B$	Two or more people's indifference curves	Societal (multi-person) Pareto efficiency

The points $\mathbf{t}^A$, $\mathbf{t}^B$, and $\mathbf{i}$ lie on the purple **Pareto-efficient curve** in Figure 4.4. The Pareto-efficient curve consists of all Pareto-efficient allocations, including Ayanda getting all of both goods, or none of either. The Pareto-efficient curve is sometimes called the "contract curve," a term we do not use because there need not be any contract involved.

Confining allocations to the Pareto-efficient curve limits the choices that Ayanda and Biko need to make. But the question is still far from answered. Moving from one Pareto-efficient allocation to another must make one of the participants better off and the other worse off. The Pareto-efficiency criterion is not going to help them decide which of the points *on* the Pareto-efficient curve they would consider to be the best.

So they face a problem and a conflict of interest.

- *The problem* is that there are still innumerable Pareto-efficient outcomes on the Pareto-efficient curve and they need some way to decide which one to choose.
- *The conflict of interest* is that Ayanda prefers points on the Pareto-efficient curve to the northeast in the Edgeworth box, while Biko prefers points to the southwest, so they will not agree on which Pareto-efficient division of the coffee and data to make.

M-NOTE 4.3 Computing the Pareto-efficient curve

Taking the case in which the two have different utility functions (that we studied in M-Note 4.2), we will use $mrs^A = mrs^B$ rule to work out the equation for the Pareto-efficient curve.

To find the Pareto-efficient curve, we set Ayanda's marginal rate of substitution equal to Biko's marginal rate of substitution. We already know that $mrs^A(x^A, y^A) = 2\frac{y^A}{x^A}$ and $mrs^B(x^B, y^B) = \frac{1}{2}\frac{y^B}{x^B}$. We also know that

continued

PARETO-EFFICIENT CURVE The points making up the Pareto-efficient curve represent all of the allocations that are Pareto efficient.

$\bar{x} = x^A + x^B = 10$, so $x^B = \bar{x} - x^A$ and $\bar{y} = y^A + y^B = 15$, so $y^B = \bar{y} - y^A$. Solutions to these equations for x^A, y^A, x^B, y^B are Pareto-efficient allocations making up the Pareto-efficient curve:

$$\begin{aligned} mrs^A(x^A, y^A) &= mrs^B(x^B, y^B) \\ 2\frac{y^A}{x^A} &= \frac{1}{2}\frac{\bar{y} - y^A}{\bar{x} - x^A} \\ \text{Substitute } \bar{x} = 10 \text{ and } \bar{y} = 15 \qquad 4\frac{y^A}{x^A} &= \frac{15 - y^A}{10 - x^A} \\ 4(10 - x^A)y^A &= x^A(15 - y^A) \\ 40y^A - 4x^Ay^A &= 15x^A - x^Ay^A \\ (40 - 3x^A)y^A &= 15x^A \\ \text{Pareto-efficient curve} \qquad y^A &= \frac{15x^A}{40 - 3x^A} \end{aligned}$$

CHECKPOINT 4.3 Conflict on the Pareto-efficient curve

Using Figure 4.4 do the following:

a. Explain Ayanda's and Biko's preference among the Pareto-efficient points $\mathbf{t^A}$, $\mathbf{t^B}$, and $\mathbf{i}$.

b. Show that they rank these points in opposite order.

c. Explain why, for any two points on the Pareto-efficient curve, Ayanda will prefer one point and Biko another point; they will never agree on which is preferable.

4.5 ADAM SMITH'S IMPARTIAL SPECTATOR SUGGESTS A FAIR OUTCOME

EXAMPLE To see how maximizing total utility might lead to unacceptable outcomes, think about two people, one who in order to minimize her carbon footprint or for other ethical reasons has cultivated a simple lifestyle and is not much interested in increasing her material consumption and the other who has cultivated a taste for luxuries and will be miserable without them. Maximizing total utility would require giving most of the goods to the second person.

Not wanting to waste time fighting over who gets more of the goods limited to $\bar{x}$ and $\bar{y}$, Ayanda and Biko have to figure out an institution or set of rules to pick an allocation. This means stepping back and looking at the problem *without* thinking about their own particular preferences. They would probably experiment with some simple rules. They could adopt:

- the "*finders keepers*" rule and allocate the goods to whoever had first discovered the discarded coffee and data; but this might not seem fair;
- the *fifty-fifty* norm of the landlords and sharecroppers in Chapter 2, and each take half the quantity of the two goods; but if they have different preferences (as is the case in panel (b) of Figure 4.4) splitting both goods equally would not even be Pareto efficient (an equal split is not on the purple Pareto-efficient curve).
- the *maximize total utility principle*; but this places no value on equality, and might result in selecting an allocation in which one person had most of the goods (and utility) and the other little of either.

To develop more satisfactory rules, they might consult Adam Smith's Impartial Spectator, a fair and impartial observer who can assist them (and us) in reasoning about what a good outcome might be. We use uppercase letters for her name to remind you that she is an entirely made-up character, not a person at all, and not a part of the game in which Ayanda and Biko are engaged. The Impartial Spectator is a thought experiment representing our conscience, allowing us to explore differing values and how they could lead us (and Ayanda and Biko) to select a particular allocation as the best.

We're going to follow the Impartial Spectator's thinking by looking at different fairness criteria that she could adopt. For example, she could ask:

- Are the *procedures* that determined the allocation fair?
- Is the *outcome* itself fair?

The first criterion is referred to as a *procedural* judgment, and therefore she judges the outcome based on the *procedure* used to acquire the goods. She would ask for example if the original endowment bundles had been acquired *fairly*, for example through hard work, or perhaps as a gift from someone who themselves had acquired the goods fairly. If they had acquired the goods through previous trade, the Spectator would go on to inquire if the process of trading had itself been fair: For example did either of them have unfair advantages in determining the price at which they would exchange.

The second criterion is called *substantive*: it asks about the *substance* of the resulting allocation—how much do each of them get—asking for example if it is fair (no matter how it came about).

Both criteria are important, but we will focus on the *substantive* judgements because it allows us to illustrate how the Impartial Spectator could select the "best" allocation by solving a constrained optimization problem. For the Impartial Spectator to make judgments among Pareto-efficient allocations that give Ayanda and Biko different levels of utility using the constrained optimization method, she needs two pieces of information:

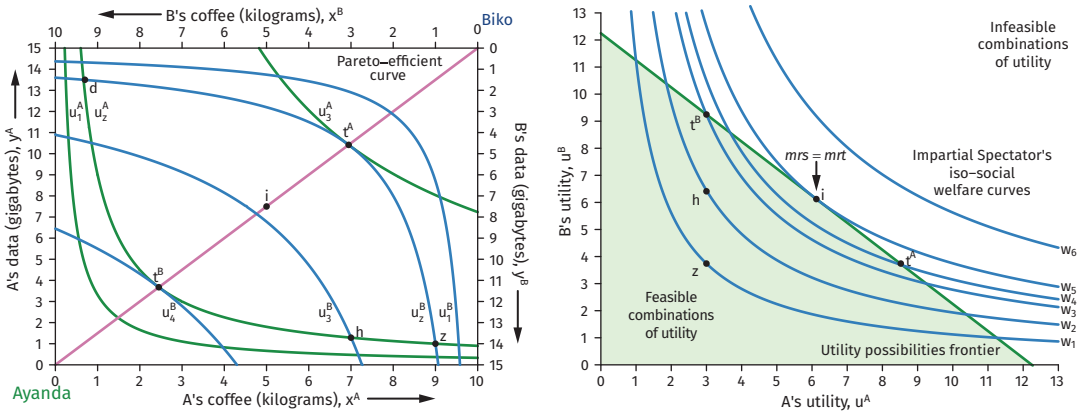
- the set of all Pareto-efficient combinations of utility levels that Ayanda and Biko could experience by allocating the goods in different ways;
- the value that she places on each of these combinations of the utility levels of the two.

! REMINDER You met the Impartial Spectator in section 3.13 in Chapter 3. There she was called on to determine a fair distribution of wealth, so the decision was one-dimensional: there was just one thing—wealth—being divided up. The problem here is two-dimensional: there are two goods and the Spectator will have to develop a rule for allocating both of them between Ayanda and Biko.

The utility possibilities frontier

Setting aside Pareto-dominated allocations, the Impartial Spectator will concentrate on the boundary of the set of feasible utility pairs of the two.

Figure 4.5 The utility possibilities frontier (UPF) and the Impartial Spectator's iso-social welfare curves (w). The utility functions of the two players used to create panel (a) are identical, with in both cases $\alpha = 0.5$. Because they both value the two goods in the same way, they consume them in the same proportions at all points on the Pareto-efficient curve. The only difference is which player has more. Each point in panel (b) corresponds to an allocation in the Edgeworth box shown in panel (a). The downward-sloping blue curves in panel (b) are the Impartial Spectator's iso-social welfare curves, corresponding to six levels in her judgment of social welfare w_1 through w_6. Social welfare is lower at points closer to the origin. The allocation given by point **i** is the social optimum determined by the $mrs = mrt$ rule.



(a) Identical preferences Edgeworth box

(b) The utility possibilities frontier (UPF) and the Impartial Spectator's iso-social welfare curves (w)

This is called the **utility possibilities frontier** (UPF) and it shows the pairs of Ayanda's and Biko's utilities associated with allocations on the Pareto-efficient curve.

In Figure 4.5 (a) we show an Edgeworth box of the two player's allocation problem in which they have identical preferences. In panel (b), we show the utility possibilities frontier for this case. For the moment, ignore the downward-sloping blue curves.

The utility possibilities frontier is downward-sloping because the participants are in *conflict*. We are considering only Pareto-efficient points, so any increase in the utility of one must be associated with a reduction in the utility of the other. The utility possibilities frontier is constructed from the Pareto-efficient curve by translating each Pareto-efficient allocation $(x^A, y^A; x^B, y^B)$ into a point $(u^A(x^A, y^A), u^B(x^B, y^B))$ that represents the utility levels of the two participants at that allocation. To construct it, take any

UTILITY POSSIBILITIES FRONTIER The utility possibilities frontier is composed of the feasible Pareto-efficient combinations of utilities of the members of a population.

point on the Pareto-efficient curve in Figure 4.4, say point $\mathbf{t^A}$, then read from the two indifference curves through $\mathbf{t^A}$ the two levels of utility of Ayanda and Biko at that allocation (namely 8.52 and 3.74 respectively), then go to Figure 4.5 (b) where those two utility levels become the coordinates in the utility possibilities graph of point $\mathbf{t^A}$ in the Edgeworth box graph.

Points $\mathbf{t^A}$, **i**, and $\mathbf{t^B}$ correspond to the same lettered points in Figure 4.4 and portray the *utilities* of each of the two at these Pareto-efficient allocations. In similar fashion, points **z** and **h** correspond to the same letters in Figure 4.4, but these allocations, being Pareto inefficient, are of no interest to the Impartial Spectator.

As in the case of other feasible frontiers, the negative of the slope of utility possibilities frontier, $-\frac{\Delta u^B}{\Delta u^A}$, is the marginal rate of transformation of B's utility into A's utility by progressively giving A more of the goods and B less. This is also the opportunity cost—in the Spectator's reasoning—of A having more utility in terms of the sacrifice in B's utility necessary to allow this. A steep utility possibilities frontier means that for A to gain one unit of utility, B must sacrifice a lot.

! REMINDER The utility possibilities frontier is similar to what we did in Chapter 1 to understand the Pareto efficiency of different game outcomes. The utility possibilities frontier is another feasible frontier introduced in Chapter 3, since it shows the feasible combinations of utility possible given the available goods and the preferences of the participants.

CHECKPOINT 4.4 The utility possibilities frontier and the Pareto-efficient curve

a. Explain why the utility possibilities frontier in Figure 4.5 is downward-sloping.

b. Explain why, if the utility functions of the two *differ*, an even split of the two goods—half of each to Ayanda and half of each to Biko—could *not* be the Impartial Spectator's choice of the best allocation.

The Impartial Spectator's social welfare function

Which point on the utility possibilities frontier—in other words which allocation of the goods between Ayanda and Biko—the Impartial Spectator ranks as best will depend on her values. She has to compare how much she values the utility of Ayanda and Biko respectively and how this varies depending on the level of utility that each are experiencing.

To do this she has to be able to compare the levels of utility for the two for each of the allocations on the utility possibilities frontier. She needs to treat the utility of each like ordinary numbers that measure the *size* not just the *rank* of something, in this case the *cardinal utility* of each.

A summary of the Impartial Spectator's evaluation of different utility distributions (u^A, u^B) is provided by her **social welfare function**, $W(u^A, u^B)$. This is similar to the utility function that expresses a person's preferences

SOCIAL WELFARE FUNCTION A social welfare function is a representation of "the common good" based on some weighting of the utilities of the people making up the society.

! REMINDER Assigning cardinal utility numbers to bundles means that we can make statements like:

- for Ayanda, the outcome (x', y') is twice as good as (x, y) but also
- the sum of the utility experienced by Ayanda and Biko is greater with outcome (x', y') than with outcome (x, y) because $u^A(x', y') + u^B(x', y') > u^A(x, y) + u^B(x, y)$.

The sum of the utilities of the two—in the second statement—is an example of a social welfare function.

over bundles of goods, (x, y), but remember the Impartial Spectator is not a person, but a thought experiment. This is why it is called the social welfare function rather than the Impartial Spectator's utility function.

An example is a social welfare function that expresses *total welfare* as the *product* of the utility of the citizens, each utility raised to some exponent.

$$\text{Example of social welfare function:} \qquad W(u^A, u^B) = (u^A)^{\lambda}(u^B)^{1-\lambda} \tag{4.2}$$

This social welfare function has the same form as a Cobb-Douglas utility function: the participants' levels of utility are the "goods" for the Impartial Spectator. When $\lambda = 0.5 = 1 - \lambda$, then the Impartial Spectator:

- weights the two people's utilities *equally*; and
- places diminishing marginal value on increases in the utility of either of Ayanda or Biko; the greater is their utility, the less they add to the Impartial Spectator's assessment of social welfare.

Because the Impartial Spectator values the two people's utilities equally, and (in the judgment of the Spectator) the marginal value of increased utility is diminishing, she will not rank highly any outcome in which one or the other gains most of both goods.

Just as we can use indifference curves to represent a person's utility function over goods, we can use **iso-social welfare curves** to represent the Impartial Spectator's social welfare function over the utility distribution between people. The level of social welfare is the same along an iso-social welfare curve, just as utility was the same along an indifference curve. A set of iso-social welfare curves, w_1 to w_6, are depicted in Figure 4.5 (b) with higher numbers in the subscripts indicating higher levels of social welfare, just as with indifference curves earlier.

Given the Impartial Spectator's social welfare function, the problem of choosing the Pareto-efficient allocation of coffee and data becomes a constrained optimization problem similar to those in Chapter 3. The feasible frontier for the Impartial Spectator is the utility possibility frontier, because it represents the levels of utility that are achievable given the amount of goods available and the preferences of the participants.

The iso-social welfare curves of the social welfare function are analogous to indifference curves for a single individual, but apply to the utilities of the two people not the two goods consumed by the single individual. They express the valuations of the Impartial Spectator, not the preferences of the individual. Similar to the individual indifference curve, the negative of the slope of the iso-social welfare curve at any point (u^A, u^B) is the Impartial

ISO-SOCIAL WELFARE CURVE Iso-social welfare curves show constant or equal ("iso") levels of welfare for different combinations of utility among those involved.

Spectator's marginal rate of substitution of Ayanda's utility in terms of Biko's utility. And we can use the $mrs = mrt$ rule to find the constrained social welfare-maximizing allocation. It is the point where the utility possibilities frontier is tangent to an iso-social welfare curve.

The social welfare maximum shown in the Edgeworth box in Figure 4.5 (a) is $x^A = 5, y^A = 7.5, x^B = 5, y^B = 7.5$ or a fifty-fifty split of each good. In the utility possibility frontier in Figure 4.5 (b) this is point **i**. If their preferences *differed*, or if the Spectator had a reason to value the utility of one of the two more than the other then the social optimum would require each getting *different* amounts of x and y.

Societies do not have an Impartial Spectator to determine how to weight the competing interests of society's members in a social welfare function. Instead, in a democratic society we debate the question of distribution and sometimes come to a consensus (and sometimes to a deadlock). Controversy about the rights and wrongs of economic policies such as the tax rates paid by wealthy people and the provision of public services to all, are often implicitly about the weights (such as λ, in Equation (4.2)) that policymakers should place on the well-being of different people.

Here we see a sharp contrast between the Pareto-efficiency criterion and the maximization of social welfare. Preferring a particular Pareto-efficient allocation over an alternative allocation in which both are worse off cannot be a matter of conflict. But maximization of some particular social welfare function subject to the constraint of the utility possibilities frontier—some gaining and some losing depending on the social welfare function used—is certain to be controversial.

The imaginary Impartial Spectator helps us understand how values dictate what we think of as better or worse allocations. These outcomes, as we have seen in previous chapters and we will now see in greater detail, depend on the rules of the game. So the Impartial Spectator will have something to say about how we evaluate which are better or worse institutions by which we organize the process of exchange.

4.6 PROPERTY RIGHTS AND PARTICIPATION CONSTRAINTS

The scenario of Ayanda and Biko enjoying their coffee and data in their student residence and deciding how to allocate them helps us understand the abstract issues of Pareto efficiency and fairness. Very similar issues arise when instead we consider Ayanda and Biko to be total strangers, interacting in a market. But in this new setting the allocation will not be determined by some imaginary Impartial Spectator. Instead, the allocation will be determined by who initially owns which goods and the rules of the game that regulate how Ayanda and Biko might exchange some of their goods with each other.

Market institutions: Property rights and participation constraints

Nobody actually owned the data and coffee that the Impartial Spectator allocated in our thought experiment, and Ayanda and Biko were not really engaged in a game. This is not how markets work. Key aspects of the rules of a market game are:

- The *rule of law* establishes that the institutions–the laws and other informal rules–governing the interaction are observed, and not violated by arbitrary acts (for example theft of the other's goods by one of the traders or confiscating by a government official).
- *Private ownership*. At any moment in the game the goods are the private property of one or the other of the players, so a point in the Edgeworth box indicates a distribution of property between the two. The ownership of coffee and data by the two at the start of the game is each player's *endowment*.
- *Fallback option*. The endowments are the next best alternative for the two, their fallback options if no exchange takes place.
- Private property and the rule of law mean that each player has the *option to refuse offers* so any exchanges that a player will agree to participate in must be Pareto improvements over the endowment.
- **Bargaining power** is determined by rules of the game that may favor one of the traders over the other and will affect the nature of the exchanges that are executed, and who captures the greater share of the gains from exchange.

HISTORY It has not always been true that one's property rights did not depend on one's identity. In many societies, some people—such as women—did not have the right to own property, and some people—such as enslaved people—were themselves treated as property.

Private property does not distinguish between the two parties: each have *identical* rights to exclude the other from their bundle of goods. This would be true even if Biko initially owned all of the goods and Ayanda had none or the other way around. In this respect private property rights provide a level playing field because the right to exclude others from the use of your goods does not depend on how many goods you have, or on your identity.

The exchange process begins with the property people "start with," that is, their endowment allocation. These endowments exist *before the exchange we are considering happens*. But we are cutting into time at a particular moment. These endowments, which are the status quo of our game, are the result of similar games played in the past, and also other games in which who owns what goods may have been determined by *force* and not by *voluntary exchange*.

BARGAINING POWER The extent of a person's advantage in securing a larger share of the economic rents made possible by an interaction.

This means that unlike the Impartial Spectator starting with a clean slate–any allocation in the Edgeworth box is up for consideration–and advising Ayanda and Biko on the division of a pile of goods they have tripped over in their student residence, market exchange starts from one particular point in the Edgeworth box: the **endowment allocation**.

The rules of the game then determine how the two can move to some other post-exchange allocation. The endowment allocation is important for two reasons:

- it is the *starting point* of the process; and
- because the exchange is voluntary, meaning they can refuse to trade, the endowment allocation is their fallback position, that is, the *worst they can do*.

The participation constraint (PC)

To see how the second bullet above will narrow down what the post-exchange allocation can be, starting at any given endowment allocation we introduce the following notation, along with panel (a) of Figure 4.6 (we will explain in panel (b) below). The *endowment bundle* of person i is (x_z^i, y_z^i) where the superscript indicates who person i is ($i = \mathrm{A}$ for Ayanda, $i = \mathrm{B}$ for Biko). This allocation is point **z** in the figure. It is identical to point **z** in previous figures, but instead of being some hypothetical allocation that the Impartial Spectator was trying out in a thought experiment, it is now something entirely different: it is what Ayanda and Biko *own* at the start of the game. It is their wealth.

From point **z** you can see that Ayanda's and Biko's endowments of coffee and data are the same as the hypothetical allocation considered by the Spectator, above:

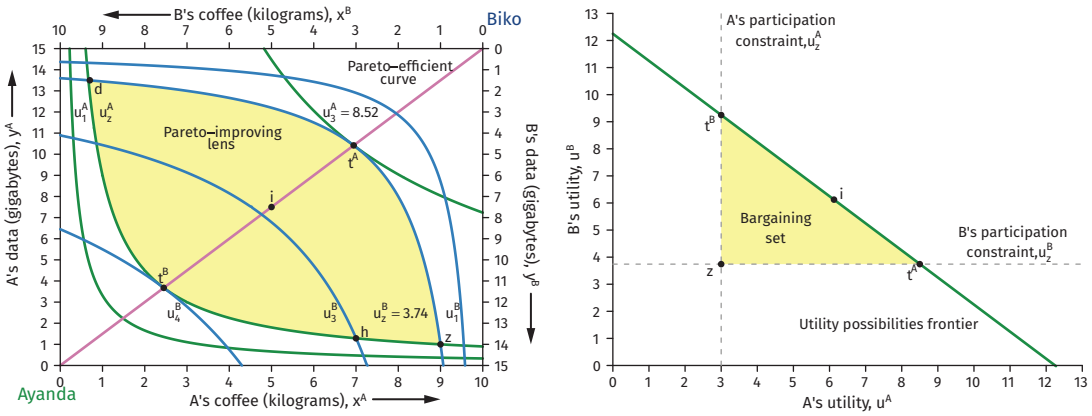
- Ayanda's endowment: $(x_z^{\mathrm{A}}, y_z^{\mathrm{A}}) = (9,1)$.
- Biko's endowment: $(x_z^{\mathrm{B}}, y_z^{\mathrm{B}}) = (1,14)$.

Introducing history in the form of initially privately owned endowments, along with the voluntary transfer requirement, limits the possible allocations that can result from exchange.

Because they can refuse any deal and therefore experience the utility from their endowment bundle, they will not accept any post-exchange

ENDOWMENT ALLOCATION The ownership of goods at the start of a game (or at the status quo) is called the endowment allocation.

Figure 4.6 Edgeworth box, the utility possibility frontier, and the bargaining set. In panel (a) u_z^A is Ayanda's utility at her endowment and is her *participation constraint* (shown by indifference curve u_z^A) and u_z^B is Biko's utility at his endowment and is his *participation constraint* (shown by indifference curve u_z^B). In panel (b) points $\mathbf{t^B}$, $\mathbf{t^A}$, and **i** show the levels of utility associated with the allocations indicated by the same lettered points in the panel (a).



(a) Edgeworth box for Ayanda and Biko

(b) The utility possibilities frontier

! REMINDER The participation constraint is also the *fallback* in the exchange scenario, the utility that a person can certainly secure if they choose not to participate in exchange at all.

bundle that makes them worse off than their fallback utilities. The indifference curves, u_z^A and u_z^B, that include the endowment point are the post-exchange bundles that yield a utility identical to their fallback position. These two indifference curves are called their **participation constraints**. They are called participation constraints because Ayanda will not *participate* in (that is she will refuse) any offer that would give her a post-exchange bundle below and to the left of u_z^A. Likewise Biko will not participate in any offer that would give him a post-exchange bundle above and to the right of u_z^B (labeled as $u_z^B = 3.74$ in Figure 4.6).

The yellow-shaded space between the two constraints—the indifference curves, u_z^A and u_z^B—including the points on these indifference curves make up the set of allocations that are Pareto superior (at least weakly) to point **z** and which therefore could be the result of voluntary modifications of the endowment allocation by means of exchange. This area is called the **Pareto-improving lens**.

PARETO-IMPROVING LENS The set of allocations that are (at least weakly) Pareto superior to the fallback options of the players is the Pareto-improving lens.

CHECKPOINT 4.5 **Pareto-improving set and the bargaining set** Explain how point **z** in panel (a) of Figure 4.4 corresponds to (meaning, is based on the same allocation as) point **z** in panel (b). Do the same for point **i**, $\mathbf{t^A}$ and $\mathbf{t^B}$. Explain why the yellow-shaded areas in the two panels represent the same sets of allocations.

4.7 SYMMETRICAL EXCHANGE: TRADING INTO THE PARETO-IMPROVING LENS

In this section we consider the case in which the two traders have identical preferences. That is, that their Cob–Douglas utility functions have $\alpha^A = \alpha^B = 0.5$.

We used a hypothetical point **z** in Figure 4.4 to show that an allocation where the indifference curves cross cannot be Pareto efficient. Our demonstration consisted of showing that at such an allocation both Ayanda and Biko could both be better off at a different allocation.

We can now use the same reasoning to illustrate how starting at point **z**, now an endowment allocation–a real distribution of ownership of two bundles–the two could trade into the Pareto-improving space, and eventually–given the right rules of the game–get all the way to the Pareto-efficient curve.

Each person has a willingness to pay for x in terms of y, their marginal rate of substitution at the endowment allocation **z**. Ayanda's maximum willingness to pay is her $mrs^A = \frac{1}{9}$ and Biko's maximum willingness to pay is his $mrs^B = 14$.

The difference between Ayanda and Biko's willingness to pay (mrs) signals an opportunity for Ayanda to trade data with Biko at a rate of exchange between her own marginal rate of substitution and Biko's marginal rate of substitution. A small exchange on these terms would move them to a *post-exchange allocation* upward and to the left of the endowment.

To stress that the game is entirely symmetrical imagine that they have agreed on a set of rules to determine the price and the amounts to be exchanged. At any allocation at which the mrs of the two differs (meaning their indifference curves intersect), take the following steps:

1. Pick a "price" midway between the mrs of the two. (This means that at point **z** the price would be $14 + \frac{1}{9}$ divided by 2, or 7.06.)
2. Ask the amounts that *each* would like to transact at the price of 7.06 gb of data for a kilo of coffee, for example how much coffee Ayanda would like to "sell" at this price, and how much coffee Biko would want to "buy" (these desired amounts will differ between the two).

3. Because the transfer has to be voluntary (nobody can be forced to buy more than they wish), transfer the amounts desired by the person who wishes to transact least.
4. At the resulting post-exchange allocation determine if the indifference curves are intersecting. If so, return to step 1 and continue.
5. If not (that is, if the indifference curves are *tangent*) end the game with this final allocation.

We can see that by this process the two will have moved, step by step from the endowment allocation at point **z** to a final post-exchange allocation that will be on the Pareto-efficient curve. We know that they will get there for two reasons:

- *Trades are Pareto-improving*: each trade they take moves them in the direction of the Pareto-efficient curve because moving in the other direction could not be a Pareto-improvement and would violate the participation constraint.
- *Trade concludes at a Pareto-efficient outcome*: by the rules of the game they have adopted they will keep on exchanging until they are at a place where their marginal rates of substitution are identical, which must be on the Pareto-efficient curve.

We can conclude then that there are no more trades to make *because* the allocation is Pareto efficient. Or, what is the same thing: *because* there are no more mutually advantageous transfers of goods possible, the allocation must be Pareto efficient.

They could have adopted a different set of rules for exchange. For example, they could have said that for step 1 above there will be two alternative prices, one just a little less than Biko's willingness to pay, and the other just a little more than the lowest price at which Ayanda would part with her coffee; and then just flipped a coin to see which of these prices they would use in that transaction. Having made that transaction and the new allocation, check to see if Biko's willingness to pay for coffee is still greater than the lowest price at which Ayanda is willing to sell. If so, flip another coin to see whose preferred price will be used, make the trade, and so on, until no further trade is possible.

Other than knowing that they would *eventually* get to the Pareto-efficient curve, we do not know which specific point on the curve they would get to. If all of the coin flips went in favor of Ayanda, they could end up close to $\mathbf{t}^{\mathbf{A}}$ in Figure 4.6 with Biko sharing very little of the gains from exchange. Or it could have gone the other way, somewhere near point $\mathbf{t}^{\mathbf{B}}$. They even could have ended up at point **i** the allocation chosen by the Impartial Spectator. But that would have been by pure chance.

The utility possibilities frontier in Figure 4.6 (b) translates these allocations and the transactions supporting them into the utilities of the two players. The Pareto-improving lens in panel (a) corresponds to the **bargaining set** in panel (b). It is called the bargaining set because it is the set of all possible pairs of the utilities of the two that could be the result of some bargain into which they entered voluntarily. Panel (a) shows all of the allocations–denominated in quantities of x and y allocated to the two–that are Pareto improvements over the endowment allocation. The second–the bargaining set–shows the utility levels associated with every allocation in the Pareto-improving set.

If we consider other rules of the game, which point to the bargaining set they implement–it may be in the interior, not on the frontier–it will depend on the rules governing how they bargain, including how the rules affect the bargaining power of the players.

CHECKPOINT 4.6 **Pareto improvements, rents, and Pareto efficiency** If point **h** is the post-exchange allocation based on the endowment allocation of point **z**, explain the following:

a. Did Biko benefit from the exchange?

b. Did Ayanda benefit from the exchange?

c. What is the rent that Ayanda receives as a result of this exchange?

d. Did the exchange result in a Pareto improvement?

e. Is the post-exchange allocation (point **h**) Pareto efficient?

4.8 BARGAINING POWER: TAKE IT OR LEAVE IT

The two examples of rules of the game for bargaining over the distribution of coffee and data above were symmetrical. Neither "split the difference between the willingness to pay of the two" nor "alternating coin flips to see whose preferred price will be used" gave any obvious advantage to either player.

But many bargaining interactions are asymmetrical. One of the players has more of the bargaining power. Bargaining power is the ability to gain a large share of the mutual gains from exchange (total rents) made possible from some interaction, as may be determined by the rules of the game governing the interaction and the skill of the players in securing a favorable agreement under these rules.

BARGAINING SET The set of all allocations that are Pareto improvements over the players' fallback (no-bargain) options and the utilities associated with these allocations is termed the bargaining set.

! REMINDER The Ultimatum Game discussed in Chapter 2 has this TIOLI structure including returning to the endowment point if the Responder rejects—both getting a payoff of zero, namely what they would have received had they not interacted. That is why it is called the Ultimatum Game, as the Proposer's offer is an ultimatum.

An example is the Ultimatum Game in Chapter 2 (whose name already suggests the asymmetry). The Proposer makes an offer of some fraction of the "pie." The Responder's strategy set is simply: accept or reject, or "take it or leave it." Being in a position to make that kind of an ultimatum is called **take-it-or-leave-it power**, or TIOLI power for short.

In the coffee-for-data-bargaining game, if Ayanda had TIOLI power, she could have said to Biko: "I'll give you 2 kilograms of coffee and you give me 9 gigabytes of data. If you refuse, I will not agree to any other trade you might propose." In other words, "either accept the allocation I impose, or we both stay at our endowment, **z**." Of course Ayanda's threat to terminate dealings if Biko refuses has to be **credible**: if Biko suspects that he could refuse and Ayanda *would* listen to a counteroffer, the threat in the TIOLI offer would be empty. A bargainer with TIOLI power can often capture most or even all of the total rents that an economic interaction provides. This is because TIOLI power allows a bargainer to specify both:

- the *price* at which the goods will be exchanged; and
- the *amount* of goods that will be exchanged.

This means that the person with TIOLI power can just pick some preferred allocation–a point in the Edgeworth box different from the endowment point–and make that the TIOLI offer.

What take-it-or-leave-it offer will Ayanda make to Biko? We have assumed that Ayanda does not care about Biko's utility, but she does care about how he will respond to her offer. If he rejects, then she gets her fallback option. She will realize that she must offer Biko a deal that Biko regards as better–or at least not worse–than his endowment. In other words, Ayanda has to take Biko's participation constraint as a limit on the kind of offer she will make. This is an example of the backward induction method that you learned in Chapter 2: Ayanda has to reason backwards from her understanding of what Biko will do *after* she has made her offer to what offer she should make now.

TAKE-IT-OR-LEAVE-IT POWER A player with take-it-or-leave-it power (TIOLI power) in a two-person bargaining game can specify the entire terms of the exchange—for example, both the quantity to be exchanged and the price—in an offer, to which the other player responds by accepting or rejecting.

CREDIBLE THREAT A threat is credible if carrying it out is the best response if the target of the threat takes the action that the threat was intended to deter.

So Ayanda has the following constrained maximization problem: find a final allocation (different from the endowments) to propose at which Biko is no worse off than at his endowment and Ayanda is as well-off as she can be. Ayanda knows that the solution to this problem must have two characteristics: It must:

- *satisfy Biko's participation constraint*, that is, be in (or on the boundary of) the Pareto-improving lens in Figure 4.6; and
- *be Pareto-efficient*, but this is not because Ayanda cares any more about efficiency than she does about Biko: if she offered an allocation that satisfied Biko's participation constraint and was *not* Pareto efficient then there would be some *other* allocation at which *she* could be better off and Biko not worse off than his fallback option, **z**.

M-CHECK Remember that in Chapter 3, a utility maximizer is often constrained by a feasible frontier. Even with TIOLI power, Ayanda is constrained by Biko's *participation constraint*, that is, $u^B(x^B, y^B) \geq u^B_z$.

Ayanda would probably offer Biko something just a tiny bit better than Biko's fallback utility to make sure he accepts. But to avoid having to keep track of that tiny amount in our thinking, here and in the rest of the book, we will assume that Biko will accept an allocation that *just meets* his participation constraint.

That solves the problem for Ayanda: to meet the two requirements bulleted above, she must find the intersection of the Pareto-efficient curve and Biko's participation constraint u^B_z. Therefore, Ayanda offers an exchange that implements point $\mathbf{t^A}$. The same result is shown in Figure 4.6 (b), where $\mathbf{t^A}$ represents the distribution of utilities resulting from the TIOLI allocation that Ayanda offered and Biko (barely and grudgingly) accepted.

! REMINDER An outcome is socially efficient when it maximizes a social welfare function; what is deemed socially efficient depends on how the utility of each member of the population is weighted in the social welfare function.

Point $\mathbf{t^B}$ in the Edgeworth box corresponds to the allocation where Biko has TIOLI power and point $\mathbf{t^B}$ on the utility possibilities frontier is the corresponding distribution of utilities.

We can see that the TIOLI allocation does not weight the two utilities identically (as did the social welfare function of the Impartial Spectator, which led to point **i**). This is why we say that allocation $\mathbf{t^A}$ is Pareto efficient but not socially efficient, where the latter term is whatever the Impartial Spectator selected based on maximizing an equally weighted social welfare function.

Two features of the TIOLI allocation are important because they arise in many social coordination problems where the constrained optimizing process is limited by a participation constraint:

1. *Inequality*: At participation-constrained outcomes the bargainer with TIOLI power gets all of the economic rent.
2. *Pareto efficiency*: The participation-constrained outcome is Pareto efficient.

The second feature is true by the definition of Pareto efficiency: an allocation in which one party cannot be made better off without making the other party worse off. Such an allocation must be the result of one person maximizing their utility subject to a constraint set by some level of utility of the other person. This is just what the person with TIOLI power does, with the minimal level of the utility of the other person being given by his fallback option.

M-NOTE 4.4 **Finding the Pareto-efficient curve**

The Pareto-efficient curve: At Checkpoint 4.3 we asked you to find the Pareto-efficient curve for Ayanda and Biko when they have identical Cobb-Douglas utility functions with $\alpha = 0.5$. The solution is that the Edgeworth box has the following Pareto-efficient curve defined over the two people's allocations of x and y:

$$y^A = \left(\frac{3}{2}\right)x^A \tag{4.3}$$

We can rewrite Equation 4.3 in terms of x^B and y^B by substituting $x^A = \bar{x} - x^B$ and $y^A = \bar{y} - y^B$ in the equation to find:

$$y^B = \left(\frac{3}{2}\right)x^B \tag{4.4}$$

As you can see, the Pareto-efficient curve is a line from the one corner of the Edgeworth box to the other.

M-NOTE 4.5 **Finding Ayanda's TIOLI offer**

We need two pieces of information to find Ayanda's TIOLI offer:

- the equation for the Pareto-efficient curve (because we know that the resulting allocation will be Pareto efficient), and
- the equation for Biko's participation constraint (because we know that Ayanda will not offer him anything better than his utility at his endowment bundle).

We calculated the Pareto-efficient curve (PEC) in M-Note 4.4. We use Equation 4.3:

$$y^B = \left(\frac{3}{2}\right)x^B$$

Biko's participation constraint: B's fallback utility (his participation constraint (PC)) at his endowment $x_z^B = 1, y_z^B = 14$ is:

$$u_z^B(1,14) = (1)^{0.5}(14)^{0.5} = 3.74.$$

So we need to find the point on the Pareto-efficient curve at which Biko has this level of utility $u = (x^B)^{0.5}(y^B)^{0.5} = 3.74$.

A's TIOLI Offer: We substitute the Pareto-efficient curve's value for x^A into B's utility function that is equal to B's fallback utility:

continued

$$u^B = (x^B)^{0.5} \underbrace{\left(\frac{3}{2}x^B\right)^{0.5}}_{\text{PEC}} = \underbrace{u_z^B = 3.74}_{\text{PC}}$$

$$\Rightarrow \left(\frac{3}{2}\right)^{0.5} x^B = 3.74$$

$$x_{TA}^B = 3.74/\left(\frac{3}{2}\right)^{0.5} = 3.05 \approx 3$$

$$\therefore y_{TA}^B = \frac{3}{2}x^B = \frac{3}{2}(3) = \frac{9}{2} = 4.5$$

$$\therefore x_{TA}^A = \bar{x} - x^B = 7$$

$$\therefore y_{TA}^A = \bar{y} - y^B = 10.5$$

So where "TA" means A had TIOLI power, the post-exchange allocation will be $x_{TA}^A = 7$, $y_{TA}^A = 10.5$, $x_{TA}^B = 3$, $y_{TA}^B = 4.5$. The post-exchange allocations imply that A made a TIOLI offer to B of 2 units of x ($x_{TA}^B - x_z^B = 3 - 1 = 2$) in exchange for 9.5 units of y ($y_{TA}^A - y_z^A = 14 - 4.5 = 9.5$). A's utility is $u_{TA}^A = 8.97$ and B remains on his participation constraint at $u_z^B = 3.74$.

CHECKPOINT 4.7 **Ayanda's TIOLI power** Explain why, when Ayanda has TIOLI power she will make an offer implementing a Pareto-efficient outcome.

4.9 APPLICATION: BARGAINING OVER WAGES AND HOURS

We illustrate TIOLI power by a case in which the two bargainers drop their student personas to take on familiar roles in what is arguably the most important market in a modern economy: Ayanda is the owner of a company interacting with Biko, a prospective employee. In a labor market with an employer and worker bargaining over wages and working conditions the employer almost always has TIOLI power, stating the wage, the job, and the hours. The worker accepts or not. We postpone until Chapter 15 the question: Why might Ayanda get to have this power and not Biko?

So, leaving the world of coffee and data behind us, we will see that the sum of the mutual gains enjoyed by the two and how these are divided between them will depend on the rules of the game governing their interaction:

- *Power*: Do the two bargain symmetrically with neither one nor the other of them having first-mover advantage? Is one of them first mover with TIOLI power?
- *Fallback*: What is each person's fallback position? How well-off are they if they do not exchange at all? Does Biko have other options than being employed by Ayanda? If Ayanda does not employ Biko, are there others she could employ?

To fill in some answers to these questions, our two actors are now:

- *Ayanda, an employer*: whose endowment bundle is a sum of money only (no employees), and who in the absence of any exchange with Biko has nobody work for her; she will make Biko a take-it-or-leave-it offer of a sum of money in return for some number of hours of work for her; and
- *Biko, a worker*: who is applying to work in Ayanda's company. His endowment bundle is free time only (no money); he has a maximum of 16 hours of (non-sleeping) time to spend, possibly working for Ayanda.

M-CHECK Quasi-linear utility functions have the form:

$$u(x,y) = ay + h(x) \quad (4.5)$$

where a is a constant. The constant a makes utility linear in y and the function $h(x)$ makes utility nonlinear in x. For example, $h(x)$ could be x^2 or the natural log of x, $\ln(x)$. We often set $a = 1$, which means that the marginal utility of y is 1 and that utility is measured in whatever units y is measured in.

We introduce a more complete model of the labor market with competition among firms for workers and customers and among workers for jobs in Chapter 11 including the ways that unemployment benefits, and the extent of competition among firms, could affect these outcomes. We will take another step toward realism by taking into account the fact that Biko has some freedom to choose how hard he is going to work while on the job.

Quasi-linear preferences for money and time

To represent the preferences of Ayanda and Biko we will introduce a new utility function, one that will simplify our analysis while still conveying the main insights. The function is called **quasi-linear** because utility is partly ("quasi") proportional to one of the arguments of the function, while being nonlinear in the other arguments. The Cobb-Douglas utility function is not quasi-linear because it is nonlinear with respect to both x and y.

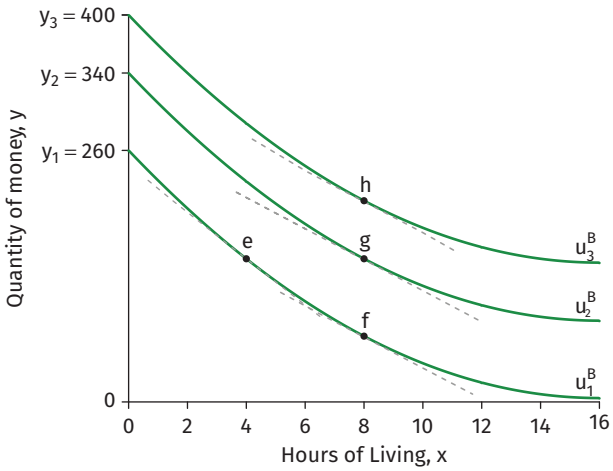
As in the case of Harriet deciding how much fish to buy in Chapter 3, we will consider the second good as "money left over" after the exchange. This may seem odd because money is not something you value for itself. But money can buy you other goods which you do value: The utility of "money left over" is the utility of the goods which the person can purchase as a result.

We now illustrate a case where one person starts off with all of one good and none of the second, so the other person starts with all of the second good, but none of the first. This could model you walking into the supermarket with money in your pocket (or more likely on a credit card) and nothing in your shopping bags, and planning to walk out with less on your credit card and some groceries in your shopping bag. So it is a model of any kind of exchange. But here we illustrate it by Ayanda (possibly) employing Biko.

The marginal rate of substitution for a person with quasi-linear preferences that are linear in "money" (y) depends only on the amount she has of

QUASI-LINEAR FUNCTION A quasi-linear function depends linearly on one variable and nonlinearly on another variable.

Figure 4.7 Marginal rates of substitution with quasi-linear preferences. With quasi-linear preferences such that utility is linear in the y-good, marginal rates of substitution depend only on the amount of the good x (here, Hours of Living for Biko), and not at all on the amount of money left over to buy other goods, y. As a result, indifference curves with different levels of utility are vertical displacements of a single curve—you can add or subtract an amount of y from the indifference curve to move it up or down. Biko's utility function is: $u^B = y^B + 32x^B - (x^B)^2$.



the good or service for which her preferences are nonlinear (x), not on the amount of money.

The reason why this is true is because:

- the marginal rate of substitution is the ratio of the marginal utility of x to the marginal utility of y;
- the person's marginal utility for y is a constant; it does not decline as she gets more y; so
- the marginal rate of substitution depends only on the marginal utility of x which varies with the quantity of x consumed because the function is nonlinear in this variable.

You can see this in Figure 4.7 by noticing that for a given amount of x the slope of the indifference curve (shown by the dashed tangent lines) is the same no matter how much y the person has, such as at $x = 8$ hours of living, as shown by points **f**, **g**, and **h**. This is because, given the quasi-linear utility function that we used to draw the figure, the willingness to pay for an additional hour of living (the marginal rate of substitution, that is, the negative of the slope of the indifference curve) does not depend on the amount of money left over that the person has. It depends only on how many hours of living they have. This means that the indifference curves u_1,

! REMINDER In Chapter 3 we analyzed a case in which the marginal utility of a person's wealth declined the more wealth she had, and asked what distribution of wealth would maximize the sum of the utilities of two individuals.

✓ FACT CHECK If Ayanda is the employer and Biko one of her prospective employees, she probably has a lot more income than him, so it might be more realistic for purposes of comparing their utilities (or adding them up) if we let her have a utility function in which the (constant) marginal utility of income were less than his. But this would complicate the model without adding any new insights.

✓ FACT CHECK In August 2020, 1 euro was equal in value to about 16 South African rand (ZAR), 1 pound sterling was equal to about 21 South African rand, and 1 US dollar was equal to 17 South African rands. In 2020, the hourly minimum wage in South Africa was ZAR 20.76.

u_2, and u_3 in the figure are just shifted up replicas (you can see the amounts by which they are shifted up by comparing the vertical axis intercepts).

We can also compare points **e** and **g** at the same level of y: Biko likes to have more living time ($u_1^B < u_2^B$) and his willingness to pay for additional hours of Living declines the more he has (the indifference curve is less steep at **g** than at **e**). It is of course unreilistic to think that anyone would have truly linear preferences in any amount imaginable of money left over, for this would require that the person did not have diminishing marginal utility in the things that money can buy. But because "money" can be considered as generalized purchasing power that can be spent on a vast array of things, and because we do not consider changes in people's bundles of money making them either billionaires or paupers, it is a useful simplifying assumption here.

Allocating money and time

Because Ayanda, the employer, and Biko, the worker, have quasi-linear utility functions their marginal utility of money is constant. So if their money left over is increased by one monetary unit, their utility rises by one unit. This means we can measure the utility of each in whatever monetary units they are using, which since their names are from South Africa, might as well be the South African rand.

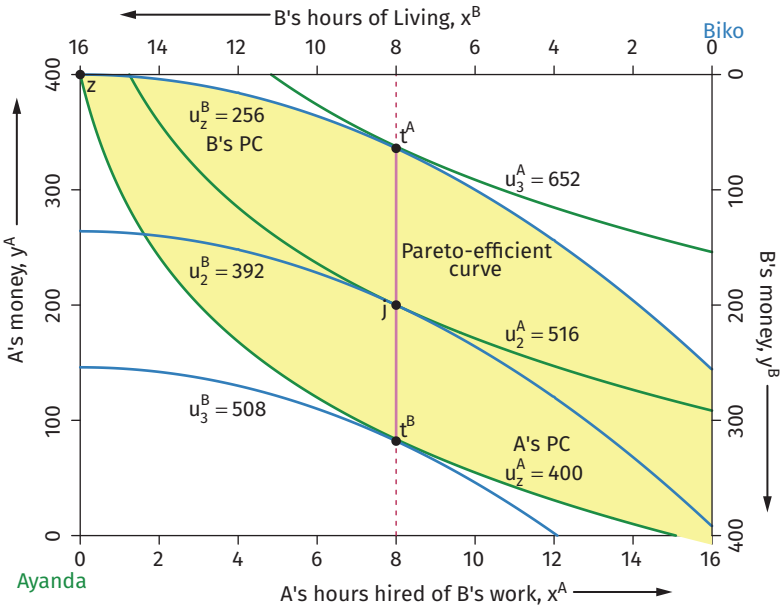
Biko values his Living (that is his 16 waking hours, minus the time he "hires out" of himself to work for Ayanda). But the marginal utility to him of free time decreases as the amount of free time he has increases, another instance of the "law of diminishing marginal utility."

Ayanda places a value, too, on Biko's free time, but it is the opposite of Biko's value: she benefits by Biko having *less* free time and her having *more* of Biko's time working for her. The positive value she places on Biko's labor—like the positive value he places on his free time—depends on how much of it she gets. The marginal utility to Ayanda of Biko's labor decreases as she hires more of his time: the value of Biko's work is high the first hour Ayanda hires, less valuable the second hour, less valuable the third, and so on.

This is because if she has just an hour of his time, she assigns him to really important tasks, but the tasks he does in later hours are less essential to Ayanda. (This is similar to why the marginal productivity of time studying diminishes as the amount of time studying increases, another instance of the diminishing marginal productivity of labor.)

Figure 4.8 shows the setting for this interaction as an Edgeworth box, with the quantities interpreted as amounts per day. Remember: Biko prefers allocations that are lower (more money for him) and to the left (more free time); Ayanda prefers allocations that are higher and to the right. The endowment point **z** is in the upper-left corner of the box showing that initially Biko has 16 hours of Living time and no money. Ayanda has 400 South African rands (ZARs) but no Labor from Biko to work in her company.

Figure 4.8 Bargaining over hours and wages. Shown are three each of Ayanda's and Biko's indifference curves and the utility that they experience at any of the allocations indicated by the points making up these curves. Point **z** is the endowment allocation which is a point on the participation constraints of each of the two. Points $\mathbf{t^A}$ and $\mathbf{t^B}$ respectively are the allocations resulting when Ayanda or Biko are first mover with TIOLI power. The yellow-shaded area is the Pareto-improving lens. The vertical line (including its dashed portions) is the Pareto-efficient curve made up of all points of tangency between the indifference curves of the two such as **j**, $\mathbf{t^A}$, and $\mathbf{t^B}$.



As before, like **z** every point in the box represents an allocation that is feasible given the amount of money that Ayanda has in her endowment bundle and the amount of free time that Biko has in his.

Three of Biko's indifference curves and three of Ayanda's are shown in Figure 4.8. For both Ayanda and Biko, their reservation indifference curve (their participation constraint) includes the endowment point **z** where Biko has 16 hours of Living (his free time) and Ayanda has \$400 per day to pay workers.

Also shown is one of Biko's indifference curves labeled u_3^B, which is tangent to Ayanda's participation constraint (u_z^A) at point $\mathbf{t^B}$. The allocation given by that tangency is a Pareto-efficient allocation (because the marginal rates of substitution of the two are equal). We also show a third indifference curve for Ayanda, labeled u_3^A, which is tangent to Biko's participation constraint (u_z^B) at point $\mathbf{t^A}$. These two tangencies are points on the Pareto-

efficient curve, which is a vertical line through these points of all the potential tangencies above each person's fallback.

The reason why the Pareto-efficient curve is *vertical* here (remember it was an upward-sloping curve or line in the previous Edgeworth boxes) is that Ayanda and Biko have *quasi-linear* utility functions. Remember: with quasi-linear utility, the marginal utility of hours depends only on the quantity of hours and not on the amount of money they have. If the two curves are tangent at 8 hours when Ayanda has most of the money and Biko little, they will also be tangent at 8 hours when Biko has most of the money and Ayanda has little.

✓ **FACT CHECK Involuntary exchanges** In the past, slavery has meant the ownership of one person by another, including the right of sale of the enslaved person to another owner. The term *modern slavery* refers to any situation in which, like historical slavery, the services or goods that one party provides for another are not voluntarily offered but are motivated by fear of severe harm. Ownership of one person by another need not be part of modern slavery. Prisoners, immigrants without legal rights of residence, residents of undemocratic countries, "sex slaves," and children are overrepresented among contemporary "modern slaves."

CHECKPOINT 4.8 Marginal rate of substitution with a quasi-linear utility function Explain why the marginal rate of substitution is the same at points **f**, **g**, and **h** in Figure 4.7, and the marginal rate of substitution is higher at point **e**.

4.10 APPLICATION: THE RULES OF THE GAME DETERMINE HOURS AND WAGES

The Edgeworth box and the indifference curves by themselves do not determine the outcome of the interaction. Without knowing more, any point in the box is a possible outcome. Knowing the endowment allocation **z** narrows down the possible post-exchange allocations but not by very much.

Employment in most modern economies is *voluntary* (but see the Fact Check), so we will require that the outcomes are limited to those that are at least as good for each participant as their fallback position given by point **z**. As a result, outcomes of bargaining between the employer and the worker must be in the yellow-shaded Pareto-improving lens in Figure 4.8.

We illustrate the importance of institutions by showing the allocations will result under four different rules of the game. Each set of rules is a specific account of four different ways that an employer and worker might interact:

- The *employer can make a take-it-or-leave-it offer* of both the wage and the hours worked.
- As members of *a trade union*, the employees (we will take Biko as a representative worker) *can make a take-it-or-leave-it offer* specifying both the wage rate and the length of the working day (hours).
- *Legislation* is passed limiting working hours per day to no more than five hours and the total pay or this period to not less than 254 rand or 50.80 rand per hour.

- The above legislation is passed, but it has a proviso that if the *two parties can agree on an alternative allocation*, their agreement can be implemented.

Employer has TIOLI power

Imagine that, like most employers, Ayanda can offer Biko a job description: work a given amount of hours for a given amount of pay (and therefore for a particular hourly wage). Biko's only choice is to accept or reject, so Ayanda has take-it-or-leave-it power. For Biko to accept, Ayanda knows the offer must be at least as good as Biko's reservation option, so the relevant constraint for her is Biko's participation constraint (as was the case for the coffee and data bargaining).

She will choose the point that she values most along this indifference curve, and therefore implement an offer indicated by point $\mathbf{t^A}$. Having TIOLI power, the employer has captured all of the economic rent, leaving Biko indifferent between taking the job and refusing it (as before in cases like this we just assume he takes the job).

What is Ayanda's rent from this transaction, meaning the excess of her utility at point $\mathbf{t^A}$ compared to at point **z**, the endowment allocation at which no trade has occurred? Reading the utility numbers from her indifference curve at point $\mathbf{t^A}$ and her reservation indifference curve through point **z** we can see that her rent is $u_3^A = 652$ minus $u_z^A = 400$ or 252. Because utility is measured vertically in terms of money this is the same thing as the vertical distance between points $\mathbf{t^A}$ and $\mathbf{t^B}$ in the graph.

› EXAMPLE Put yourself in Biko's shoes if the allocation is point $\mathbf{t^A}$. How do you think he feels about his employer and his job? Would he be motivated to work hard, not to steal from his employer, and otherwise contribute to the profitable operation of her firm? These are serious problems and a reason why extreme allocations—like Ayanda getting all of the economic rent from the interaction and Biko being indifferent between his job and being fired—are not commonly observed. If Ayanda has an interest in Biko's goodwill and hard work, she may have to share at least a bit of the gains from exchange with Biko so that he receives a rent. This fact will become important when we consider the labor market.

Employees and their trade union have TIOLI power

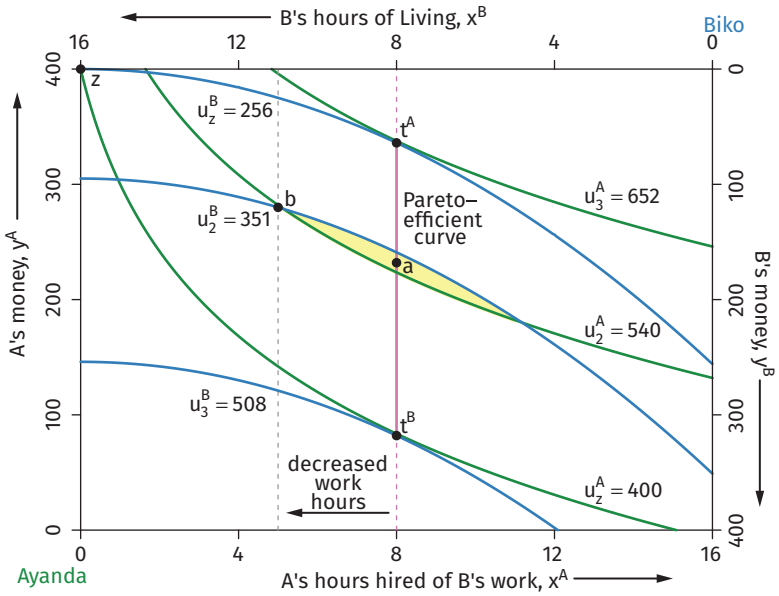
Turning to the opposite case, Biko, through his trade union, is now first mover with TIOLI power. The offer he will make (and she will accept) is the opposite of $\mathbf{t^A}$ the allocation resulting when Ayanda had TIOLI power. Biko will recognize Ayanda's participation constraint–he has to make her an offer she will not refuse. And he will choose the allocation indicated by point $\mathbf{t^B}$ in which his post-exchange bundle gives him all of the economic rents of 252.

This is the most that Biko could demand without Ayanda's simply going out of business or more realistically, seeking to move her business to a place without trade unions. This constraint on the demands that workers can make on employers in a market and profit-based economy will be a major theme in the chapters to come.

Legislation imposes hours and pay limitations

The legislation described above imposes on both Ayanda and Biko the allocation at point **b** in Figure 4.9, which is Pareto inefficient. It sets a new status quo, a new fallback position that, if they cannot come to some

Figure 4.9 Allocations with legislation and bargaining. The legislation stipulating hours and pay results in the allocation indicated by point **b**. Because **b** is preferred to the no-exchange option **z** by both of them, they will definitely make an exchange. But they both can do better than at **b**. Taking the allocation at **b** as their new fallback position, they could bargain to point **a** or any other allocation in the new yellow Pareto-improving lens.



agreement about some different allocation will be the new post-exchange allocation.

Both Ayanda and Biko can see that at **b** they could both do better by agreeing that Biko should work more than five hours, and Ayanda should pay him more. The small yellow Pareto-improving lens shows the space for their possible bargains.

Bargaining to override the legislation: more work and more pay

They could bargain to agree upon any point in the Pareto-improving lens, possibly agreeing on the Pareto-efficient allocation at point **a**. Where they ended up in or on the boundary of the Pareto-improving lens would depend on the rules of the game governing that bargaining process. They might even fail to agree on any bargain—as is often the case with players in the Ultimatum Game—and remain at point **b**.

Figure 4.10 Rents under differing rules of the game, with Ayanda as employer and Biko as worker. The rents and gains from exchange of each set of rules are shown in the figure. That is, the figure shows each player's utility under each set of rules minus that player's fallback option ($u_z^A = 400$ and $u_z^B = 256$ respectively). The gains from exchange are the sum of the rents received by Ayanda and Biko.

Source: Authors' calculations described in the text.

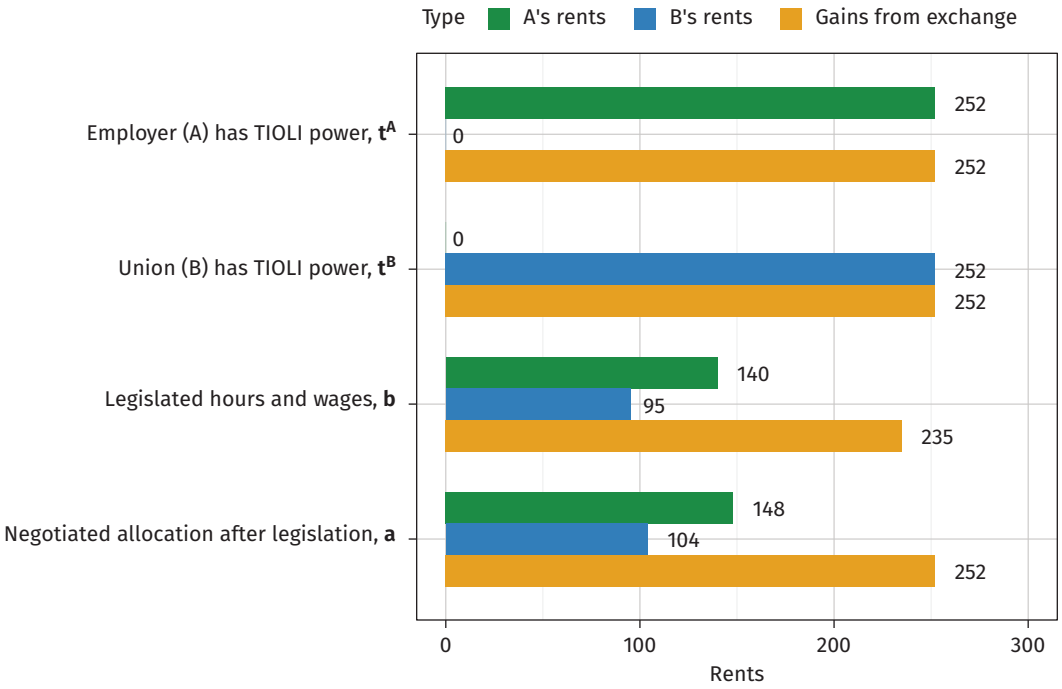


Figure 4.10 shows how Ayanda and Biko do under these differing rules of the game as indicated by the rents they enjoy in the Nash equilibrium of each game, that is the excess of their utility over the utility associated with their fallback options of 400 and 256 respectively.

Introducing a historically realistic set of rules of the game—making the employer the first mover with TIOLI power—has two effects: it generates 252 units of utility in gains from trade, and it makes the final allocation more unequal than the endowment allocation (because the employer captures all of the mutual gains made possible by exchange). Biko's share of the total utility (not shown) falls from two-fifths to one-third.

In many countries during the twentieth century the response to the unequal allocations implemented when the employer has TIOLI power was the formation of trade unions.

And you can see from the figure that if the union were powerful enough for it to have TIIOLI power (a not very realistic scenario), then Biko (and his trade union colleagues) capture the entire rent, Ayanda getting nothing

more than her reservation utility. Biko's share of the total utility (not shown) jumps from two-fifths at the endowment allocation to well over half.

Even before workers had the right to vote and before trade unions were legal, political movements mobilized to pressure governments to regulate working conditions. In the model the introduction of hours and wage regulations implemented an outcome in which both Biko and Ayanda captured some of the gains from trade. The reforms implemented a Pareto-inefficient allocation, but the shortfall from the maximum possible joint rents was minor (from 252 to 235).

The final case—bargaining up from the regulated hours and wages—describes labor markets in many countries today. Government regulations establish a *fallback position*, and then employers and workers (either individually or in trade unions) seek bargains that improve on that allocation.

Though they differ radically in their distributional aspect, all of the scenarios are Pareto superior to the endowment allocation. We can also see that the negotiated allocation after legislation is Pareto superior to the allocation implemented by the legislation.

We cannot say which of the three Pareto-efficient allocations is preferred from a fairness standpoint without knowing more about Ayanda and Biko's other wealth, their needs, and other aspects that might affect their ethical claims on the benefits of their interaction.

CHECKPOINT 4.9 **Bargaining over hours and wages** Using Figure 4.10, explain how the following two things (taken separately) would affect the outcome under the four different rules of the game above (start by explaining how the endowment point **z** would be affected):

a. If Biko does *not* exchange his time with Ayanda and is unemployed, he receives what is called an unemployment benefit, that is, a payment from the government equal to ZAR100, and this is financed by a tax on Ayanda equal to ZAR100.

b. Ayanda now has free access to a robot that will at no cost do work equivalent to two hours of Biko's time.

! REMINDER A player has a first-mover advantage when the institutions, history, or structure of a game give the player the opportunity to make an offer or move *before* the other players in the game can take action. The opportunity to move first can confer an advantage that results in a greater share of *economic rents* in the outcome of an interaction.

4.11 FIRST-MOVER ADVANTAGE: PRICE-SETTING POWER

Returning to Ayanda and Biko with their former personas as students exchanging coffee and data, we will now see that while first movers typically have advantages, these advantages need not be due to TIOLI power. Ayanda may be first-mover but be unable to commit to a take-it-or-leave-it offer that stipulates an exchange of a specific amount of coffee for a specific amount of data.

Price-setting power

She may have what is called **price-setting power** (PS power) if she can specify a price—either a monetary price or the ratio at which the two will exchange goods—but not how much (the quantity) of her good Biko will buy.

Ayanda might say, for example: "I will give you 1 kilogram of coffee for every 3 gigabytes of data you give me. You can decide how much data you would like to exchange for coffee at that ratio, but the ratio is not going to change. Of course you are free to buy nothing."

We saw that owners of companies typically have TIOLI power when hiring employees; but in their interactions with their customers they typically have only price-setting power. They set a price at which they will sell their product, and sell as much to each customer at that price as the customer wants to buy.

The incentive-compatibility constraint (ICC)

If Ayanda has price-setting power she must find a way to determine the price when it is the price alone that makes up her offer. So her constrained optimization problem is not the same as it was when she had TIOLI power.

When Ayanda had TIOLI power she had only to satisfy Biko's participation constraint. Of course whether she has TIOLI power or just price-setting power, if Ayanda wants to exchange with Biko, she will have to satisfy his participation constraint.

But there is now a second constraint she must satisfy called the **incentive-compatibility constraint**: whatever post-exchange bundle Ayanda would like to implement, she must provide Biko with incentives so that his best response will be to exchange the amount that will allow her to "move" from her endowment bundle to her desired post-exchange bundle.

This is called the incentive-compatibility constraint because she must provide Biko with *incentives* that motivate Biko to act in a way that is compatible with (meaning, that implements) her desired outcome. The incentive-compatibility constraint is based on Biko's best response—the amount of coffee he is willing to buy—to the price Ayanda offers.

PRICE-SETTING POWER A first mover with price-setting power can commit to a price—or in the case of barter, the ratio at which goods will be exchanged—but not the quantity that will be transacted at that price.

INCENTIVE-COMPATIBILITY CONSTRAINT The incentive-compatibility constraint, ICC, describes the limits on the outcomes that a first mover may implement by showing how a second mover will respond to each of the choices that the first mover might make, also known as the second mover's best-response function.

M-CHECK A continuous variable can take on any value over some interval. So, a variable that can take the value of any number between 0 and 5 is a continuous variable; a variable that is restricted to the integers between 0 and 5, namely, 1, 2, 3, or 4 is discrete. The number of your sisters or brothers is discrete; the height of any one of them is continuous.

! REMINDER Backward induction is a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen by other players in subsequent steps in response to her choice.

! REMINDER The method is identical to how we derived Keiko's price-offer curve—offering money in return for fish—in Chapter 3, except that here Biko is not "buying" coffee using money, he is exchanging data for coffee. As a result the "price" is not in terms of dollars per kilogram of coffee, but *gigabytes of data* per kilogram of coffee.

You have encountered best responses in Chapter 1. There the strategy sets were particular actions and therefore best responses were limited to actions like "Plant Late," or "Fish 12 hours." Options like "Plant a little earlier" or "Fish 10 hours and 15 minutes" were not possible. Sometimes discrete strategy sets and best responses like this make sense (think: "Drive on the left if you are in the UK, or Japan").

But sometimes the strategy sets for players are continuous, as for example in setting a price for a good or when choosing the amount of time for an activity, like fishing. When this is the case—as with Ayanda's decision to set a price—we consider the players' best responses as *continuous variables* and describe them by *best-response functions*.

As was the case when she had TIOLI power, Ayanda will reason backwards from her understanding of how Biko will respond to each of her possible offers and how that will affect her utility. That is, she will use backward induction.

To determine how Ayanda can maximize her utility subject to Biko's incentive-compatibility constraint (the price-setting case) is a somewhat more complex problem than maximizing her utility subject only to Biko's participation constraint (the TIOLI power case). The reason is that in the TIOLI case there are just two things that Biko can do: accept or reject her offer. But when Ayanda has price-setting power only, Biko can choose from the entire range of possible amounts that he might be willing to exchange with her, depending on the price.

As a result, Ayanda has to think in two stages when choosing a price ratio.

First stage: *What will Biko do*? How much coffee will Biko buy at each price ratio Ayanda offers? This is Biko's *price-offer curve*, which is his best response.

Second stage: *What should I do, given what he will do*? Given her estimate of Biko's best response, which price ratio maximizes Ayanda's utility? That is, which price ratio coupled with Biko's response to it will result in a post-exchange allocation with the highest possible utility for Ayanda?

Best response and incentive compatibility

For the first stage, that is, determining how Biko will respond to each price she might offer, Ayanda uses whatever information she might have, such as her experience in the past with Biko's response to offers, her best guess as to Biko's utility function, or her experience with other people she thinks are similar to Biko.

Just as in Chapter 3 there is a budget constraint limiting the exchanges he can undertake, but this is now a line giving feasible combinations of data and coffee available to him through exchange at some given price. If the price is p—the number of gigabytes of data per kilogram of coffee—and his

post-exchange bundle is denoted as (x^B, y^B), then Biko's budget constraint requires that the value of his post-exchange bundle must be the same as the value of his endowment bundle, or:

$$px^B + y^B = px_z^B + y_z^B \tag{4.6}$$

$$\text{or } p(x^B - x_z^B) = y_z^B - y^B \tag{4.7}$$

The second version of the budget constraint means that the value of the coffee that he acquires (at the price p) or $x^B - x_z^B$ must be equal to the value of the data that he gives up $y_z^B - y^B$.

We can rearrange Biko's budget constraint another way to show that the price p must be equal to the ratio of the amount of data he gives up to the amount of coffee he gets:

$$p = \frac{y_z^B - y^B}{x^B - x_z^B} \tag{4.8}$$

We show the derivation of Biko's best-response function in Figure 4.11. We start, in panel (a) by showing Biko's best response to *one particular price*.

We know that given the price p_4 Biko will choose how much data to transfer to Ayanda in return for her coffee in order to maximize his utility subject to his budget constraint. In panel (a) we show his feasible set with his budget constraint for that particular price, p_4. The budget constraint includes the point **z** because one of the feasible choices he could make while respecting the budget constraint is to exchange nothing.

In Figure 4.11 the slope of the p_4 line is the amount of data that Biko gives up (Δy^B) divided by the amount of coffee that he gets (Δx^B), when the price is p. So:

$$p = -\frac{\Delta y^B}{\Delta x^B} = \frac{y_z^B - y^B}{x^B - x_z^B}$$

$$\text{marginal rate of transformation } (mrt) = -\text{slope of the price line}$$

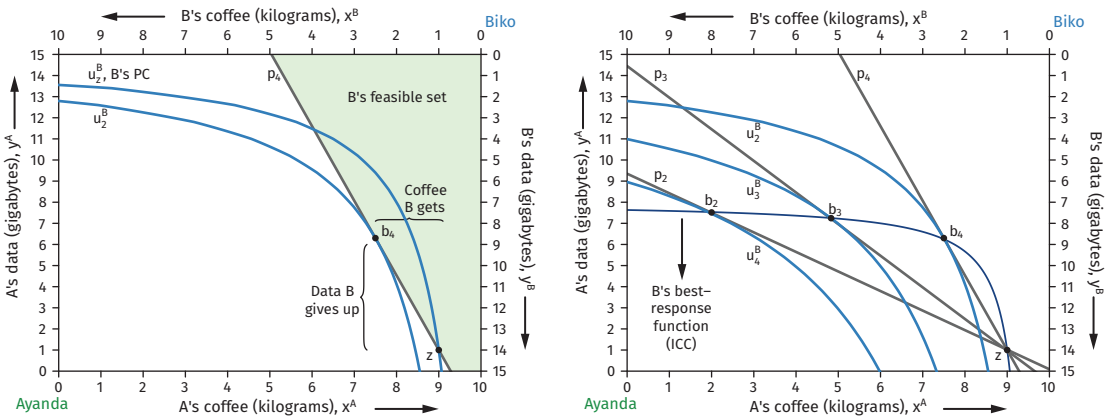
For any given price this is the kind of individual utility maximization problem that you studied in Chapter 3 in which the solution is to find the allocation at which the $mrs = mrt$ rule holds. You can see in panel (a) that the highest indifference curve that Biko can reach, consistent with his budget constraint (labeled u_2^B) is tangent to his budget constraint at point $\mathbf{b_4}$. This result expresses the principle of constrained optimization that you have already learned. It is a point equating:

- the slope of his indifference curve, which is the negative of the *marginal rate of substitution*; and
- the slope of the feasibility frontier—in this case the budget constraint—which is the negative of the *marginal rate of transformation* of coffee into data.

! REMINDER In Chapter 3 Harriet chose between two goods—fish and money left over. The price of fish was p and the "price of money" was 1 (1 Rupee is worth one Rupee.) Here we have two goods—coffee and data—but as before we have just one price, p which is the value of gigabytes of data expressed in kilograms of coffee instead of money. This is the same as if we had just set the price of coffee at 1, so as to focus on the relative price of the two goods.

M-CHECK $\Delta y^B = y^B - y_z^B$, implying $-\Delta y^B = y_z^B - y^B$. Therefore, $-\frac{\Delta y^B}{\Delta x^B} = \frac{y_z^B - y^B}{x^B - x_z^B}$.

Figure 4.11 Constructing B's best-response function (ICC). In panel (a), B's feasible set is in the upper-right corner of the Edgeworth box because, as we explained in Figure 4.3, the upper-left corner of the box is the origin for him (indicating zero of both goods). In panel (a), when the price p_4 is equal to 3.53 Biko reaches his highest feasible indifference curve (u_2^B) by giving up 5.3 gb of data in return for 1.5 kg of coffee. In panel (b) he chooses post-exchange bundles indicated by points $\mathbf{b_3}$ and $\mathbf{b_2}$ in response to prices $p_3 < p_4$ and $p_2 < p_3$. B's best-response function (ICC) connects these and similar points, all of them B's utility-maximizing bundle, for different prices.



(a) B's best response to a price = p_4

(b) B's best-response function (ICC)

The *mrt* is the price, p, set by Ayanda, that tells Biko how many gb of data he has to give up to get 1 kilo of coffee. Biko's best response is to choose a post-exchange bundle that satisfies the two conditions:

$$mrs = mrt \text{ tangency: } mrs^B(x^B, y^B) = mrt = p \tag{4.9}$$

$$\text{and, budget constraint: } px^B + y^B = px_z^B + y_z^B \tag{4.10}$$

Equation 4.9 expresses the *optimizing* part of Biko's choice, while Equation 4.10 expresses the *constraint*. The utility Biko enjoys at $\mathbf{b_4}$ in the figure is the best he can do at that price and it is also greater than the utility of his endowment bundle ($u_2^B > u_z^B$). From this we conclude that if the price is p_4 Biko will choose the post-exchange bundle given by point $\mathbf{b_4}$. This gives us one point on Biko's best-response function.

In panel (b) we construct Biko's best-response function, by repeating the analysis in panel (a) but for differing prices, tracing out a curve in the (x, y) coordinates. This is his *best-response function* because, by construction, points on the curve show for each the value of p the post-exchange allocation that maximizes his utility if he could buy any amount of Ayanda's coffee at the price p. Ayanda now has all the information she needs to set the price.

M-NOTE 4.6 The incentive-compatibility constraint

Here we show the derivation of the incentive-compatibility constraint for Ayanda's utility choice of a utility-maximizing price to offer Biko. This equation will show, for every price that Ayanda could offer, the amount of goods that Biko will be willing to exchange.

To do this we use the two conditions that Biko's response must satisfy. Given the price p offered by Ayanda, Biko's budget constraint is Equation 4.10.

$$px^B + y^B = px_z^B + y_z^B, \tag{4.11}$$

where y^B is a function of x^B. That is Equation 4.9:

$$y^B(x^B) = -px^B + px_z^B + y_z^B$$

To maximize his utility $u^B(x^B, y^B(x^B))$, Biko will choose the bundle $(x^B, y^B(x^B))$ that satisfies:

$$\frac{du^B}{dx^B}u_x^B + u_y^B\frac{dy^B}{dx^B} = u_x^B - u_y^B p = 0$$

That is

$$mrs^B(x^B, y^B) = -\frac{u_x^B}{u_y^B} = -p = mrt \tag{4.12}$$

Suppose that $u^B = (x^B)^{\frac{1}{3}}(y^B)^{\frac{2}{3}}$, we can derive the incentive-compatibility constraint using Equations 4.11 and 4.12. From M-Note 4.2, we have

$$mrs^B(x^B, y^B) = \frac{1}{2}\frac{y^B}{x^B}$$

Moreover, the budget constraint can be rewritten as Equation 4.8, that is,

$$p = \frac{y_z^B - y^B}{x^B - x_z^B}$$

Therefore, we have

$$\frac{1}{2}\frac{y^B}{x^B} = -\frac{y_z^B - y^B}{x^B - x_z^B} \tag{4.13}$$

which defines the incentive-compatibility constraint shown in the Edgeworth box.

CHECKPOINT 4.10 First-mover advantage Explain why the person with TIOLI power is constrained by the participation constraint while the person with price setting power is constrained by the incentive compatibility constraint.

4.12 SETTING THE PRICE SUBJECT TO AN INCENTIVE-COMPATIBILITY CONSTRAINT

Biko's best-response function is the incentive-compatibility constraint for Ayanda's optimizing problem, shown in Figure 4.12. Because Ayanda always has the option of simply discarding some of the data she gets from Biko,

we can think about the green-shaded area under Biko's best-response function as her *feasible set*. The slope of Biko's best-response function is (from Ayanda's viewpoint) the marginal rate of transformation of coffee into data, given how Biko responds to each of the prices she could set.

You can see that starting at the endowment allocation, the best-response function is initially steep, so a modest amount of coffee that she gives up can be transformed—through exchange—into a substantial amount of data. But the more data she wishes to acquire—moving up on the best-response function—the less favorable to her the *mrt* becomes. For each additional kilogram of coffee that she gives up, she gets fewer and fewer gigabytes of data.

Notice that the incentive-compatibility constraint is more limiting to Ayanda than is Biko's participation constraint in Figure 4.12 (a) labeled: u_z^B, B's PC. This means that there are some allocations (between the participation constraint and the incentive-compatibility constraint) which would make Biko better off than at his endowment bundle, and which Ayanda would prefer to any point in her feasible set, but which Ayanda could not implement when she has price-setting power but not take-it-or-leave-it power.

Ayanda's choice of what price to set is a familiar constrained optimization problem. It proceeds in two steps:

1. Determine the *final allocation* she would like to implement by finding the point in the feasible set that is associated with the highest utility. To do this she uses the $mrs = mrt$ rule and selects point **n** in the figure, with its associated utility u_N^A. This is where her indifference curve is tangent to Biko's best-response function. This is shown in Panel (a) of Figure 4.12.
2. Determine the *price* that will implement this outcome. Every allocation on the best-response function corresponds to some particular price that will implement it. Price p^N shown in Figure 4.12 (b) implements point **n**.

We have given the price that Ayanda sets a superscript N because the allocation that it implements is a Nash equilibrium. To confirm that this is the case we ask two questions:

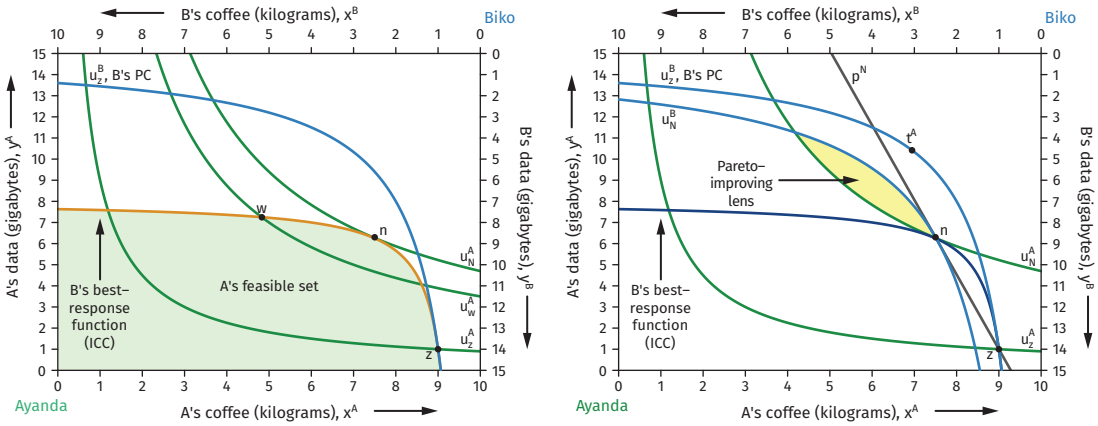
- Given the strategy that Ayanda has adopted—that is, setting the price p^N—is there any way that Biko could do better than he does by trading with her so as to implement her chosen allocation (point **n**)? The answer is no, because **n** is a point on his best-response function, which tells us that if she offers the price p^N the best he can do is to trade with her so as to implement her desired point.
- Given the strategy that Biko has adopted—his best-response function—is there any way that Ayanda could do better than she does by setting the price p^N? The answer is no, because she found point **n** exactly by doing the best she could given his best-response function.

There are two important aspects of the Nash equilibrium (allocation **n**) of this game.

First, the Nash equilibrium is *not Pareto efficient*. Ayanda's and Biko's indifference curves are **not** tangent at **n**, they intersect, and you know from the $mrs^A = mrs^B$ rule any allocation at which the indifference curves intersect is not Pareto efficient (because then the rule is violated). The reason why Ayanda implemented a Pareto-inefficient allocation is that the constraint she faced was not Biko's PC (the slope of which is mrs^B) but instead his best-response function (the slope of which is the mrt). So she implemented $mrs^A = mrt \neq mrs^B$ violating the Pareto-efficiency rule. The allocations that are Pareto superior to **n** are shown by the yellow lens between the indifference curves through **n**.

Second, the person who is *not* the first mover (Biko) receives a rent in the Nash equilibrium: as you can see from Figure 4.12 at **n** he is better off (on a higher indifference curve) than with his endowment bundle (which is his fallback option, namely no trade) indicated by the indifference curve labeled u_z^B, B's PC.

Figure 4.12 A sets the price subject to B's best-response function (ICC). Ayanda's utility-maximizing post-exchange bundle is indicated by point **n** where her indifference curve is *tangent* to Biko's best-response function (his price-offer curve or incentive-compatibility constraint). The negative of the slope of the solid gray line through both **n** and the endowment point **z** is equal to the price Ayanda chooses, p^N. Biko's budget constraint given by Ayanda's choice of p^N is tangent to Biko's indifference curve through **n** by construction, that is, because **n** is on Biko's best-response function. To interpret the lower-shaded area as a feasible set, it must be the case that A could choose *not* to consume the data or coffee she has in that area (that is, some of it could be thrown away).



(a) Maximizing utility subject to the incentive-compatibility constraint

(b) Pareto-superior alternatives to Ayanda's choice of **n**

There is an important lesson here: when one of the two parties has price-setting power, but not TIOLI power, she may use that advantage to advance her distributional interests in a way that implements an inefficient outcome.

This explains why when Ayanda has the power to set the price but *not* to stipulate the amount that Biko is to purchase at that price—when she has price-setting power but not TIOLI power—she uses her power to get a *larger* piece, but of a *smaller* pie. When she had TIOLI power she knew that she would get the *whole* pie—the entire economic rent—because the only constraint she faced was Biko's participation constraint. So for Ayanda with TIOLI power, her slice and the entire pie were the same thing. Doing the best she could do and implementing a Pareto-efficient allocation therefore coincided.

The takeaway is: when a person is maximizing their utility:

- *constrained by the other's participation constraint*, then the Nash equilibrium allocation will be Pareto efficient because the best they can do is to implement the $mrs^A = mrs^B$ rule; but if they are
- *constrained by the other's ICC* (best-response function) the result will not be Pareto efficient because instead they will implement the $mrs = mrt$ rule.

We will see that in many economic interactions—including credit markets, labor markets, and markets for goods of variable quality—it is the incentive-compatibility constraint that constrains the actor setting the price (or wage, or interest rate) not the participation constraint. So the Nash equilibria in these markets will be Pareto inefficient, even if the market in question is highly competitive.

Moreover, when firms face limited competition either in selling outputs or buying inputs, we will see that the same principle is at work. It is *not* the participation constraint that constrains the profit-making process, and so the resulting allocations will be Pareto inefficient.

CHECKPOINT 4.11 PS power vs. TIOLI power

a. Using Figure 4.12, by reading the relevant points on the x and y axes, say what the post-exchange allocations for Ayanda and Biko will be (how much coffee for each, how much data for each). Compare this to the post-exchange allocations when Ayanda has TIOLI power, calculated in M-Note 4.5.

b. Test your understanding of the first-mover case by explaining the outcome when Biko is the first mover and has price-setting power. Draw a new version of Figure 4.12.

4.13 APPLICATION: OTHER-REGARDING PREFERENCES—ALLOCATIONS AMONG FRIENDS

Ayanda and Biko are about to experience one final change in their identities, along with a personality transplant: they have become friends and they care about each other. Both are altruistic: they place some positive weight on the well-being of the other. This means, as you will recall from Chapter 2, that they are other-regarding: when evaluating an allocation they take account not only of their bundle but also the other person's bundle.

They still have a decision to make: how to divide up their coffee (still 10 kilos of it) and the data (15 gigabytes of it as before). But we will assume now as it was when you first met them that neither of them own any of either good—so there is no endowment allocation like our interpretation of point **z** so far.

To see how the Edgeworth box helps us to understand their decision problem and because this involves some unusual indifference curves, we first treat a hypothetical case in which Ayanda is altruistic and Biko is as before entirely self-regarding. (We do not imagine that Ayanda would put up with this, it is just a first step along the way to seeing how two other-regarding friends would look at the problem.)

An altruistic utility function

Altruistic Ayanda cares not only about her bundle at an allocation, but also what Biko gets. Ayanda's utility therefore depends not only on x^A and y^A but also on x^B and y^B. We measure how much she cares about what Biko gets—her *degree of altruism*—by λ ("lambda") a number that varies from 0, if she is entirely self-regarding, to one-half if she places as much weight on what Biko gets as what she herself gets, in which case she would be called a *perfect altruist*.

M-NOTE 4.7 An altruistic utility function

Remember if Biko did not exist so that Ayanda were making her choice of an allocation in isolation, her utility would be

$$u^A(x^A, y^A) = x_A^{\alpha} y_A^{(1-\alpha)} \tag{4.14}$$

But interacting with Biko and dividing goods with him, for $\lambda > 0$ we have Ayanda's utility function as an altruist:

$$u^A(x^A, y^A, x^B, y^B) = \left(x_A^{\alpha} y_A^{(1-\alpha)}\right)^{(1-\lambda)} \left(x_B^{\alpha} y_B^{(1-\alpha)}\right)^{\lambda} \tag{4.15}$$

To see why we say that λ is a measure of how much Ayanda cares about what Biko gets we can take the natural logarithm of equation 4.15

$$\ln(u^A) = (1-\lambda)\ln\left(x_A^{\alpha} y_A^{(1-\alpha)}\right) + \lambda \ln\left(x_B^{\alpha} y_B^{(1-\alpha)}\right) \tag{4.16}$$

continued

Equation 4.16 says that the natural logarithm of A's utility is $(1-\lambda)$ times the natural logarithm of her valuation (if made in isolation) of her own bundle plus λ times the natural logarithm of B's evaluation (if made in isolation) of his bundle.

If Biko is also altruistic, then his utility function has the same structure as Ayanda's but the interpretation of λ is the opposite. In Biko's utility function λ is the exponent of Ayanda's bundle, and $(1-\lambda)$ is the exponent on his own bundle, the opposite of where these terms appear in Ayanda's utility function. The totally self-regarding person, Biko in this case, recalling that he places no weight on the bundle of the other person; his degree of altruism, $\lambda = 0$. So self-regarding B's utility function would be:

$$\begin{aligned} u^B(x^A, y^A, x^B, y^B) &= \left(x_A^{\alpha} y_A^{(1-\alpha)}\right)^0 \left(x_B^{\alpha} y_B^{(1-\alpha)}\right)^1 \\ &= x_B^{\alpha} y_B^{(1-\alpha)} \end{aligned} \tag{4.17}$$

which is just his previous utility function *before* we introduced λ. Any term raised to a zero exponent (as in Biko's utility function) has a value of 1.

CHECKPOINT 4.12 Spite and love

a. What would it mean in the utility function 4.15 if we had $\lambda < 0$? Can you give an example of someone acting as if they had preferences like this?

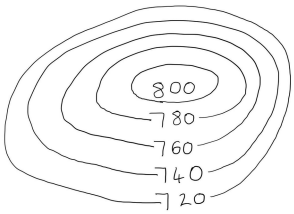
b. Can you imagine a person having a value of λ greater than one-half, what would this mean? Can you think of situations in which people have acted on preferences of this type?

An altruistic indifference map

To draw her indifference map, we will give Ayanda some particular value of λ. Figure 4.14 (a) shows an Edgeworth box representing a not-perfectly-altruistic Ayanda with $\lambda = 0.4$.

Ayanda's indifference curves look like the contours on a topographic map of a mountain. We described the constrained optimization process in Chapter 3 as a kind of hill climbing, where both elements in the bundle were a "good" and over the entire map, the mountain rose to higher levels if you moved in the "northeast" direction, that is more of both goods. In those figures you never saw the top of the mountain, because there was not any top. There was no such thing as "too much" of either good.

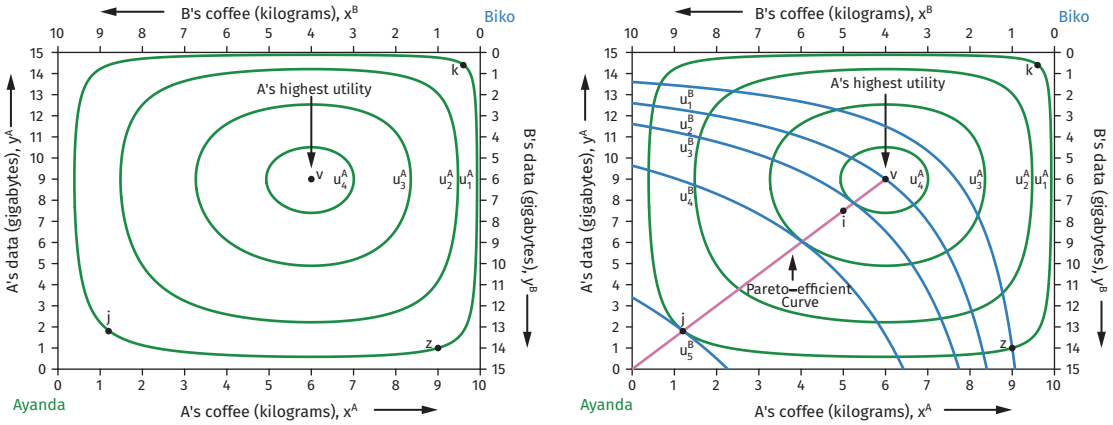
Figure 4.13 Reminder: we used the lower-left contours of a hill in Chapter 3. Here the indifference curves are themselves full contours of a hill.



But Ayanda's indifference map has a definite peak at the allocation indicated by point **v**. The reason is that from her other-regarding perspective she can have "too much" of a good when that means that Biko (who, she cares about) has too little. This is why Ayanda's indifference curves are oval shaped.

Notice that when she has little of either good (close to her origin in the lower left of the box) her indifference curves look as you have seen before. In this situation both coffee and gigabytes are "goods" so more of each is better, even if this necessarily means less for Biko. So the indifference curves slope downward, as you would expect. Moving up or to the right

Figure 4.14 Allocation and distribution with one altruistic person and one self-regarding person. In panel (a) the green oval-shaped curves labeled u^A are the indifference curves based on Ayanda's utility function. In both panels, points **z** and **i** are the same allocations here as in Figure 4.5. Notice that in panel (a) because Ayanda values what Biko gets she regards the allocation at point **j** as equivalent to the allocation **k**, despite the fact that she receives less of both goods at **j** than she does at **k**. For the same reason, Ayanda's utility reaches a maximum at the allocation **v** indicated in the figure. The Pareto-efficient curve now does not include **k**, because Biko is so deprived of both goods at the point that Ayanda prefers **v** to **k**.



(a) Altruistic Ayanda's indifference curves

(b) Altruistic Ayanda and self-regarding Biko

brings you to a higher indifference curve. In this part of the figure "more is up."

But beyond a certain point "more" for Ayanda is no longer "up." If she has most of both goods, then getting even more is not something she values, so moving up and to the right leads her to lower not higher indifference curves.

To understand the upward-sloping parts of Ayanda's indifference curves, remember that if one of the axes represents a good and the other a bad, then the indifference curve slopes *upward*, as in the case of study time (a bad) and expected grades (a good). In the upper right of the box for example near point **k** where she has *most* of both goods and Biko has *little* of either the indifference curves slope downward because for Ayanda having more of either good (and Biko having less) *reduces* her utility: both her coffee and her gigabytes are "bads" not goods.

M-CHECK The tangencies at points above and to the right of **v** in Figure 4.14 (not shown in the figure) that we have excluded from the Pareto-efficient curve illustrate the cases in which the $mrs^A = mrs^B$ rule fails.

In Figure 4.14 (b) we add Biko's conventional (self-regarding) indifference curve, so we now know how both of them evaluate every feasible allocation given by the dimensions of the box. To do this we use Biko's self-regarding

utility function with the value he places on Ayanda's utility being zero that is $\lambda = 0$ because he is entirely self-regarding (that is, zero altruism).

The Pareto-efficient curve is, as before, made of points of tangency between Ayanda's and Biko's indifference curves. But now we *exclude* tangencies at allocations for which Ayanda places a *negative* value on having more of one or both of the goods, above and to the right of her "utility peak" at **v**. As a result the Pareto-efficient curve in Figure 4.14 looks different from the one in Figure 4.5 as it does not extend upward and to the right beyond Ayanda's maximum **v**. Ayanda does not want more of either good than she gets at her maximum **v**, while Biko prefers **j** to any allocation in which she gets less of either or both of the goods.

CHECKPOINT 4.13 Altruistic comparisons Consider Figure 4.14

a. Where is Biko's utility peak in the figure (analogous to Ayanda's allocation at point **v**)?
b. Where would point **v** be if $\lambda = \frac{1}{2}$ (or as close to $\lambda = \frac{1}{2}$ as possible)?
c. What happens if Ayanda is self-regarding and Biko is an altruist? How would the Edgeworth box change?

Efficiency and fairness among altruists

With these analytical tools we can now look at the decision problem faced by the friends Ayanda and Biko both with other-regarding social preferences. Figure 4.15 shows for the same Edgeworth box, the indifference maps of the two. We assume their levels of altruism toward each other to be the same, that is, λ.

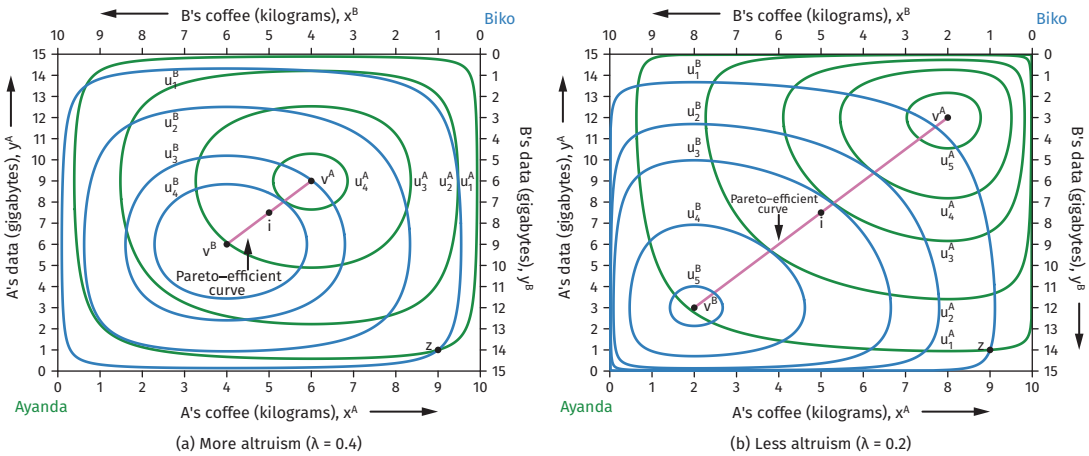
Unlike the case of one altruistic actor, now both participants have preferred allocations in the interior of the Edgeworth box. They both would like to avoid "too much of a good thing."

Each of their preferred allocations are shown in the figures by the allocations, $\mathbf{v}^{\mathbf{A}}$ for Ayanda and $\mathbf{v}^{\mathbf{B}}$ for Biko. Around each person's preferred allocation, their iso-social welfare curves move outward and downward in all directions, corresponding to lower and lower levels of utility.

As you can see from Figure 4.15 (a) the Pareto-efficiency curve is a line between their two preferred "utility peaks" $\mathbf{v}^{\mathbf{A}}$ and $\mathbf{v}^{\mathbf{B}}$. By comparing panels (a) and (b) depicting greater and lesser degrees of altruism, you can see that the more altruistic they are, the shorter the Pareto-efficient curve is, because greater altruism eliminates more of the extremely unequal allocations.

There is still a conflict of interest, however. At Ayanda's preferred allocation Biko has a level of utility less than the utility he enjoys at this own preferred allocation. The same is true of Ayanda: she does much better at her preferred allocation than at Biko's.

Figure 4.15 Altruistic indifference maps. The two panels depict two different levels of altruism: high ($\lambda = 0.4$) in panel (a) and low ($\lambda = 0.2$) in panel (b). The allocations indicated by the points $\mathbf{v^A}$ and $\mathbf{v^B}$ are respectively A's and B's preferred allocation. The Pareto-efficient curve is composed of all allocations at which both own coffee and own data are "goods" rather than "bads" to both A and B, and where their marginal rates of substitution are equal, that is, their indifference curves are tangent.



Along the Pareto-efficient curve movements in one direction or the other necessarily involve one gaining and the other losing. As always the Pareto-efficient curve is a conflict region even among altruists. The fact that the "utility peaks" are closer together in panel (a) illustrating a greater degree of altruism means that the conflict of interest between them is lesser the more altruistic they are. This is one of the reasons why agreeing on a set of rules of the game may be less of a challenge among friends or neighbors than among total strangers.

How might they resolve their remaining conflicts of interest? Here, to make a decision, they need to go beyond their own utilities (even taking account of their altruistic nature) to bring in some additional way of making a judgement. They might adopt:

- a *social norm* that they both share, for example if one of the two found the coffee and the data they could go by "finders keepers"; in this case whichever of them who found the goods could make the decision, presumably implementing his or her preferred allocation;
- a *procedural rule of justice*, for example flipping a coin to see whose preferred allocation $\mathbf{v^A}$ or $\mathbf{v^B}$ would be implemented; or

- *a substantive rule of justice*, for example picking an allocation on the Pareto-efficient curve midway between their two utility peaks.

Point **i** in the figures is a reference point showing the allocation that the Impartial Spectator (who weights Ayanda's and Biko's utilities equally) would implement. This is the same allocation that they would have *both* preferred had they been perfect altruists.

CHECKPOINT 4.14 Altruism and rents Why does altruism reduce the conflict over which allocation to implement?

4.14 CONCLUSION

From the silent trade that Ibn Battuta and Herodotus described centuries ago to eBay, Amazon, and Alibaba today, people have exchanged goods to their mutual advantage and engaged in conflicts over who would get the lion's share of the gains from exchange. The four scenarios we have introduced have made it clear that the outcomes of these exchanges and conflicts depend on the institutions under which they take place, and the preferences of the people involved.

We have examined several institutional approaches to resolving the conflict between Ayanda and Biko over allocations of available goods. They all illustrate the dilemma posed in social interactions between:

- The goal of reaching an allocation that is Pareto superior to the endowment and possibly even Pareto efficient.
- The goal of resolving the conflict over the distribution of the resulting economic rents in a way that is fair.

Table 4.2 summarizes some of the key aspects of the cases we have discussed. Which of the scenarios in the table are relevant in any particular case depends on the rules of the game for the society of which the players are a part.

Table 4.2 **The rules of the game: Objectives, constraints, and the characteristics of the resulting allocations.**

Case	Constraints implied by the rules of the game	Objectives of the actor(s)	Characteristics of the resulting allocation
Impartial Spectator	The available goods (dimensions of the Edgeworth box)	Social welfare equally weighting the utility of each	Pareto efficient and fair (by the standard of the social welfare function)
Symmetrical bargaining with no first-mover advantage	Endowment allocation (private property) Each player's participation constraint (PC) at each stage of the bargaining	Utility of the two traders	Pareto efficient if no impediments to bargaining, Pareto improvements over the endowment allocation
Take-it-or-leave-it power	Endowment allocation (private property) Second-mover's PC	Utility of the first and second movers	Pareto efficient, first-mover's rent is all of the gains from exchange
Price-setting Power	Endowment allocation (private property) Second-mover's incentive compatibility constraint (ICC)	Utility of the first and second movers	Pareto inefficient; first mover gets most of the gains from exchange but second mover gets some
Legislation	The available goods (money and time)	Whatever the legislators were seeking to accomplish (possibly the social welfare optimum)	Possibly Pareto inefficient, could be improved upon by private bargaining
Bargaining away from legislated hours and wages	The new participation constraints given the fallback position implemented by the legislation	Utilities of the two players	Pareto efficient if no impediments to bargaining, otherwise possible Pareto improvements over the new fallback
Altruism	The available goods (dimensions of the Edgeworth box)	Utilities of both (taking account of how much they value the other's bundle); fairness	Pareto efficient and (if they can agree on a fairness principle) fair

MAKING CONNECTIONS

Constrained optimization in strategic interactions: The constrained optimization techniques developed in Chapter 3 are used to better understand strategic interactions introduced in Chapters 1 and 2.

Optimization rules: In addition to the $mrs = mrt$ rule which we developed in Chapter 3 for individual optimization we also have the $mrs^A = mrs^B$ rule defining a Pareto-efficient outcome, both of which are used in strategic interactions.

Mutual gains from trade: If the endowment allocation (status quo) is not Pareto efficient, then mutual gains are possible by implementing some different allocation of the goods which people may be able to agree to voluntarily.

Rents and conflicts: These improvements over the fallback option accruing to the players are *rents*, made possible by the gains from exchange; there will be conflicts over how the total rent is distributed among the players.

Institutions (rules of the game) and bargaining power: The distribution of these rents in the Nash equilibrium allocation depends on the players' preferences and the initial endowment, as well as on the property rights in force, other institutions, and the forms of bargaining power that each participant can exercise.

Pareto efficiency, institutions: Some rules of the game will result in Pareto-efficient outcomes. Examples are the allocation implemented by the imaginary Impartial Spectator, symmetrical bargaining with no barriers to trading as long as mutual gains are possible, and take-it-or-leave-it power exercised by one player. Price-setting power by one person, however, results in a Pareto-inefficient outcome.

Self-regarding and social preferences: Among the set of Pareto-efficient allocations there will generally be conflict of interest among the participants. But, the extent of these conflicts may be reduced by social preferences such as altruism or a commitment to fairness.

IMPORTANT IDEAS

utility function	marginal rate of substitution	Cobb-Douglas utility
Edgeworth box	Pareto criterion	Pareto-improving lens
Pareto efficiency	Pareto-efficient curve	utility possibilities frontier
endowment	post-exchange allocation	Impartial Spectator
social welfare function	$mrs^A = mrs^B$ rule	iso-welfare curve
bargaining	$mrs = mrt$ rule	allocation
altruism	private property	first-mover advantage
take-it-or-leave-it power (TIOLI power)	price-setting power	participation constraint (PC)
incentive-compatibility constraint (ICC)	price-offer curve	institutions/rules of the game
gains from trade	economic rent	bundle
willingness to pay	voluntary exchange	best-response function (ICC)

MATHEMATICAL NOTATION

Notation	Interpretation
α	exponent of good x in the Cobb-Douglas utility function
$u()$	utility function
$\bar{x}, \bar{y}$	total amounts of x and y available
p	price of coffee (gb of data per kilo of coffee)
W	social welfare function
λ	extent of altruism (value placed on the other's bundle)
$h()$	nonlinear term of quasi-linear utility function
a	parameter in the linear term of quasi-linear utility function

Note on superscripts and subscripts: A, B, and i: people; z: endowment point; $\mathbf{t}^i$: outcome with a take-it-or-leave-it offer by player i.

CHAPTER

5

COORDINATION FAILURES AND INSTITUTIONAL RESPONSES

Right now, my only incentive is to go out and kill as many fish as I can . . . any fish I leave is just going to be picked by the next guy.

John Sorlien
Rhode Island (US) lobsterman[1]

DOING ECONOMICS

This chapter will enable you to:

- Understand how the external effects of our actions on others that are not taken into account when people make choices lead to coordination failures.
- Represent social interactions with graphical and algebraic indifference curves, feasible sets, best-response functions, and Nash equilibria.
- To see how at the Nash equilibria, the extent of both inequality and unrealized potential mutual gains will depend on the rules of the game.
- To explain the dynamic process by which a Nash equilibrium may be attained.
- Understand how government policies such as taxes or direct regulation, the exercise of ownership rights or power by private individuals, and cooperation based on social preferences can help to avert a coordination failure.
- See that the Pareto improvement made possible in each of these cases occurs because (in very different ways) the rules of the game in each case induce actors to internalize the external effects that their actions have on others.

5.1 INTRODUCTION: TRAGEDY AVERTED

Don't get him wrong: John Sorlien, the lobsterman, is not the kind of self-interested and amoral *Homo economicus* you might find in an economics textbook. He is actually an environmentalist of sorts, and as President of the Rhode Island Lobstermen's Association he was up against a serious problem of incentives, not a shortcoming of human nature. When he started lobstering at the age of 22, he set his traps right outside the harbor at Point Judith, within a few miles of the beach, and made a good living. But the inshore fisheries have long since been depleted, and now his traps lie 70 miles offshore. He and his fellow lobstermen are struggling to make ends meet.

Across the world in Port Lincoln on Australia's south coast, Daryl Spencer, who dropped out of school when he was 15 and eventually drifted into lobstering, has done much better. During the 1960s the Australian government assigned licenses—one per trap—to lobstermen working at the time, and from that time on, any newcomer seeking to make a living trapping lobsters off Port Lincoln had to purchase licenses.

Spencer purchased his start-up licenses for a modest sum and by 2000 his licenses were worth more than one million US dollars (in 2000 prices); considerably more valuable than his boat. More than giving Spencer a valuable asset, the policy has limited the Australian lobstermen's work: Spencer has 60 traps, the maximum allowed, at the same time at Point Judith, John Sorlien was pulling 800 traps and making a lot less money.

Regulating the amount of lobsters trapped is a coordination problem. Point Judith and Port Lincoln represent extremes along a continuum of failure and success; with the lobstermen of Port Lincoln reaping the mutual gains made possible by a joint decision to limit the number of traps. One may wonder why the Point Judith fishermen do not simply copy the Australians. This is especially surprising since one of Sorlien's friends and a fellow Point Judith lobsterman visited Port Lincoln, returning with tales of millionaire fishermen living in mansions. But getting the rules right is a lot more difficult than the Port Lincoln story may suggest, and good rules often do not travel well.

One of the common obstacles to successful coordination is that the rules that address the coordination problem also implement a division of the gains to cooperation. In Port Lincoln, those who were awarded the licenses benefited; others did not. Had the young Daryl Spencer not agreed one day to help out a lobsterman friend and then decided to become a lobsterman himself, someone else would be a millionaire, and Spencer might still be painting houses and complaining about the high price of lobsters.

Even if policies to address coordination failures could result in benefits for everyone affected, how a group coordinates, and what policies they coordinate on will affect how these benefits will be distributed. And this

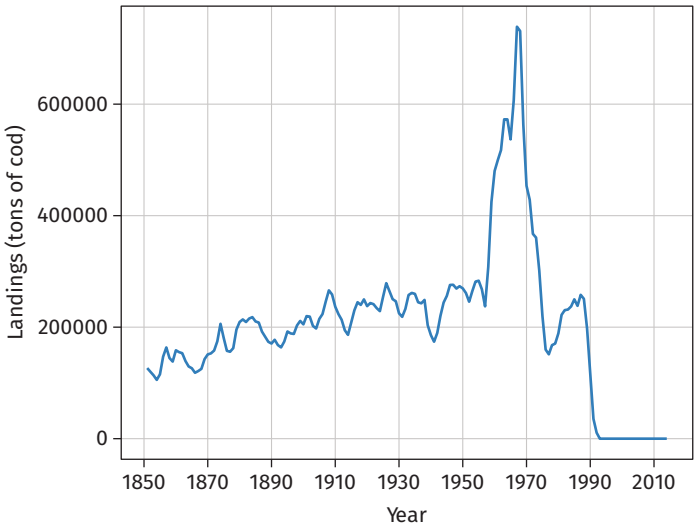
Figure 5.1 Sounding the alarm on climate change, a coordination problem. Greta Thunberg, then 16 years old, speaking at the United Nations in 2019 about what is probably the most serious coordination problem that humanity has ever faced. She said: "We are in the beginning of a mass extinction, and all you can talk about is money and fairy tales of eternal economic growth. How dare you!"[2]

Photo by Jemal Countess/UPI/Alamy Live News.

! REMINDER A coordination problem is a situation in which people could all be better off (or at least some be better of and none be worse off) if they were able to jointly decide how to act—that is, if they could coordinate their actions—than if they act independently.

Figure 5.2 The Grand Banks (North Atlantic) fisheries: cod landings in tons (1851–2014). In the 1960s new fishing technologies allowed a dramatic increase in cod fish caught ("landings") far outpacing the capacity of the fish to reproduce. This led to a partial collapse of the fishery in the 1970s and a total collapse in 1992 when the Canadian government banned fishing entirely. Restoration of fishing stocks to the sustainable levels of the past may occur by the 2030s.

Sources: Frank et al. (2005); Rose and Rowe (2015).



makes it difficult to agree on a policy. An example is the Ultimatum Game experiment, in which conflicts over the size of the Proposer's and Responder's "slice of the pie" sometimes result in neither getting any piece of the pie at all.

A far more momentous coordination problem is climate change: conflicts between richer and poorer nations, conflicts between those who make their living in carbon-intensive industries and others who bear the costs of that production are prominent among the reasons for the failure to address the climate emergency. Depleting a fishing stock is little different in the structure of its incentives and its consequences from many other social interactions. In Chapter 9, for example, using exactly the model we develop here of the coordination problem that fishermen face "overharvesting fish," that is depleting fish stocks, we will study how firms compete on markets "overharvesting customers," each firm attempting to sell more, and as a result reducing the market for the products of competing firms.

In the case of overfishing or "overharvesting" consumers, when one person fishes more, or a firm cuts prices, the external effects—on the catch of the other fishermen or the profits of other firms—are negative. But

external effects can also be positive, for example if you find a way to reduce your carbon footprint, benefiting others including future generations.

In this chapter we develop tools to understand the nature of coordination problems like the tragedy of the commons. We use these tools to analyze some of the policies (changes in the rules of the game) that improve the Nash equilibrium outcome when external effects are present.

We will illustrate how coordination failures occur and how policies might address them with the example of common property resource problems (or common pool resource problems). The "common property" or "common pool" is the stock of fish available for catching or the pool of customers who might purchase the goods sold by the firms.

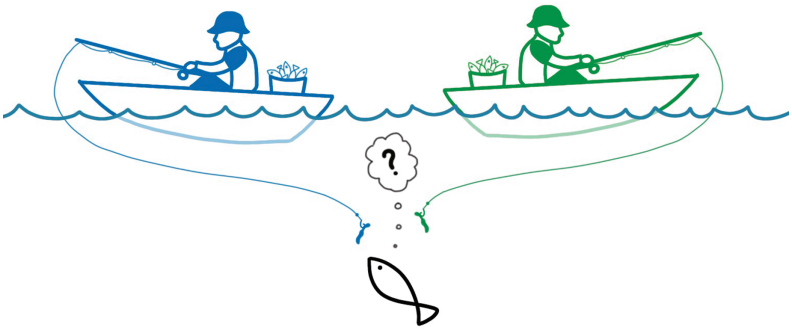
Remember from Chapter 2 that common property resources are non-excludable and rival, people who use them impose *external costs* on each other. The "problem" is that self-regarding people will *overexploit* the resource because they will not place any value on the negative external effects of their actions on others. Just such a pattern of exploitation is shown in Figure 5.2, which displays the catches of cod fish in the North Atlantic fisheries.

5.2 A COMMON PROPERTY RESOURCES PROBLEM: PREFERENCES

Let's consider a specific example of a common property resource problem: the overexploitation of an environmental resource. It could be the oceans, or forests, or a livable planet, but we'll stick to the problem of overharvesting

Figure 5.3 Abdul and Bridget trying to catch the same fish. The lake is a common pool resource, so the benefits are rival and each person's fishing imposes a negative external effect on the other. We call them fisher*men* which they *were* when it was Bob and Alfredo in Chapter 2; but neither fisher*people* nor fishers seemed right.

Picture credit: Anmei Zhi.



fish. We will look at the ways that the rules of the game and the preferences of the actors determine what we expect to happen in these situations.

Preferences over fishing time and fish consumed

We turn now to the problem confronted by two fishermen, called Abdul (A) and Bridget (B). We model just two fishermen as a way of representing how a large number of them might interact. They fish in the same lake, using their labor and their nets. To start, we assume they consume the fish they catch (what we call their "catch") and do not engage in any kind of exchange. As a benchmark for comparison with later changes in the rules of the game, we will begin by assuming that they do not make any agreements about how to pursue their economic activities. (Recall that this means that they are engaged in a *noncooperative game*.)

M-CHECK Abdul's utility function in fish and fishing time is *quasi-linear*: it is linear in fish—he derives a positive and constant marginal utility from consuming fish—but is *negative* (it is a *dis*utility) and nonlinear in fishing time. His marginal disutility of fishing time is not constant; it is greater the more time he spends fishing.

Each derives well-being from eating fish and experiences a loss of well-being (disutility) with additional fishing time. We represent their preferences when they are engaged in some amount of fishing with the following quasi-linear utility functions:

Fisherman's utility = Fish consumption – Disutility of fishing

Abdul's utility $$u^A(h^A, y^A) = y^A - \frac{1}{2}(h^A)^2 \quad (5.1)$$

Bridget's utility $$u^B(h^B, y^B) = y^B - \frac{1}{2}(h^B)^2 \quad (5.2)$$

The utility function given by Equation 5.1 tells us four things about Abdul's preferences:

- *Consumption* (y^A) measured in pounds (0.454 of a kilogram) of fish is a "good"; Abdul derives utility from obtaining more consumption (consuming more fish) which is why y^A has a positive sign.
- *Time spent fishing* (h^A) measured in hours is a "bad": the second term has a negative sign.
- *Utility* (u^A) is increased by one unit if he is able to consume one more pound of fish, so the units in which we can measure utility are pounds of fish.
- *Marginal utility of fish consumption* is not diminishing but instead is a constant (equal to 1, because the coefficient of y in the utility function is 1).

Bridget's utility function Equation 5.2 is interpreted in the same way as Abdul's. Both of them refer to some given time period, such as a week. So *output and consumption* are pounds of fish caught and eaten in a week, while *time spent fishing* is hours fished over the course of a week.

If they do not fish at all, they are able to find work yielding them a utility (income minus the disutility of work on that job) equal to $u_z = y_z$. These are

labeled with the subscript z because this is their fallback position (as the endowment allocation was in Chapter 4).

To decide how much time to fish, people like Abdul have to balance their disutility of hours of work with the utility of consumption that they get from consuming the fruits–or the fish–of their work time. To understand this process, we look at Abdul's indifference curves. His indifference curves provide a comparison of all possible combinations of fishing time and fish consumption, even if many of those combinations are not available to Abdul.

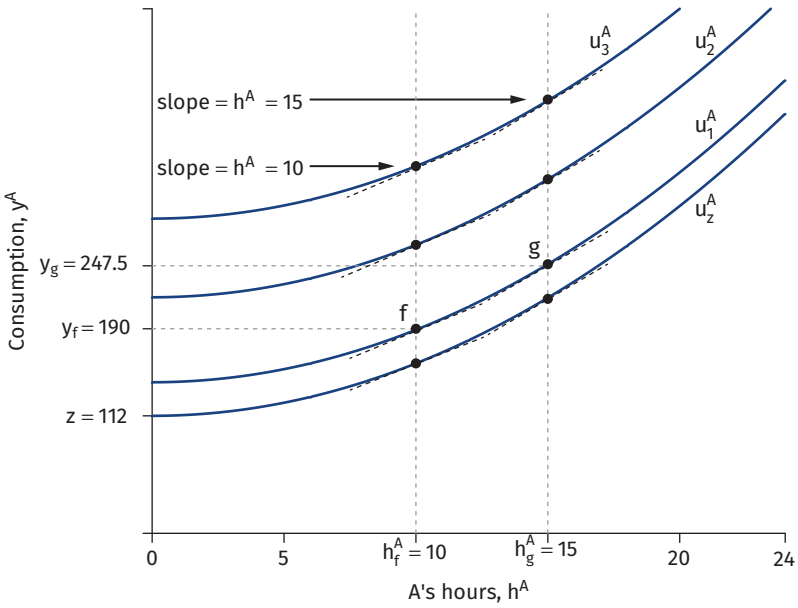
REMINDER The indifference map provides information on how he evaluates all of the imaginable combinations of fishing time and fish caught. It says nothing about the actions and outcomes that are *feasible* for him.

The marginal rate of substitution and the marginal cost of fishing time

Four indifference curves derived from Abdul's utility function, Equation 5.1, are presented in Figure 5.4. Notice the following:

- The higher numbered (meaning more preferred) indifference curves are above (more fish) and to the left (less work).
- The curves slope upward because fish is a good and fishing time is a bad, so comparing points **f** and **g** he is indifferent between fishing less and consuming less (point **f**) and fishing more and consuming more (point **g**).

Figure 5.4 Abdul's indifference curves over output (y^A) and fishing time measured in hours (h^A). Output (fish) (y^A) is a "good" and provides Abdul with positive utility, whereas fishing time (h^A) is a "bad." Notice that Abdul's indifference curves in fishing hours and output are *upward-sloping*, similar to the indifference curves over money (income, a good) and working time (a bad) in Chapter 4.



- The lowest indifference curve is labeled u_z^A and its vertical axis intercept is point **z** or the level of utility measured in fish per week, y_z^A that he will receive if he does not fish at all.
- For any given level of y^A the indifference curve is steeper the more hours Abdul works: the more he works, the greater is his dislike of working more compared to how much he likes eating more fish.
- For any of the indifference curves, u_1^A, u_2^A, and u_3^A, the vertical intercept is the amount of utility (in pounds of fish) that, if they were not working at all, would be the same as the utility at every other point on that indifference curve.

! REMINDER From M-Note 3.2 you know that the slope of an indifference curve is $-\frac{u_x}{u_y}$ the negative of the ratio of the marginal utility of the variable on the x-axis divided by the marginal utility of the variable on the y-axis.

The negative of the slope of his indifference curve is Abdul's marginal rate of substitution between fish (y^A) and fishing time (h^A). This is the ratio of his marginal utility of fishing time to his marginal utility of fish. This quantity takes a particularly simple form in this case. Abdul's marginal utility of fish is 1 and (as is shown in M-Note 5.1) his marginal utility of fishing time is $-h^A$. So, the marginal rate of substitution of fish consumption for fishing time is:

$$mrs^A(h^A, y^A) = -h^A \tag{5.3}$$

Or, what is the same thing:

$$\text{slope of indifference curve} = h^A$$

Abdul's marginal rate of substitution of fish consumption for fishing time is $-h^A$, and this is also his marginal utility of fishing time, which is negative, because he regards fishing time as a "bad."

The slope of his indifference curve is his marginal *dis*utility of fishing time (just the marginal utility with the sign reversed). This is Abdul's maximum willingness to pay (in forgone consumption) to work less. If he were working 12 hours, then his disutility of hours of fishing $h^A = 12$ is the *greatest* amount of fish he would be willing to give up in order to be able have an hour more free time. This can also be seen as the *marginal cost of fishing* more, if he is already fishing 12 hours.

M-NOTE 5.1 The *mrs* and the marginal cost of fishing time

When Abdul's utility is given by Equation 5.1:

$$\text{Abdul's utility} \qquad u^A(h^A, y^A) = y^A - \frac{1}{2}(h^A)^2$$

We have:

$$\text{Marginal utility of fish consumed (holding } h^A \text{ constant)} = \frac{\partial u^A(h^A, y^A)}{\partial y^A} = 1$$

$$\text{Marginal utility of fishing time (holding } y^A \text{ constant)} = \frac{\partial u^A(h^A, y^A)}{\partial h^A} = -h^A$$

continued

The marginal *utility* of fishing time is *negative* (it *reduces* Abdul's utility and is equal to $-h^A$). We use the term marginal *dis*utility of fishing time for the same quantity but with a positive sign (it *increases* Abdul's *dis*utility).

The marginal rate of substitution of output for hours of work $(mrs^A(h^A, y^A))$ is the negative of the slope of the indifference curve, which is the ratio of the marginal utilities:

$$mrs^A(h^A, y^A) = \frac{-h^A}{1} = -h^A$$

This is Equation 5.3.

CHECKPOINT 5.1 The lake as a common property resource

a. Explain why the lake that Abdul and Bridget are fishing in is a common property resource. What are its characteristics? Explain.

b. Return to Chapter 1 and the choice of strategies that the fishermen had in the Fishermen's Dilemma to Fish 10 hours or Fish 12 hours. Substitute these values into the utility functions to see what the payoffs in the corresponding game table would be if the fishermen could only choose these two strategies. Find the Nash equilibrium of the game.

5.3 TECHNOLOGY AND ENVIRONMENTAL LIMITS: THE SOURCE OF A COORDINATION FAILURE

A coordination problem arises because Abdul or Bridget fishing more reduces the amount of fish the other catches in an hour of fishing.

These external effects are part of the **technology** of fishing. A technology is a description of the relationship between inputs—such as fishing time, equipment, and fish in the wild—and outputs—in this case caught fish. A technology is often described mathematically as a production function. You already used a production function in Figure 3.8 where the input was time spent studying and the output was learning.

! REMINDER A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows on the other.

We depict Abdul's production function in the top panel of Figure 5.5. The higher of the two green curves represents the relationship between his labor input and his fish output when Bridget is *not* fishing at all, that is: $h^B = 0$. The lower black curve shows how his output varies with his time fishing when Bridget fishes $h^B = 12$ hours.

Below is Abdul's production function and a similar one for Bridget, where x^A for Abdul and x^B for Bridget represent the number of fish caught by each of them in a week and h^A and h^B are the hours of fishing time they work during the week. Thus, production functions translate the actions taken by

TECHNOLOGY A technology is a description of the relationship between inputs—including work, machinery, and raw materials—and outputs.

the two—their fishing hours (h^A and h^B)—into the amount that each catches (x^A and x^B) and consumes (y^A and y^B).

$$\text{A's catch \& consumption:} \quad y^A = x^A(h^A, h^B) = h^A(\alpha - \beta(h^A + h^B)) \tag{5.4}$$

$$\text{B's catch \& consumption:} \quad y^B = x^B(h^A, h^B) = h^B(\alpha - \beta(h^A + h^B)) \tag{5.5}$$

The two parameters of the production function are:

- α (Greek *alpha*) is the fisherman's *maximum average productivity*, that is, total catch divided by time spent fishing which would occur if one of them fished some small amount of time and the other did not fish at all. We let $\alpha > 0$; otherwise they could not ever catch any fish.
- β (Greek *beta*) measures the decrease in average productivity for each hour fished in total by the two. We consider the case in which $\beta > 0$ to reflect their interdependence and the negative external effect that each fishing has on the other's catch.

The parameter β expresses three important aspects of the technology:

- *Decreasing average productivity*: If Abdul spends more time fishing, his catch will be larger, but his average productivity—the size of the catch per hour fished—decreases. (You can see this by dividing his output shown in Equation 5.4 by his time fishing h^A to get his average productivity.)
- *Decreasing marginal product of work time*: If Abdul already fishes a lot, then the additional amount of fish that he catches were he to fish a little more will be less than if he were initially fishing a lesser amount. You can see this from Equation 5.6 in M-Note 5.2.
- *Interdependence*: The fact that h^A appears in Bridget's production function and h^B in Abdul's represents the *external effects* and therefore the *interdependence* between the fishermen. The fact that the sign of these terms is *negative* means that the external effect is *negative*.

Abdul's production function in the top panel of Figure 5.5 is increasing but becomes flatter the more time Abdul fishes. The slope of this production function is the marginal product of time fishing, indicating for each level of h^A the increase in the amount of his catch that would result if he increased his fishing time a little. The slope of the production function provides two related pieces of information; the slope is both:

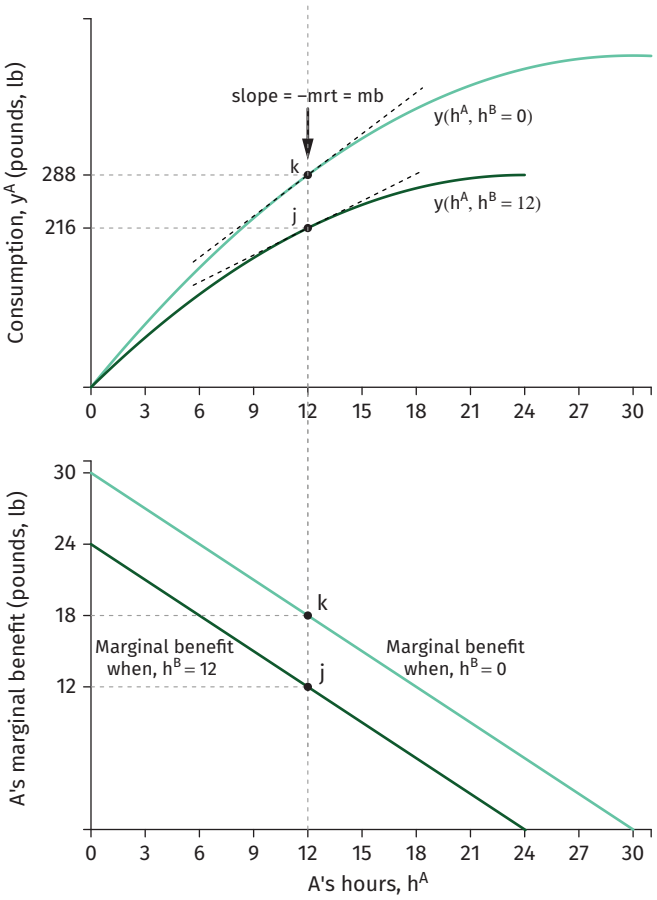
- *the marginal benefit of fishing time* because it indicates how much he benefits if he fishes a little more (how much the larger catch from additional fishing time raises his utility); and
- *the opportunity cost* (in lost fish consumption) *of working less*.

The italicized expressions are different ways of stating *the negative of the marginal rate of transformation*.

In the lower panel of Figure 5.5 we show the marginal product of an hour of fishing based on the production function shown in the top panel, labeled

! REMINDER When x is a bad (like fishing time) rather than a good (like free time) the *mrt* is still the negative of the slope of the feasible frontier, but the opportunity cost is the amount of the y-good (fish caught in this case) that will be sacrificed by fishing *less*.

Figure 5.5 Abdul's production of fish with hours of fishing and marginal benefit of hours spent fishing. In the top panel, Abdul's light-green total product line corresponds to when Bridget does not fish ($h^B = 0$) and Abdul's dark-green total product line corresponds to when Bridget fishes 12 hours ($h^B = 12$). Similarly, in the lower panel, Abdul's light-green marginal benefit line corresponds to when Bridget does not fish ($h^B = 0$) and Abdul's dark-green marginal benefit line corresponds to when Bridget fishes 12 hours ($h^B = 12$). Points **k** and **j** in the upper and lower figure present the same information in different ways. The slope at **k** in the upper panel, for example, is the height of the marginal benefit curve in the lower figure.



the *marginal benefit* of hours of fishing. When Abdul fishes 12 hours a week (and Bridget does not fish), his catch is 288 but when she also fishes 12 hours (the lower green curve) his catch is just 216 lbs. Equally important, when Bridget is not fishing, and Abdul is fishing 12 hours, his marginal product is 18. The fact that the marginal benefit curve shifts downward when Bridget fishes 12 hours reflects the fact that in the top figure for any given amount

M-CHECK We adopt parameters for the production functions so that Bridget and Abdul cannot work so many hours that their average productivity becomes negative, so that fishing more would reduce their total catch. This is why we do *not* extend the lower of the two production function curves in Figure 5.5 beyond 24 hours, the point after which the function turns downward.

of fishing time by Abdul, the total product curve is flatter if Bridget fishes more.

Summarizing, by comparing the two curves in the upper and lower panels of Figure 5.5 we can see the effect of Bridget fishing more on Abdul. Compared to when she does not fish, his production function is:

- *lower*: This is a negative external effect, reducing the utility that Abdul can attain for any level of fishing time that he does; and
- *flatter*: Its slope is also less, so when Bridget fishes more the marginal benefit to Abdul of his fishing more declines. This *reduces* Abdul's incentive to fish.

M-NOTE 5.2 The *mrt* and marginal benefits of fishing time

We begin with Equation 5.4:

$$\text{A's catch \& consumption:} \quad y^A = x^A(h^A, h^B) = h^A(\alpha - \beta(h^A + h^B))$$

This is the production function shown in Figure 5.5 for two different values of h^B. We know that the cost to Abdul is the disutility of fishing time. The benefit is the fish he catches and consumes, so differentiating equation Equation 5.4 with respect to h^A we can find:

$$\text{Marginal productivity (benefit) of fishing time} \quad y^A_{h^A} = \alpha - \beta(h^A + h^B) - \beta h^A \tag{5.6}$$

Equation 5.6 is the slope of the production function in Figure 5.5, and also known as the marginal benefit of fishing time. The mrt is the negative of the slope of the feasible frontier, so:

$$-mrt = y^A_{h^A} = \alpha - \beta(h^A + h^B) - \beta h^A = \text{slope of feasible frontier} \tag{5.7}$$

M-NOTE 5.3 Numerical examples for productivity and external effects

Throughout the chapter, we'll cover a worked example where Abdul and Bridget have the same level of productivity and external effect on each other. As an illustration, we set $\alpha = 30$ and that $\beta = \frac{1}{2}$.

Abdul and Bridget's utility functions therefore become the following:

$$\text{Abdul's utility:} \quad u^A(h^A, h^B) = h^A\left(30 - \frac{1}{2}(h^B + h^A)\right) - \frac{1}{2}(h^A)^2 \tag{5.8}$$

$$\text{Bridget's utility:} \quad u^B(h^A, h^B) = h^B\left(30 - \frac{1}{2}(h^A + h^B)\right) - \frac{1}{2}(h^B)^2 \tag{5.9}$$

In the case where the fishermen fished *alone*, that is the other fishermen had zero hours fishing, Abdul's utility would therefore be: $u^A = 30h^A - \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^A)^2 = 30h^A - (h^A)^2$.

When Abdul and Bridget both spend time fishing, the external effect reduces Abdul's utility; therefore he would have:

$$\begin{aligned} u^A &= 30h^A - \frac{1}{2}h^A h^B - \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^A)^2 \\ &= 30h^A - \frac{1}{2}h^A h^B - (h^A)^2 \end{aligned}$$

CHECKPOINT 5.2 The meaning of β (Greek *beta*) The negative external effect of one's fishing on the other is represented by β. Why does β affect the person who is fishing's own utility even when no one else fishes? Why does it make economic sense?

! REMINDER A player's best-response function gives, for every possible strategy chosen by other players, the strategy that maximizes the player's utility. A strategy profile in which all players are playing a best response, is a Nash equilibrium.

5.4 A BEST RESPONSE: ANOTHER CONSTRAINED OPTIMIZATION PROBLEM

To understand the Nash equilibrium of the interaction between Abdul and Bridget we will need to know how each will best respond to any of the possible levels of fishing chosen by the other. To do this we will derive the *best-response function* of each. We begin, as we did in Chapter 3, with a simpler problem: here we show how one of them, Abdul, will choose how many hours to fish, when Bridget is fishing at some given number of hours.

Choosing a level of fishing time

This problem is set out in Figure 5.6, which combines Abdul's indifference curves from Figure 5.4 with his production function (when $h^B = 0$) from Figure 5.5.

Abdul might first consider fishing six hours, with results indicated by points **f**, **g**, and **h** in the two panels of Figure 5.6. To determine if he should fish six hours he would compare:

- *the marginal cost of working more*: namely the marginal disutility of working time, which is the slope of the indifference curve at **f** shown as point **h** in the lower panel with
- *the marginal benefit of working more*: namely, the marginal benefit of his fishing time, which is the slope of the production function at **f** shown as point **g** in the lower panel.

From the figure we see that at point **f**:

$$24 = \text{slope of feasible frontier} > \text{slope of indifference curve} = 6$$

$$\text{marginal benefit of fishing more} > \text{marginal cost of fishing more}$$

So Abdul would see that he would increase his utility by working more than six hours.

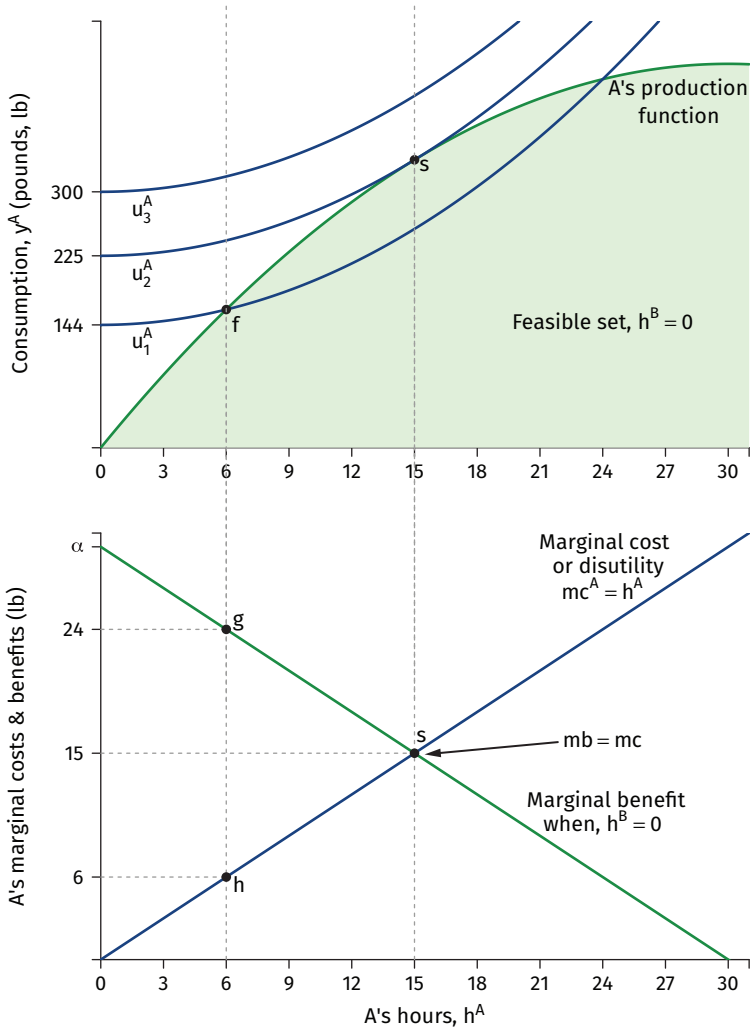
How much more? He will best respond if he follows some simple rules:

- $mb > mc$ If the marginal benefit exceeds the marginal cost as at point **f**, then fish more.
- $mb < mc$ If the marginal cost exceeds the marginal benefit, then fish less.
- $mb = mc$ *If the marginal benefit equals the marginal cost, do not change how much you fish.*

Table 5.1 Three rules: individual constrained optimization, societal Pareto efficiency, and firm cost minimization.

Rules	Tangency of	Rule for what
$mrs = mrt$	An individual's feasible frontier and indifference curve	Individual constrained optimization
$mb = mc$	Restatement of $mrs = mrt$ using marginal costs and benefits	Individual constrained optimization
$mrs^A = mrs^B$	Two or more people's indifference curves	Societal (multi-person) Pareto efficiency

Figure 5.6 Abdul maximizes his utility subject to the constraint of his production function when Bridget does not fish. A's production function is $x^A(h^A, h^B = 0)$ and it defines his feasible consumption, y^A, when B is not fishing at all.



He will therefore use the rule: choose the level of fishing time such that:

$$\text{marginal benefit} = \text{marginal cost}$$

$$\text{slope of feasible frontier} = \text{slope of indifference curve}$$

$$\text{or} \qquad mrt = mrs$$

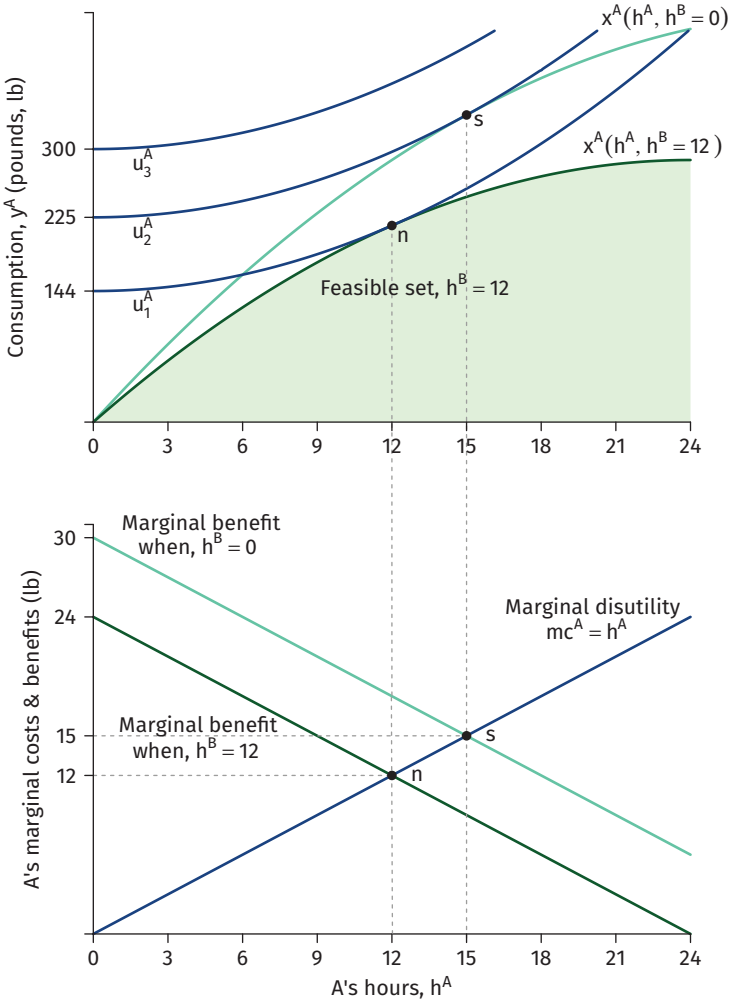
A best-response function: Interdependence recognized

You can confirm from the figure that following the $mrs = mrt$ rule, Abdul will fish 15 hours if Bridget is not fishing, indicated by point **s** in the figure

(**s** for "solo" because Bridget is not fishing). This gives us just one point on his best-response function $h^A(h^B = 0) = 15$ hours.

What about when Bridget *is* fishing, for example, 12 hours? This case is illustrated in Figure 5.7 where the new feasible set constraining Abdul is smaller, because his catch for any amount of time that he spends fishing is *reduced* by Bridget also fishing.

Figure 5.7 Abdul maximizes his utility subject to the constraint of the production function when Bridget spends 12 hours fishing. The feasible set is now smaller because of the negative external effect that her fishing imposes on Abdul. In the top panel, at point **n** his indifference curve labeled u_1^A is tangent to his production function, meaning in the lower panel, that the marginal disutility of fishing time is equal to the marginal productivity of fishing time, or the marginal cost of fishing more is equal to the marginal benefit.



Abdul knows that the level of fishing that will maximize his utility under these new conditions is that which, following the rule, equates:

- the slopes of his production function and the slope of his indifference curve so that the two are tangent in the top panel;
- or, to put it another way, the marginal benefit and the marginal cost of more fishing in the bottom panel.

This gives us a second point on Abdul's best-response function, $h^A(h^B = 12) = 12$. It is no surprise that Abdul fishes less when Bridget fishes more because we already know that Bridget's fishing more reduces the marginal benefit to Abdul's fishing.

What about Abdul's response to Bridget fishing other hours? We do not have to go through the above process, tediously making a separate figure for each level of fishing time she might choose. Instead we can use mathematical expressions for the marginal costs and benefits of fishing to determine Abdul's best response not as a discrete point, but as a continuous function, giving us his fishing time for any level of fishing Bridget might do.

Using the rule that the best response is the number of hours that equates marginal benefits to marginal costs we have a general rule that can be expressed mathematically and which allows us to isolate h^A as a function of h^B and the parameters α and β. A best response is a value of h^A that satisfies the following rule:

$$\text{Slope of production function} = \text{Slope of indifference curve}$$

$$\text{Marginal benefit of fishing more} = \text{Marginal cost of fishing more}$$

Using Equations 5.3 and 5.7:

$$\alpha - \beta(2h^A + h^B) = h^A \tag{5.10}$$

Rearranging Equation 5.10 to isolate h^A and to express his utility-maximizing fishing hours as a function of Bridget's hours $h^A(h^B)$, we have:

$$\text{Abdul's best-response function:} \qquad h^A(h^B) = \frac{\alpha - \beta h^B}{1 + 2\beta} \tag{5.11}$$

How does Abdul's fishing time h^A change when the variable (h^B) and the parameters (α and β) change?

- *Change in Bridget's fishing time* (h^B): If Bridget *decreases* her fishing time, Abdul's marginal benefit curve shifts *up*, and Abdul's best response is to *increase* his fishing time to balance his marginal cost with the higher marginal benefit. Abdul's best-response function does not shift, he chooses a different level of fishing due to the change in Bridget's fishing time.
- *Change in maximum productivity* (α): If Abdul's basic productivity *increases*, and nothing else changes, this shifts his marginal benefit curve *up* and *independently of any change in Bridget's fishing time*, he

will *increase his fishing time* to balance his marginal cost with the higher marginal benefit. This is a *shift* in Abdul's best-response function itself, not just a movement from one point on it to another as in the bullet above.

- *Change in the external effect* (β): If the external effect *increases*, Abdul's marginal benefit curve pivots *downward* with a corresponding decrease in fishing time (β changes the *slope* of his marginal benefit curve). Like the increase in α, in this case Abdul changes his fishing time due to a *shift* in this best-response function.

The best-response function for Bridget can be derived in the same way we derived Abdul's. Therefore her best-response function (BRF) is:

$$\text{Bridget's BRF:} \qquad h^B(h^A) = \frac{\alpha - \beta h^A}{1 + 2\beta} \tag{5.12}$$

M-NOTE 5.4 Marginal benefits, marginal costs, and finding the best responses

In M-Note 5.3, we used the example of $\alpha = 30$ and $\beta = \frac{1}{2}$ to provide utility functions for Abdul and Bridget, as represented in Equations 5.8 and 5.9. We now use those parameters to identify the first-order condition for Abdul's utility maximization where his marginal benefits equal his marginal costs and therefore to provide a best-response function.

$$u^A(h^A, h^B) = h^A\left(30 - \frac{1}{2}(h^B + h^A)\right) - \frac{1}{2}(h^A)^2$$

$$u^A_{h^A} = \frac{\partial u^A}{\partial h^A} = \underbrace{\left(30 - \frac{1}{2}h^B - h^A\right)}_{\text{Marginal benefit}} - \underbrace{h^A}_{\text{Marginal cost}} = 0$$

We can isolate Abdul's hours of work, h^A, to find his best response to Bridget's hours of work:

$$\text{Abdul's BRF:} \qquad h^A(h^B) = \frac{30 - \frac{1}{2}h^B}{2} = 15 - \frac{1}{4}h^B \tag{5.13}$$

$$\text{Bridget's BRF:} \qquad h^B(h^A) = \frac{30 - \frac{1}{2}h^A}{2} = 15 - \frac{1}{4}h^A \tag{5.14}$$

Each of them therefore has a best-response function that is a function of the other person's time spent fishing: $h^A(h^B)$ for Abdul and $h^B(h^A)$ for Bridget.

M-NOTE 5.5 Mathematics of the best-response function

To understand each player's response to the other, it is useful to understand their marginal utilities of hours of fishing. We do this for Abdul, in the understanding that Bridget will have symmetrical results. We will therefore find $u^A_{h^A}$, Abdul's marginal utility of his own hours of fishing, $u^A_{h^B}$, marginal utility to Abdul of Bridget's hours of fishing, and $h^A(h^B)$, Abdul's best response to Bridget's choice of hours. We start with Abdul's utility function:

continued

$$u^A(h^A, h^B) = h^A(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^A)^2$$

We can differentiate Abdul's utility function with respect to his own hours (h^A) to find his marginal utility of his own hours of work. We also differentiate his utility function with respect to Bridget's hours of work to find how his utility changes when Bridget changes her hours (h^B).

Let us first find A's marginal utility of h^A:

$$\begin{aligned} u^A_{h^A} = \frac{\partial u^A}{\partial h^A} &= \alpha - \beta h^B - 2\beta h^A - h^A \\ &= \alpha - \beta h^B - h^A(1 + 2\beta) \end{aligned} \tag{5.15}$$

And now, the marginal effect on A's utility of B's hours (h^B):

$$u^A_{h^B} = \frac{\partial u^A}{\partial h^B} = -\beta h^A \tag{5.16}$$

Using Equation 5.15, if we set Abdul's marginal utility $u^A_{h^A} = \frac{\partial u^A}{\partial h^A} = 0$, then we can find his best response to Bridget's hours of work:

$$u^A_{h^A} = \frac{\partial u^A}{\partial h^A} = \alpha - \beta h^B - h^A(1 + 2\beta) = 0$$

Isolate h^A term $\quad h^A(1 + 2\beta) = \alpha - \beta h^B$

A's BRF: $\quad h^A = \dfrac{\alpha - \beta h^B}{(1 + 2\beta)}$

Which is what we found from setting marginal benefit equal to marginal cost to find Abdul's best-response function in Equation 5.11.

CHECKPOINT 5.3 How the BRFs change Make a figure similar to Figure 5.8 but using Abdul's and Bridget's best-response functions with the values of $\alpha = 24$ and $\beta = 1$. How and why do they differ?

5.5 HOW WILL THE GAME BE PLAYED? A SYMMETRIC NASH EQUILIBRIUM

! REMINDER We began our analysis of Ayanda and Biko trading data and coffee in a similar way, with the two being symmetrical traders with neither of them having any particular advantage in the bargaining process.

We do not have enough information to answer the question in the section title. To do this we need answers to other questions. Is one of them powerful enough determine the allocation unilaterally, stating: I fish 15 hours, and you are excluded from fishing? Is there a government that can place a tax on fishing to discourage overharvesting the stock? Can Abdul and Bridget agree to fish less? If they did, can they count on their agreement being enforced? In other words, we need to know more about the *rules of the game*.

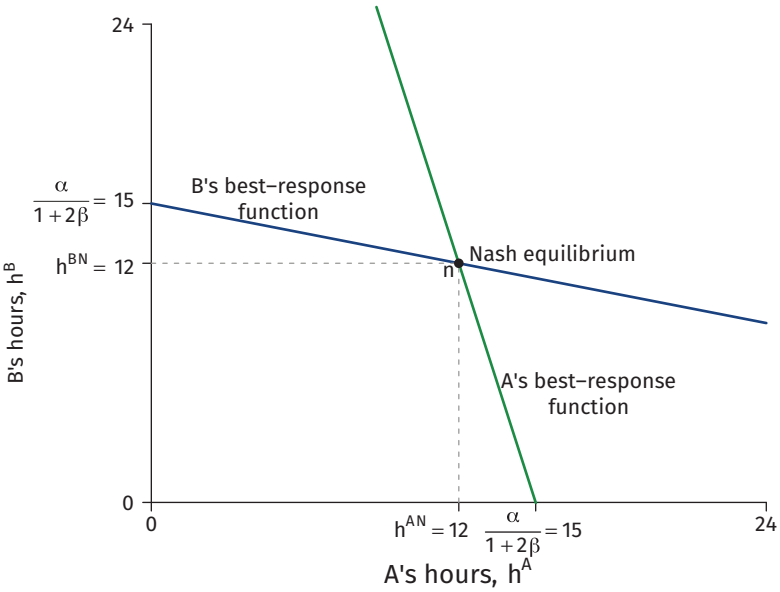
One possibility is that the two act non-cooperatively (they do not make agreements with each other) and neither has any particular advantage in their interaction. So they simply try to do the best that they can, given what the other is doing and given the information they have. We will investigate other rules of the game later.

Figure 5.8 Nash equilibrium: mutual best responses for Bridget and Abdul. The equations for the best-response functions are:

$$h^B(h^A) = \frac{\alpha - \beta h^A}{1+2\beta}$$

$$h^A(h^B) = \frac{\alpha - \beta h^B}{1+2\beta}$$

If $\alpha = 30$ and $\beta = 0.5$, the parameters we used in the previous figures, then we can see that when Bridget does not fish (the intercept of Abdul's best-response function with the horizontal axis) he fishes 15 hours. The point at which their best-response functions intersect is the Nash equilibrium of the interaction. Using these same parameters, we can see that the Nash equilibrium given by Equation 5.17 is that they both fish 12 hours.



A stationary allocation among symmetric players

To study this case, we graph the two best-response functions in Figure 5.8. This gives us all the information we need to determine the Nash equilibrium of their interaction.

A Nash equilibrium is a mutual best response, so Abdul's choice of fishing hours must be a best response to Bridget's choice of fishing hours, which must in turn be a best response to Abdul's choice of fishing hours. This sounds complicated but with a little help from the mathematics we have already done, it is not: a Nash equilibrium is a point that is on both of the players' best-response functions.

We label the point **n** and define the hours that they work at the Nash equilibrium as (h^{AN}, h^{BN}), where point **n** is the Nash equilibrium in the

figure and the superscript N indicates each player's Nash equilibrium hours. A Nash equilibrium is a pair of fishing times (h^{AN}, h^{BN}) that satisfy each fisherman's best-response function.

We show in M-Note 5.6 how to find the Nash equilibrium hours of fishing for each person. At the Nash equilibrium, the two fishermen will spend the same amount of time fishing.

$$h^{AN} = h^{BN} = \frac{\alpha}{1+3\beta} \tag{5.17}$$

Equation 5.17 shows that each fisherman's hours spent fishing is defined by the parameters α and β, capturing the effects on their best response of their maximum productivity, their decreasing marginal productivity, and the negative external effect each has on the other.

The Nash equilibrium fishing hours, h^{AN} and h^{BN}, are equal because Abdul and Bridget have identical utility functions and production functions (other than reversing the superscripts), and they are determined by the parameters α and β. The greater is the maximum productivity, α, the greater will be their equilibrium hours of fishing. The larger is the negative effect each has on their own productivity and on the other person's productivity, β, the lower their equilibrium hours will be.

M-NOTE 5.6 **Finding Nash equilibrium fishing time**

By definition of the Nash equilibrium, Abdul's Nash equilibrium fishing time must be a best response to Bridget's Nash equilibrium fishing time, and Bridget's Nash equilibrium fishing time must be a best response to Abdul's Nash equilibrium fishing time. A Nash equilibrium is therefore a pair of fishing times (h^{AN}, h^{BN}) that satisfy the following equations:

$$h^{AN} = h^{A}(h^{BN}) = \frac{(\alpha - \beta h^{BN})}{(1+2\beta)} \tag{5.18}$$

$$h^{BN} = h^{B}(h^{AN}) = \frac{(\alpha - \beta h^{AN})}{(1+2\beta)} \tag{5.19}$$

Equations 5.18 and 5.19 are two linear equations in two unknowns. We can solve the equations for the unknowns, which are the fishing times at the Nash equilibrium.

There is a particularly simple way to do this in our case because:

1. the two fishermen have identical utility functions (they are mirror images of each other); so
2. we know that it must be that $h^{AN} = h^{BN}$; and
3. we can therefore set the Nash equilibrium level of fishing of the one equal to the best-response function of the other.

So substituting $h^{B} = h^{A}$, into Abdul's best-response function is:

$$h^{A}(h^{B}) = \frac{\alpha - \beta h^{A}}{1+2\beta}$$

continued

Multiplying out and isolating h^A :

$$\begin{aligned} h^A + 2\beta h^A &= \alpha - \beta h^A \\ h^A + 3\beta h^A &= \alpha \\ h^A(1+3\beta) &= \alpha \\ h^{AN} &= \frac{\alpha}{1+3\beta} = h^{BN} \end{aligned}$$

CHECKPOINT 5.4 **Storms and sustainability** Imagine that the external effect increased, as it would, for example, if greater climate volatility produced storms that caused the two fishermen to fish in the same limited part of the lake.

a. Use the equation for the best responses of the two to redraw the figure. Why do the fishermen best respond by fishing less?

b. Use the equation for h^{AN} and h^{BN} to show that the Nash equilibrium level of fishing will decline.

c. Use what you have learned to explain how the best-response functions and the Nash equilibrium would change if the fishermen jointly adopt a strategy to let go of young fish to make the fish population more sustainable and reduce the external effect they have on the other fisherman.

5.6 DYNAMICS: GETTING TO THE NASH EQUILIBRIUM

When we used the equation for the Nash equilibrium level of hours of fishing (Equation 5.17) to say what the effect of a change in α or β would be, we used what is called **comparative static** analysis.

When using comparative statics we compare the status quo Nash equilibrium *before* the change with the Nash equilibrium *after* the change.

! REMINDER When we say "other things equal" we are using the *ceteris paribus* assumption which allows us to compare what happens when *one* variable of interest changes.

- The word *static* refers to the Nash equilibrium because at a Nash equilibrium there are no reasons for the actors to change what they are doing.
- The process is *comparative* because we compare two or more equilibrium states before and after a change.

We did the following:

- We started with Abdul and Bridget at the Nash equilibrium, each fishing 12 hours.

COMPARATIVE STATICS A method which analyzes the process of change by comparing the status quo Nash equilibrium with a new equilibrium after some change in the underlying data of the problem.

- We then assumed that other things (like the weather) that might affect their fishing time are held constant (this is the *ceteris paribus* assumption or 'other things equal').
- Then we compared the two Nash equilibria, one before the change in α or β and the other after the change.
- We assumed that after the change Abdul and Bridget would be at the new Nash equilibrium, working some different number of hours.
- Finally we considered the difference in work hours between the two Nash equilibria to be the effect of the change on work hours.

EXAMPLE Static analysis is like inferring movement by looking at two photos taken at some time interval; dynamic analysis is like watching a film of the same process.

This type of analysis is called *static* because it compares two *static* (unchanging) situations *without* looking at *how* the change takes place, that is, how they get to the new Nash equilibrium. This is an essential method of economic analysis, simplifying the matter by a shortcut. The shortcut is that we *did not* explore who did what to implement the move. In contrast, an analysis that includes the *process of change* rather than focusing exclusively on equilibria is called dynamic (the term **dynamics** refers to change, it is the opposite of static).

We did not even explain why Abdul and Bridget would have been at the original Nash equilibrium in the first place. Fortunately, the way we have derived our best responses provides a method to fill in the necessary dynamic analysis.

Remember, when Abdul was selecting a best response he adopted a simple checklist based on the marginal *benefits* of fishing more (mb) and the marginal *costs* of fishing more (mc): *if* $mb > mc$, then *fish more*; *if* $mb < mc$, then *fish less*; if $mb = mc$ don't change how much you are fishing.

M-CHECK Abdul might adopt the instruction: in any period, close half of the difference between the hours I am now working and the hours indicated by my best-response function, given how many hours Bridget is now working. For example, if Abdul were fishing six hours while Bridget fished 12 hours, he would increase his hours by $(12 - 6)/2 = 3$ hours.

When we introduced this checklist we focused on the last line, because that is the equality that determines the utility-maximizing level of fishing for Abdul; that is, it is a point on his best-response function.

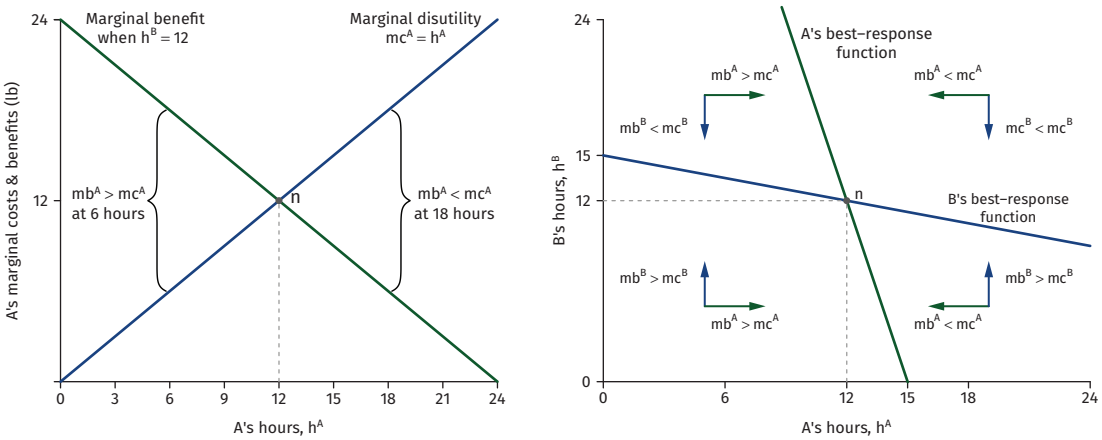
The first two "if" statements tell Abdul what to *change* when they are not at an equilibrium because he is *not* fishing the optimal utility-maximizing amount given what Bridget is doing, that is, when he is "off" his best-response function.

As Figure 5.9 shows, these first two lines of Abdul's checklist tell us that starting at any allocation (that is any combination of fishing hours of each of them) in which direction he should move, shown by the arrows (called vectors). The dynamic analysis gives the following simple instruction: if you are not on your best-response function, move toward it.

Abdul's arrows are green and horizontal (when he changes his fishing hours he moves left or right). The same reasoning allows us to show the dynamic arrows for Bridget, they are blue and horizontal, because when she changes her hours that moves the allocation point up or down.

DYNAMICS Refers to the study or process of change.

Figure 5.9 How players can get to the Nash equilibrium: a dynamic analysis. Panel (a) shows the marginal costs and benefits of Abdul's fishing if Bridget fishes 12 hours. Panel (b) shows the dynamics of the choices in terms of the fishermen's marginal benefits and marginal costs. The horizontal arrows show the direction Abdul will move if he is initially at the base of the arrow. The vertical arrows show the same for Bridget. The inequalities involving marginal benefits and costs (*mb*, *mc*) are the reason for the movement shown in the arrows (which are called "vectors").



(a) Marginal benefits and costs when B fishes 12 hours

(b) Dynamic analysis of A and B's choices

For example, in Figure 5.9 (a), if Abdul is fishing six hours the marginal benefits of fishing more exceed the costs (the bracketed term on the left). So in Figure 5.9 (b), the horizontal green arrows show that he will fish more. Similar reasoning (in reverse) applies to the case where he is fishing, for example, 18 hours. The extent by which the benefits differ from the costs depends on how much fishing Bridget is doing. Figure (a) shows the case for when she is fishing 12 hours. You can also work out how Bridget will adjust her hours if she is fishing more or less than the amount indicated by her best-response function.

You can see from the figure that unless the allocation is at point **n** one or both of them will have an incentive to move (horizontally for Abdul, vertically for Bridget) in ways that will lead them toward the Nash equilibrium.

This explains why, if the Nash equilibrium shifted because of some change in either α or β, we would *expect* the two to alter their fishing hours to move toward the *new* Nash equilibrium. It also explains, if enough time had passed since the change to which they are responding, then we would expect both Bridget and Abdul to be at (or very close to) the Nash equilibrium. We now introduce a way that we can evaluate all of the possible equilibria of this game by the standards of Pareto efficiency and the resulting level of inequality.

M-CHECK They could have moved away from the equilibrium rather than towards it, had their best-response functions been different. Whether and how actors may get to an equilibrium depends on their best-response functions, and the details of how they act when they are not at an equilibrium.

M-NOTE 5.7 Numerical Nash equilibrium

In M-Note 5.4, we found the best responses for Abdul and Bridget given by Equations 5.13 and 5.14. Using the method we outlined above and the same parameter values ($\alpha = 30$ and $\beta = \frac{1}{2}$), we set Abdul's Nash equilibrium hours of fishing (h^{AN}) equal to Bridget's best-response function to find the Nash equilibrium level of fishing time:

$$\begin{aligned} \text{Abdul's Nash equilibrium hours}: \quad & h^{AN} = 15 - \frac{1}{4}h^{AN} = \text{Bridget's BRF} \\ \text{Collect terms} \quad & h^{AN} + \frac{1}{4}h^{AN} = 15 \\ & \left(\frac{5}{4}\right)h^{AN} = 15 \\ \text{Multiply by } \frac{4}{5} \quad & h^{AN} = \left(\frac{4}{5}\right)15 = 12 = h^{BN} \end{aligned} \tag{5.20}$$

As a result, we see that each will fish 12 hours at the Nash equilibrium. Therefore they each obtain the following Nash equilibrium utility (by substituting h^{AN} and h^{BN} into their utility functions):

$$\begin{aligned} u^{AN}(h^{AN}, h^{BN}) &= h^{AN}\left(30 - \frac{1}{2}(h^{BN} + h^{AN})\right) - \frac{1}{2}(h^{AN})^2 \\ &= 12\left(30 - \frac{1}{2}(12 + 12)\right) - \frac{1}{2}(12)^2 \\ &= 216 - 72 = 144 = u^{BN} \end{aligned}$$

Each of them has a utility of 144 at the Nash equilibrium and the total welfare (sum of utilities) is $W^N = u^{AN} + u^{BN} = 288$.

CHECKPOINT 5.5 Dynamics In Figure 5.9 imagine the status quo is at the right-angle roots of the two arrows in the lower right of Figure 5.9 panel (b). How would each of the fishermen change their hours of fishing? Do the same for the other three right angle roots of the arrows.

5.7 EVALUATING OUTCOMES: PARTICIPATION CONSTRAINTS, PARETO IMPROVEMENTS, AND PARETO EFFICIENCY

Because the symmetrical interaction is just one of many possible rules of the game that Bridget and Abdul might engage in, we need to go beyond the Nash equilibrium for that game and find a way to evaluate all of the possible allocations that they might experience.

To do this, as in Chapter 4, we use the indifference maps of the two players superimposed on the same set of outcomes. Recall that in the previous chapter every point in the Edgeworth box indicated an allocation composed of a bundle of goods for Ayanda and another bundle of goods for Biko. We will see that the same is true in this case if we plot an allocation as the pair of fishing hours of the two, (h^A, h^B). We start with Abdul's preferences.

Because the utility of each depends on their own fishing time and the fishing time of the other, that is, because

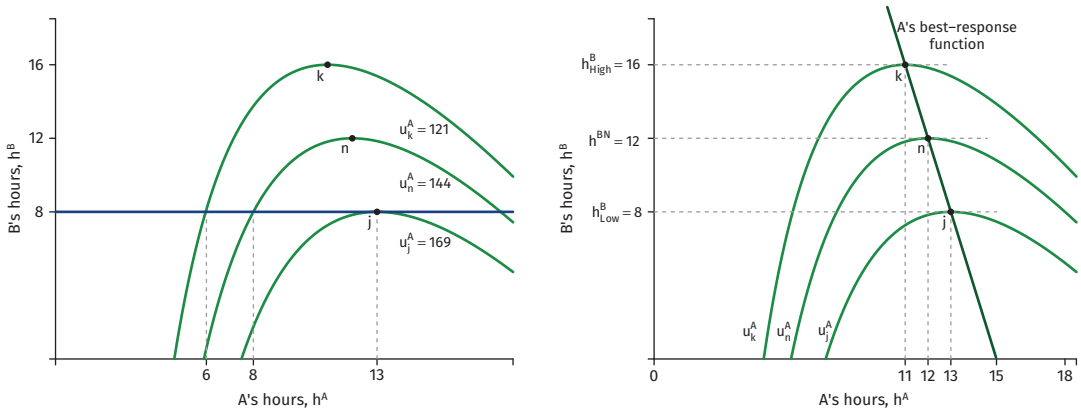
$$\text{Abdul's utility:} \quad u^A(h^A, h^B) \tag{5.21}$$

$$\text{Bridget's utility:} \quad u^B(h^A, h^B) \tag{5.22}$$

we can plot indifference curves with fishing time on each axis: Bridget's fishing time (h^B) on the vertical axis and Abdul's fishing time (h^A) on the horizontal axis.

We do this in panel (a) of Figure 5.10, where every point in the figure is a particular allocation of fishing times (h^A, h^B). Using these allocations we can calculate the utility that Abdul would experience were that allocation to occur. On this basis we can calculate Abdul's indifference curves based on his hours of fishing and Bridget's hours of fishing. Abdul prefers curves labeled with higher numbers, $u_j^A > u_n^A > u_k^A$ (169 > 144 > 121). Notice two things about the indifference curves:

Figure 5.10 A new look at Abdul's constrained optimization problem for selecting his fishing time depending on Bridget's fishing time. To illustrate the construction of Abdul's best-response function, in panel (a) we consider Abdul's decision about how many hours to work, given that Bridget has (hypothetically) decided to work eight hours. The horizontal blue line is the constraint on Abdul's utility-maximizing process. In panel (b), we consider three hypothetical levels of Bridget's fishing time. The horizontal lines represent Bridget's fishing time at each of these levels, and are the constraint on Abdul's maximization process. One of these horizontal lines is tangent to each of Abdul's indifference curves $h^B = 8$ tangent to u_j^A at point **j**, $h^B = 16$ tangent to u_k^A at point **k**, and $h^{BN} = 12$ tangent to u_n^A at point **n**. Abdul's entire best-response function is made of points like **j**, **n**, and **k**, for each of Bridget's possible levels of fishing hours.



(a) Abdul making a choice constrained by Bridget's hours

(b) Abdul's best-response function

- *The vertical dimension, or the effect of Bridget fishing more.* Abdul's preferred indifference curves are lower. This is because the less Bridget fishes the better it is for Abdul.
- *The horizontal dimension, or the effect of Abdul fishing more.* If Bridget is fishing at the "low" level indicated in the figure and supposing that Abdul initially does not fish at all but considers fishing a little, he will start by finding himself at successively higher indifference curves as he fishes more, crossing the indifference curves labeled u_k^A, and then u_n^A and up to u_j^A. But if he spends too much time fishing he will then cross from u_j^A back down to indifference curve u_n^A and again go back down to u_k^A.

Another perspective on a best-response function

We can use the horizontal dimension of the figure to identify a point on Abdul's best-response function, associated with Bridget hypothetically fishing just eight hours. We take this thought experiment as a constraint on Abdul's utility maximization. Remember Abdul prefers indifference curves that are lower down (indicating Bridget fishing less). The most preferred indifference curve that is feasible is the one tangent to the constraint, at point **j**, which is therefore a point in Abdul's best-response function.

It may help to think of his indifference map as showing the contours of the shoulder of a hill, and Abdul as walking along a horizontal line at eight hours toward point **j**, trying out different amounts of time he might devote to fishing. This is exactly what he did in Figure 5.6, comparing the marginal benefit and marginal cost of fishing more. At first he is climbing–crossing contours indicating ever-higher altitudes–higher utility. When he fishes six hours, he achieves utility $u_k^A = 121$, proceeding on to fish eight hours, he achieves $u_n^A = 144$, and finally fishing 13 hours, he achieves $u_j^A = 169$. At point **j** his path levels off and if he continues to increase his fishing time he will descend to lower altitudes–lower utility–once more.

Figure 5.10 (b) illustrates how two additional points on Abdul's best-response-function are derived. The best-response function is constructed by considering all of the possible levels of fishing that Bridget could hypothetically do, and then reason as we did for point **j**.

Notice that Abdul's best-response function intersects the indifference curves where the indifference curves are flat. If the indifference curve is flat, then the *mrs* must be zero. In M-Note 5.8 we show why this must be true. In the right panel you can see that as Bridget's fishing time increases from eight to 16 hours, Abdul's fishing time declines from 13 to 11 hours. He identifies his best-response hours of fishing by finding the point on his best-response function that corresponds to the number of hours Bridget fishes.

M-NOTE 5.8 Why is the best-response function made up of points where the indifference curves are flat?

Abdul's utility function defined over his and Bridget's hours of fishing is $u^A(h^A, h^B)$. To find the slope of his indifference curves in (h^A, h^B) space, we proceed as we did in M-Note 3.2. We totally differentiate his utility function and set the result equal to zero:

$$du^A = u^A_{h^A}\, dh^A + u^A_{h^B}\, dh^B = 0$$

$$\frac{dh^B}{dh^A} = \frac{-u^A_{h^A}}{u^A_{h^B}}$$

which, using Equations 5.15 and 5.16, becomes:

$$\frac{-u^A_{h^A}}{u^A_{h^B}} = \frac{-(\alpha - \beta h^B - (1+2\beta)h^A)}{-\beta h^A} \tag{5.23}$$

$$\text{or, multiplying by } \frac{-1}{-1} = \frac{\alpha - \beta h^B - (1+2\beta)h^A}{\beta h^A} \tag{5.24}$$

The numerator of this expression is the marginal effect on Abdul's utility of fishing more, which for modest amounts of fishing time is positive. This is why the indifference curve slopes upwards when he is fishing six or eight hours in Figure 5.10.

Abdul's best-response function gives the values of h^A and h^B for which the derivative of Abdul's utility with respect to his fishing time is equal to zero or:

$$u^A_{h^A} = \alpha - \beta h^B - (1+2\beta)h^A = 0$$

If $u^A_{h^A} = 0$, then the numerator of Equations 5.25 and 5.23 is zero, so the slope of the indifference curve is equal to zero, which means that it is flat.

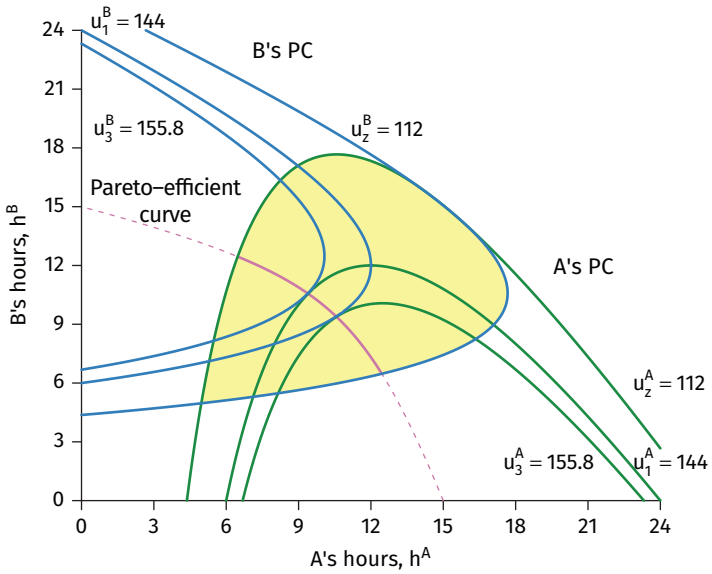
CHECKPOINT 5.6 The marginal rate of substitution In Figure 5.10, suppose B is fishing 8 hours. Explain why, starting from a small number of A's fishing hours, as A increases his hours, his utility first increases and then decreases. Why is his indifference curve horizontal at point **j**?

Fallback positions and the Pareto-improving lens

We have said that rules of the game other than symmetrical interaction will lead to different Nash equilibrium allocations. As long as the interaction between the two is voluntary—there are no "offers you cannot refuse"—we can limit the possible outcomes by thinking about the alternatives that the two would have, should they decide *not* to fish at all. Any allocation in which they both fish and that makes either of them (or both) worse off than how they would do if they did not fish at all will not occur for the simple reason that they will not fish if they could do better by not.

Recall that if they do not fish at all, they both have a fallback option yielding them a utility $u_z = y_z$. This is their fallback position (like the allocation **z** in the Edgeworth box of the previous chapter). But they only receive their fallback if they *do not* fish, so the opportunity cost of fishing—what they cannot have if they fish—is y_z. In Figure 5.11 we show both Abdul's and

Figure 5.11 Abdul's indifference curves and Bridget's indifference curves showing their fallback levels of utility (their participation constraints), u_z^A and u_z^B. At their fallback positions, they are not fishing at all and have their fallback utility $u_z = u_2 = y_z = 112$.



Bridget's indifference maps. We see from the numbering of the utility labels on the curves that Bridget's indifference curves give greater values the closer they are to the vertical axis (as Abdul's did with the horizontal axis).

For each of them, one of their indifference curves is particularly important: it is labeled u_z^A and u_z^B. These two curves show all of the allocations (h^A, h^B) that yield, for Abdul and Bridget respectively a level of utility equal to the utility of their fallback position namely $u_z = y_z$. This is the *participation constraint* for each of them: they will not participate in fishing unless they can do at least this well. Any point between these indifference curves is a Pareto improvement over their fallback position: in the Pareto-improving yellow-shaded lens both are better off than their fallback option.

The Pareto-efficient curve

There is another important curve in Figure 5.11: the purple solid and dashed Pareto-efficient curve. We know that Pareto efficiency requires that the fishermen's indifference curves be *tangent*, that is, for their marginal rates of substitution to equal. You can see two of these tangencies in the interior of the Pareto-improving lens. The other tangencies defining the Pareto-efficient curve are not shown. The Pareto-efficient curve is made up of all points representing allocations that satisfy the $mrs^A = mrs^B$ rule:

$$mrs^A = \frac{u^A_{h^A}}{u^A_{h^B}} = \frac{u^B_{h^A}}{u^B_{h^B}} = mrs^B \tag{5.25}$$

The figure clarifies the difference between *Pareto improvements* and *Pareto efficiency*:

- The points on the purple Pareto-efficient curve that are indicated by a dashed line outside the yellow Pareto-improving lens are *Pareto efficient but not Pareto improvements* over the fallback *no fishing* option.
- The points in the yellow Pareto-improving lens that are *not* on the purple Pareto-efficient curve are *Pareto improvements over the no-fishing option but not Pareto efficient*.

CHECKPOINT 5.7 Understanding the parameters

a. Draw the lens of Pareto improvements over the no-fishing option, if the fallback option of both improved to $u_z = 144$.

b. Do the same if Abdul's fallback option improved (to 144) but Bridget's remained unchanged.

5.8 A PARETO-INEFFICIENT NASH EQUILIBRIUM

We return now to the symmetrical interaction between Bridget and Abdul, in which the Nash equilibrium is the allocation at the intersection of their best-response functions. And we ask: Is that allocation Pareto efficient?

To answer, we combine two figures we have already introduced: Figure 5.11 showing the two fishermen's indifference curves and Figure 5.8 showing their best-response functions. The combination of these figures results in Figure 5.12.

Figure 5.12 shows that the Nash Equilibrium is not Pareto efficient: at the Nash allocation (point **n**) the indifference curves of the two *intersect* rather than being *tangent*. So allocation **n** *cannot* be Pareto efficient.

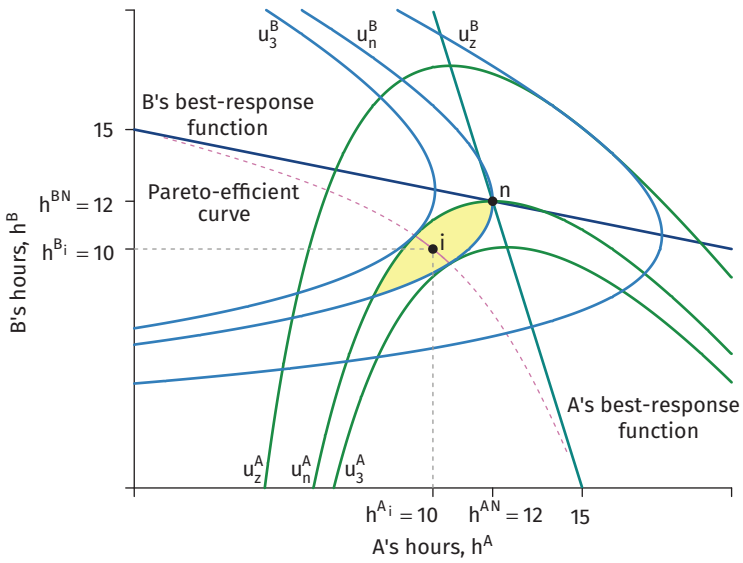
How do we know that their indifference curves cannot be tangent at that point that is, how do we know that

$$mrs^A = \frac{u^A_{h^A}}{u^A_{h^B}} \neq \frac{u^B_{h^A}}{u^B_{h^B}} = mrs^B \tag{5.26}$$

The answer is that the Nash equilibrium is a point on both best-response functions, defined by $u^B_{h^B} = 0$ for Bridget's best response and $u^A_{h^A} = 0$ for Abdul's best response. At their best responses each fisherman adjusts their own fishing time to maximize utility so that these two terms will be zero. If we substitute the zeroes for the marginal utilities in Equation 5.25, we find the following:

! REMINDER To understand Figure 5.12 it will help to remember that for Abdul *down* is better (his indifference curves have higher utility the lower they are) because the less Bridget fishes the better it is for him. Similarly, Bridget is better off on the indifference curves further to the left.

Figure 5.12 The Nash equilibrium and the Pareto-improving lens. The Pareto-improving fishing times (in which both fish less) are in the pale-yellow lens. Notice that Abdul's indifference curve at the Nash equilibrium is flat, and Bridget's at the same point is vertical (their marginal rates of substitution are not equal). This being the case there must be a Pareto-improving lens and the Nash equilibrium cannot be Pareto efficient.



- the first expression is now zero divided by $u^A_{h^B}$ so the slope of Abdul's indifference curve is zero; it is flat (as in Figure 5.12, and as we show in M-Note 5.8); and
- the second expression is now $u^B_{h^A}$ divided by zero, so the slope of Bridget's indifference curve is infinite; it is vertical (as in Figure 5.12 too).

A flat line cannot be tangent to a vertical line, so the condition for Pareto efficiency is violated and the Nash equilibrium is not Pareto efficient.

A view from a Pareto-inefficient status quo Nash equilibrium

We now imagine Abdul and Bridget, fishing 12 hours each as indicated by the Pareto-inefficient Nash equilibrium. They realize they could both do better. And they consider the options. Each fisherman might propose some different allocation. To agree on an alternative level of fishing, the proposal would have to implement a Pareto improvement. The Pareto improvement would need to be over the Nash equilibrium, *not* over their no-fishing fallback option. Remember that the Nash equilibrium is already better than their fallback positions.

With allocation **n** the new fallback for the agreement, we now have a new yellow-shaded Pareto-improving lens. There are two things to notice about Pareto improvements over the Nash allocation:

- both fishermen spend *less* time fishing and both are *better off* (have higher utility than at the Nash equilibrium); and
- the new Pareto-improving lens is much smaller than the lens of Pareto improvements over the no-fishing fallback option.

The reason why there exist allocations that are Pareto improvements over the Nash is as follows.

- *Reason* 1: each of them would benefit a lot if the other were to fish less; and
- *Reason* 2: at the Nash equilibrium each of them would experience very little lost utility by themselves fishing a little less.

Reason 1 concerns each fishermen's marginal utility with respect to the other's hours of fishing, which is negative in both cases because each fisherman's fishing time reduces the other's productivity.

Concerning Reason 2, suppose that, at the Nash equilibrium level of the fishing times, Bridget decided she would try to *bribe* Abdul to fish less. How much would she have to give him to fish a tiny bit less? The answer is "almost nothing" because at the Nash equilibrium, changes in his fishing time have no effect on his utility. The reason is that the marginal benefits of fishing a little more equal the marginal costs of fishing a little more (that is how he chose that level of fishing to do).

So Abdul's fishing a little less would not matter much to Abdul but it would definitely benefit Bridget. A similar result is true for Bridget: Abdul could bribe her to fish a little less for a tiny portion of his fish. This being the case if they both could agree to fish less (and just forget about the bribes) they would both be better off.

The conclusion is that Bridget and Abdul need not lament their sorry condition at the Nash equilibrium. If a deal can be enforced—an agreement to limit fishing, maybe along with a bribe—there's a deal to be made that benefits them both.

We turn now to considering changes in the rules of the game that might reduce fishing times, keeping in mind that we are thinking about not just two people, but an entire community of people—perhaps the entire world's population if we are considering coordination problems such as climate change or the spread of epidemic diseases.

M-NOTE 5.9 The Nash equilibrium cannot be Pareto efficient

To show that the Nash equilibrium is not Pareto efficient we ask: if they could agree each to fish an arbitrarily small amount less, then would they both be better off? If the answer is "yes," then the Nash equilibrium cannot be Pareto efficient. We know that $u^{A}_{h^{B}} < 0$ and $u^{B}_{h^{A}} < 0$, so each would be better off if the other fished less. We also know that $u^{A}_{h^{A}} = 0$ and $u^{B}_{h^{B}} = 0$ because these

continued

equalities define Bridget's and Abdul's best-response functions, and the Nash equilibrium they are trying to improve on is a pair of strategies each of which is a best response to the other.

So for any change dh^A and dh^B, representing an agreement to change their fishing time, we can evaluate the change in each utility associated with change in the fishing times of each.

$$du^A = u^A_{h^A} dh^A + u^A_{h^B} dh^B$$
$$du^B = u^B_{h^A} dh^A + u^B_{h^B} dh^B$$

Eliminating the terms equal to zero in the expressions above, namely those involving $u^A_{h^A}$ and $u^B_{h^B}$ we have:

$$du^A = u^A_{h^B} dh^B < 0$$
$$du^B = u^B_{h^A} dh^A < 0$$

or, rearranging

$$\frac{du^A}{dh^B} = u^A_{h^B} < 0$$
$$\frac{du^B}{dh^A} = u^B_{h^A} < 0$$

Both expressions are negative: the utility of each would be enhanced by an agreement to fish a little less. The Nash equilibrium allocation of fishing times is not Pareto efficient.

CHECKPOINT 5.8 **Pareto efficiency** Explain why the Nash equilibrium level of fishing hours is not Pareto efficient.

! REMINDER In Chapter 4, we gave the Impartial Spectator Cobb-Douglas preferences over the two players' utilities. Here we have the Impartial Spectator sum the utilities of the two fishermen, as the Impartial Spectator did in Chapter 3.

5.9 A BENCHMARK SOCIALLY OPTIMAL ALLOCATION

To provide a benchmark or standard against which we might evaluate the various rules of the game that might improve on the Nash equilibrium of the symmetric interaction above, we will reintroduce the Impartial Spectator, who we relied on for the same purpose in Chapter 4. The Impartial Spectator wishes to determine fishing time and distribute fish so as to maximize a social welfare function, which, because she values the utilities of the two equally, is just the sum of the utilities of the two:

$$\text{Total social welfare} = \text{Abdul's utility} + \text{Bridget's utility}$$
$$W = u^A(h^A, h^B) + u^B(h^A, h^B) \tag{5.27}$$

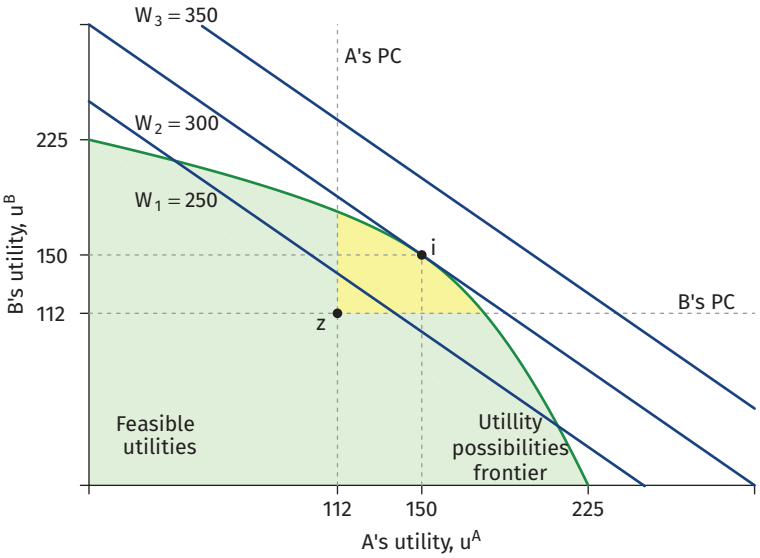
She knows that the solution to this problem must be Pareto efficient, because if it were not, then one of the two could be made better off without worsening the condition of the other, so this could not be the optimum for the Impartial Observer, who values the well-being of both. This means that the socially optimal allocation must be somewhere along the Pareto-efficient curve in Figure 5.12. But where?

A socially optimal allocation

To answer the question, we transform the view of the problem in Figure 5.12, where the space in the figure is defined for hours of fishing, into a new graph, Figure 5.13, which presents the same information in terms of the utilities of the two. The Pareto-efficient curve in Figure 5.12 appears in Figure 5.13 as the dark-green curve that is the utility possibilities frontier. The negative of its slope is the marginal rate of transformation of Bridget's utility into Abdul's utility. This provides the answer to the question: Along the utility possibilities frontier, how much does Bridget's utility have to fall in order for Abdul's to increase by one unit?

! REMINDER This is exactly what the Impartial Spectator did in the previous chapter, for example, Figure 4.5.

Figure 5.13 Feasible utilities, the utility possibilities frontier, and the Impartial Spectator's iso-social welfare indifference curves. Here we show the utility possibilities frontier and feasible utilities for the Impartial Spectator. All points on the frontier are Pareto efficient. The points above and to the right of the fishermen's participation constraints constitute the bargaining set, that is the outcomes that are Pareto superior to their fallback options, $u_z^A = u_z^B = 112$. The Impartial Spectator's iso-social welfare indifference curves show her equal valuation of the utility of the two and the negative of the slope of her iso-social welfare curves is her marginal rate of substitution. The slope of her iso-social welfare curves is −1 indicating that she values the two utilities equally. The negative of the slope of the utility possibilities frontier is the marginal rate of transformation of Bridget's utility into Abdul's. That is, it is the opportunity cost of Abdul having more utility in terms of the utility that Bridget forgoes as a result. The Impartial Spectator will therefore choose point **i** where $mrs = mrt$ to maximize social welfare given the constraint of the utility possibilities frontier.



When Bridget has almost all of the feasible utility then it does not "cost" Bridget much for Abdul to have a little more (the frontier is not very steep); but the marginal rate of transformation rises (the curve steepens) as Abdul gains more utility. The reason is that when Bridget has most of the utility, she is working long hours (almost 15) and incurring a *substantial disutility* of working time as a result. Taking account of both her reduced disutility and her smaller catch, fishing a little less would *not reduce* her utility much. But for Abdul fishing a little *more* would substantially *increase* his utility. So when Bridget is doing most of the fishing (and gaining most of the utility) the opportunity cost of increasing Abdul's utility (in terms of Bridget's forgone utility) is small.

The Impartial Spectator's values are expressed by her indifference curves (the blue lines), their slopes, the negative of her marginal rate of substitution, are a constant, namely –1, because she values the utility of the two equally.

The point **z** represents the fallback utilities of the two (namely 112), and the yellow-shaded area is the set of feasible Pareto improvements over this fallback position. You can see that the Impartial Spectator's social optimum point, **i**, is found where the highest feasible Impartial Spectator's indifference curve is tangent to the utility possibilities frontier (the frontier of the feasible set). So this is another case of the $mrs = mrt$ rule, but now for the Impartial Spectator, rather than Abdul or Bridget.

Rules that implement the social optimum

We know that acting on the basis of their best-response functions Abdul and Bridget *overexploit* the resource. They could both do better if they adopted a different rule for deciding how much to fish. Before turning to institutions that might implement such a new rule for their decisions, let's think about a rule that would exactly implement point **i** in the figure. The Impartial Spectator reminds the two fishermen that the coordination failure occurring at the Nash equilibrium occurs because when they chose their fishing time they did not take account of the external costs that their fishing imposed on the other fisherman. Rules that would avert (or lessen) this coordination failure will get each of them to internalize these costs.

To find the optimum, point **i** in Figure 5.13, the Impartial Spectator proposes the following rules that, if followed, will maximize her social welfare function (Equation 5.27). We show in M-Note 5.10 how these are derived.

$$\begin{aligned} h^A &= \alpha - 2\beta h^A - 2\beta h^B \\ h^B &= \alpha - 2\beta h^A - 2\beta h^B \end{aligned} \tag{5.28}$$

Focusing on the equation for Abdul, the socially optimal condition looks very similar to Abdul's best-response function when he was maximizing his own utility, except for one big difference. We show Abdul's best-response

function again below as his *own* utility maximizing condition, Equation 5.29:

Marginal private costs = Marginal private benefits

A's own u-maximizing condition $$h^A = \alpha - 2\beta h^A - \beta h^B \quad (5.29)$$

Comparing Equations 5.28 and 5.29 we can find the following:

Marginal social costs = Marginal private benefits

A's social optimality condition $$h^A + \beta h^B = \alpha - 2\beta h^A - \beta h^B \quad (5.30)$$

we see that the difference is that there is an extra $-\beta h^B$ in the socially optimal condition (Equation 5.28). This is the negative external effect of Abdul's fishing on Bridget's utility. In Equation 5.30 we have moved the βh^B term to the left-hand side of the equation, adding it to the marginal private cost of fishing more (namely the disutility of hours of fishing, h^A). Together, these are the *marginal social cost*. So Equation 5.30, the condition for Abdul's fishing time to implement a social optimum, says the following:

marginal private cost + marginal external cost = marginal private benefit

The left-hand side is called the *marginal social cost*. The **private cost** (marginal or average) is the cost that the decision maker bears as a result of some action that he or she takes. The **social cost** is the private cost plus any costs imposed on others as negative external effects. The Impartial Spectator's rule causes Abdul to act as if he is taking account of this cost—treating the costs he imposes on Bridget no differently than his own disutility of labor—when deciding on how much to fish.

Imposing the same condition on Bridget, the Impartial Spectator provides a rule where each fisherman *internalizes the negative external effect* of their hours of fishing on the other. As a result, we arrive at the levels of socially optimal fishing time for Abdul and Bridget, denoted by as h_i^A and h_i^B:

Abdul's socially optimal fishing time: $$h_i^A = \frac{\alpha}{1+4\beta} \quad (5.31)$$

Bridget's socially optimal fishing time: $$h_i^B = \frac{\alpha}{1+4\beta} \quad (5.32)$$

Because $\beta > 0$, we see that each of the player's Nash equilibrium levels of fishing time are higher than the socially optimal levels:

! REMINDER In Chapter 1 (section 1.16), we showed how making a defector in the Prisoners' Dilemma game responsible for paying for the harm they inflict on a cooperator internalizes these external effects and replaces the Pareto-inefficient dominant strategy equilibrium with a Pareto-efficient outcome.

M-CHECK There is a simple way to calculate the socially optimal hours. Using the fact that they are identical, their socially optimal hours will also be the same. So, dropping the superscripts in Equation 5.30, we have: $h = \alpha - 4\beta h$ $h + 4\beta h = \alpha$ And therefore: $h = \frac{\alpha}{1+4\beta}$ as shown in Equations 5.31 and 5.32.

PRIVATE COST The private cost (marginal or average) is the cost that the decision maker bears as a result of some action that he or she takes.

SOCIAL COST The social cost is the private cost that the decision maker bears plus any costs imposed on others as negative external effects.

$$h^{AN} = h^{BN} = \frac{\alpha}{1+3\beta} > \frac{\alpha}{1+4\beta} = h_i^A = h_i^B \tag{5.33}$$

These socially optimal levels of fishing time correspond to point **i** (for impartial) in Figure 5.12 where Abdul and Bridget have the same fishing time (ten hours) and the same level of utility.

The job description of the Impartial Spectator is not to figure out how this optimal allocation might be implemented. We leave that to a second (also imaginary) person who we will introduce in the next section.

M-NOTE 5.10 **The Impartial Spectator's choice**

We know that the Impartial Spectator has the following social welfare function, and we can substitute the fishermen's utility functions into W as follows:

$$\begin{aligned}
&\text{Total Social Welfare} = \text{Abdul's Utility} + \text{Bridget's Utility}\\
&W = u^A + u^B\\
&\quad = h^A\left(\alpha - \beta(h^A + h^B) + h^B(\alpha - \beta(h^A + h^B)\right) - \frac{1}{2}(h^B)^2 - \frac{1}{2}(h^A)^2\\
&\quad = \alpha h^A + \alpha h^B - 2\beta h^A h^B - \beta(h^A)^2 - \beta(h^B)^2 \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^B)^2
\end{aligned} \tag{5.34}$$

Next, we need to find the social welfare maximum, or optimal social welfare, for the Impartial Spectator. To do this, we partially differentiate W with respect to the hours of fishing of each fisherman, h^A and h^B, as follows to find the first-order conditions for the social welfare optimum:

$$\begin{aligned}
W_{h^A} = \frac{\partial W}{\partial h^A} &= \alpha - 2\beta h^B - 2\beta h^A - h^A = 0\\
h^A &= \frac{\alpha - \beta h^B - \beta h^B}{1+2\beta}
\end{aligned} \tag{5.35}$$

$$\begin{aligned}
W_{h^B} = \frac{\partial W}{\partial h^B} &= \alpha - 2\beta h^A - 2\beta h^B - h^B = 0\\
h^B &= \frac{\alpha - \beta h^A - \beta h^A}{1+2\beta}
\end{aligned} \tag{5.36}$$

Notice that Equations 5.35 and 5.36 look similar to the best-response functions we found previously, but that each incorporates an additional $-\beta h^B$ for Abdul in the numerator of Equation 5.35 and $-\beta h^A$ for Bridget as shown in the numerator of Equation 5.36. These terms correspond to the cost of the external effect that each fisherman imposes on the other.

M-NOTE 5.11 **Numerical choice of the Impartial Spectator**

We substitute each fisherman's utility function (Equations 5.8 and 5.9) into the social welfare function and use the parameter values of $\alpha = 30$ and $\beta = \frac{1}{2}$.

$$\begin{aligned}
W &= u^A + u^B\\
&= h^A\left(30 - \frac{1}{2}(h^B + h^A)\right) - \frac{1}{2}(h^A)^2 + h^B\left(30 - \frac{1}{2}(h^A + h^B)\right) - \frac{1}{2}(h^B)^2\\
&= 30h^A + 30h^B - h^A h^B - (h^A)^2 - (h^B)^2
\end{aligned} \tag{5.37}$$

As in M-Note 5.10, the Impartial Spectator needs to determine how much each person should work. To determine this, we differentiate the social welfare

continued

function defined by Equation 5.37 with respect to the two hours of work:

$$W_{h^A} = \frac{\partial W}{\partial h^A} = 30 - h^B - 2h^A$$

$$\therefore h^A = \frac{30 - h^B}{2} = 15 - \frac{1}{2}h^B \tag{5.38}$$

$$W_{h^B} = \frac{\partial W}{\partial h^B} = 30 - h^A - 2h^B$$

$$\therefore h^B = \frac{30 - h^A}{2} = 15 - \frac{1}{2}h^A \tag{5.39}$$

We can solve for the Impartial Spectator's choice of work hours for Abdul and Bridget by substituting Equation 5.39 for h^B into Equation 5.38. Following the process of substitution as usual, we find:

$$h_i^A = 15 - \frac{1}{2}\left(15 - \frac{1}{2}h_i^A\right)$$

$$h_i^A = 15 - 7.5 + \frac{1}{4}h_i^A$$

Collect terms $$\frac{3}{4}h_i^A = 7.5$$

Multiply by $$\left(\frac{4}{3}\right)h_i^A = \frac{4}{3}(7.5) = \frac{30}{3} = 10$$

$$h_i^A = 10 = h_i^B$$

Each of them will work ten hours. At the allocation (h_i^A, h_i^B), each has a utility of 150, which is higher than they had at the Nash equilibrium (i.e. 144) and the total social welfare is $W_i = u_i^A + u_i^B = 150 + 150 = 300$.

CHECKPOINT 5.9 **Pareto-improvements** Explain why the allocations in the yellow ellipse in Figure 5.12 are the same as those in the yellow-shaded area in Figure 5.13.

Remedies: Preferences, power, and policy

Stories about two fictional people such as Abdul and Bridget and their textbook lake are light years away from real fishermen in Point Judith, US or Port Lincoln, Australia. John Sorlien, the lobsterman who we quoted at the start of the chapter, is in competition not with a single other fisherman, but with hundreds. Unlike Abdul and Bridget (as you have known them so far), real-world fishermen and lobstermen do sometimes cooperate to pursue common objectives (Sorlien headed the Rhode Island Lobstermen's Association).

A friendly conversation and a handshake might be enough for Bridget and Abdul. But how might such an agreement be arrived at, and how might it be enforced in Rhode Island or South Australia? An even greater challenge is how to design and enforce similar agreements–for example to burn less carbon in order to mitigate climate change–that span not thousands of actors but billions living under the jurisdiction of hundreds of independent governments.

But the parable of Abdul and Bridget has provided an important insight that has illuminated the basic source of coordination failures: the negative effect of their own fishing on the other person ($u_{h^B}^A$ and $u_{h^A}^B$, that is) is *not*

HISTORY In 1968, Garrett Hardin wrote that "freedom in the commons means ruin to all" and as a result he advocated—"mutual coercion mutually agreed upon." But Hardin's pessimism overlooked the many non-coercive ways that local communities have prevented the tragedy.[3]

part of the utility-maximizing process by which each chooses how much to fish.

Addressing these external effects is where institutions come in, meaning changes in the rules of the game. There are three basic approaches whether the common property resource that is being overexploited be fish stocks in a lake or the limited carbon emissions carrying capacity of the earth's atmosphere.

- *Regulation* of the exploitation of the resource by a government.
- *Private ownership* of the resource so that private incentives will deter overexploitation.
- *Management* of the resource through local interactions among the resource users.

These three approaches are sometimes referred to as *states* (meaning governments), *markets*, and *communities*, or similar terms.[4] In Chapter 16 we return to these three approaches and the different ways that each may contribute to a well-organized society.

5.10 GOVERNMENT POLICIES: REGULATION AND TAXATION

To underline the fact that we are not modeling actual governments but instead an ideal of what well-informed public officials seeking to implement better allocations might do, we will introduce a second hypothetical character, the Mechanism Designer. He is tasked with finding policies to implement outcomes that would be recommended by the Impartial Spectator according to her values of efficiency and fairness. He is an economist, with an engineering mentality; she is a philosopher and her job is to identify good, better, or best outcomes.

✓ **FACT CHECK** In response to the rapid depletion of fish stocks in 1992 the Canadian government simply banned fishing for cod in the Grand Banks region of the North Atlantic. Stocks have been slowly recovering since then.[5]

The Mechanism Designer's job then is to implement the best that can be done. The Mechanism Designer is the main character in Chapter 16 which is about public policy. You can think about him as an economist advising a government about how to design and implement its policies.

A government has many options to address coordination failures such as common property resource overexploitation, including educating the public about the costs and promoting basic research to find ways of making the resource more sustainable. But we focus on just two, both relying on a government's capacity to enforce compliance with its policies:

- *Fiat*: Governments can order the implementation of an allocation–for example, reduced fishing–that they or the voters who elected them prefer. A fiat is an order.
- *Taxation:* Governments regularly implement taxes as incentives: taxes make activities more costly and can therefore discourage these activities, while still allowing each person or firm to choose how much of the activity to engage in given the increased costs.

Fiat power

With respect to fiat, the government, if it knew all the relevant information, could select $h^A = h_i^A$ and $h^B = h_i^B$ to maximize total utility. The government might implement this outcome by direct regulation, simply issuing a fishing permit allowing each fisherman a certain number of hours. Any deviation from the permitted hours would result in revocation of the permit and the fishermen would have to revert to their no-fishing fallback positions.

Point **i** in Figure 5.12 is the Impartial Spectator's efficient fishing time allocation that corresponds to what the government would choose. Assuming the government had no reason to favor one fisherman over the other from the standpoint of fairness, a Pareto-efficient and equal distribution of fishing times would be the fiat allocation.

Optimal taxes: Internalizing external effects

Rather than implementing the efficient fishing time plan by fiat, however, the government might want to let the fishermen each decide how much to fish, but change their incentives in order to address the coordination failure. The government would levy what is called a Pigouvian tax on fishing designed to eliminate the discrepancy between the social and private marginal costs and benefits of fishing.

The problem for the government is to select a tax rate on fishing time that as an *intended* byproduct will motivate the fishers to implement an allocation that maximizes total utility while at the same time maximizing their own utility. This means bringing the fishermen's private incentives (the utility function that each maximizes) into alignment with the conditions laid out by the Impartial Spectator.

The problem can be posed this way: find the tax rate that would transform the utility functions of the two fishermen so that their individual best-response functions are identical to those implied by the problem solved by the Impartial Spectator: maximizing total utility and internalizing the costs of the negative external effects.

To internalize the cost means to require each of them to pay (in taxes) for the reduction in the catch of the other that their additional fishing time imposes. We know that each additional hour that Bridget fishes means that Abdul catches βh^A pounds less of fish. So to get Bridget to take account of this negative external effect, she must be taxed at a rate of βh^A for every hour she fishes.

Bridget's socially optimal tax rate depends on Abdul's fishing time because the external effect of Bridget's fishing on Abdul's well-being depends on how much Abdul fishes. If Abdul is not fishing at all, for example, there is no need to tax Bridget's fishing, because it has no external effect.

The tax that induces the fishermen to choose the socially optimal levels of fishing time is just equal to the negative *external effect* they impose on others at the Pareto-efficient levels of fishing time. Such a tax is a change in

HISTORY Imposing taxes on particular behaviors which the government wants to discourage because they impose negative external effects on others—over-fishing, smoking—the government is taking an approach pioneered by the early twentieth-century economists Alfred Marshall (1842–1924) and A. C. Pigou (pee-GOO) (1877–1959). In recognition of his contribution to the field of what is called welfare economics, these are sometimes called Pigouvian taxes.

EXAMPLE The "golden rule" is a common ethical principle that people should treat each other as they themselves would like to be treated. A Pigouvian tax is designed to accomplish the same result by imposing on each decision maker the costs that their decisions impose on others.

the rules of the game that has the effect of *internalizing the external effect* that is the cause of the coordination problem. The tax is an indirect form of coordination: the fishermen as citizens elect a government which they delegate to impose on them a set of incentives to overcome the overfishing problem.

M-NOTE 5.12 **The best-response function of fishing with taxes**

Equation 5.40 shows Bridget's utility function when her fishing time is taxed at the rate of τ per hour fished:

$$u^B(h^A, h^B, \tau) = h^B(\alpha - \beta(h^A + h^B)) - \tau h^B - \frac{(h^B)^2}{2} \tag{5.40}$$

To obtain Bridget's best-response function conditional on the tax rate and Abdul's fishing time, we differentiate Equation 5.40 and set the result equal to zero:

$$\frac{\partial u^B}{\partial h^B} = u^B_{h^B} = \alpha - \beta h^A - 2\beta h^B - \tau - h^B = 0$$

We can rearrange this first-order condition to say that (on the left-hand side of the equation below) the marginal benefits of fishing more (in fish caught) must be equal to (on the right-hand side) the marginal *costs* including the disutility of additional fishing time plus the taxes incurred by fishing more:

$$\alpha - \beta h^A - 2\beta h^B = \tau + h^B$$

Re-arranging this to isolate h^B we have:

$$(1 + 2\beta)h^B = \alpha - \tau - \beta h^A$$

$$h^B(h^A, \tau) = \frac{\alpha - \tau - \beta h^A}{1 + 2\beta} \tag{5.41}$$

M-NOTE 5.13 **Implementing the Impartial Spectator's choice**

We know from M-Note 5.10 that the first-order condition for Bridget's fishing time in the social optimum allocation proposed by the Impartial Spectator is:

$$h^B = \frac{\alpha - \beta h^A - \beta h^A}{1 + 2\beta} \tag{5.42}$$

The Mechanism Designer's job is to find the tax rate per hour of Bridget's fishing time that will induce her to act as if that were her private (self-regarding) first-order condition too. We know from M-Note 5.12 that Bridget's best-response function, including taking account of the tax, is the following:

$$h^B(h^A, \tau) = \frac{\alpha - \tau - \beta h^A}{1 + 2\beta} \tag{5.43}$$

The question that the Mechanism Designer must now solve is: what is the level of the tax rate τ that will make Equation 5.43 look like Equation 5.42 so that Bridget's private incentives will lead her to implement the Impartial Spectator's social optimum?

Comparing the two equations you can see that setting the tax rate that Bridget pays $\tau^B = \beta h^A$ will make the two equations identical. So that is the

continued

optimal tax rate. The tax rate for Abdul would, by the same reasoning, be $\tau^A = \beta h^B$.

Then we can calculate the tax rate that Bridget pays at the Nash equilibrium. Because we know that the optimal tax implements the social optimum recommended by the Impartial Spectator, we substitute the value for h_i^A into the expression for the tax rate so we have, $\tau = \beta h_i^A$ or

$$\text{Bridget's tax rate} \quad \tau = \beta h_i^A$$

$$\text{Abdul's hours worked (Nash)} \quad h_i^A = \frac{\alpha}{1+4\beta}$$

$$\text{Bridget's tax rate (Nash)} \quad \tau = \frac{\alpha\beta}{1+4\beta}$$

CHECKPOINT 5.10 **External effects and taxes** What does "internalize the external effect" of one's fishing time mean. How does the tax on fishing time accomplish this?

5.11 PRIVATE OWNERSHIP: PERMITS AND EMPLOYMENT

But government policies are not the only change in the rules of the game that might address the overfishing coordination problem. Suppose that the property rights over the lake are changed such that the lake is no longer a common pool resource, but is *privately owned*. As a result, the lake is no longer non-excludable; now, as a resource that is both excludable and rival, it is a *private good*. The person who owns the lake, say Bridget, could exclude Abdul entirely (remember that is what private property means). But as an owner she now has bargaining power over Abdul, and may be able to do better by letting him fish under conditions favorable to her.

How do these new rules of the game change the Nash equilibrium?

- *Permits*: Bridget might sell Abdul a fishing **permit** allowing him to catch not more than a given amount of fish, setting the highest possible *fee* for the permit consistent with Abdul being willing to fish under those terms.
- *Employment*: Bridget might offer Abdul an employment contract under which Abdul would fish a given amount of time; the fish caught by Abdul would be Bridget's. Abdul's compensation would be a *wage* (paid in the fish caught by the two of them) which would be sufficient to offset the disutility of Abdul's fishing time and the opportunity cost of his fishing (and therefore to satisfy Abdul's participation constraint).

In both cases the bargaining power she has as the owner of the lake allows Bridget to make take-it-or-leave-it offers. Abdul's only choice is to accept or reject.

PERMIT A permit allows a firm or person to engage in an activity: it gives permission.

! REMINDER While a single owner will take account of the costs of overfishing and restrict fishing accordingly, the owner may also be a monopolist in selling the fish to others, and will restrict fishing even more than is socially optimal, so as to sustain a high price of fish. We do not include the consumers of fish other than the owner and those fishing on the lake. But we will include them when we return to these effects of a monopolist in Chapter 9.

Selling permits to fish

To understand what Bridget will do as the owner of the lake, let us return to her utility function:

$$\text{Bridget's utility} \qquad u^B(h^A, h^B) = y^B - \frac{1}{2}(h^B)^2 \tag{5.44}$$

In Equation 5.44, when Bridget was one of the two fishermen and could *not* charge anyone to access the lake, her *production* of fish limited her *consumption* of fish (y^B). She ate what she caught, and no more.

Now, though, as she will be charging Abdul to access the lake, she can consume more than she catches. So, we distinguish her *catch* of fish, $x(h^A, h^B)$, from her *consumption* of fish, y^B. When she is the owner charging a fee (paid in fish) to Abdul for her permission that he fish in the lake, her consumption of fish equals her *own* catch plus the fee she charges, F. Her utility therefore becomes:

$$\text{Owner's (B's) utility} \qquad u^B(h^A, h^B, F) = \underbrace{\overbrace{x(h^A, h^B)}^{\text{Bridget's catch}} + F}_{y^B\text{, Bridget's consumption}} - \overbrace{\frac{1}{2}(h^B)^2}^{\text{Bridget's disutility}} \tag{5.45}$$

Equation 5.45 tells us that Bridget, as the owner, now has three variables to determine not just one (her own fishing time) as she had when she interacted with Abdul in the symmetric game:

- Her own fishing time, h^B.
- Abdul's fishing time, h^A.
- The cost F to Abdul of the permit allowing him to fish h^A hours.

Bridget, being self-regarding, will want to know: "what is the largest fee that I can charge Abdul?" Remember Abdul has the option of not fishing at all, that is, taking his fallback option with associated utility $u_z^A = y_z$. Agreeing to fish in Bridget's lake means foregoing the fallback option, so $u_z^A = y_z$ is the *opportunity cost* of fishing.

This is Abdul's participation constraint, limiting how much Bridget can charge for the permit: the fee plus the opportunity cost of fishing cannot be larger than Abdul's utility from fishing, $u^A(h^A, h^B)$:

Abdul's utility from fishing ≥ Permit fee + Foregone fallback

$$\text{Abdul's participation constraint} \qquad u^A(h^A, h^B) \geq F + y_z \tag{5.46}$$

Because Bridget would never consider charging Abdul less than she could, we can assume that Equation 5.46 will be satisfied as an equality and so (rearranging the equation) we have:

$$\text{Abdul's participation constraint (PC)} \qquad F = u^A(h^A, h^B) - y_z \tag{5.47}$$

Bridget's constrained optimization problem is now to vary F, h^A, and h^B to maximize her utility from fishing plus the fee she charges Abdul, or:

$$\text{Maximize:} \quad u^B(h^A, h^B, F) = x^B(h^A, h^B) + F - \frac{1}{2}(h^B)^2 \tag{5.48}$$

subject to Abdul's participation constraint (Equation 5.47). This means that we can use Equation 5.47 to replace the F in Equation 5.48, so that now Bridget's objective is to choose h^A and h^B to:

$$\text{Maximize:} \quad u^B(h^A, h^B) = x^B(h^A, h^B) - \frac{1}{2}(h^B)^2 + u^A(h^A, h^B) - y_z \tag{5.49}$$

Once we have found the h^A and h^B that maximize Equation 5.49, we can insert those values of h^A and h^B into Equation 5.47 to determine F, the cost of the permit to charge Abdul.

What will these new values of (h^A, h^B) be? We will find that, maybe surprisingly at first, they are just the same as the values introduced by the Mechanism Designer tasked with implementing the Impartial Spectator's social optimum.

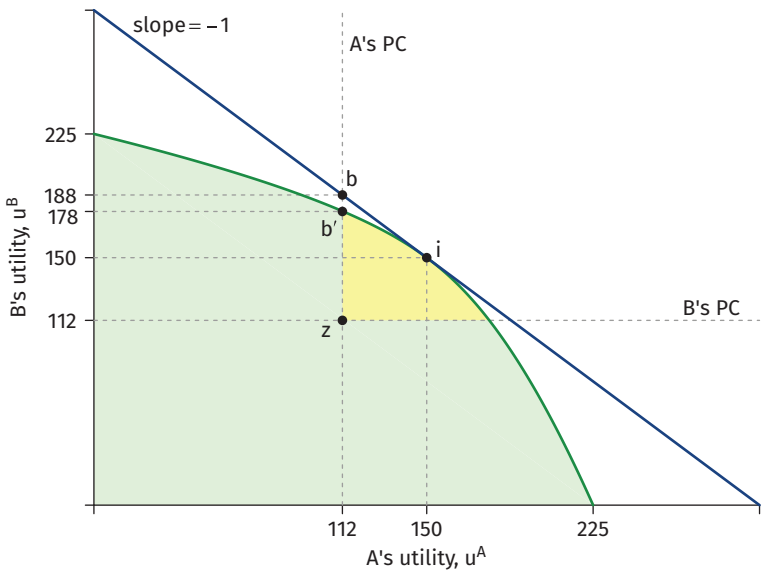
Looking carefully at Equation 5.49 we see that the first term is her own utility if no fee is paid and the second is Abdul's utility if no fee is paid. So far this is the same quantity that the Impartial Spectator maximized, namely the sum of the utilities of the two, except that now y_z is subtracted. But because y_z is a constant (112 pounds of fish in our numerical examples) the solution of these two optimizing problems—the values of h^A and h^B chosen—must be the same.

This means when Bridget is the owner the hours worked, h_i^A and h_i^B will be *equal*, ten hours each in our numerical example—but the levels of utility realized will be maximally *unequal*. Abdul will get exactly 112, his fallback position, and Bridget will get 188 as shown by point **b** in Figure 5.14.

Because the allocation was determined by Bridget maximizing her utility subject to a constraint on Abdul's level of utility (that is the participation constraint) it has to be Pareto efficient, by definition. In fact, she implements exactly the level of fishing by the two that would be socially optimal in the absence of the fee.

To see this, imagine that Bridget as the owner had implemented a plan in which she reduced Abdul's work hours and increased her own so that she would obtain an allocation on the utility possibilities frontier at point **b**′. She can do better than this if she implements the socially optimal number of hours ($h_i^A = h_i^B = 10$) at point **i** and then requires Abdul to pay her for the right to fish. Then the total rents available are 300. She will charge a fee such that Abdul will receive just a bit more than his fallback, $u_z^A = 112$ and she will get $300 - 112 = 188$. The payment of the fee corresponds to a movement *along* the blue line with slope $= -1$ from point **i** to point **b**. The blue line therefore indicates movements from point **i** to alternative feasible trades when the fishermen are able to trade fish between them as payments.

Figure 5.14 Payments in fish takes the fishermen to allocations outside of the original feasible set. The blue line with slope –1 shows the allocations of utility that are possible if the two fish at the socially optimal times indicated by point **i** followed by a transfer of fish from one to the other. The slope is –1 because the opportunity cost of, say Bridget having a kg more fish, is that Abdul has one kg less. If Bridget as the owner implemented a plan in which she reduced Abdul's work hours and increased her own, she would obtain an allocation on the utility possibilities frontier at point **b**′. However, if she implements the optimal number of hours ($h_i^A = h_i^B = 10$) at point **i** and has Abdul pay her for the right to fish with a fee, then the total rents available are 300. She will charge a fee such that Abdul will receive just a bit more than his fallback, $u_z^A = 112$ and she will get $300 - 112 = 188$. This corresponds to a movement *along* the blue line with slope $= -1$ from point **i** to point **b**.



The economic reason Bridget implements the socially optimal level of fishing times–namely the times that maximize the total rents–follows directly from the fact that Bridget knew in advance that by charging the largest fee that would satisfy Abdul's participation constraint she would capture all of the feasible rents. Given that fact, she had every interest in making the total rents as large as possible.

But why wouldn't Bridget select $h^A = 0$, and have exclusive access to the lake? The reason is that the marginal cost of compensating Abdul's fishing time is *very small* when he is not fishing much, or at all. So it is to Bridget's *advantage* to let Abdul fish in the lake and pay her for the privilege, rather than doing all the fishing herself.

Employing others to fish

Instead of issuing a permit, Bridget might hire Abdul to work for her. Employment differs from the permit system in that when Bridget employs Abdul, she owns all of the fish caught by Abdul, but must devote some of this to paying a wage w to Abdul that is sufficient to satisfy his participation constraint.

From our reasoning in the permit case we know that the participation constraint will be satisfied as an equality. This allows us to use the fact that the total wage paid (w) must offset Abdul's disutility of fishing time and the opportunity cost of fishing (namely his fallback option, y_z that he gives up if he fishes) or:

$$\text{Abdul's PC as an employee} \qquad w = \frac{(h^A)^2}{2} + y_z \tag{5.50}$$

Bridget then must choose w, h^A, and h^B to maximize her utility:

$$\begin{aligned} u^B(h^A, h^B, w) &= x^A(h^A, h^B) + x^B(h^A, h^B) - \frac{1}{2}(h^B)^2 - w \\ &= \text{A's catch} + \text{B's catch} - \text{B's disutility} - \text{A's wage} \end{aligned}$$

Then substituting Equation 5.50 for Abdul's wage, what Bridget maximizes when she employs Abdul is:

$$u^B(h^A, h^B) = h^B(\alpha - \beta(h^A + h^B)) - \frac{(h^B)^2}{2} + h^A(\alpha - \beta(h^A + h^B)) - \frac{(h^A)^2}{2} - y_z \tag{5.51}$$

Equation 5.51 can be understood as follows:

- it is identical to what Bridget maximized in the permit case, namely Equation 5.49 and
- identical to what the Impartial Spectator maximized, namely, Equation 5.27, minus the constant y_z.

In both the permit and the employment cases, the outcome is Pareto efficient, but Abdul gains an amount equal only to his disutility of fishing time plus the opportunity cost of his fishing at all. The allocation proposed by the Impartial Spectator and that implemented by Bridget as owner of the lake does not differ in the fishing times of each, or the degree of exploitation of the fishing stock. In this sense private ownership of the lake has addressed the Pareto inefficiency of the overexploitation of the lake as a common property resource.

The only difference is that in the private ownership case there is transfer of rents (amounting to 88 pounds of fish in both cases) from the nonowner to the owner:

- In the permit case the transfer took the form of the fee for the permit to fish ($F = 38$) that Abdul paid to Bridget.

- In the employment case the transfer occurred because Bridget owned all of the fish that Abdul caught (200), 38 pounds of which she retained for her own consumption after paying him the wage ($w = 162$), which, minus his disutility of fishing of 50, equals his fallback utility of $u_z = 112$.

This is a general feature of social coordination problems. When one actor is sufficiently powerful to maximize their utility subject to the participation constraint of other actors, then the powerful actor will maximize total utility and get all of the economic rents.

CHECKPOINT 5.11 How private ownership can eliminate the coordination failure Explain how private ownership of the entire lake by Bridget would both eliminate the over-fishing problem and result in a very unequal distribution of the benefits of her interaction with Abdul.

5.12 COMMUNITY: REPEATED INTERACTIONS AND ALTRUISM

Here and in previous chapters we have used two-person games to represent economic interactions among a very large number of people. But some of our interactions really *are* with small numbers of people, for example, in our neighborhoods, families, and workplaces, and even, in some cases, in exploiting a local common property resource like a forest or fishery.

These small communities often address coordination problems in ways not possible when the number of people interacting is very large. This is possible because members of small communities:

- often have *information* about each other that is not available to governments or private owners who are not part of the community;
- *interact repeatedly* with each other so that there are opportunities to retaliate against members who violate social norms or informal agreements; and
- often care about each other, and these *social preferences* can reduce conflicts of interest (as we saw in the previous chapter) and can provide the basis for addressing coordination problems.

These characteristics of small communities give them capabilities in solving coordination problems that are unavailable to purely government- or market-based approaches. As we have seen in Chapter 2, public goods experiments show that people are willing to punish fellow group members whose behaviors violate norms, even when inflicting the punishment is costly to the punisher.

Let's see how a small community of fishing people—illustrated by Bridget and Abdul—might address the overexploitation of the common property resource.

Repeated interactions

In Chapter 1 we showed how the Pareto-inferior (Defect, Defect) outcome of a Prisoners' Dilemma Game could be averted by a change in the rules of the game requiring players to compensate others for the costs that their actions (in this case defecting) imposed on others. The objective of internalizing the external costs imposed on others was directly built into the change in the rules of the game.

Here we accomplish the same thing in the same manner—internalizing the external costs by changing the rules of the game—but here we do it by expanding the set of strategies the players are able to use. We do this by repeating the game over many periods.

In the one-shot games we have introduced so far the strategies available to the players are limited: select some amount of fishing hours. One way to make the game more realistic is to let the interaction be repeated over possibly many periods with the same players. Then more complicated strategies are possible, even if in every period there are just two actions one can take, for example, fish ten hours or fish twelve hours. Importantly, strategies can now be conditional on what the other player has done in previous play.

One strategy might be to play the strategy that would implement the social optimum (fish ten hours) in the first round and on the next and successive rounds of the game, play whatever the other player played on the previous round. This strategy is called "nice tit for tat": nice because it begins with a strategy that could be mutually optimal, but "tit for tat" because it punishes the other player if she takes the overfishing option.

Consider a repeated game between Abdul and Bridget with the following properties:

- *Actions*: In each period of the game, they may fish either ten hours, the socially optimal amount or 12 hours, the overfishing level at the Nash equilibrium of the symmetric game in which they do not coordinate in any way.
- *Duration of the game*: After every period that the game is played, it is continued with some probability $0 < P < 1$.
- *Payoffs*: In each period the payoffs are given in panel (a) of Figure 5.15. The cell entries are from the analysis of their interaction we have carried out so far, with the parameters used in our numerical examples. Payoffs for the game are the sum of payoffs for each period the game is played.
- *Strategies*. Each may choose either to Fish 12 Hours in every period of the game (called "Defect") or Fish 10 Hours in the first period of the game and every subsequent period until the other plays Fish 12 Hours, in which case Fish 12 Hours as long as the game lasts. This strategy is called "Grim Trigger": the term *trigger* is used because an act of defection by

Figure 5.15 Repeated interactions can convert a one-shot Prisoners' Dilemma into an Assurance Game allowing for coordination on a socially optimal allocation. Panel (a) is the payoff matrix for the one-shot (stage) game between Bridget and Abdul, that is played once only. Inspection of the payoffs shows that it is a Prisoners' Dilemma. You can confirm using the method introduced in Chapter 1 that each player fishing 12 hours is a Nash equilibrium (the circles and dots show that this is also a dominant strategy equilibrium). Panel (b) gives expected payoffs for the game, if at the end of each period with probability $P = 0.9$ the game is played again (with the same payoffs per period as shown in panel (a). In this case the circles and dots indicate that the repeated game has two Nash equilibria: the Nash equilibrium of the one-shot game with payoffs to each of 1,440, and the socially optimal allocation with payoffs of 1,500.

		Bridget	
		10 hours	12 hours
Abdul	10 hours	150 / 150 a	156 / 140 d
	12 hours	140 / 156 b	144 / 144 c

(a) Stage game

		Bridget	
		Grim Trigger	Defect
Abdul	Grim Trigger	1500 / 1500	1452 / 1436
	Defect	1436 / 1452	1440 / 1440

(b) Repeated interaction

the other sets off a punishing defection by the actor; it is *grim* because the defections go on as long as the game lasts.

M-CHECK We show in the appendix that the expected duration of an interaction is the inverse of the probability that at the end of a period the interaction will be terminated.

How will this repeated game be played? We will assume that both players are entirely self-regarding. You can see that the single-period payoff matrix—called the stage game—has the structure of a Prisoners' Dilemma. The repeated game will continue following each period with probability P, which for concreteness we set as $P = 0.9$. This means that the expected duration of the game is $\frac{1}{1-0.9} = 10$ periods. If the game were played among total strangers, it is unlikely that it would be repeated with such a high probability. But if the two players are neighbors or coworkers they are very likely to continue interacting.

The payoffs in the repeated game are derived as follows.

- *Playing Defect against Defect*: The expected payoff will be the payoff (144) of the mutual defect option of the stage game times the expected duration of the game (10 periods), or 1,440.

- *Playing Grim Trigger against Grim Trigger*: The expected payoff will be 150 (fishing 10 hours) times the expected duration of the game (10), or 1,500.the stage game payoff to
- *Payoff to playing Grim Trigger against Defect*: In the first period, the Grim Trigger player fishes ten hours, while the defector fishes 12 hours, and so receives 140 that period; and then he defects in the next period (should it occur) and until the game ends. The probability that the second period happens is P and if it does it can be expected to continue for ten *more* periods, so the payoffs from period two to the end of the game are $0.9 \times 10 \times 144$ or 1,296 which, adding the first period's payoffs totals 1,436.
- *Payoff to playing Defect against Grim Trigger*: This is calculated exactly as in the case immediately above. The defector gets 156 in the first period and then the mutual defect payoff as long as the game lasts, totaling 1,452.

Looking at panel (b) of Figure 5.15 you can see that mutual Defect is still a Nash equilibrium in the repeated game: it is a best response to itself as the circles and dots show.

But in this case the repeated game is not a Prisoners' Dilemma: the best response to Grim Trigger is not Defect, but Grim Trigger itself. You can confirm using the circle and dot method introduced in Chapter 1 that each player fishing 12 hours is a Nash equilibrium (the circles-and-dots show that this is also a dominant strategy equilibrium). Panel (b) gives expected payoffs for the game, if, at the end of each period, the game is played again with probability $P = 0.9$ (with the same payoffs per period as shown in panel (a)).

In this case the circles and dots indicate that the repeated game has two Nash equilibria: the Nash equilibrium of the one-shot game with payoffs to each of 1,440, and the socially optimal allocation with payoffs 1,500. And so if the two had decided to play Grim Trigger they would continue doing so until the game ended, implementing the social optimum.

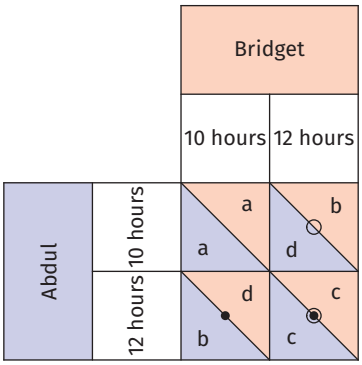
This means that what was a Prisoners' Dilemma if played as a one-shot game can become an Assurance Game when played repeatedly if the game is repeated with a sufficiently high probability. Under these conditions, entirely self-regarding actors acting independently and without government regulation can avoid the Prisoners' Dilemma.

M-NOTE 5.14 Cooperation without agreements in a repeated game

The key to how game repetition converts a Prisoners' Dilemma stage game into an Assurance Game that can implement the socially optimal level of fishing is that Defect should not be a best response to Grim Trigger. This requires that the payoff to playing Defect against Grim Trigger should be less

continued

Figure 5.16 Payoffs for the one-shot Prisoners' Dilemma Game shown in Figure 5.15 (a) repeated here to accompany M-Note 5.14. For both players: $b > a > c > d$.



than the payoff to playing Grim Trigger against itself, or (using the letters in the payoff matrix of Figure 5.16):

$$b + c\left(\frac{P}{1-P}\right) < \frac{a}{1-P}$$

which, rearranged to isolate the P gives us:

$$P > \frac{b-a}{b-c}$$

This means that (in the one-shot game) the probability that the game will be continued after each round must be greater than the payoff advantage of defecting on a cooperator $(b-a)$ relative to the payoff advantage to coordinating on 10 Hours rather than 12 Hours $(b-c)$. This means that repeating the game is more likely to result in the Pareto-superior symmetric outcome (both fishing less) if:

- the incentive to exploit the cooperation of the other is less,
- the joint benefit of mutual restricting fishing hours is more, and if
- the interaction will be repeated with high probability.

For the payoffs in Figure 5.15 (a) this condition is satisfied for any $P > 0.5$

Social preferences: Altruism

In addition to the greater likelihood that the interaction will be repeated, the fact that the community of fishermen is *small* means two additional important facts about their context are likely to hold. First, people in small communities can more easily access information about one another. Second, in small communities people often have a concern about each other's well-being, such as altruism, fairness concerns, or reciprocity.[6]

Because we have used u^A and u^B to refer to the utility each gets from fishing, we now introduce the value functions v^A and v^B, which include their concern for the other's well-being. Because the two are other-regarding, their evaluation of the outcomes they believe their actions will produce are based on these social preference utility functions, v^A and v^B. To see how social preferences might help solve coordination failures, suppose that in choosing an action each participant puts some weight (which we represent by λ) on the utility of the other as in Equation 5.52, so that Abdul and Bridget have social preferences like those we introduced in Chapters 3 and 4. We let λ be the same for Abdul and Bridget (they are equally altruistic) and take any value between 0 (no altruism) and 1 (perfect altruism, caring about the other's well-being as much as one's own).

M-CHECK Notice here that $0 \leq \lambda \leq 1$ rather than $0 \leq \lambda \leq \frac{1}{2}$ in Chapters 3 and 4. This is because in those chapters the utility functions are Cobb-Douglas whereas in this chapter we are *adding up* the utility functions to display altruism.

$$\text{Altruistic utility function: } = \text{Own utility} + \lambda\text{Other's utility} \tag{5.52}$$

$$\text{Altruistic A:} \quad v^A(h^A, h^B) = u^A + \lambda u^B$$

$$\text{Altruistic B:} \quad v^B(h^A, h^B) = u^B + \lambda u^A$$

We show in M-Note 5.15 that altruism alters the marginal costs and benefits of fishing:

$$\text{marginal benefit of fishing more} = \text{marginal cost of fishing more}$$
$$\alpha - \beta(2h^A + h^B) = h^A + \lambda\beta h^B \tag{5.53}$$

Except for the last term on the right, this is exactly the benefits and cost of fishing more when Abdul was not altruistic in Equation 5.10. The final term, $\lambda\beta h^B$, is new; it shows how altruism affects Abdul's decision. Given his degree of altruism λ, $\lambda\beta h^B$ is how much his utility is *reduced* by his additional fishing imposing the cost βh^B on Bridget.

As a result, the two Nash equilibrium fishing times are as follows (shown in M-Note 5.15):

$$h^{A\lambda} = \frac{\alpha}{1+\beta(2+(1+\lambda))} = h^{B\lambda} \tag{5.54}$$

This expression shows that when $\lambda = 0$ we have the Nash equilibrium fishing times before we introduced altruism h^N, that is $(\lambda^{AN}, \lambda^{BN})$. And when $\lambda = 1$ the Nash equilibrium is the allocation selected by the Impartial Spectator h_i, that is $(\lambda_i^A, \lambda_1^B)$. So for some positive degree of altruism λ we have:

$$h^N = \frac{\alpha}{1+3\beta} > h^\lambda = \frac{\alpha}{1+\beta(2+(1+\lambda))} \geq \frac{\alpha}{1+4\beta} = h_i \tag{5.55}$$

These equations make it clear that in order for altruism to implement the social welfare-maximizing allocation proposed by the Impartial Spectator, the two would have to be the perfect altruists that we defined in Chapter 2, caring as much for the other person as she does for herself (namely $\lambda = 1$). The difficulty of sustaining this level of altruism may suggest why most successful communities do not rely entirely on goodwill, but supplement it with mutual monitoring and punishment for transgression of norms.

EXAMPLE Even the most utopian communities, such as the contemporary Amish or Hutterite communities in the US, supplement altruistic preferences with social sanctions and monitoring of peers (which is easier to do in small communities).

M-NOTE 5.15 **Nash equilibrium with altruism**

Now, Abdul cares about Bridget. His value function incorporating altruistic preferences is:

$$\begin{aligned} v^A(h^A, h^B) &= u^A + \lambda u^B \\ &= h^A\left(\alpha - \beta(h^A + h^B)\right) - \frac{1}{2}(h^A)^2 + \lambda\left(h^B\left(\alpha - \beta(h^A + h^B)\right) - \frac{1}{2}(h^B)^2\right) \end{aligned}$$

To obtain his best-response function, we differentiate Abdul's utility function with respect to fishing time and set it equal to zero:

$$\frac{\partial v^A}{\partial h^A} = \alpha - \beta(2h^A + h^B) - h^A - \lambda\beta h^B = 0 \tag{5.56}$$

$$\text{Rearranging} \qquad (1+2\beta)h^A = \alpha - (1+\lambda)\beta h^B \tag{5.57}$$

We shall use the superscript λ to indicate what is special about this case, namely altruism. Then, rearranging Equation 5.57 we have the best-response function of altruistic Abdul:

continued

$$h^A = \frac{\alpha - (1+\lambda)\beta h^B}{1+2\beta} \tag{5.58}$$

To see how altruism affects the Nash equilibrium we use the fact that, because the two people are identical, their fishing times will be the same in equilibrium. So we can just solve the equation below for $h^{A\lambda}$:

$$h^A = \frac{\alpha - (1+\lambda)\beta h^A}{1+2\beta}$$

from which, rearranged to isolate $h^{A\lambda}$, we have:

$$h^{A\lambda} = \frac{\alpha}{1+\beta(2+(1+\lambda))} = h^{B\lambda} \tag{5.59}$$

We use the superscript λ here to indicate the Nash equilibrium when the degree of altruism is given by $\lambda > 0$.

CHECKPOINT 5.12 **External effects and altruism** How does altruism (at least partially) internalize the external effect of one's fishing on the other? Why is $\lambda = 1$ required to fully internalize the external effect?

5.13 APPLICATION: IS INEQUALITY A PROBLEM OR A SOLUTION?

Recall that there are two standards that we use to evaluate policies to address coordination failures:

- Does it result in a Pareto improvement over the status quo, that is, does it improve efficiency by making at least one of the participants better off and none worse off?
- Is the resulting allocation more fair than the status quo, that is, are the rents (improvements over the status quo) that the players receive fair?

In some cases the two objectives can be jointly realized; in others they are in conflict.

The distribution of the economic rents resulting from coordination depends on the particular transformation of the game which makes some degree of coordination possible. Unequal solutions to local social coordination problems are generally based on the wealth or power of one of the fishermen. If one of the fishermen has a much larger net than the others, and so can be assured of catching most of the fish, his best response will approximate the allocation of a single owner of the lake. In this case, inequality in wealth among the fishermen would lessen the coordination failure.

Important inequalities may exist even among otherwise identical fishermen. To see this consider two possibilities:

1. *Take-it-or-leave-it (TIOLI) power*: When one player has substantially more power than the other, allowing them to make a TIOLI offer of

both their own and the other's fishing time, then they may implement an allocation where they obtain all the economic rents and the other remains on their participation constraint. As you already know from the previous chapter, the outcome is Pareto efficient.

2. *First-mover advantage:* When one player can credibly commit to a fishing time such that the other must simply respond, they obtain more of the economic rent than the other player, but not as much as if they had TIOLI power. This is similar to the price-setting power in Chapter 4. The outcome is Pareto inefficient.

We start with first-mover advantage: the power to commit to one's own fishing time.

First-mover advantage: fishing time-setting power

HISTORY Heinrich von Stackelberg (1905–1946) used this model to represent price-setting among duopolists (two firms in a market).

Suppose that Abdul can announce a level of fishing time and commit to it in such a way that Bridget understands that nothing she can do will alter Abdul's fishing activity. Bridget will then select her level of fishing to maximize her utility given what Abdul has committed to. In this situation, Abdul is the first mover and has fishing time-setting power similar to the *price-setting power* (as in Chapter 4). Economists call Abdul in this situation the *Stackelberg leader*.

The big difference between Abdul having fishing time-setting power and our previous analysis is that the game is now *sequential* and the order of play matters: who gets to go first is important.

! REMINDER In Chapter 4, when she had price-setting power, A maximized her utility subject to B's price-offer curve. B's price-offer curve was A's *incentive compatibility constraint*, which is exactly what a best-response function is: a best-response function shows what action would be your *best response* to the action (e.g. price or fishing time level) the first mover commits to.

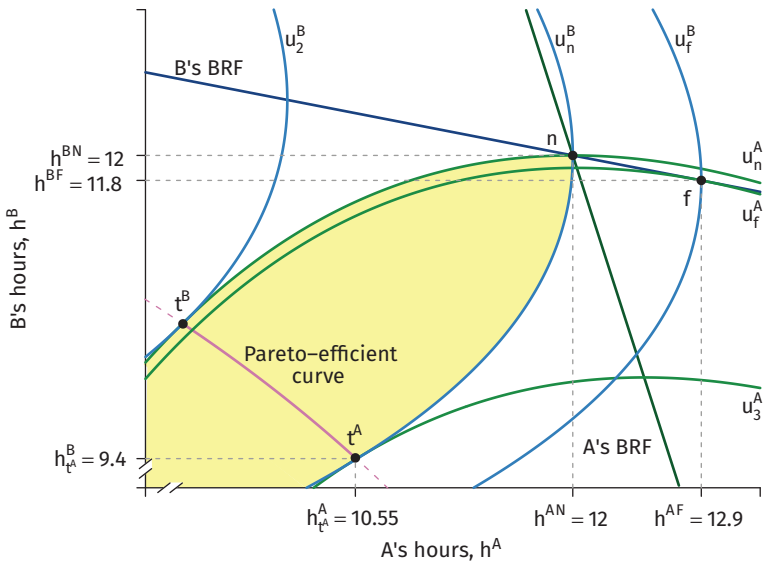
How would Abdul decide what level of fishing time to commit to as the fishing time-setter? As the first mover, he will begin by determining what the second mover will do in response to each of his actions, and then select the action that maximizes his own utility given the best-response function. The second-mover's best-response function is the *incentive compatibility constraint*.

Abdul maximizing his utility subject to Bridget's incentive compatibility constraint is a simple but important change in the assumed behavior of the fishermen: Abdul now recognizes and takes advantage of the fact that by choosing various levels of fishing time he can affect the level of fishing time Bridget chooses. Abdul's behavior is *strategic* because it takes account of Bridget's reaction to his action.

In this first-mover case shown in Figure 5.17, Abdul is constrained not by a given level of Bridget's *utility*, that is, by her participation constraint, but by Bridget's *maximizing behavior* as given by her best-response function. His choice of hours will be the point on her best-response function where it is tangent to his indifference curve, shown by point **f** in the figure.

The first-mover outcome will not be Pareto-efficient because Bridget's indifference curve intersects Abdul's indifference curves at point **f**, and so the marginal rates of substitution cannot not be equal. Abdul's first-

Figure 5.17 First-mover advantage: Fishing time-setting power. In the figure, Abdul is the leader with fishing time-setting power (first-mover power) and Bridget is the follower. Abdul takes Bridget's best-response function as his *incentive compatibility constraint*. He maximizes his utility subject to satisfying the ICC, finding the point at which his indifference curve is tangent to her best-response function as occurs at point **f** where Abdul exerts fishing time h^{AF} and Bridget exerts fishing time h^{BF} and the two fishermen obtain utilities $u_2^A = u_f^A$ and $u_1^B = u_f^B$. The first-mover-based outcome is contrasted with the Nash equilibrium outcome of the simultaneous interaction where Bridget had higher utility $u_2^B = u_n^B$ and Abdul had lower utility $u_1^A = u_n^A$. Point $\mathbf{t^A}$ indicates the allocation Bridget would implement if she had TIOLI power over Abdul (similarly, $\mathbf{t^A}$ indicates would Abdul would implement if he had TIOLI power over Bridget).



mover advantage allows him to improve his position by comparison to the Nash equilibrium, in this case at the expense of Bridget, whose outcome as second mover is worse than the Nash equilibrium.

Take-it-or-leave-it power

! REMINDER When either trader had TIOLI power in Chapter 4, the exercise of power led to a Pareto-efficient outcome.

Let us switch roles now and consider what would happen if Bridget had more power than Abdul. She has enough power to make Abdul a take-it-or-leave-it offer, specifying not only how much she would fish, but how much Abdul is to fish, too, along with the threat that should Abdul *not* accept the offer, then Bridget would simply fish at the level of the Nash equilibrium of the simultaneous move game.

Because Abdul will refuse her offer if it is worse for him than the Nash outcome, Bridget must make an offer to Abdul that satisfies Abdul's participation constraint. If she does so, the outcome will be Pareto efficient.

In Figure 5.17, Bridget's TIOLI offer is shown at point $\mathbf{t^B}$. At $\mathbf{t^B}$, her indifference curve u_2^B is tangent to Abdul's indifference curve u_n^A, so the allocation she chooses for her TIOLI offer is Pareto efficient. If Abdul had TIOLI power over Bridget, he would implement the allocation at point $\mathbf{t^A}$ and this allocation would also be Pareto efficient.

Summing up, our model shows that the effect of unequal power will always benefit the more powerful and may, but need not, result in a Pareto-inefficient outcome.

Positive effects on Pareto efficiency occurred when Bridget had take-it-or-leave-it power and the outcome was Pareto efficient, but probably regarded as unfair by Abdul or an Impartial Spectator. The model is *hypothetical* but the problem is not.

CHECKPOINT 5.13 TIOLI power and Pareto efficiency In the case where Abdul moves first and sets the price (but he does not have TIOLI power) why is the outcome Pareto-inefficient?

Evidence from field studies

Abdul's exercise of TIOLI power illustrates an important point: giving one actor all the power can sometimes allow efficient coordination. But a field experiment among forest commons users in rural Colombia underlines how inequality may be an impediment to achieving more satisfactory outcomes through coordination. Juan Camilo Cardenas implemented common pool resource behavioral experiments among villagers who rely for their living on the exploitation of a nearby forest.[7] So the subjects in the experiment were in real life playing the same game that the experimenter invited them to play.

In Cardenas's game, the subjects choose to withdraw a number of tokens from a common pool (these represented exploitation of the common property resource), and after all subjects had taken their turn, the tokens remaining were multiplied by the experimenter and then divided equally among the players, the tokens then being exchanged for money. (This is similar to the Public Goods Game experiment in Chapter 2 except that subjects decide how much to *withdraw* rather than how much to *contribute* to the pool.) For an initial set of rounds of the game, no communication was allowed. But in the final rounds of the game, subjects were invited to converse for a few minutes before making their decisions.

Cardenas expected that communication would reduce the level of withdrawals from the common pool (as has been the case in similar experiments) despite the fact that it does not alter the material incentives of the game. Communication was indeed effective among groups of subjects with relatively similar wealth levels (measured by land, livestock, and equipment ownership): their levels of cooperation increased dramatically in the communication rounds of the experiment. But this was not true of the

✓ **FACT CHECK** In a recent study of participation in church, local service, and political groups, as well as other community organizations providing local public goods by Alberto Alesina and Eliana La Ferrara it was found that participation in these groups was substantially higher where income is more equally distributed, even when a host of other possible influences are statistically "held constant."[8]

groups in which there were substantial differences in wealth among the subjects.

In one group, one of the wealthiest subjects tried in vain to persuade his fellow participants (who in real life were his tenants and employees) to restrict their withdrawals to the socially efficient amount, in order to maximize their total payoff. But the wealthy subject's advice fell on deaf ears.

"I did not believe Don Pedro," one of the less well-off women in his group later explained, "I never look him in the face." She was right not to trust him: Don Pedro (not his real name) had withdrawn the maximal amount despite his contrary advice to the other players.

This is not an isolated example.

- A study of water management in 48 villages in the Indian state of Tamil Nadu found lower levels of cooperation in villages with high levels of inequality in landholding. Moreover, lower levels of compliance were observed where the rules governing water supply were perceived to be chosen by the village elite.
- A similar study of 54 farmer-maintained irrigation systems in the Mexican state of Guanajuato found that inequality in landholding was associated with lower levels of cooperative effort in the maintenance of the field canals.[9]

In other cases, inequalities based on traditional hierarchies have made a positive contribution.

- Another study of Mexican water management, for example, found that increased mobility of rural residents undermined the top-down relationships that had been the foundation of a highly unequal but environmentally sustainable system of resource management.[10]

5.14 OVEREXPLOITATION OF A NON-EXCLUDABLE RESOURCE

We stated at the beginning of the chapter that we would illustrate the problem of the common pool resource problem by the example of just two people. This was despite the fact that, as a non-excludable resource, there would be *no limit* on the number of people who could, if they wished, fish on the lake and compete with Abdul and Bridget for the available fish.

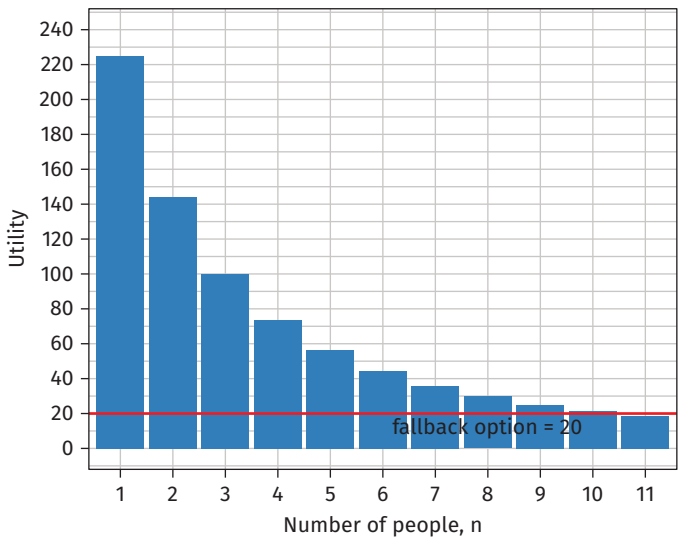
We have so far studied just one of the two aspects of the coordination failure resulting in over fishing: the fact both Abdul and Bridget fished more hours than was Pareto efficient. They both could have been *better off* had they been able to agree to fish *less*.

Now we introduce the second aspect of the problem: many more people could fish the lake. Because the lake is a common pool resource, its non-excludablity property means that there is open access.

How many people would use the lake? To answer the question we add more context to the initial problem. Abdul and Bridget are part of a large community of people who may make their livelihood fishing on the lake, or if not that, then doing some other kind of work yielding a utility of u_z (their fallback option). People will decide to make their living fishing as long as the utility they gain from fishing exceeds u_z.

In Figure 5.18 using the same values for the parameters as in the other numerical examples in this chapter, we calculate the maximum utility that could be attained by one person fishing alone, two (as in the case of Bridget and Abdul), three, and so on up to 11. However many there are fishing, they will receive the same utility in the Nash equilibrium because they

Figure 5.18 The dynamics of over exploitation of a common property (non-excludable) resource. All of the people who might fish on the lake have the same utility functions as Abdul and Bridget with the values of $\alpha = 30$ and $\beta = \frac{1}{2}$. The height of the bar for a given number on the x-axis is the utility of each of the fishermen when there are the indicated number fishing on the lake. The fallback utility is u_z = 20. You can see from the figure that if the lake is a common property resource, so that no one can be excluded, the Nash equilibrium number fishing on the lake is 10 with each receiving a utility of 21.2. If the 11th person fished on the lake, she—and all of the rest of the fishermen there—would receive a utility of 18.4, that is, less than their fallback option. The mathematics on which this figure is based are shown in M-Note 5.16.



are identical, and we have so far assumed that none has any advantage in bargaining with the others. The height of each bar is the utility attained.

When there are just two people fishing as in our previous examples involving Abdul and Bridget each receives a utility of 144, as we found in M-Note 5.7. The more people that fish in the lake, the lower the utilities each of them receive will be. When there are ten they all have a utility of 21.2, barely greater than their fallback options.

M-CHECK The equilibrium number of people fishing must be a whole number because the entry of a "fractional fisherman" would not make much sense unless we allowed people to split their day between fishing and the fallback option, which we do not.

Now think about some other member of the community who is *not* currently fishing but is thinking of doing so. Those fishing are doing better than the fallback options. But if the eleventh person decided to fish they would *all* receive a utility of 18.4 (including the new fisherman). That is, they would all receive *less* than their fallback options. So the eleventh person would decide *not* to fish.

Generalizing from this example, the Nash equilibrium number of people fishing is the largest whole number of people fishing such that the utility of those fishing is greater than or equal to the utility they *would* have at their fallback option.

As a result, the Nash equilibrium of this game is that we have:

- $n^N = 10$ the number of people fishing; and
- $h^N = 4.62$ the number of hours each of them works.

We use the N superscript for each of these quantities because *both* are Nash equilibria (but under different rules of the game):

- $n^N = 10$ is a mutual best response because none of those fishing could do better by *not* fishing, and none of those *not fishing* could do better by fishing; and
- given that ten people are fishing, then $h^N = 4.6$ is also a mutual best response because for each person fishing this is a utility-maximizing choice of hours, given the hours that everyone else is fishing.

The Nash equilibrium is Pareto inefficient for two reasons: too many people are fishing too many hours each. Just as was the case with Abdul and Bridget, if each fished a little less they all would be better off.

And if fewer of them fished, all ten of them could be better off. Figure 5.18 shows that if three people fished they would each have a utility of 100. Suppose the status quo were that just three people were fishing, and that there were another seven people who might consider joining them to fish in the lake. We call people who are already doing an activity, such as fishing or owning a firm, *incumbents*. We therefore call the existing three fishermen, the *incumbents* or *incumbent fishermen*.

Suppose this was the case, and that the incumbents could somehow agree to bribe the other seven *not* to fish. Notice we have just changed the rules of the game to allow the incumbent three to *coordinate*.

The incumbent three would have to give the other seven potential fishermen an amount of fish sufficient that each would be as well-off as their fallback. This amount would be $u^N - u_z$ or $21.2 - 20$, or 1.2 each. The total payments by the three to the other seven would be $7 \times 1.2 = 8.4$, leaving each of the three better off (each receiving $300 - 8.4)/3 \approx 97$). If the incumbent three *increased* the 'bribe' just a little bit then all ten would be better off than at the Nash equilibrium with open access.

M-NOTE 5.16 Nash equilibrium number of people fishing

Because access to the lake is open to all, the number fishing there will be the largest whole number (because we cannot have fractions of people fishing) such that the utility of those fishing is greater than the fallback option (their utility if they are not fishing in the lake), which is $u_z = 20$. To determine this number, we first derive $h^N(n)$ the hours of fishing that each will do as a function of the numbers fishing, and use this result to determine $u^N(h^N(n))$ the utility of those fishing as a function of how many there are.

To determine $h^N(n)$ we study the utility maximization problem of person 1, where h_i is the hours of fishing the ith fisherman:

$$\text{Vary } h_1 \text{ to maximize} \qquad u_1 = h_1\left(\alpha - \beta\sum_{i=1}^{n} h_i\right) - \frac{1}{2}h_1^2$$

$$= h_1\left(\alpha - \beta h_1 - \beta\sum_{i=2}^{n} h_i\right) - \frac{1}{2}h_1^2 \qquad (5.60)$$

To find the hours of fishing that maximizes the utility of person 1 we differentiate Equation 5.60 with respect to h_1, and set the result equal to zero. This gives us the first-order condition:

$$\alpha - 2\beta h_1 - \beta\sum_{i=2}^{n} h_i - h_1 = 0$$

$$\text{marginal benefit} = \alpha - 2\beta h_1 - \beta\sum_{i=2}^{n} h_i = h_1 = \text{marginal cost} \qquad (5.61)$$

Rearranging Equation 5.61 we get person 1's first-order condition giving the utility maximizing amount of fishing time:

$$(1+2\beta)h_1 = \alpha - \beta\sum_{i=2}^{n} h_i$$

$$h_1 = \frac{\alpha - \beta\sum_{i=2}^{n} h_i}{1+2\beta} \qquad (5.62)$$

All face the same first-order condition so in the Nash equilibrium all fish the same amount of hours: $h_1 = h_2 = \ldots = h^N$. Equation 5.62 becomes:

$$h^N = \frac{\alpha - \beta(n-1)h^N}{1+2\beta}$$

$$(1+2\beta)h^N = \alpha - \beta(n-1)h^N$$

$$h^N + 2\beta h^N + \beta n h^N - \beta h^N = \alpha$$

$$(1+\beta+\beta n)h^N = \alpha$$

$$h^N = \frac{\alpha}{1+\beta+\beta n} \qquad (5.63)$$

M-NOTE 5.17 Nash equilibrium number fishing (numerical example)

Using the analysis of M-Note 5.16, and as in the rest of the chapter, letting $\alpha = 30$ and $\beta = \frac{1}{2}$, we can illustrate the endogenous determination of the number of people fishing. Starting with Equation 5.63, you can verify that, if $n = 10$, then $1 + \beta + \beta n = 6.5$ and so $h^{\mathrm{N}}(10) = \frac{30}{6.5} = 4.62$. The utility of each fisher would be:

$$\begin{aligned} u^{\mathrm{N}}(10) &= 4.62 \times \left(30 - \frac{1}{2} \times 10 \times 4.62\right) - \frac{1}{2} \times 4.62^2 \\ &= 21.2 \end{aligned} \tag{5.64}$$

If one more enters and starts fishing so that $n = 11$, then the hours of fishing would be $h^{\mathrm{N}}(11) = \frac{30}{7} = 4.29$. The new utility of each fisher would now be:

$$\begin{aligned} u^{\mathrm{N}}(11) &= 4.29 \times \left(30 - \frac{1}{2} \times 11 \times 4.29\right) - \frac{1}{2} \times 4.29^2 \\ &= 18.37 \end{aligned} \tag{5.65}$$

But this ($n = 11$) cannot be a Nash equilibrium, because everyone—including the new entrant—would then be worse off than with the fallback option, $u_z = 20$. So the eleventh person would not enter (or if she did, others would leave). So the Nash equilibrium is $n^{\mathrm{N}} = 10$. This is illustrated in Figure 5.18.

CHECKPOINT 5.14 Pareto efficiency and open access

a. Explain why the open-access Nash equilibrium outcome with ten fishermen is *not* Pareto efficient. What alternative, if any, is Pareto superior to it?

b. Given your reasoning for (a), do you think there are alternative outcomes that are Pareto superior to, say, three fishermen bribing the other ten not to fish? Explain what the dynamics for the situations you describe would be? How many fishermen? How many hours spent fishing? And so on.

5.15 THE RULES OF THE GAME MATTER: ALTERNATIVES TO OVEREXPLOITATION

The new rules of the game allowing the incumbent three to bribe the others is just a thought experiment demonstrating that the Nash equilibrium with open access is *not* Pareto efficient. But commonly observed real-life rules of the game–like inequalities in bargaining power, cooperative management of the lake, or private ownership–could also address the overfishing problem.

TIOLI bargaining power

To see that the institutions governing the interactions among them matter, think about the case in which one of the ten people fishing on the lake has the power to make a take-it-or-leave-it offer to all the rest. Here we have

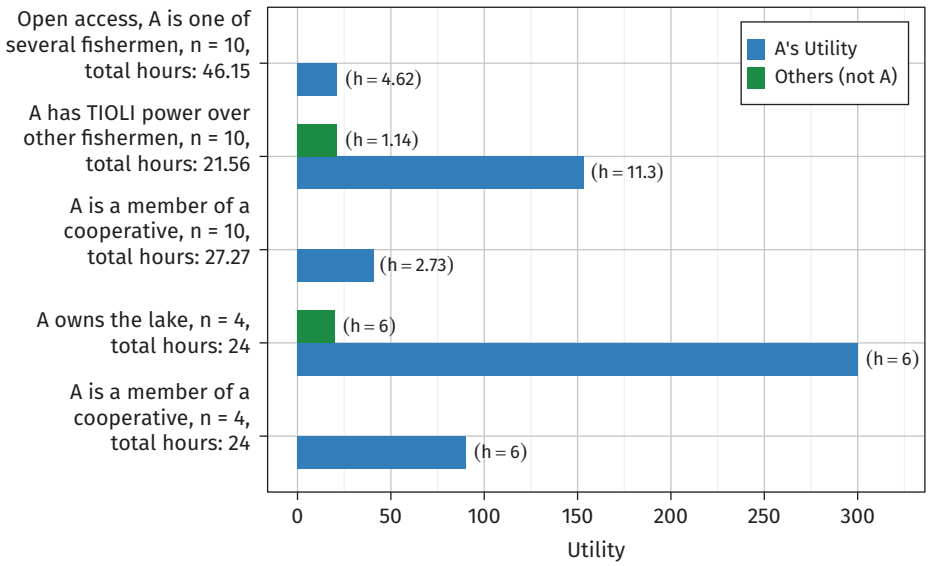
changed the rules of the game by giving one of the fishermen TIOLI power (which allows a kind of coordination). But the lake is still open access, so there are ten fishermen there.

The one with bargaining power—suppose it is Abdul—can now say to the others "each of you will fish x number of hours, and I will fish as many hours as I wish." This is the "take it" part of the offer. The "leave it" part is: "and if you refuse, then I will return to fishing 4.61 hours." That is, return to the former Nash equilibrium hours that occurred when there was no coordination among the fishermen.

The other fishermen would know that without coordination the best they could do is to all fish 4.62 hours, gaining a utility of 21.2. This is the others' fallback option to the TIOLI offer. If accepting Abdul's offer made them worse off than their fallback they would refuse, and just fish 4.62 hours. This is their participation constraint; if it is violated—so that they would receive a utility of less than 21.2—the others will not accept ("participate in") Abdul's offer.

Abdul would know, therefore, that he needs to find the hours of all ten of them (his and the rest) that maximize his utility subject to the participation constraint on the minimum utility the others can receive. Figure 5.19 shows the offer Abdul would make, and the utility that he and the others would

Figure 5.19 The rules of the game: Cooperation, bargaining power, and ownership. The bars show the utility of the fishermen (it is identical for all fishermen in the first and third row). The numbers at the end of the bars show the hours fished, where h is person indicated in the figure label on the left.



experience. The first row of the figure shows, from the previous figure, the result for the unlimited access case *without* coordination.

When Abdul has TIOLI power the other fishermen work fewer hours (just 1.14 each, rather than 4.62 before), but get exactly what they had under the uncoordinated open-access case. This is so because that level of utility—21.2—is the participation constraint on what Abdul can offer them.

Abdul himself works 11.30 hours and enjoys utility equal to 153.2. Notice that the total number of hours is reduced sharply compared to the uncoordinated Nash equilibrium: from 46.15 to 21.56 hours. This reduction in total hours is the reason why the others are able to fish less but still attain the same utility: they catch more fish in an hour due to the lesser total hours of fishing. This is another case in which inequality in power allowed coordination, reducing the inefficiency of the case when there were no differences among the fishermen.

A democratic fishing cooperative

An entirely different set of rules of the game—a democratic **cooperative** of the fishermen—would implement a correspondingly different set of results. Suppose that *none* of the ten fishermen has any bargaining power advantage and that they jointly own the lake. They can decide jointly—democratically by unanimous consent—on the same number of hours that each of them will fish.

To figure this out they would think in the same way the Impartial Spectator did when she maximized total utility. They will maximize the *sum* of their utilities because this will also maximize the utility of each fisherman. The result is shown on line 3 of Figure 5.19. When there are ten fishermen, they would each fish 2.73 hours and attain utilities of 40.9 each. Because their utility as co-op members is now double their fallback option, others who are experiencing the fallback utility of $u_z = 20$ would wish to join the cooperative. But it might be difficult to persuade the members to admit others, as this would reduce the utility of the incumbent fishermen.

Their total fishing hours (27.3) is substantially greater than under the TIOLI power of Abdul, and so is their total utility (409.1 compared to 344.9).

We can conclude two things from this last fact:

- Suppose Abdul still had TIOLI power. If the other fishermen could coordinate their actions, they could "bribe" Abdul to give up his bargaining power and join their cooperative; they could have offered him the 153.2 that he received under his TIOLI power and still be better

COOPERATIVE A cooperative is a business organization or other association whose members together own the assets of the organization; they share the income resulting from their activities and jointly determine how the organization will be run (possibly through the democratic election of a manager).

off dividing up the rest of their utility (fish) among themselves. They would each receive (409.1 – 153.2)/9 = 28.4, far better than the 21.2 they had when Abdul had bargaining power. So shifting from Abdul holding TIOLI power to the democratic cooperative (with the bribe for Abdul) is a *Pareto improvement*.

- The reason why this is the case is that under Abdul's TIOLI power they were as a group *under*exploiting the fishing stocks. Abdul forced them to do this because the less the other people fished, the more fish he could catch, and that was the only way he could increase his utility.

The reason why the cooperative's decision results in a greater total utility than the TIOLI case is that the members of the cooperative were pre-committed to *sharing* the total utility equally. And so they each had an interest in making total utility as large as possible.

Things would have been very different if Abdul had had the power to *take* some of the fish caught by the others (as in the employment and fee cases we dealt with earlier). In this case he would have done the same as the cooperative. He would have implemented the fishing times that maximized total utility. And then he would have taken fish from the others, leaving them just enough fish so that they did not decide to stop fishing. The TIOLI allocation was inefficient because Abdul's bargaining power was limited.

The TIOLI case was not inefficient because Abdul had some bargaining power. It was inefficient because he did not have enough power. As you will suspect from the two-person case studied earlier, the allocation would have been Pareto efficient if Abdul had had all of the powers of a private owner of the lake. We now show this.

! REMINDER In Chapter 1 (section 1.17), we pointed out that a shortcoming of the Nash equilibrium concept is that it requires only that a *single individual acting alone* could not benefit by adopting a different strategy given the strategies adopted by others. It does not require that a *group of individuals acting jointly* could not benefit by changing their strategies. The fishing cooperative illustrates this shortcoming: if the fishermen can jointly manage the lake then they could also have together bargained with the TIOLI actor as a group. For example the other nine could have done better than taking the TIOLI offer if they together decided to form a cooperative then, and as described here bribed the TIOLI actor to join them.

Private ownership

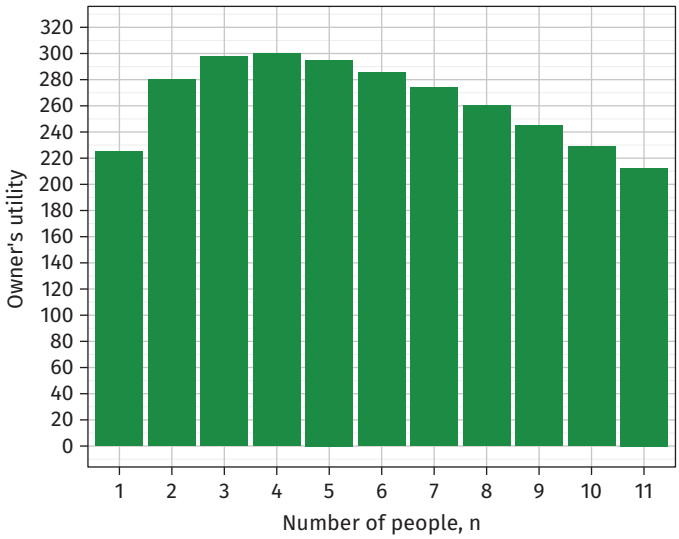
Under these new rules—private ownership of the lake—the lake is no longer a common property resource because ownership means that Abdul can exclude anyone he wishes from fishing. Abdul would make three decisions:

- How many other fishermen should I *allow* to fish in the lake?
- How many *hours* should I allow them to fish?
- If I *employ* them, then what *wage* should I pay them? Or if I charge them a fee for fishing, how large should the fee be?

You know how to answer the second and third questions from the case earlier in the chapter when Abdul was the owner with just one other person Bridget on the lake.

The first question is similar to that asked in Figure 5.18 but the answer is very different. The number of people fishing on the lake is no longer based on the fishermen's own decisions about where they can make a better

Figure 5.20 Utility of the owner when the lake is privately owned. On the horizontal axis are the total number of people fishing in the lake, including the owner. So, for example, where $n = 2$ we have Abdul as the owner and there is one other person, Bridget, the case we analyzed earlier in this chapter. The height of each bar is the utility gained by the owner of the lake when he can both determine how many people fish there and dictate the terms under which they will work (as long as they receive utilities no worse than their fallback positions).



livelihood. This is not their decision to make. The *owner* determines the number of others so as to maximize his utility.

How he would do this is shown on line 4 in Figure 5.19. Abdul will allow three other fishermen to access the lake (so that means $n = 4$ including himself). Going back to Figure 5.19, remember there are now just four fishermen, *not* ten as before. All four fishermen fish six hours with the owner receiving a utility of 300 and each of the others receiving the same utility as their fallback option, that is, 20.

The last line in Figure 5.19 shows what happens if Abdul is not the owner and instead if all four of the fishermen were members of a democratic cooperative. The members of the cooperative would implement exactly the same allocation of work time as occurred under private ownership: six hours of work time each. But the distribution of utility would be radically different, each of the four would receive 90.

We know that each of the four working six hours is the allocation that maximizes total utility. The reason why private ownership of the lake implements this outcome is that the owner is limited only by the participation constraints of the others, and this is a constant (their fallback position of 20). So he implements an allocation to maximize the total utility, from which he must subtract the amounts required to keep the three others "participating."

In sum, we can say the following:

- Open access leads to a Pareto-inefficient overfishing outcome in which all the fishermen receive the same utility.
- TIOLI power leads to a highly unequal and Pareto-inefficient underexploitation of the lake.
- Private ownership implements a Pareto-efficient and highly unequal outcome.
- A democratic cooperative implements a Pareto-efficient and equal outcome.

CHECKPOINT 5.15 **Choose your game** Compare the outcomes under two different rules of the game: a democratic cooperative of fishermen and private ownership of the lake. How and why do the outcomes differ?

5.16 CONCLUSION

In practice, none of the approaches to addressing the common property resource coordination problem could be expected to work perfectly as:

- no government is likely to have the information about the people's preferences, production functions, and fishing times necessary to implement Pareto-efficient fishing levels by fiat, or to design the optimal taxes that would achieve the same result;
- private owners face some of the same problems as government due to lack of information, and moreover, while private ownership of a small lake or other common property resource is conceivable, for many important common property resources private ownership is infeasible (owning the oceans or the atmosphere for example) or undesirable (think of the unaccountable power that a private owner of such a vast common property resource could wield); and
- altruism toward close family and loved ones will lead us to take at least some account of the effects of our actions on their well-being, but we are less likely to know or care deeply about how our actions affect total strangers or even yet unborn generations who will benefit or suffer the external effects of what we do.

The conclusion is *not* that the approaches to addressing coordination problems introduced here are ineffective. Variants of each approach can contribute to making economic outcomes more efficient and fair.

In contrast to the vast diversity and complexity of rules that we observe, the models we introduced simplify the rules of the game that regulate how we interact with each other in exploiting a common property resource. What the models have done is *not* to represent the world as it is, but to identify key aspects of *how the world works* to provide a lens for understanding them better.

✓ **FACT CHECK** Elinor Ostrom and her colleagues' field research in different parts of the world from Colombia to Switzerland uncovered 27 different local rules for excluding others from access to common property resources. These were based on such things as residency, age, caste, clan, skill level, continued use of the resource, and use of a particular technology.[11]

MAKING CONNECTIONS

Social interactions and external effects: Economists study buying and selling in markets, but we also study nonmarket interactions, sometimes called 'social interactions' ("social" here means simply nonmarket). The social interactions studied here include an external effect: the fact that one person's fishing reduces the catch of another, and this effect is not taken into account when each of the fishermen decide how many hours to fish.

Public goods, common pool resources, and club goods: All of these have the property such that each person's actions have *external effects* on others, and, in the absence of social preferences or policies that internalize the external effects, these are not taken into account when people decide how to act, resulting in outcomes in which some potential mutual gains remain unexploited.

Policy: Government policies and institutions may be designed so that people take account of *external effects* when they act. An example is a *tax* on fishing that imposes on anyone fishing the marginal costs that their fishing imposes on the others inducing each to choose their hours of fishing as if they cared about these external effects borne by others.

Property: Converting a *common property resource* into a *privately owned resource* may result in a *Pareto-efficient outcome* in which the owner captures all of the potential *mutual gains (rents)*.

Power: When a single person has all of the bargaining power and so can make a binding *take-it-or-leave-it offer*, they may implement an outcome in which there are no unrealized mutual gains, and all of these gains (*rents*) go to the powerful person. Lesser forms of power—to commit to a particular fishing time, to which the other must respond, for example—advantage the powerful and result in inefficient outcomes.

Mutual benefits from coordination and conflicts over their distribution: Policies to address coordination failures differ in how the resulting rents are distributed; the resulting conflicts may make it difficult to agree on any policy.

Inequality: Differences in wealth, political connections, and other sources of power can be both a source and a consequence of inefficient and unfair outcomes among people facing coordination problems. In some cases, these differences can also mitigate the inefficiencies arising from coordination failures.

Models and relevance: Models, we wrote in Chapter 3, are like maps—a simplified guide to the territory, not the territory itself. But the model of social interactions introduced here, though quite abstract can be directly applied to very concrete economic actions such as firms competing for customers and, suitably extended, can illuminate *global social interactions* and *coordination problems* such as climate change and the spread of epidemic diseases.

IMPORTANT IDEAS

utility	private property	interdependence
common property resource	excludability	reciprocity
rivalness	altruism	decentralized
Impartial Spectator	asymmetrical interactions	implementation
symmetrical interactions	first mover	public good
TIOLI power	government policy	sequential game
$mrs = mrt$ rule	$mrs^A = mrs^B$ rule	fallback
perfect	participation constraint	social preferences
altruist	best-response function	common property resource
incentive compatibility constraint	private good	Mechanism Designer
club good	repeated game	democratic cooperative
Grim Trigger strategy	external effect	
disutility	coordination failure	

MATHEMATICAL NOTATION

Notation	Interpretation
h	fishing times
α	parameter regulating the productivity of fishing times
β	external effect of fishing time on the other's productivity
$u(\ )$	utility function
$v(\ )$	value function expressing an altruistic concern for the utility of another person
W	Impartial Spectator's social welfare function
w	wage in the employment solution
F	permit fee in the permit solution
τ	per unit tax in the government policy solution
a, b, c, d	payoffs in the repeated interactions game
λ	extent of altruism (valuation of the other's utility relative to one's own)

Note on superscripts and subscripts: A and B: people; N: Nash Equilibria; i: outcome selected by the Impartial Spectator; F (as a superscript): outcome with a first mover; λ (as a superscript): the case with altruistic actors; z: fallback (reservation) option.

"Enter the Royal [Stock] Exchange of London, that place more respectable than many a court; you will see there agents from all nations assembled for the utility of mankind. There the Jew, the Mohammedan, the Christian deal with one another as if they were of the same religion. There the Presbyterian confides in the Anabaptist, and the Churchman depends on the Quaker's word . . . They give the name infidel only to those who go bankrupt."

Voltaire,

Mélanges (1734)[1]

PART 2

MARKETS FOR GOODS AND SERVICES

When you hear the word "market" what other word do you think of? "Competition" probably is what comes to mind. And you would be right to associate the two words.

But you might have also come up with "cooperation." That is what impressed Voltaire about the London Stock Exchange: mutually advantageous interactions, even among total strangers "from all nations assembled for the utility of mankind." Markets allow us, each pursuing our private objectives, to work together producing and distributing goods and services in a way that, while far from perfect, is in many cases better than the alternatives. Markets accomplish an extraordinary result: unintended cooperation on a global scale, although often with a highly unequal distribution of the benefits.

To better understand what markets do and how they work, begin with two everyday facts: we acquire skills as we produce things and, for this and other reasons, producing a lot of the same thing is often more effective in terms of time and other inputs per unit than producing just one or a few of many different things. This is called learning-by-doing.

Because of learning-by-doing and other advantages of large-scale production, people do not typically produce the full range of goods and services on which they live. Instead we specialize, some producing one good, others producing other goods, some working as welders others as mothers, teachers, or farmers.

There are huge advantages to this pattern of specialization—called the division of labor. Those who are naturally better at some task, or have learned to be good at it by experience, or are in an environment in which it can be most productively done can devote themselves entirely to what they are relatively good at.

This is part of the explanation of why as a species we are so productive. The limited number of species that have adopted a highly developed division of labor—humans, ants, and other social insects, for example—have outcompeted other species. The total biomass of humans and the livestock we have domesticated, for example, is estimated to be 23 times the weight of all the other mammals on earth. And throughout most of human history the biomass of ants—one of the most cooperative of species—has exceeded that of humans by a considerable amount.[2]

But the division of labor poses a problem for society: Once they are produced by specialized labor, how are the goods and services to be distributed from the producer to the final user? In the course of history this has happened in a number of distinct ways from direct government requisitioning and distribution as was done in the US and many economies during World War II to gifts and voluntary sharing as we do in families today and was practiced among even unrelated members of a community by our hunting-and-gathering ancestors.

In a modern capitalist economy, the institutions that govern how the goods and services are distributed from producer to user include markets, firms, families, and governments. In Part II of our book we study markets and the actors who make up markets: the owners (and managers) of firms and other people.

To understand how markets facilitate specialization in Chapter 6 we study the production process and how the division of labor and the exchange of products can be advantageous to all concerned. Then to understand the workings of markets we explain how individuals' valuation of goods and services is expressed in market demands (Chapter 7). Then, along with these market demands, we explain how firms' costs of production are expressed in their owners' and managers' decisions about how much to produce and supply to the market (Chapter 8).

We then study the process of competition among sellers and buyers, each seeking to enlarge their share of the mutual gains made possible from the division of labor and exchange. And we show how this so called rent-seeking process affects the movement of prices and the quantities produced (Chapter 9).

Taking these four chapters as a whole poses a tension that can be expressed by the following contradiction:

- The models and evidence on the advantages of large-scale production provide a reason for why we specialize.
- Competitive markets are essential to the processes by which specialization can be organized in ways that allow the mutual benefits of the division of labor to be widely shared, as Voltaire said "for the utility of mankind."
- But the advantages of large-scale production can also promote the emergence of giant firms and a winner-take-all process that appears to be making markets *less* competitive.

Making market competition sustainable given the advantages of large-scale production will have to be addressed by public policy.

CHAPTER

6

PRODUCTION

TECHNOLOGY AND SPECIALIZATION

The division of labor is limited by the extent of the market.

Adam Smith,
The Wealth of Nations (1776)[1]

DOING ECONOMICS

This chapter will enable you to:

- Explain how learning-by-doing and economies of scale are reasons for the division of labor and specialization.
- Understand how markets allow specialization according to the principle of comparative advantage.
- See how, in the presence of economies of scale and learning-by-doing, an economy can benefit by specializing, but also may specialize in ways that perpetuate its low income as a result of a poverty trap similar to some of the coordination failures studied earlier.
- Manipulate some commonly used production functions to study marginal and average products of labor and capital goods and derive a production possibilities frontier.
- Describe the main dimensions on which technologies differ—the extent of substitution among inputs, productivity, factor intensity, and economies of scale—and how these are represented in different production functions.
- Understand how the owners of a firm can determine a set of inputs and a way of combining them to produce output that will minimize the costs of a given level of output.
- See that owners of a firm will try to innovate to reduce the inputs required to produce a given output and therefore lower costs and receive innovation rents (at least until the competition catches up).

6.1 INTRODUCTION: DREAM LIFTERS

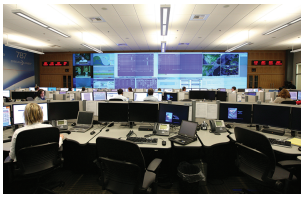
A technician glances quickly from one to the other of her three monitors. Around the huge room many other technicians do the same. Occasionally a technician looks up at gigantic blue video screens on which news reports flash, international weather reports display, and flight conditions stream live. Twenty-four hours a day, translators stand ready to facilitate conversations in 28 languages.

What is this command center?

It's the Production Integration Center that coordinates the global production of the Boeing 787 Dreamliner four stories above the production floor at the company's plant in Everett, Washington, US. There the superjumbo airplanes are being assembled from components being flown in from around the world: parts of the wing from Japan, wing tips from South Korea, the center fuselage and the horizontal stabilizer from Italy, passenger doors from France, cargo doors from Sweden, landing gear from the UK, and the list goes on and on. In 2015, Boeing contracted with over 26,000 suppliers around the world.[2]

Figure 6.1 Boeing's Production Integration Center in Everett, Washington, US.

Photo by Kevin P. Casey/Bloomberg via Getty Images.



Boeing selected Rolls-Royce, Mitsubishi, Saab, Fuji, and other companies to design and build the components because they were—in Boeing's estimation—simply the best companies to do the job, anywhere in the world. The Japanese companies were global leaders in aircraft construction. The Italian partner Alena had critical intellectual property rights (patents) that Boeing would otherwise not have had access to.

Boeing modified four 747-400 aircraft—renaming them "Dreamlifters"—to deliver the wings, body, and other parts of the plane to Everett where American machinists and others assembled the planes. In the Production Integration Center above them the engineers at the Production Integration Center kept minute-by-minute track of the movement of the components around the world.

6.2 THE DIVISION OF LABOR, SPECIALIZATION, AND THE MARKET

Boeing's globally integrated Dreamliner production process illustrates an important economic idea: the *division of labor*.

The division of labor is an expression for the fact that people, organizations, or geographical regions specialize in particular tasks or the production of a limited range of goods or services. For Boeing, purchasing components of the Dreamliner from hundreds of specialized firms was more cost-effective than producing the entire plane in-house at their plant in Everett, Washington.

There are two consequences of specialization.

The first is *increased productivity*. The specialization allowed by the division of labor increases productivity for three reasons:

- *Comparative advantage*: Specialization enables people, firms, and regions to focus on the tasks and products that they are comparatively good at (we will take up comparative advantage in section 6.6).
- *Learning-by-doing*: People learn better ways of working by developing individual skills and discovering better ways to organize production among members of a team. (Figure 8.2 (b) in Chapter 8 provides a dramatic example of learning-by-doing.)
- *Economies of scale*: By allowing the production of a few things on a large scale rather than many things on a small scale, specialization raises the amount of output that is available for a given amount of inputs. (Figure 8.6, also in Chapter 8, presents some physical evidence for economies of scale based on engineering studies.)

A second consequence of specialization is the *need for integration*. The advantages of the division of labor can only be realized if there are institutions to *coordinate* the many distinct production activities that take place when people specialize. The Boeing example illustrates the need for integration: somebody has to put the parts together to produce a Dreamliner.

This can be summarized: production is *specialized*, but the use of goods and services is *generalized*. Specializing in consumption is not biologically sustainable. As a result, the goods and services somehow have to get from specialist producers to generalist users.

To grasp the scope of this problem, imagine a 3D map of the world showing the amounts of goods produced annually in each location. There would be a hundred or so Dreamliners piled up in Everett, Washington, billions of square meters of cloth stacked in Bangladesh and other textile-producing countries, well over a billion barrels of oil piled over tiny Kuwait, mountains of computer components and other consumer electronics rising from coastal China where Dell (the computer company) is located, and so on.

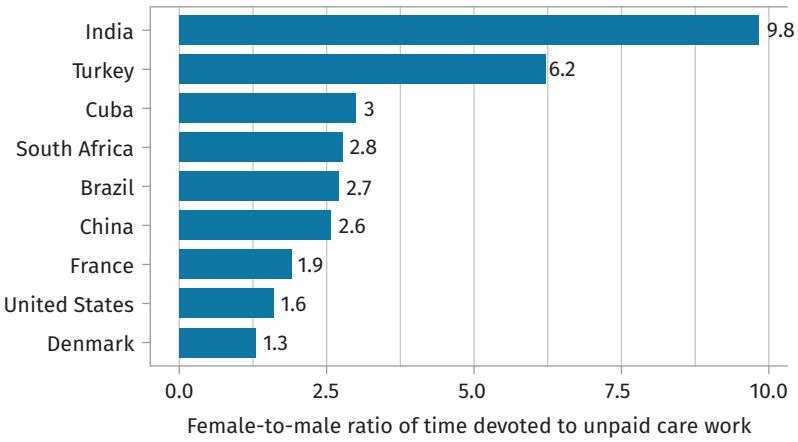
Now imagine the same map, but showing where all of the goods are *used*. The second map would be different from the first in two ways:

- it would be much flatter, the goods would have been spread around to the entire population of the world; and
- in any location there would be an assortment of a great many products, not towering stacks of a single product.

The coordination of specialized producers and generalist users is accomplished by a set of institutions that differ in importance both over time and across the economies of the world today. These include:

- *Market exchanges*: Selling the goods that specialized producers have made provides the budget for purchasing the general market basket of goods and services we live on.
- *Government acquisition and provision:* Publicly provided services are based on the integration of the specialized producers' goods and services to provide education, security, and other government-provided services to generalist users of these services.
- *Families and other face-to-face communities*: Many families exhibit a division of labor by age and gender: adult women, for example, bearing children and spending more time on raising them and caring for other family members relative to men (for example preparing meals). The goods produced and tasks performed by adult men and women and by children are shared within the family or some other larger consumption unit. Figure 6.2 shows the international variation in this care work.
- *Business organizations*: Boeing's own employees work at highly specialized tasks; owners and managers direct both the assignment of tasks and the integration of the results of the work of Boeing's employees to construct airplanes (along with the parts supplied by specialist workers in other companies).

Figure 6.2 Female-to-male ratio of time devoted to unpaid care work. Unpaid care work refers to all unpaid services provided within a household for its members, including care of persons, housework, and voluntary community work. A ratio of 1 would mean equal hours spent on care work. In India, for every hour of care a man does, a woman on average does 9.8 hours of care work. In Denmark for every hour of care work a man does, a woman on average does 1.3 hours of care work. Data are from 2014 and measured for a representative sample of 15- to 64-year-olds, except for China where the data are for 15- to 74-year-olds.
Source: OECD (2014).



This chapter's head quote by Adam Smith tells us that markets play a critical role in allowing the division of labor to expand to global proportions, leading to ever greater specialization. Here in Part II of this book and also in Chapter 14, we explain how markets work to coordinate the division of labor. We begin with the production process, and some aspects of it that favor *specialization*.

M-CHECK To understand the marginal product curve in the diseconomies of scale case remember that $d(ln(z))/dz = 1/z$

CHECKPOINT 6.1 Specialized production and generalized consumption
Explain in your own words what it means to say that production is specialized but the use of goods and services is generalized.

6.3 PRODUCTION FUNCTIONS WITH A SINGLE INPUT

! REMINDER A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand (the arguments of the function) and the maximum quantity of output.

In Chapter 3, we used information on the way that Keiko's study time translated into her learning to ask how much time she will choose to study. In Chapter 5 information about the relationship between fishing time and the amount of fish caught was a key idea in posing and then addressing the common property resource coordination problem. In both cases we were using production functions.

To better understand the division of labor and specialization we now need to look more carefully at the properties of production functions. Think about another person, Alex, who has to choose how much of his time to spend producing one or more goods. Alex can devote more or fewer hours of labor (l) to production and he can observe how his output (x) varies as he changes the number of hours he works. Alex may use a computer and a desk, or a given plot of land and farming equipment, or other inputs, but for now we assume that the only input is his labor time.

The relationship between the input of his labor and the output of the goods is described in a production function, $x = f(l)$, taking the form:

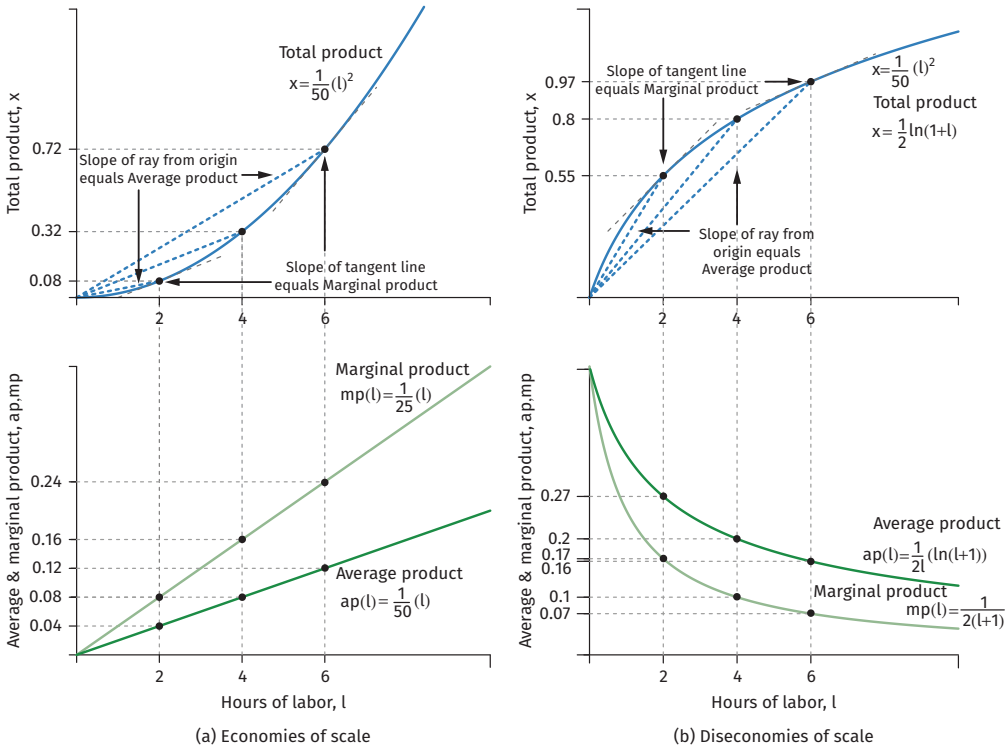
$$x(l) = ql^{\alpha} \tag{6.1}$$

The exponent α measures the responsiveness of output to a change in the hours of labor. The positive constant q measures the overall productivity of the production process, which will be greater the more skilled or hard-working Alex is (or what we could call the *quality* of his labor hours). The top panel in Figure 6.3 (a) illustrates this production function (equation 6.1) with $q = \frac{1}{50}$ and $\alpha = 2$.

The top panel in 6.3 (b) shows a different production function: $x(l) = \frac{1}{2}ln(l+1)$ (remember to read "ln" as "the natural logarithm of"). In the top portion of both panels more hours of labor result in more output, but the panels differ in *how much* output is obtained for given inputs.

- **Economies of scale** (Figure 6.3 (a)): when Alex doubles all of the inputs—in this case that means just his labor input—the output *more than doubles*.
- **Diseconomies of scale** (Figure 6.3 (b)): when Alex doubles his labor input (assumed to be the only input), the output *less than doubles*.

Figure 6.3 Production functions with economies and diseconomies of scale. With *economies of scale* (panel (a)) doubling labor input more than doubles output, as can be seen by going from two hours of labor to four hours. Average product and marginal product increase with additional labor input as you can see from the slope of the production function (the marginal product), which *steepens* as the labor input increases and the slope of the ray from the origin to a point on the production function (the average product), which also steepens. The average and marginal products given by these slopes in the upper figure are shown in the lower figure of panel (a). With *diseconomies of scale* (panel (b)), the opposite occurs: the average and marginal products of labor are both *decreasing*.



ECONOMIES OF SCALE When production exhibits economies of scale, multiplying all inputs by some number greater than one increases output by a factor greater than that number.

DISECONOMIES OF SCALE When production exhibits diseconomies of scale, multiplying all inputs by some number greater than one increases output by a factor less than that number.

The term **constant returns to scale**, not shown in the figure, refers to the case where when inputs double, output doubles, so the production function is just a straight line (as shown in the lower right panel of Figure 6.4).

In the lower figures of both panels we show two important statistics describing aspects of the two production functions in Figure 6.3. The *ratio* of the amount of output to the amount of the input involved in producing the output is the **average product** of that input (also called average productivity). Average product is measured by the *slope* of a line from the origin (called a ray) to a point on the production function. The *bottom* panels in Figure 6.3 (a) and (b) show the average product of labor associated with the production function shown in the *top* panels of the figures. When production exhibits economies of scale, average product *increases* as the scale of production increases through an increase in inputs.

The ratio of the increase of output to a small increase in labor input is the **marginal product** of labor (also called marginal productivity).

M-NOTE 6.1 **The average and marginal product**

Here we summarize the concepts of total product, average product, and marginal product. Because the marginal product is the slope of the production function, it is also the derivative of the production function with respect to a particular input, e.g. for a production function using only labor, $x_l = \frac{df(l)}{dl}$ is the marginal product.

If there is just a single input, labor, and the marginal product of labor is greater than the average product of labor, then the average product must increase as more labor is used. We use Equation 6.1 to illustrate why this must be the case.

We start by calculating the average product (ap) and marginal product (mp) of the production function $x(l) = ql^{\alpha}$:

$$\text{Average product:} \qquad ap = \frac{x(l)}{l} = \frac{ql^{\alpha}}{l} = ql^{\alpha-1} \tag{6.2}$$

$$\text{Marginal product:} \qquad mp = \frac{dx(l)}{dl} = \alpha ql^{\alpha-1} \tag{6.3}$$

To analyze how the average product changes as more labor is put into production, we calculate the derivative of the ap function with respect to labor hours (l):

$$\frac{dap(l)}{dl} = (\alpha - 1)ql^{\alpha-2} \tag{6.4}$$

continued

CONSTANT RETURNS TO SCALE In the case of constant returns to scale, increasing all inputs by some constant increases output proportionally.

AVERAGE PRODUCT The average product of an input is total output divided by the total input.

MARGINAL PRODUCT The marginal product of an input is the change in total output associated with a small change in the input.

If $\alpha > 1$, Equation 6.4 is *positive*, which means that the ap *increases* with more labor as in Figure 6.3 (a). If $\alpha < 1$, it is *negative*: the ap is *reduced* if we add labor shown in Figure 6.3 (b).

In summary:

- If $\alpha > 1$, then $mp > ap$, and so $\frac{dap(l)}{dl} > 0$ and
- If $\alpha < 1$, then $mp < ap$ and so $\frac{dap(l)}{dl} < 0$

CHECKPOINT 6.2 **Production and labor inputs** Consider a production function: $x(l) = 10l^{0.5}$:

a. Sketch the production function.

b. Calculate $ap(l)$ and $mp(l)$ if $l = 9$. Sketch the functions.

c. Does the production function exhibit *economies* or *diseconomies* of scale?

6.4 ECONOMIES OF SCALE AND THE FEASIBLE PRODUCTION SET

Suppose that Alex can spend his time fishing and making shirts in some combination, including complete specialization (spending all of his time on one or the other). He prefers to have more of *both* shirts and fish: both are goods. He needs at least some of each to survive. His labor time, as in Chapter 5 is a "bad" but we will set aside his choice of total hours of work by saying that he can work any amount up to 10 hours a day, and that given how productive his labor is and how much he values the goods, he will choose to work the full ten hours.

As a result, the more time Alex devotes to producing one good, the more of that good he will have, but because Alex's time is limited, the less he can produce of the other good. Therefore the opportunity cost of more fish is the amount of shirts he will have to forego if he shifts his work time from shirt-making to fishing. In order to pose the question—How much time will he spend on each?—we need two pieces of information:

- the *feasible set* of combinations of fish and shirt amounts that are available to him, given his labor time and the production functions at his disposal; and
- the *indifference map* representing his valuation of each of the combinations of the two goods.

We derive the feasible set in this section and introduce the indifference map in the next.

The feasible set with economies of scale

We will assume that there are *economies of scale* in shirt-making (as is common in manufacturing processes; see Chapter 8) but *constant returns to scale* in fishing. This means that the fish production function is linear: it is a straight line. As a result for fishing both the average and marginal products of labor are constant and equal to each other and also equal to the slope of the blue line through points **f** an **c** in Figure 6.4. In Figure 6.4 we derive Alex's feasible set of fish and shirts (shown in the upper-right quadrant) based on:

- the *total amount of time he will work* (shown in the lower-left quadrant of the figure); and
- the *production functions* for fish and shirts (shown in the lower-right and upper-left quadrants, respectively).

In Figure 6.4, the horizontal axis to the left of the origin represents positive amounts of labor devoted to making shirts, and the vertical axis below the origin is positive amounts of labor devoted to fishing.

Alex needs to make a choice between three different ways to allocate his time in production to two types of output (fish and shirts):

a. allocate ten hours of work to producing only shirts;
b. allocate some hours of work to producing shirts and some to fish; and
c. allocate ten hours of work to producing only fish.

Option (a) is shown as point **a** in the figure where Alex dedicates all of his ten hours of work time to producing shirts. We extend a line up to his production function for shirts and see that ten hours of labor results in Alex producing ten shirts and no fish as a result of dedicating all his labor to shirts.

We could follow the same process for option (c) corresponding to point **c** in the figure. He would produce five fish and no shirts by dedicating all his time to fishing.

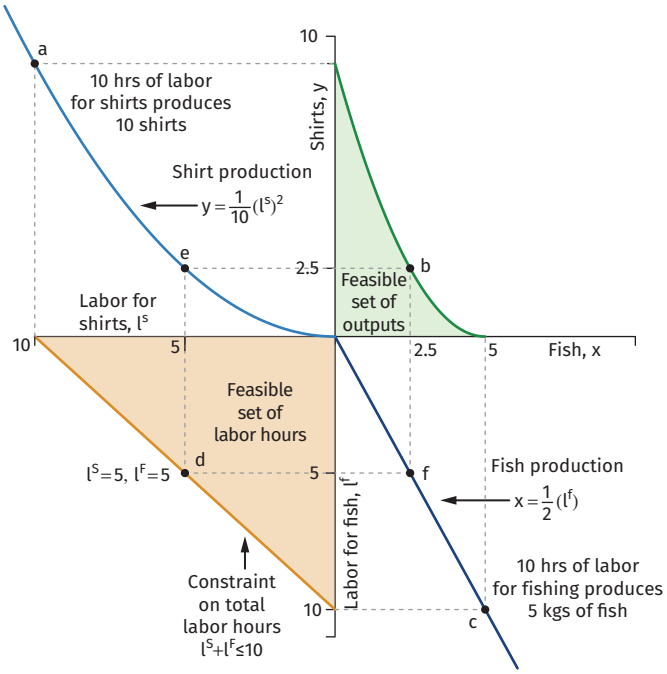
Option (b) (that is, point **b**) on the other hand, shows what Alex would produce by dedicating half his time to shirts and half to fishing. Because production in fishing is *linear*, if he dedicates five hours, he simply gets half of what he would have produced at ten hours (2.5 fish). But, because there are economies of scale in shirt production, from dedicating half his time to shirts, he only gets a quarter of the output relative to ten hours of labor for shirts (2.5 shirts vs. ten shirts).

The top-right quadrant of Figure 6.4 illustrates an important consequence of economies of scale. The result is that Alex's production possibilities frontier is bowed *inward* toward the origin. This reflects the fact that, with economies of scale in one or both production functions, dividing your work time between the production of both goods is not as beneficial (it

! REMINDER The average product is the slope of a ray from the origin to a point on the production function, and the marginal product is the slope of the production function. So if the production function is just a straight line, both of these are equal and do not vary as more labor is devoted to production.

M-CHECK The production possibilities frontier with economies of scale shown is *convex to the origin*, that is, it is *bowed in* to the origin. With diseconomies of scale, the production possibilities frontier is *concave to the origin*; that is, it is bowed out.

Figure 6.4 Deriving the production possibilities frontier with economies of scale. The lower-left quadrant shows the constraint: a given number of hours of labor is available. The upper left and lower right show how the available labor can produce shirts and fish respectively. In the upper-left panel labor hours are measured from right to left; we illustrate an economies-of-scale production function; similar to panel (a) in Figure 6.3, but just rotated clockwise 180 degrees. In the lower-right panel, labor hours are measured from the top down and output from left to right. The production function has constant returns to scale. Points **d**, **e**, **f**, and **b** illustrate the production of shirts and fish that are possible if the labor time is divided equally among the two sectors.



is closer to the origin) than devoting *all* your time to just one or the other. But specializing in one good works only if there is some way of acquiring the other good through exchange.

The (negative of the) slope of the **production possibilities frontier** shows the opportunity cost of acquiring more fish by shifting labor from shirt-making to fishing, in terms of the amount of shirts that must be foregone

PRODUCTION POSSIBILITIES FRONTIER The production possibilities frontier (PPF) for two goods shows the maximum amount of one good that can be produced given the output of the other good. The production possibilities frontier is the boundary of the producer's feasible set and is an alternative name for the feasible frontier when we study production.

as a result. This is the marginal rate of transformation of shirts into fish accomplished by shifting labor time from shirt-making to fish-catching. With economies of scale, as Alex shifts his labor input from producing shirts to producing fish, $mrt(x,y)$ *declines*, so he gives up smaller and smaller amounts of shirts to get larger and larger amounts of fish.

This reflects the fact that, with economies of scale, the marginal product of his labor *decreases* the less labor he devotes to production of a good. This means that when he is doing little shirt-making, his marginal productivity in that activity is low, so doing a little less (so as to allow him to do more fishing) does not result in a large reduction in shirts produced.

CHECKPOINT 6.3 Production possibilities with economies of scale

a. To check that you understand how Figure 6.4 works, find the point on the feasible frontier associated with devoting eight hours to fishing and two to shirt-making, (trace out the new points, **d'**, **e'**, **f'**, and **b'**).

b. You can see from the figure that if Alex devotes all ten hours of his labor to producing shirts he produces ten of them. How many shirts and fish would he produce if he devoted nine hours to shirt-making and one to fish-catching?

c. From the calculation you have just done, what is the marginal rate of transformation of shirts into fish?

! REMINDER Remember that production functions with economies of scale are *convex* (output increases at an increasing rate with input) and production functions with diseconomies of scale are *concave* (output increases at a decreasing rate with input).

6.5 SPECIALIZATION AND EXCHANGE

In Figure 6.5 we combine the feasible set from Figure 6.4 with an indifference map. Recall that, as indicated by the numbering of the indifference curves, farther away from the origin is *better* in Alex's evaluation of outcomes because both shirts and fish are goods.

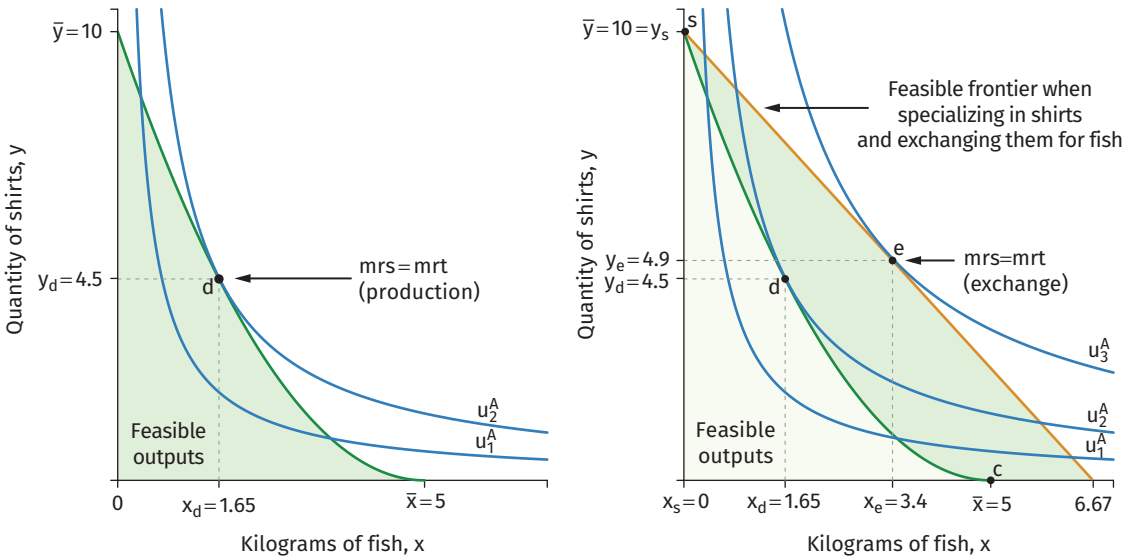
"The division of labor is limited by the extent of the market"

In the case shown the marginal utility of both goods is diminishing the more he consumes, so having some of both is superior to having lots of one kind of good and none of the other. Remember this is why his indifference curves are bowed *inward* toward the origin.

If Alex *cannot* exchange goods with other people, he does the best he can by following the $mrs = mrt$ rule and finding the point on the production possibilities frontier that is tangent to the highest indifference curve, that is, at point **d** (for **d**iversified production), and consuming x_d and y_d.

But, if the producer can exchange the goods with others, then there is a second way that he can "transform shirts into fish." He does *not* do so by reallocating his time from shirt-making to fishing. Instead he can spend *all* of his time making shirts and then exchange some shirts for some fish if he can find a willing buyer for his shirts.

Figure 6.5 Diversification and specialization with economies of scale. In panel (a), we present the producer's choice when they do not have the opportunity to trade. Using the $mrs = mrt$ rule, the producer chooses the point at which their indifference curve is tangent to his production possibilities frontier at point **d**. In panel (b), we show what happens if the producer can exchange shirts for fish at some constant price ratio (p). In this case, he can do better by specializing in shirt production (good y) at point **s**, and then acquiring the fish he desires through exchange, not by producing the fish. He produces shirts only at point **s**, then exchanges the shirts on the market for fish, taking him to his higher indifference curve, u_3^A at point **e**.



(a) Economies of scale without trade

(b) Economies of scale and trade

Suppose such a trader is found, and she is willing to buy any amount of his shirts at a given price: in return for p shirts she is willing to provide 1 kilogram of fish. This is the second way of transforming shirts into fish, and the marginal rate of transformation is p: the quantity of shirts that Alex has to give up in exchange for a kilogram of fish.

We do not yet ask what determines p. What determines the price will occupy the next three chapters, in which we introduce the demand for various products and how this along with costs of production and the nature of competition among buyers and sellers determines prices.

This opportunity for exchange alters the feasible set constraining what Alex can do, as shown in Figure 6.5 (b). The orange line with the y-intercept at $\bar{y}$ (which is the maximum amount of shirts Alex can produce) represents his *new* feasible frontier with exchange. Its slope is $-p = -2/3$, the (negative

of the) opportunity cost of acquiring more fish in terms of the shirts foregone, or what is the same thing, the marginal rate of transformation of shirts into kg of fish: giving up 1 shirt gets him 2/3 of a kilogram of fish, giving up 10 shirts gets him 6.67 kilograms of fish.

Movement along the price line shows *exchange* in varying amounts, *not* shifting Alex's own labor from making shirts to fishing. The darker green-shaded area is the enlargement of his set of feasible levels of consumption of the two goods made possible by the opportunity to exchange his shirts for fish.

How does the expansion of the feasible set change Alex's choice about how to use his labor? Before he could trade he did his best by diversifying at point **d**. Now, with the ability to trade, he can choose whether to diversity as previously, or to specialize in fish (point **c**) or shirts (point **s**).

He will consider the exchange option. If he specialized in shirt-making, for example, his willingness to pay to get some fish would be very high. And he could transform one shirt into two-thirds of a kilogram of fish at the given price, $p = 2/3$, as in Figure 6.5 (b). So he would think as follows:

- His *willingness to pay* for more fish (*mrs*), shown by the steepness of the indifference curve if he has little or no fish, *exceeds* the *opportunity cost* of acquiring more fish by reallocating some of his labor to fishing (the flatter slope of his production possibility frontier);
- but the *opportunity cost* of acquiring fish by exchange (the negative of the even flatter slope of the price line) is even *less*.

So if he specializes in shirt-making he will exchange some of the shirts he produces for fish at the price p to reach point **e** (for **e**xchange) on the highest indifference curve in his new feasible set, u_3^A. This is where:

$$mrs = p = mrt \text{ (by exchange)}$$

In the next Checkpoint we ask you to explain how Alex would know that, at the price shown in the figure, specializing in shirts and trading goods is better for him than specializing in fish and trading.

Our example—Alex choosing what to produce—demonstrates two general truths:

- if one or more production functions with economies of scale is available, it may make sense to *specialize*; but
- this will be true only if there are others producing different goods and there are opportunities for exchange, integrating specialized producers with generalist users to coordinate the *division of labor*.

This is the basis of the interdependence of different producers within the division of labor.

EXAMPLE In modern economies a household may specialize in providing labor with some particular mix of skills, training, and experience to an employer.

CHECKPOINT 6.4 The choice of what to specialize in

a. Using Figure 6.5 show that Alex could do better than at point **d** by specializing entirely in fish, and exchanging fish for shirts at the price p. Why does he do even better if he specializes in shirts?

b. Redraw Figure 6.5 with a higher relative price of fish (so p, the number of shirts that one must give up to get a kg of fish is now larger).

c. Show that if p is sufficiently high, Alex will do better specializing in fish than specializing in shirts (indicate by a point on your redrawn figure the amount he will produce, and the amount he will exchange).

d. What is the lowest price, p (number of shirts he gets per kg of fish) that would make it worthwhile for him to specialize in fish?

Diseconomies of scale, diversification, and exchange

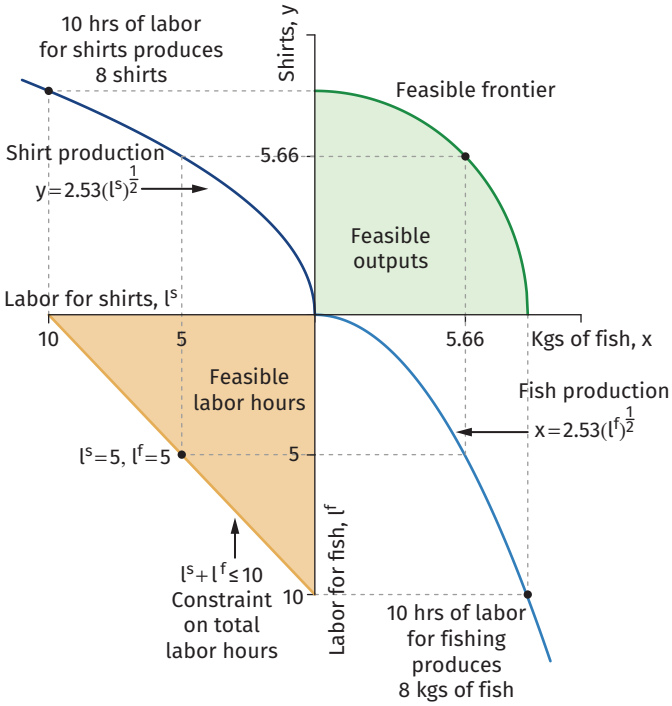
In contrast with the economies of scale illustrated in Figures 6.4 and 6.5, Figure 6.6 illustrates the case of diseconomies of scale, in both fishing and shirt-making. With diseconomies of scale, as Alex shifts his labor input from producing shirts to producing fish, he gives up larger and larger amounts of shirts to get smaller and smaller amounts of fish. This reflects the fact that, with diseconomies of scale, the marginal product of his labor *decreases* the more labor he devotes to production of a good. As a result of diseconomies of scale in production of the goods, Alex's production possibilities frontier is *bowed outward* from, or concave to, the origin.

With diseconomies of scale and the corresponding bowed-out production possibilities frontier, the marginal rate of transformation ($mrt(x,y)$) increases as labor is reallocated from shirts to fish reflecting the idea of *increasing opportunity costs*. That is, along the production possibilities frontier, as Alex produces more fish, he must give up more and more shirts to do so.

To understand what his utility-maximizing choices will be when he cannot exchange and when he can exchange his goods, we combine the production possibilities frontier from Figure 6.6 with his indifference map, as shown in Figure 6.7. If Alex did not have opportunities for trading goods, he would select point **d** in Figure 6.7 (a) with utility u_1^A at bundle (x_d, y_d), at which his marginal rate of transformation in production equals his marginal rate of substitution.

But Alex can do better if he decides what to produce *knowing in advance* that he will be able to *exchange* the goods he produces. To see why this is so, imagine that once Alex was given the opportunity to trade, he continued producing at point **d**, and then looked at his new feasible set, the frontier of which is defined by the orange price line through point **d** with as before a slope of $-p$, shown in Figure 6.7 (a). The new feasible set gives him options that he prefers to remaining at point **d**: he would trade some fish for some

Figure 6.6 Production possibilities frontier (PPF) with diseconomies of scale. The graph is constructed exactly as Figure 6.4. It shows in the top-right quadrant a production possibilities frontier with diseconomies of scale in production. The diseconomies of scale depicted in the production possibilities frontier arise from a relationship in the production technologies from the two different sectors: fishing (bottom-right quadrant) and shirt-making (top-left quadrant).



shirts (the tangency of the price line with his highest indifference curve is not shown in the figure).

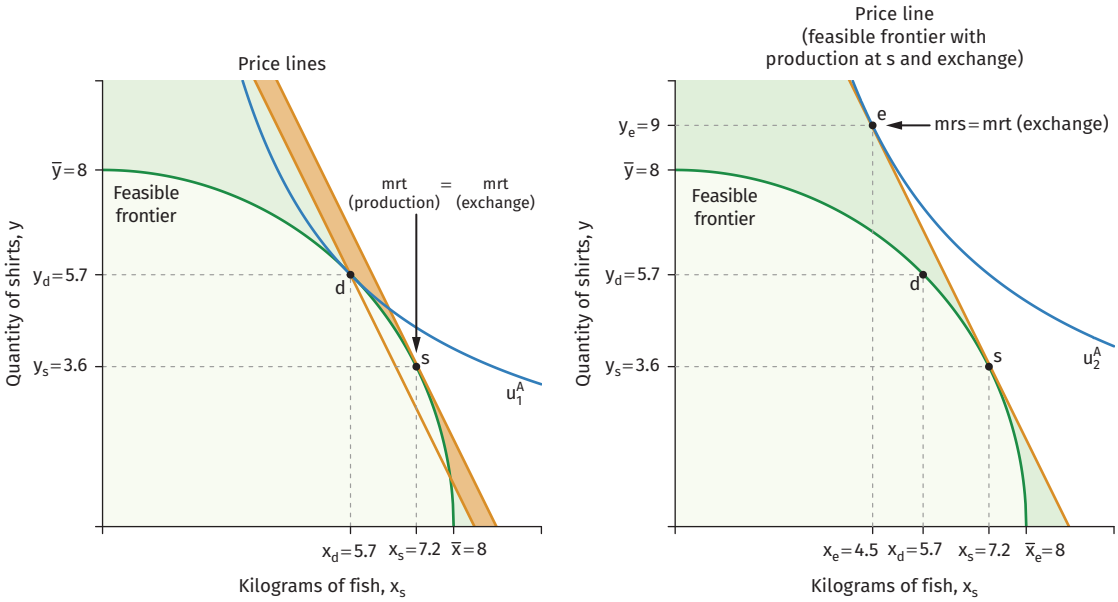
But he can do better: he can "drag" the price line to include any point on the feasible frontier, *not* changing its slope but moving it farther away from the origin (good for him) or closer to the origin. Where would he want to drag the price line? The largest feasible set he could have would be if he shifted his production to point **s**.

The possibility of exchange *expands* his feasible set: the orange line through point **s** is the feasible frontier with production at point **s** and exchange possible at price p, and the expansion of the feasible set beyond what he could have done if he had produced at **d** and then considered exchange is shown by the area shaded in orange.

Is there a simple rule that he could follow to know which point on the feasible frontier to pick in deciding what to produce? There is: *maximize the value of the goods you produce valued at the price p.*

M-CHECK The production possibilities frontier with diseconomies of scale in both production functions is *concave* toward the origin, or bowed outward from the origin.

Figure 6.7 A two-step optimizing process with diseconomies of scale and opportunities for exchange. If the producer cannot exchange the goods with others, he does the best he can by finding the point on the production possibilities frontier that is tangent to the highest indifference curve ($mrs = mrt$) at point **d**, and consuming x_d and y_d. But, suppose the producer can exchange the goods with others, at the rate of selling one kg of fish for every p shirts bought. If he continued producing at point **d** his new feasible set by means of exchange would be points on or to the left of the orange price line through point **d**. But the price line is not fixed, its location depends on what he produces. First, he will "drag" the price line to the point with the highest value at the given price (located as the point where the mrt in production (along the feasible frontier) is equal to the mrt in exchange: point **s**. The possibility of exchange at **s** expands his feasible set, shown by the area shaded in orange. Then, in panel (b), he exchanges fish for shirts, moving along the price line, which is his new feasible frontier, to point **e** on u_2 where the mrt in exchange is equal to the mrs (from his indifference curve).



(a) Expanding feasible set by producing at **s**

(b) Production at **s** and consumption at **e**

This works because at the price p, 1 kg of fish is worth the same as p shirts, then all of the combinations of quantities of fish and shirts along an orange price line in the figure have the same value, that is: $y + p \cdot x$. This has to be true because in any exchange along that line, the value (expressed in number of shirts) of the fish sold—p times the kgs of fish sold—must be equal to the value of the shirts purchased (which is just the number of shirts).

This means that a price line farther away from the origin represents a higher value of the various combinations of x and y making up the line. The price line passing through the production point (x_s, y_s) on the

production possibilities curve is now a budget constraint indicating all of the combinations of x and y that are attainable for Alex given how he has allocated his labor. Moving the price line to include the point s by selecting the output bundle (x_s, y_s) to produce maximizes the budget that constrains his exchange options.

The constrained optimization problem that Alex is facing thus comes in two steps. But to take the first step Alex must anticipate that after he chooses how much to produce he will have the opportunity—through exchange—to change the bundle he has produced so as to consume a different bundle.

- *Step 1: Maximize the value of the output bundle subject to the constraint set by the frontier of the feasible production set.* To do this find the distribution of labor time between fishing and shirt-making that *maximizes the value of one's output*. In step 1, the price line is similar to an indifference curve: the objective is to choose the point on the production possibilities frontier that is also a point on the price line that is as far as possible from the origin. So the choice in the first step is the point at which the marginal rate of transformation on the production possibility frontier is equal to the marginal rate of substitution (the negative of the slope of the price line or what is the same thing, p).
- *Step 2: Maximize utility subject to the budget constraint given by the price line that includes the production point you have chosen.* In step 2, the price line is now the budget constraint not the objective. The point chosen will equate the marginal rate of transformation along the price line with the marginal rate of substitution along the indifference curve.

The two cases—economies of scale and diseconomies of scale—differ in important ways, economies of scale leading to specialization followed by exchange, and diseconomies of scale leading to diversification, also followed by exchange. But they have in common a simple rule: from all the feasible combinations of the production of the two goods, choose the one that maximizes the value of your output, whether diversified or specialized.

M-NOTE 6.2 **Two-step constrained optimization with diseconomies of scale and exchange**

We explain how to determine the maximum utility that Alex can attain given his utility function $u = u(x, y)$, the feasible set of production of two goods that he could implement, and his opportunities to exchange one good for another at the price ratio $p = \frac{p_x}{p_y}$.

Step 1: Maximize the value of the output subject to the constraint set by the frontier of the feasible production set.

continued

The total value of the goods produced by Alex (or the price line) is the price of good y multiplied by the amount of y plus the price of good x multiplied by the amount of x:

$$\text{Value:} \qquad v = p_y y + p_x x$$

For simplicity, we will let $p_y = 1$, divide through by p_y, and let $p = \frac{p_x}{p_y}$, giving us:

$$v = y + px$$

Alex's constraint is his feasible frontier:

$$\text{Constraint:} \qquad y = g(x), \qquad \frac{dg}{dx} < 0$$

Thus, his first step in the maximization problem is as follows:

$$\text{Vary } x \text{ to maximize} \qquad v = g(x) + px$$

$$\text{First-order condition:} \qquad \frac{dv}{dx} = \frac{dy}{dx} + p = 0$$

$$\text{Rearranging:} \qquad \text{mrs of iso-value} = p = \frac{-dy}{dx} = \text{mrt of feasible frontier}$$

This gives Alex point **s** in Figure 6.7, or a combination (x_s, y_s) for him to produce where his marginal rate of transformation is equal to the relative price p.

Step 2: Maximize utility subject to the budget constraint given by goods he has produced (x_s, y_s), and the price line:

$$\text{Utility:} \qquad u = u(x_e, y_e)$$

Alex's budget constraint given y_s, x_s, and p is his price line (what he maximized in step 1):

$$\text{Budget constraint:} \qquad v_s = y_s + px_s$$

The values of y consistent with the constraint, are $y = v_s - px_s$ and we can use this to eliminate y from his utility function. Thus, his maximization problem becomes:

$$\text{Vary } x \text{ to maximize} \qquad u = u(x_e, v_s - px_e)$$

$$\text{First-order condition (chain rule):} \qquad \frac{du}{dx} = u_x - pu_y = 0$$

$$\text{Rearranging:} \qquad \text{mrs (utility function)} = \frac{u_x}{u_y} = p = \text{mrt (budget constraint)}$$

This gives Alex point **e** in Figure 6.7, or a combination (x_e, y_e) on the highest indifference curve in his feasible set with exchange.

CHECKPOINT 6.5 Maximize the value of your output In Figure 6.7.

a. What is the price of fish, p (that is, how many shirts are required to buy one kg of fish)?

b. Calculate the value of his output (sum of price times quantity of both goods) at points **s**, **e**, and **d**.

c. Explain why two of them are the same and both larger than the other.

d. What would be the value of his production if Alex had specialized in fish? What about if he had specialized in shirts?

6.6 COMPARATIVE AND ABSOLUTE ADVANTAGE

The question that Alex faced, "What should you specialize in?" seems to have an obvious answer: "Specialize in what you are best at." The same would seem to go for countries: they should specialize in what they are best at producing. But what, exactly, does that mean? "Better" than other people (or countries)? What if you are *not* better than others at anything? Should you not specialize in anything?

Differing opportunity costs and comparative advantage

Or, does "better" mean "better than you are at *other things* that you could do"? If that's what it means—at least *compared to* how good others are in those same things—then we are talking about comparative advantage.

To see what this means, suppose that a recent graduate, Brett, has started a data science business. When Brett writes reports for his business, there are two tasks: entering data in some digital format and generating graphs to some detailed specifications using the digitized data. Let's say that each graph requires 1,250 keystrokes of data.

Brett has the option of doing both tasks, or doing either one of them himself and getting the other done for pay on Mechanical Turk (which calls itself "the *online marketplace* for work"). There are many people like Brett, some of them offering their services on MTurk, as it is called, and with their pay purchasing other services from MTurk. But we will consider just one of these people named April (there are lots of people like her too).

ONLINE MARKETPLACES
Mechanical Turk is one of many online marketplaces for work tasks for pay. Others are Clickworker, Fiverr, UpWork, among many others. People can be paid for small, short tasks like data input (which is more typical for sites like MTurk and Clickworker) or for more advanced jobs like Fiverr and UpWork.

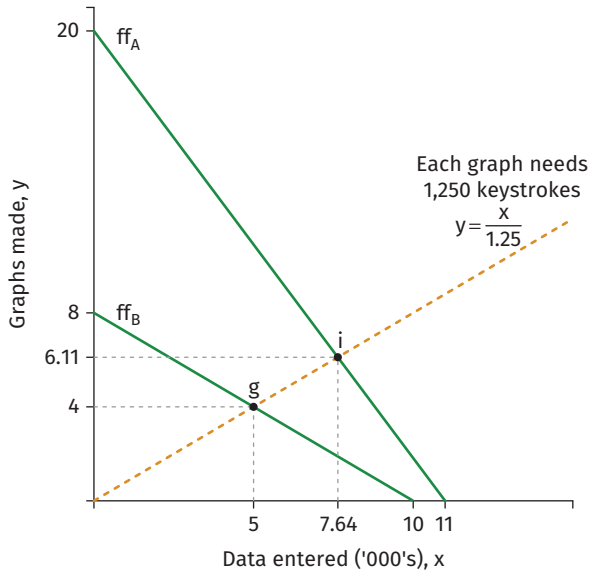
Figure 6.8 shows how good April and Brett are at the two tasks, as indicated by their feasible frontiers if they are restricted to working for just one hour. The points making up their respective frontiers are the combinations of outputs from data entry and graph making that use up one hour of their working time. For example, in an hour Brett can produce eight graphs and enter no data, or ten (thousand) keystrokes of data entry, and no graphs, or four graphs and five (thousand) keystrokes, and so on.

Figure 6.8 illustrates the concepts of both absolute and comparative advantage. April has an **absolute advantage** in producing both goods: data entry in thousands (x) and graphs (y). A person has an absolute advantage in the production of a particular good if, given the set of available inputs, she can produce *more* of it than some other person. M-Note 6.3 shows how the feasible frontier is derived from the data on the productivity of April and Brett at the two tasks.

In our case, absolute advantage means that, in one hour, if April devoted all of her time to data entry, she could enter more thousands of keystrokes

ABSOLUTE ADVANTAGE A person or a nation has an absolute advantage in the production of a particular good if, given a set of available inputs, they can produce more of it than another person or country.

Figure 6.8 Feasible frontiers: Absolute and comparative advantage. April has an *absolute advantage* in the production of both goods because her feasible frontier is outside Brett's. Brett has *comparative advantage* in data entry because his feasible frontier is *flatter* than hers (lower opportunity cost of data entry). Without the possibility of exchange Brett completes four graphs (at point **g**). Remember: each graph requires 1,250 keystrokes of data (also remember that the horizontal axis of Figure 6.8 is measured in thousands of keystrokes). This means that they must be on the dashed orange line from the origin. The question is how far out they can get. The dashed line from the origin for $y = \frac{x}{1.25}$ lets us see how many complete graphs (data entry and graph preparation combined) each person could get by themselves in one hour.



of data than Brett (11 rather than ten) and likewise for making graphs (20 rather than eight). As her feasible set includes Brett's entire feasible set (her feasible frontier is farther from the origin), she can produce more in an hour than Brett can in any combination—complete specialization in one or the other or some ratio of data entry to figure making.

M-NOTE 6.3 Opportunity costs, feasible frontiers, and comparative advantage

To understand how production functions, opportunity costs, and the feasible frontier determine absolute and comparative advantage we use the following notation:

- a_x = time (fraction of an hour) required to input 1,000 keystrokes of data ($\frac{1}{11}$ for April and $\frac{1}{10}$ for Brett, so April has an absolute advantage in data entry.)

continued

- a_y = time (fraction of an hour) required to produce one graph ($\frac{1}{20}$ for April and $\frac{1}{8}$ for Brett, so April has an absolute advantage in graph-making)
- x = thousands of keystrokes of data entered
- y = number of graphs made
- T = total time is 1 hour

Total labor time is composed of time spent entering data plus time spent producing graphs, so the feasible set is defined by:

$$\text{Time constraint} \qquad a_x x + a_y y \leq T \tag{6.5}$$

The equation for the feasible frontier is Equation 6.5 expressed as an equality and rearranged with y as a function of x:

$$\text{Feasible frontier} \qquad y = \frac{T}{a_y} - \frac{a_x}{a_y}x \tag{6.6}$$

(Equation 6.6 is similar to the equation for a budget constraint like we saw in Chapter 3.) Equation 6.6 means that:

$$-\frac{dy}{dx} = \frac{a_x}{a_y} = mrt \tag{6.7}$$

The term $-\frac{dy}{dx}$ is the negative of the slope of the feasible frontier (which can be seen from Equation 6.6). In other words, the marginal rate of transformation is the number of graphs that one has to *give up* by reallocating time to data entry, or the *opportunity cost* of data entry.

We can use the numbers to find their opportunity costs.

For April:

$$mrt(x,y) = -\frac{dy}{dx} = \frac{a_x}{a_y} = \frac{\frac{1}{11}}{\frac{1}{20}} = \frac{20}{11} = 1.82 \tag{6.8}$$

For Brett:

$$mrt(x,y) = -\frac{dy}{dx} = \frac{a_x}{a_y} = \frac{\frac{1}{10}}{\frac{1}{8}} = \frac{8}{10} = 0.8 \tag{6.9}$$

Because $0.8 < 1.82$, Brett's comparative advantage is in data entry.

Different opportunity costs: The basis of specialization and exchange

This raises the question: If April is better at both data entry and graph-making, why would she want to trade with Brett at all? This is where the concept of **comparative advantage** comes in. A person has a comparative advantage in the production of a particular good if their opportunity cost of producing that good is *lower* than it is for another person.

COMPARATIVE ADVANTAGE A person or a nation has a comparative advantage in the production of a particular good if the opportunity cost to them of producing it (in terms of foregone output of another good) is lower than it is for another person or country.

For Brett, spending the hour that it would require to enter 10,000 more keystrokes of data would mean that he could *not* make eight graphs. So eight graphs is his opportunity cost of 10,000 keystrokes of data entry. Translating this to be in the units of the figure, 0.8 graphs is the opportunity cost to Brett of 1,000 keystrokes of data entry (which is the negative of the slope of his feasible frontier).

By contrast, for April, 10,000 keystrokes of data entry requires just 55 minutes (she enters 11,000 keystrokes per hour) and in that period of time she could have made 18.2 graphs. So for April the opportunity cost of 10,000 keystrokes is 18.2 graphs, or translating this to the quantities in the figure, the opportunity cost of 1,000 keystrokes is 1.82 graphs.

Brett's comparative advantage is in data entry. This is not because he is so good at data entry; April is better at data entry than him. It is because he is so *unproductive* in producing graphs, so the opportunity cost of taking time *away* from graph-making to do data entry (the graphs he otherwise could have made) is *low*. It can similarly be seen that April's comparative advantage is in producing graphs. Table 6.1 summarizes Brett and April's absolute and comparative advantage in these tasks.

Here is a simple way to remember the difference between absolute and comparative advantage:

- If for a given axis (horizontal or vertical) the *intercept* of one person's feasible frontier is outside (farther from the origin than) the other's, then that person has an *absolute advantage* in the good on that axis.

Table 6.1 Absolute and comparative advantage: Number of bits of data and graphs created in one hour's work. The entries in blue show that April has the absolute advantage in producing *both* data entry and making graphs. The entries in red show that Brett has a comparative advantage in producing data entry ($0.8 < 1.82$) and similarly April has a comparative advantage in making graphs ($0.55 < 1.25$). Remember, if they are working alone, they would never produce only graphs or only data, because they need a combination of graphs and data for the project. For both of them the two opportunity costs are simply the inverse of one another, e.g. for Brett $1.25 = \frac{1}{0.8}$.

	Brett	April
Maximum possible data entry (thousands of keystrokes per hr)	**10**	**11**
Maximum possible graphs (graphs per hr)	**8**	**20**
Opportunity cost of 1,000 keystrokes of data entry (in graphs)	**0.8**	**1.82**
Opportunity cost of making 1 graph (in thousands of keystrokes)	**1.25**	**0.55**

- *Comparative advantage is determined by the slope of the feasible frontier*: The comparative advantage of the person with the flatter feasible frontier is in the good on the horizontal axis. This is because the (negative of the) slope of the feasible frontier (how steep it is) is the *opportunity cost* of the good on the x axis. The comparative advantage of the person with the steeper feasible frontier is in the good on the vertical axis.

The second bullet says that unless the two feasible frontiers have *the same* slope, the comparative advantage of the two people will differ. Even though one of them may *not* have an absolute advantage in either good (like Brett) each will have a comparative advantage in one of the goods.

So Brett has to be *comparatively* good at something. The data show that the opportunity cost of data entry is less for Brett than it is for April. We now show why this provides the basis for Brett specializing in data entry and exchanging some of his data for graphs, April spending all her time making graphs and exchanging some of them for data entry.

CHECKPOINT 6.6 Comparative and absolute advantage

a. If April could make only seven graphs in an hour (Brett's productivity remaining unchanged) which of them would have an absolute advantage in which of the goods? In which good would April's comparative advantage be?

b. If in an hour Brett could enter only 4,400 keystrokes of data, who would have comparative advantage in data entry?

6.7 SPECIALIZATION ACCORDING TO COMPARATIVE ADVANTAGE

Brett and April could of course exchange with each other. But they would have to find a *price* at which the number of keystrokes that Brett was willing to sell was also the number of keystrokes that April wished to purchase. And similarly for April's graphs.

It is more realistic to think of them selling their goods on a market in which many "Bretts" are selling data entry, and many "Aprils" are selling graphs. When we say April, we really mean "people like April ready to sell graphs on MTurk, in return for data entry sold by people like Brett." And a similar statement goes for Brett. The reason is that if we had some particular Brett exchanging with a particular April, then in agreeing on a price they also would have to agree on the amounts to be exchanged (as Ayanda and Biko did in Chapter 4).

This would be an unnecessary complication, so we avoid it by assuming that both Brett and April can purchase and sell as much as they like at whatever price is posted on MTurk. Then the amount that any particular Brett wanted to sell would not have to be equal to what some particular April

! REMINDER The two ways that April may get the necessary data entered if she specializes in graph-making—entering them herself, or exchanging some graphs with another person who will enter the data—are exactly the two ways that Alex had to acquire fish if he specialized in shirt-making: divert some of his time to fish-catching, or exchange some of the shirts he made for fish.

wanted to buy. So we will imagine Brett and April considering producing some of the product in which they specialize as a task on MTurk and with the proceeds of the sale of these tasks, then buying what they lack (also on MTurk).

In Figure 6.8 we can see that if Brett produced both data entry and graphs himself, he would get to point **g**—four completed graphs based on the required 5,000 keystrokes of data. Similarly, April working by herself could produce 6.11 graphs along with the necessary data at point **i**.

Will the two be able to do better by *specializing* in the production of just a single task each, and then exchanging graphs for data entry, so that each would have the required keystrokes of data for each graph?

The opportunity to exchange expands the feasible set

To see that they will do better with exchange, think about two ways that April can get data entered: she can do this herself, with every 1,000 keystrokes entered bearing an opportunity cost of 1.82 graphs not made. Or she could pay someone else to enter data, paying with some of her graphs. What is the most she would be willing to pay for 1,000 keystrokes of data entered? The answer is 1.82 graphs, which is what she would have to "pay" in graph-making *foregone*, if she did the data entry herself. This is her *maximum willingness to pay for data entry*. This is also the negative of the slope of her feasible frontier.

Would Brett be willing to sell her data entry for a price less than 1.82 graphs per 1,000 keystrokes? The lowest price at which he would sell 1,000 keystrokes of data entry is 0.8 graphs because this is his opportunity cost of data entry. His opportunity cost is the number of graphs he *gives up* producing if he enters 1,000 more keystrokes. This price is called Brett's minimum *willingness to accept* (giving up data in return for graphs). This is the negative of the slope of his feasible frontier.

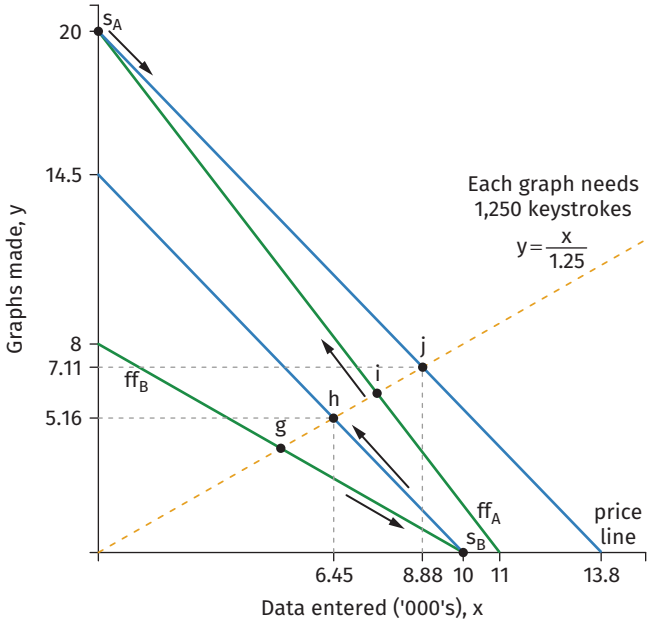
Because April's willingness to pay is greater than Brett's minimum willingness to accept—the lowest price at which he would sell keystrokes—each of them can benefit by specializing and then entering into an exchange.

In Figure 6.9 we show what Brett can accomplish when he specializes in data entry and then exchanges some data entry for the graphs he needs to complete his project. The exchange opportunities are shown by the price lines (which are parallel because *both* people face *the same* relative prices). The (negative of the) slope of the price line is the number of graphs that can be 1,000 purchased with keystrokes of data entry, or a price equal to 1.45 in our example illustrated in the figure. A *steeper* price line is better for Brett, a *flatter* price line is better for April.

Specialization and mutually beneficial exchange

To see why specialization and exchange will be mutually beneficial, you can think of the following:

Figure 6.9 Feasible frontiers and relative prices for exchange. Without the possibility of exchange Brett completes four graphs (at point **g**). The two arrows show that instead, he could move to point $\mathbf{s_B}$, specializing entirely in data entry, and then exchange some of his data entry with April in return for her making 5.16 graphs for him. The arrows at the top show how April could specialize and exchange.



- The (negative of the) slope of the price line as the *marginal rate of transformation of keystrokes into graphs by means of exchange.*
- The (negative of the) slope of the feasible frontier is the *marginal rate of transformation of keystrokes into graphs by means of devoting more time to graph-making and less to data entry.*

The possibility of exchange gives Brett a *new* feasible set, with the frontier being the price line passing through any point on his "working alone" feasible frontier, indicating his exchange opportunities when he can buy 1.45 graphs with 1,000 keystrokes. With this new opportunity he could move in two steps from point **g** to point **h**. He could do this if he, first, specialized at point $\mathbf{s_B}$ and then, second, engaged in exchange, moving *up* the price line to point **h**.

Similarly, if April were at point **i** producing both goods, she could move to point **j** if she specialized at point $\mathbf{s_A}$ and engaged in exchange to take her down her price line to point **j**. In Table 6.2 we compare their situation when producing both goods with the outcome when they specialize and trade.

Table 6.2 Specialization and exchange according to comparative advantage. The price of 1,000 keystrokes of data entry is 1.45 graphs. Remember Brett is exchanging his data entry for graphs with people like April (not just April herself). And the same goes for April's exchanges. This is why it is possible for the number of graphs that our particular Brett purchased (5.16) to differ from the number of graphs that April sold (12.89). The different colors in the table each correspond to a point for each Brett and April in Figure 6.9. For example, the entries in orange in Brett's column correspond to Brett's initial point, **g**, where he is working alone. The entries in purple correspond to his first step in the exchange process where he specializes in data entry, point $\mathbf{s_B}$. The entries in green correspond to his final position after the trade, point **h**. The colored entries in April's column are interpreted in a similar manner.

	Brett	April
	Working independently	
Data entry (000s)	**5** (0.5 hours)	**7.64** (0.695 hours)
Making graphs	**4** (0.5 hours)	**6.11** (0.305 hours)
Graphs submitted for the project	**4**	**6.11**
	Specializing and trading	
Data entry (000s)	**10** (1 hour)	**0**
Making graphs	**0**	**20** (1 hour)
Work produced for own project	**6.45** (000) data entry	**7.11** graphs
Work for pay to exchange with others	**3.55** (000) data entry	**12.89** graphs
Others' work purchased	**5.16** graphs	**8.89** (000) data entry
Project submitted	**5.16** graphs	**7.11** graphs

The reason why a mutually beneficial exchange would be possible is that the price at which the exchange took place (1.45 graphs per 1,000 keystrokes) was *greater* than Brett's opportunity cost of keystrokes and *less* than April's opportunity cost of keystrokes. Or returning to Figure 6.8 the slope of the price line was greater than the slope of Brett's feasible frontier and less than the slope of April's.

We have not explained why this particular price was the one at which they traded (as this would have been distraction from introducing comparative advantage). But any price between 0.8 and 1.82 would have allowed mutually beneficial exchange to take place.

What made specialization possible in this case is two things:

- *Differences*: Brett and April *differed* in their comparative advantage so there was some price–1.45 per 1,000 is just one example–at which they could both benefit form an exchange.
- *Opportunities for exchange*: There was a way to exchange one's completed tasks with others so as to obtain the right mix of data entry and graph making.

CHECKPOINT 6.7 The distribution of the gains from specialization and exchange

a. Using Figure 6.8, determine the price (graphs per 1,000 keystrokes) such that Brett would *not* benefit at all from specializing and trading.

b. Do the same for the price ratio such that April would not benefit.

c. Use the required number of thousands of keystrokes per graph (1.25) to say how many graphs Brett could make if the price at which he could sell 1,000 keystrokes fell from 1.45 graphs to one graph.

6.8 APPLICATION: HISTORY, SPECIALIZATION, AND COORDINATION FAILURES

Brett and April simply decided to specialize in the tasks in which each had a comparative advantage. The existence of the online marketplace for tasks made this possible, and both benefited by comparison to producing their reports without specializing. In the earlier example, Alex simply chose to produce shirts rather than fish, and he was able to feed himself because he could exchange shirts for fish. These personal examples have important lessons about specialization.

But comparative advantage is more often applied to what *countries* do, not to what people do. When it comes to countries, we cannot say that, for example, Germany "decided" to specialize in machine tools and Bangladesh in textiles. What countries specialize in is the result of decisions made by vast numbers of people independently choosing what kinds of skills they will learn, the jobs they will take, what kinds of products the firms they own will produce and similar decisions. Countries—unlike Brett and April—can sometimes end up specializing in such a way that they remain poor, while had they specialized in something else, they would have been rich.

To see how countries can specialize and stay poor, return to the feasible frontier in Figure 6.5. But now think about the figure as applying to an entire country, *not* just choices that Alex might make between fishing and making clothing.

In this case, the economies of scale in the production of shirts occur because in every shirt-making business labor is more productive in producing shirts the more shirts are being produced *in all of the clothing industry*. Industry-wide (rather than firm-level) economies of scale are called **economies of agglomeration**. Economies of agglomeration means

ECONOMIES OF AGGLOMERATION The advantages that firms may enjoy when they are located close to other firms in the same or related industries.

that the productivity of labor is greater the larger is the total output of the *many* firms producing *similar* goods in one country or region.

Economies of agglomeration contribute to the geographical concentration of particular industries, for example:

- software engineering in Bangalore (Bengaluru), India
- finance in Hong Kong, London, and New York City
- information technology and IT-related production in Silicon Valley, California
- machine tools and motor vehicles in the Stuttgart-Munich region of Germany

Economies of agglomeration occur because, when large numbers of people are employed in producing the same product, the skills and other knowledge particular to that industry are widely diffused in the population, resulting in higher levels of productivity across the board. Public policies implemented by governments favoring a locally dominant industry also reduce costs.

If the relevant economies of scale do not pertain to an individual firm, but instead are economies of agglomeration, then a single firm, even if capable of operating at a large scale, will have little reason, for example, to introduce a truck-manufacturing plant in a finance agglomeration such as Hong Kong, or an IT region such as the Silicon Valley.

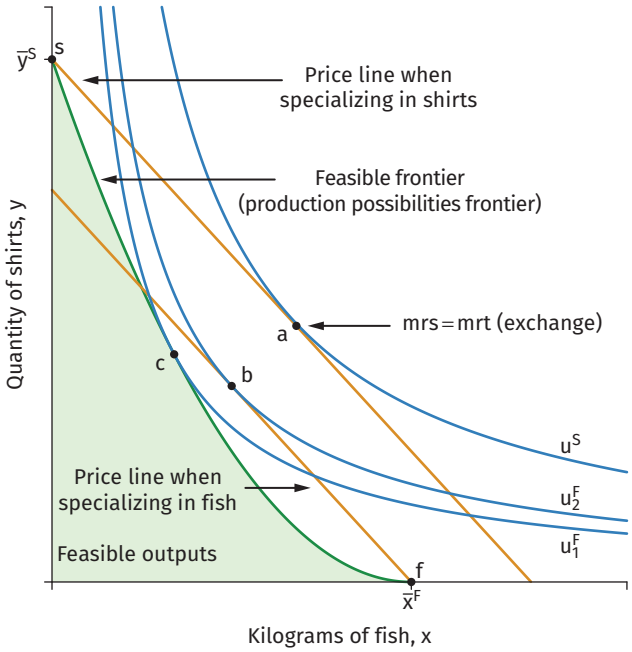
In Figure 6.10, a country could find itself at point **c** on indifference curve u_1^F where they have diversified production or at **b** where they have specialized production in fish, trade some of the fish along the price line at price p and arrive at bundle **b** on indifference curve u_2^F. They are better off specializing in fish than they would be if they produced a diversified set of goods as shown by the specialized production resulting in higher utility at **b** than at **c**.

But they could do a lot better. If the very same country specialized in producing shirts and then *traded* some of their output to acquire fish, they would consume at bundle **a** on indifference curve u^S. Suppose that for as long as anyone can remember they have specialized in fishing. Why don't they just change that? In the case of April deciding to specialize in data entry, or Brett in making graphs, the two people would have quickly realized that they were ignoring their comparative advantage: they would quickly switch their specialization to what they have a comparative advantage in. But in the case of an entire country how would they switch?

Economies of scale and poverty traps as an assurance game

Suppose that in the fish-producing country a few people realized that everyone would be better off (be on a higher indifference curve) if they switched to specializing in shirts. What could they do? If people decided *individually* to produce shirts they would be much worse off than

Figure 6.10 Production possibilities frontier poverty trap. In the figure, the country is specialized in fish at point **f** and producing $\bar{x}^F$ and obtaining welfare (utility) of u_2^F at point **b**, but would like to produce shirts (which would take them to a higher indifference curve). The country is better off specializing in fish than they would be if they produced a *mix* of goods, but they would prefer to be producing $\bar{y}^S$ shirts at point **s** where they would be on the higher price line and therefore obtaining higher utility at u^S at point **a** ($u^S > u_2^F > u_1^F$). The country cannot simply shift inputs to produce new outputs as that would require a dramatic repurposing of production.



specializing in producing fish. You can see this because getting more shirts by producing them—that is, moving along the feasible frontier away from point $\bar{x}^F$—rather than producing fish and trading them for shirts is a losing proposition. With specialization, people would produce at $\bar{x}^F$ and trade to point **b** on indifference curve u_2^F. If they chose to produce a few shirts and a bit less fish—moving to the left along their feasible frontier—they could still engage in some exchange to acquire additional shirts. But as you can see the price line passing through a point on the feasible frontier a bit to the left of point **f** would shrink their set of feasible consumption. The new price line would shift to the left, indicating that the value of the goods they produced had *fallen*. No business acting independently would consider this to be a good plan.

A country specializing in fish in this model is *locked in* to lower income. If they all decided to switch (specializing in shirts at point **as**) then they could

all be better off if they could then trade downward along the same price line. But as long as the decision about what each person will produce is taken independently, people would not specialize in shirt production. They are facing a coordination problem similar to those discussed in Chapters 1, 4, and 5.

To see this, imagine that the population of the country we have been modeling is composed of just two people: Anjali and Budi. Budi's parents have urged him to take up fishing, and Anjali's parents, too, have urged her to continue with the family's traditional livelihood.

To determine if each will take up fishing or shirt-making they will engage in the noncooperative game shown in Figure 6.11. Assuming that each spends five hours a day working, we have calculated their output depending on their choice and the choice of the other, using the production functions in Figure 6.4. So:

- Fishing for five hours will produce 2.5 kg independently of what the other does, and the price of fish is 1, so the value of their output if either of them fish is just 2.5.
- If both produce shirts, that is ten total hours of shirt production resulting in ten shirts, or five for each of them; at the price 0.67, they both receive a value of output of 3.33 (5×0.67).
- If one produces shirts and the other does not, the production function for shirts tells us that the output will be $\frac{1}{10}(l^s)^2$, with $l^s = 5$; this results in 2.5 shirts, with a value of 1.67 (2.5×0.67).

Figure 6.11 Two people in a country have to choose whether they will play "Shirts" or "Fish." At given prices of $p = \$1$ for fish and $p = \$0.67$ for a shirt, they confront the following payoffs. The game is an Assurance Game with two Nash equilibria (Shirts, Shirts) and (Fish, Fish) where (Shirts, Shirts) is Pareto superior to (Fish, Fish) and Pareto efficient.

		Player B	
		Shirts	Fish
Player A	Shirts	3.33 / 3.33	1.67 / 2.5
	Fish	2.5 / 1.67	2.5 / 2.5

You can use the circle-and-dot method (introduced in Chapter 1) to identify players' best responses and therefore the Nash equilibria of the game. There are two: both fish or both produce shirts. Producing shirts is Pareto superior to fishing.

How would the two play the game? That would depend on their beliefs. Budi might think that taking up shirt-making is risky because if Anjali does *not* make the same choice then there will be no economies of agglomeration and his payoff would be 1.67. Fishing by contrast is a sure thing as it results in a payoff of 2.5 with certainty. Anjali might well think the same way. Based on the traditions of their society they would probably believe that the other would take up fishing. And so they would both fish. This is similar to the Planting in Palanpur game in Chapter 1.

Of course if they could have agreed to both produce shirts, then they would have benefited from the economies of agglomeration and each produced twice as many shirts (five) as one of them working singly could do. But we are letting Anjali and Budi represent an entire population who are mostly strangers to one another, not two neighbors who could agree on a course of action.

So they have no way of coordinating their actions. Like the farmers of Palanpur—all planting late when they could all be better off by planting

early—they will be less well-off because of the poverty trap which they cannot escape because they lack institutions that would allow them to coordinate a joint decision.

This kind of self-perpetuating specialization is part of the reason why so much of the world remained poor while other parts became wealthier. The labor force of Africa, Asia, and Latin America engaged in agriculture and other low-productivity sectors. Europe and its offshoots (North America, Australia, and New Zealand) became wealthier starting in the early nineteenth century in some measure by producing shirts and other manufactured goods.

In the late nineteenth century and into the twentieth century other countries shifted their specialization to sectors with higher labor productivity. This began with Japan, and continued with South Korea, Singapore, China, and Vietnam. These countries shifting to manufacturing as a higher labor productivity activity mirrors our example of switching to shirts from agriculture. The modern manufacturing in these countries includes electronics, ship-building, and automobile production.

In all of these cases the change in specialization occurred in large part as a *deliberate* government project (which addressed the coordination failure), not as the result of countless people deciding to produce commodities like shirts rather than fish.

✓ **FACT CHECK** Consider South Korea as an example. In 1960, farming produced more than three times as much as manufacturing. In 2019, manufacturing produced 15 times as much as South Korea's farming. This kind of transformation is what motivated our choice of "fishing" (agriculture, forestry, and fishing) vs. "shirt-making" (manufacturing). Source: World Bank Development Indicators (2020).

CHECKPOINT 6.8 **Poverty traps** Explain why a country specializing in fish and acquiring shirts by exchanging fish with a shirt producing country (as shown in Figure 6.10) would

a. be better off specializing in shirts instead, but

b. might find it difficult or impossible to switch from fish to shirts.

6.9 APPLICATION: THE LIMITS OF SPECIALIZATION AND COMPARATIVE ADVANTAGE

Economies of scale and opportunities to exchange are common in modern societies, and, as a result, we live with an extensive (even global) division of labor in which many individual households and firms specialize in producing only one or a narrow range of products and meet their needs by exchanging these products through monetary transactions.

When we think of specialization, we often conjure images of Silicon Valley's engineering and technology hub or the City of London's financial center. But, India is home to one of the most developed and specialized information technology industries in the world based in Bangalore. The Bangalore-based IT firms InfoSys and Wipro exemplify the dynamics of an

industry that grew from nothing in the early 1980s to become major global players by the early 2000s.[3]

Specialization occurs, too, in older industries. Manufacturers in Bangladesh export a lot of shirts and hats, and very few bedsheets, whereas firms in Pakistan export a great number of bedsheets, but very few hats.[4] Neither is a particularly skill-intensive kind of production and there is no reason for us to expect that one of them *ought* to be better at bedsheets than hats. But they have specialized due to the advantages of learning-by-doing and economies of scale.

In contrast with this specialization, however, many households do still remain *diversified* rather than *specialized*. Many households cannot achieve the benefits of economies of scale due to insufficient wealth to sustain the training and investment required for specialization, and also because of the riskiness of starting businesses or engaging in just a single kind of work. As a result, many poor Indian households diversify rather than specialize.

Figure 6.12 A street-side dosa.

Photo ©Ashish611/Shutterstock.

The women spend time in the morning selling *dosas* (a rice and bean breakfast food), they make small amounts of money collecting trash, they gather firewood to sell, they sell fruit, vegetables, and clothing (mostly saris), they make and sell pickles, or they work as short-term laborers. Similar patterns of diverse occupations occur in Côte d'Ivoire, Guatemala, Indonesia, Pakistan, Nicaragua, Panama, Timor-Leste, and Mexico.

An example from India shows one extreme: a survey by Nirmala Banerjee in West Bengal showed that the average family had three people who worked, sharing seven occupations among them.[5]

The economic analysis of these two different configurations of production–specialization or diversification–is based on the same fundamental concept–doing the best you can given a set of constraints. But as the examples above show, whether a person or family specializes or diversifies is not simply a matter of technology–economies or diseconomies of scale for example, or learning-by-doing or differential skills. For a family with limited or no wealth and exposed to uncertainty of their incomes in any single pursuit, risk mitigation becomes an important priority. As a result, diversification may be the best they can do. We show in Chapter 13 how this very common combination of limited wealth and exposure to uncertainty may contribute to the perpetuation of poverty.

6.10 PRODUCTION TECHNOLOGIES

In modern economies, **production** takes place in families, in governments, in privately owned firms, and in other settings, each distinguished by a

PRODUCTION Production is the process by which we transform the resources of the natural world using already produced tools, facilities, and other inputs to meet human needs.

characteristic set of rules of the game determining who owns the goods produced, who directs the production process, and so on.

Private owners of the buildings, machinery, intellectual property, and other assets making up a firm aim to sell the output (which they also own) for more than their inputs cost, the difference between sales revenues and costs being the owners' *profit*. The owners (or managers) of the firm choose the methods of production, the amounts of inputs it hires (hours of labor, number of machines), and the level of output to *maximize their profits*, given the production methods available, the prices they pay for inputs, and the market prices for their output. For this reason owners of firms want to *minimize the costs* that they incur to produce any given level of output that they decide to produce. Here we explain how the owners of firms choose cost-minimizing *technologies* to use in converting raw materials or unfinished goods into some given level of output of products for sale. In Chapters 8 and 9 we turn to the owners' decision about how much to produce, given the technology they have chosen to use.

Inputs and outputs

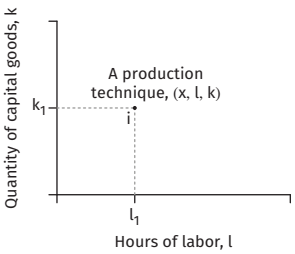
Consider a firm producing an *output*, for example, cars, smartphones, or English-language lessons. To produce its output, x, the firm needs to hire labor with the skills necessary for the production tasks and provide the workers with raw materials, tools, and facilities. Think of the inputs as a list describing the amounts of labor of each kind and of all the different raw materials (wood, steel, plastic, glass), tools (dies, drill presses, computer forges), and facilities (factories, offices, labs) required to produce the output. These inputs to the production process are termed **factors of production**. We would measure all these inputs over the same time period as output: so many hours of each kind of employee per month, so much steel per month, so much office space per month, and so on.

It's easy to see how the process works if we look at two dimensions on horizontal and vertical axes in Figure 6.13. The labor hired is l (on the horizontal axis), and k (on the vertical axis) is the quantity of capital goods that the firm uses. Its capital goods are the machines, tools, and facilities that the firm needs to produce their output over the relevant time period. Each point in Figure 6.13 is a different list of factors of production.

We can describe one way of producing a particular level of output by indicating a level of the labor input, l, and the capital goods input, k, that will produce the specified output, x. This combination (x, l, k) describes

FACTOR OF PRODUCTION Any input into a production process is called a factor of production. In the past economists often referred to land, labor, and capital goods as primary factors of production, but this usage is overly narrow given the essential role of other production inputs such as our natural environment beyond "land" and knowledge.

Figure 6.13 A production technique, (x, l, k) producing some amount of output x using labor l_1 and capital goods k_1.



one of the firm's possible **techniques of production**. For a given level of x, we can describe the technique of production—which is a particular way of producing some given amount of output—as a point in (l, k) space, as in Figure 6.13.

Technology and feasible production

The firm is constrained by the available technology, which describes what techniques it can in fact carry out, given its state of knowledge, the skills of workers, and the conditions of work (health, safety, and intensity) that the firm can legally and socially impose on its workers. Technology is therefore *not* just a question of engineering or scientific knowledge, but also involves relations between workers and management and among workers themselves, as well as the legal and institutional framework within which the firm operates.

Figure 6.14 displays the feasible set for producing a hundred units of output. The green-shaded region shows the set combinations of capital goods and labor, sufficient to produce a given output, $\underline{x} = 100$, of the good x. The dark-green lines are the border of the feasible set called an **isoquant**, meaning equal (iso) quantity. There is a set of isoquants, called an isoquant map (not shown), each one of them associated with a different level of output.

Each of the four lettered points in Figure 6.14 is a particular combination of labor and capital goods that are sufficient to produce 100 units of good x. But the owners of a firm seeking to produce that amount would not be equally happy to use any of the four.

Technique **i** *dominates* technique **h** because it uses *less* of both inputs to produce the same output. Similarly techniques **f** and **g** are dominated by point **i**: each uses the same amount as does technique **i** of one input and more of the other input to produce 100 units. Technique **i** is called **technically efficient** because (considering the alternatives, **f**, **g**, and **h**) there is *no other technique* that produces the required amount of output ($x = 100$) with less of at least one input and not more of any other.

The production isoquant derived from a production function is analogous to the indifference curves based on the utility functions in Chapter 3. But,

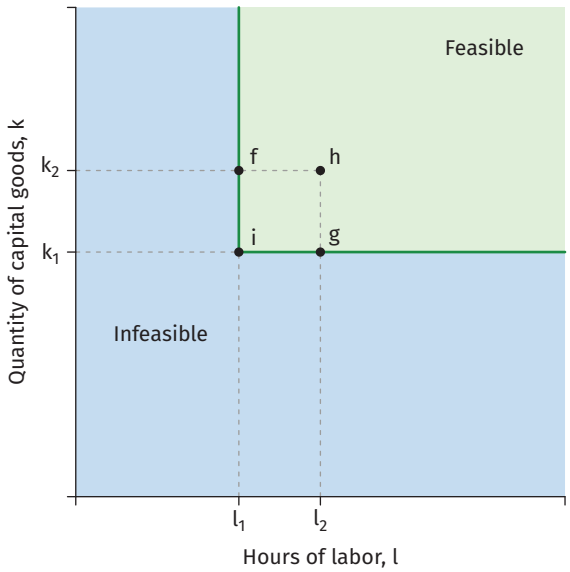
! REMINDER A Pareto-efficient allocation is one that is not dominated by any alternative, so there is no other allocation that is preferred by at least one person and not "dis-preferred" by any person. The definition of technical efficiency is similar but applying to techniques and inputs used rather than allocations of goods, and people's utilities.

TECHNIQUE OF PRODUCTION A technique of production is a particular way—bundle of inputs—of producing some given amount of output.

ISOQUANT An isoquant gives the combinations of two inputs that are just sufficient to produce a given level of output.

TECHNICAL EFFICIENCY A technique of production is technically efficient if there is no other technique with which the same output can be produced with less of at least one input and not more of any input.

Figure 6.14 Production techniques. For a given level of output $\underline{x} = 100$, we can describe any technique of production as a point showing the amount of labor, l, and the quantity of capital goods, k, required to produce output x. The area shaded in green shows the *feasible* combinations of labor and capital goods that can produce output x, which is equivalent to the feasible set introduced in Chapter 3. The area in blue shows the *infeasible* combinations of capital goods and labor to obtain an output of $x = 100$.



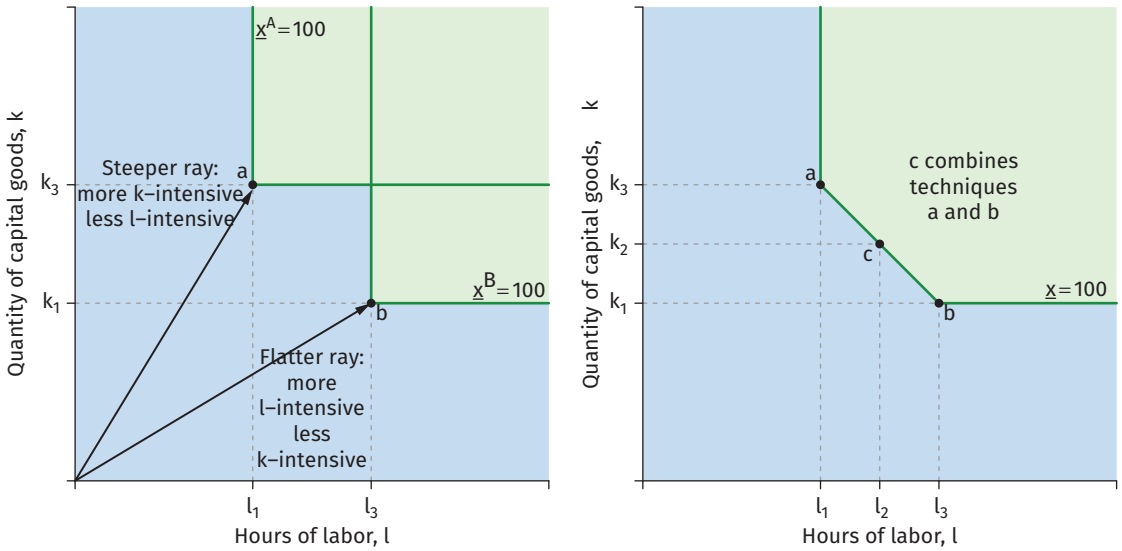
it is important to remember that a production isoquant is a *constraint* on the choice of inputs required to produce a particular level of output, rather than something to be maximized. A utility maximizer wants to get to the highest possible indifference curve given the set of feasible options. The cost-minimizing firm wants to get to the *minimum cost point* on the production isoquant for any given level of output.

Figure 6.15 (a) shows two production isoquants representing two techniques of production one of which uses more capital goods and less labor than the other. The production isoquant in Figure 6.15 (b) includes the points representing the two techniques in panel (a) and the line joining them, representing the possibility of doing some of the production with one technique and some with the other. The isoquant in panel (b) provides some possibility of *substitution* of one input for the other by switching from one technique to the other, but this substitution is limited because there are only two techniques.

As shown in Figures 6.14 and 6.15, a production isoquant can be thought of as points corresponding to the various techniques of production, and the lines connecting those points (which correspond to mixing the techniques of production). The production isoquant is equivalent to the idea of the

! REMINDER A production function $x = f(l, k)$ describes a firm's available set of techniques of production as a mathematical relationship. Here we present production functions with just two inputs—labor and capital goods—but production functions may describe the relationship between output and any number of inputs, labor with different skills, for example, or different kinds of capital goods (buildings, machines, and so on).

Figure 6.15 A production isoquant combining two techniques. For a given level of output x, there may be more than one feasible technique of production. The production isoquant in this figure consists of the two techniques, and the line between them, representing production with combinations of the two techniques (one shown by point **c**). The availability of more than one technique implies that *substitution* of one input for the other is possible by shifting some production from a more capital-intensive technique to a more labor-intensive technique.



(a) Comparing two techniques (b) Combining two techniques

! REMINDER We have already examined production functions, but so far they have only involved one input, such as labor as an input into studying in Chapter 3 or labor as an input into either fishing or shirt production in section 6.3.

feasible frontier in earlier chapters as it defines what combinations of labor and capital goods can *feasibly* produce the level of output, x.

6.11 PRODUCTION FUNCTIONS WITH MORE THAN ONE INPUT

The techniques of production available are often described in a production function, which is a mathematical expression giving the least quantity of inputs—such as capital goods (k) and labor (l)—that are sufficient to produce any given level of output, x. The production function can also be thought of as specifying the *maximum* level of output attainable for each combination of inputs:

$$\text{Production function } x = f(l, k) \tag{6.10}$$

You have already seen examples of the simplest production function in the *Leontief production function* in which there is but a single technique available for a given level of output x as in the production isoquants in Figures 6.14 and 6.15.

As in those figures, what are called Leontief production isoquants are rectangular because the technology specifies a given ratio of capital goods

to labor (at the point of the rectangular isoquant closest to the origin, such as point **i** in Figure 6.14, or points **a** and **b** in Figure 6.15). If that particular ratio of inputs is in use, then adding more labor or more capital goods has no effect on production: their marginal products are *zero*. There are therefore no possibilities of substituting one factor of production for another.

To clarify what this "no substitution" characteristic means with an extreme example, think about nuts and bolts: if you have n nuts and n bolts, then having $n+1$ bolts is no better than having n bolts. A bolt is useless without a nut, and a nut is useless without a bolt. You need to use the inputs in fixed proportion to each other to get "a nut and a bolt."

HISTORY Wassily Leontief (1906–1999) was a Russian-American Nobel Laureate in economics. He modeled the whole economy as what became known as an input–output system, with each industry being represented by a Leontief production function. His work is valued by economists because it allows a mathematical representation of the whole economy that can be estimated empirically (for example, engineers can determine how many tons of coal are needed to produce a ton of steel).

M-NOTE 6.4 Leontief production function

The output of good x, is produced with l the amount of labor input used and k the amount of capital goods used. Let a_l and a_k be the minimum amounts of labor and capital goods required to produce a single unit of output.

Noting that $min(m,n)$ means m and/or n, whichever of m or n is *lowest* (or both of them if they are equal), the Leontief production function can be written:

$$x = f(l,k) = \min\left(\frac{l}{a_l}, \frac{k}{a_k}\right) \tag{6.11}$$

The equation can be read: "The number of units of x produced is the smaller ("min.") of the ratio of the amount of the input used (the numerator in the two fractions) to the input required for a single unit of production (the denominator)." Any capital goods input in excess of the minimum amounts required is of no use in production in producing a single unit.

EXAMPLE Leontief's input–output models are today used, for example, to calculate the amount of CO_2 emissions produced per unit of output of each industry, taking account of both the direct and the indirect inputs. That is counting for example not only the coal used to produce a ton of steel, but the coal used in producing the machinery and all of the other inputs required for a ton of steel.[6]

Cobb–Douglas production function

Another representation of how inputs are combined to produce outputs is the *Cobb–Douglas production function*.

$$\text{Cobb–Douglas production function} \qquad x(l,k) = ql^{\alpha}k^{\beta} \tag{6.12}$$

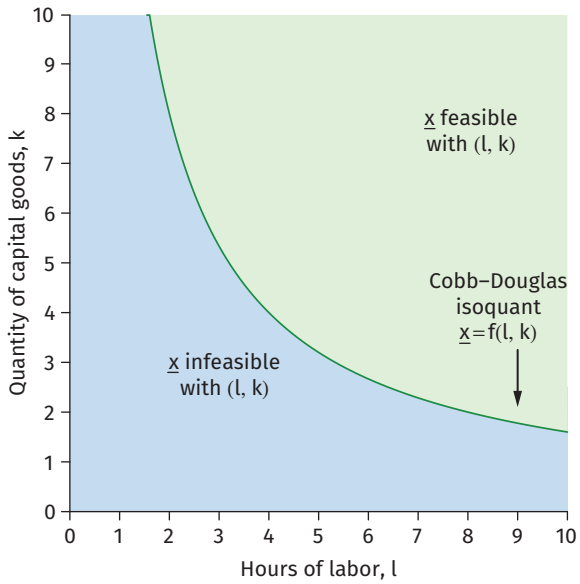
The Cobb–Douglas production function requires that $l > 0, k > 0$ for production to take place: some of both inputs are essential, but their proportions used can vary. The parameters of the Cobb–Douglas production function convey the following information.

- $0 < \alpha$ and $0 < \beta$ capture the contribution of labor and capital goods, respectively to producing output;
- the sum of α and β tells us how output responds to proportional increases in both of the inputs indicating whether the firm experiences economies of scale ($\alpha + \beta > 1$), diseconomies of scale ($\alpha + \beta < 1$), or constant returns to scale ($\alpha + \beta = 1$);
- $q > 0$ is a positive constant that captures a level of productivity of the specific technology.

REMINDER The Cobb–Douglas production function has the same structure as the Cobb–Douglas utility functions we studied in Chapter 3.

A Cobb–Douglas isoquant for $\underline{x} = 100$ is shown in Figure 6.16. The green-shaded area shows the *feasible* combinations of capital goods and labor that

Figure 6.16 Production isoquant given a Cobb–Douglas production function. The feasible set of production for a given x is the set of techniques of production, combinations of labor input, and capital goods, (l, k) that permit the firm to produce some given amount of the output, $\underline{x}$.



EXAMPLE Constant returns to scale. Given the production function $x(l,k) = l^{\alpha}k^{\beta}$ suppose $\alpha = \frac{1}{2} = \beta$ so the sum of the exponents is equal to 1. Compute the output if both k and l are equal to 2. Output is equal to 2. Now double the inputs so that both are now 4. The output is now $x(l,k) = 4^{\frac{1}{2}}4^{\frac{1}{2}} = 4$. So doubling both inputs doubled the output: this is constant returns to scale.

are more than enough or just enough (the dark-green line) to produce $\underline{x}$. The blue-shaded area shows the combinations of capital goods and labor that are insufficient to produce $\underline{x}$, that is, they are *infeasible*.

The marginal rate of technical substitution

The **marginal rate of technical substitution** is the negative of the slope of the production isoquant and equal to the ratio of the marginal product of the input on the x-axis to the marginal product of the input on the y-axis. It shows how much more of the y-axis input (k in our figures) must be added to compensate for the withdrawal of one unit of the x-axis input (l in our figures) so that output is unchanged. You can think of the isoquant as another feasible frontier, the boundary of the feasible set of inputs sufficient to produce the particular level of output.

MARGINAL RATE OF TECHNICAL SUBSTITUTION The marginal rate of technical substitution is the negative of the slope of the production isoquant and is equal to the ratio of the marginal product of the input on the x-axis to the marginal product of the input on the y-axis. It shows how much more of the y-axis input must be added to compensate for the withdrawal of one unit of the x-axis input so that output is unchanged.

Figure 6.17 A production isoquant for the Cobb–Douglas production function $x(l,k) = l^{0.5}k^{0.5}$ for the output level $\underline{x} = 4$. The marginal rate of technical substitution is the ratio of the marginal products, $mp_l/mp_k = x_l/x_k$, which is the negative of the slope of the production isoquant $\Delta k/\Delta l$. Three points along the isoquant curve are shown: **a**, **b**, and **d** illustrating how the marginal rate of technical substitution decreases comparing points on the production isoquant as the ratio of labor to capital goods inputs increases. Three gray dashed lines are tangent to the production isoquant, the slopes of which are the negative of the marginal rate of technical substitution, $mrts(l,k)$, at each point.

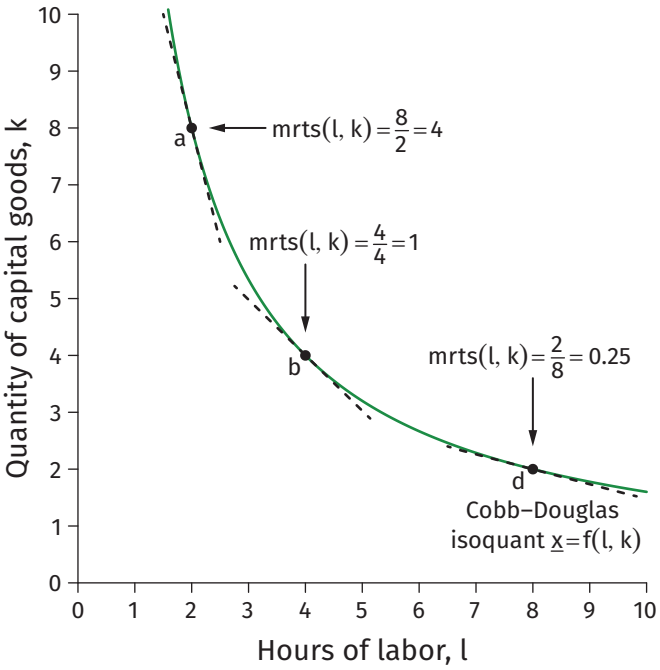


Figure 6.17 shows a Cobb–Douglas isoquant with the combinations of capital goods and hours of labor that can produce $\underline{x} = 4$. The figure also shows the value of the marginal rate of technical substitution at three points **a**, **b**, and **d**. Comparing the points, we can see that the $mrts(l,k)$ decreases as the ratio of the labor input to the capital goods input increases.

In M-Note 6.5 we show that:

$$\text{mrts} = \frac{\text{Marginal product of labor}}{\text{Marginal product of capital}} \equiv \frac{x_l(l,k)}{x_k(l,k)} \tag{6.13}$$

To see what this means, think about the technique given by the input bundle (l_d, k_d), point **d** in Figure 6.17. The reason why the *mrts* is low (=0.25) at point **d** is that the substantial amount of labor used at that point has diminished the marginal product of labor (x_l), which is the numerator in the expression for the *mrts*. And the very limited amount of capital goods in use at point **d** also means that the marginal product of capital goods (x_k, the denominator of Equation 6.13) is high.

HISTORY Paul Douglas (1892–1976) developed the function with his colleague at Amherst College, Charles Cobb. Though a Quaker, Douglas was fiercely anti-fascist and during World War II volunteered for the US Marine Corps as a private at the age of 50. He later won two purple hearts in recognition of the battle wounds he suffered in the Pacific theater. He went on to be a prominent member of the Democratic Party and a US Senator serving from 1949 to 1967.

M-NOTE 6.5 **The marginal rate of technical substitution and marginal products**

The production isoquant is defined as the combination of inputs that can produce a given output, $x(l,k) = f(l,k)$.

To find the slope of an isoquant we proceed as we did when finding the slope of an indifference curve. We use the property of the isoquant that the points on it made up of different amounts of l and k result in the same level of output x. So for small changes in l and k the following is true:

$$dx = \frac{\partial x}{\partial l}dl + \frac{\partial x}{\partial k}dk = 0 \tag{6.14}$$

$$dx = x_l(l,k)dl + x_k(l,k)dk = 0 \tag{6.15}$$

Where x_l is the marginal product of labor and x_k is the marginal product of capital. Because along a production isoquant the difference in output is zero (just like along an indifference curve the difference in *utility* is zero), Equation 6.15 can be understood as follows:

$$\underbrace{x_l(l,k)dl}_{\substack{\text{Change}\\ \text{in x as}\\ \text{l changes}}} + \underbrace{x_k(l,k)dk}_{\substack{\text{Change}\\ \text{in x as}\\ \text{k changes}}} = 0$$

Rearranging $$mrts(l,k) = -\frac{dk}{dl} = \frac{x_l(l,k)}{x_k(l,k)} \tag{6.16}$$

Equation 6.16 can be stated as:

$$\text{Marginal rate of technical substitution} = \frac{\text{Marginal product of labor}}{\text{Marginal product of capital}}$$

This is the negative of the slope of the production isoquant.

M-NOTE 6.6 **Cobb–Douglas economies of scale**

To study economies of scale, start with the following Cobb–Douglas production function:

$$x(l,k) = ql^{\alpha}k^{\beta}$$

and then increase both inputs by some proportion, S. Therefore, increase l and k by the proportion S. That is, multiply each input by S before raising the input to the relevant power:

$$x(Sl,Sk) = q(Sl)^{\alpha}(Sk)^{\beta}$$

Now, take each S out of the parentheses:

$$\begin{aligned} x(Sl,Sk) &= qS^{\alpha}l^{\alpha}S^{\beta}k^{\beta} \\ &= S^{\alpha+\beta}ql^{\alpha}k^{\beta} \\ &= S^{\alpha+\beta}x(l,k) \end{aligned}$$

continued

The final step occurs because we know that $ql^{\alpha}k^{\beta}$ is our original production function with output, $x(l,k)$. If $\alpha+\beta$ is *greater* than one, then the output increased *more* than proportionally with an increase of l and k by the proportion S. Therefore, the production function has *increasing returns to scale*. If $\alpha+\beta$ is *less* than one, the production function has *decreasing returns to scale*. If $\alpha+\beta$ is *equal to one*, it has *constant* returns to scale.

CHECKPOINT 6.9 **Cobb–Douglas Constant returns to scale** In the "Example: Constant returns to scale" margin note you saw that if the sum of the exponents of the Cobb–Douglas production function is equal to one, then doubling both inputs from 2 to 4 doubled output. Show what happens to output for the same doubling of inputs, but when the exponents sum to *more* than one, i.e. $\alpha = 2 = \beta$.

Diminishing marginal products of inputs

It is important not to confuse *economies and diseconomies of scale*, which describe what happens when *all* inputs are changed proportionally with *diminishing or increasing marginal productivity* of one input when the other inputs are held constant (say, increasing labor, holding capital goods inputs constant).

A production function may have diminishing marginal productivity to any one input when the others are held constant, and still exhibit economies of scale when all the inputs change together.

M-CHECK For the Leontief production function we cannot compute the marginal rate of technical substitution from the slope of a production isoquant because its slope is undefined at the kink in the isoquant. But at the "kink" in the isoquant, adding more capital goods or more labor has no effect on output, so we could view the Leontief production isoquant as representing an extreme form of diminishing marginal products.

M-NOTE 6.7 **Diminishing marginal productivity**

To compute the marginal product of labor in the Cobb–Douglas production function we start with the production function:

$$x(l,k) = ql^{\alpha}k^{\beta}$$

To find the marginal product of labor, we calculate the first partial derivative of the production function with respect to labor, which gives us the effect on total output of a small change in the labor input, holding constant the level of capital goods input:

$$\begin{aligned}\frac{\partial x(l,k)}{\partial l} = x_l &= \alpha q l^{\alpha-1}k^{\beta}\\ &= \frac{\alpha q l^{\alpha}k^{\beta}}{l}\\ &= \frac{\alpha x(l,k)}{l}\end{aligned}$$

For $\alpha > 0$ and $l > 0$, the marginal product of labor is positive: as you can see from the equation immediately above, it is equal to α itself times the average product.

To work out whether the marginal product of labor is diminishing, we need to know whether the derivative of the marginal product of labor with respect to the labor input itself is positive, zero, or negative:

continued

$$\frac{\partial^2 x(l,k)}{\partial l^2} = x_{ll} = \alpha(\alpha-1)ql^{\alpha-2}k^{\beta}$$
$$= \frac{\alpha(\alpha-1)ql^{\alpha}k^{\beta}}{l^2}$$
$$= \frac{\alpha(\alpha-1)x(l,k)}{l^2}$$

The sign of x_{ll} depends on the size of α.

Diminishing If $\alpha < 1$, then $\alpha - 1 < 0$, so $x_{ll} < 0$ which means *diminishing* marginal productivity of labor.

Constant If $\alpha = 1$, then $\alpha - 1 = 0$, and $x_{ll} = 0$, that is, *constant* marginal productivity of labor.

Increasing If $\alpha > 1$, then $\alpha - 1 > 0$, that is, *increasing* marginal productivity of labor.

CHECKPOINT 6.10 Production functions and factor inputs Check your understanding by doing the following:

a. Explain why, in Figure 6.15 at point **a**, adding more labor has no effect on production.

b. Explain why in Figure 6.16 the isoquant is downward sloping.

c. We have the following Cobb-Douglas function: $x(l,k) = 2l^{0.7}k^{0.5}$. Does this production exhibit economies or diseconomies of scale? Does it exhibit diminishing or increasing marginal productivity of labor?

6.12 COST-MINIMIZING TECHNIQUES

Having introduced a description of the production process—the production function—we now consider the firm as a profit-maximizing entity. To determine the level of output that will yield the greatest profit for the owners of the firm, consider two pieces of information that the owners of the firm would need:

- *Cost minimization*: for every possible level of output, given the costs of using the inputs to the production function, find the technique of production that minimizes the costs of production.
- *Profit maximization*: using the resulting cost curve (describing the least cost at which each level of output can be produced) and the demand curve for the firm's product, determine the level of output to produce.

Here we describe cost minimization. We describe the profit maximization step in Chapters 8 and 9. We call any particular combination of labor and capital goods used (l,k) a bundle of inputs. Finding the minimum cost bundle for producing each level of output the firm's owners might want to produce requires three steps:

- *Step 1*: Calculate the cost of every input bundle that the firm might use.
- *Step 2*: Identify bundles that cost the same, and use the resulting isocost line to distinguish between more costly and less costly bundles.
- *Step 3*: Use the isoquants based on the available production functions to determine, for each level of output, the least costly bundle.

In this process the owners of the firm seek to minimize the cost of producing a given level of output so the isoquant is the *constraint* not the firm's objective. It tells the owners of the firm what combinations of inputs will produce a given level of output.

! REMINDER Isoquants look similar to an individual's indifference curves which represent the person's objectives; but isoquant curves represent constraints on the firm's cost minimization process, not the firm's objectives.

Isocosts: Equally costly bundles of inputs

We consider a case in which:

- the capital goods used by the firm are rented (for example, buildings and equipment) rather than owned; and
- the firm's own demand for labor and capital goods *does not* influence the price it pays for these inputs. This assumption would be true, for example, if the firm were small in relation to the markets for its inputs so that hiring more inputs does not raise market-wide demand sufficiently to increase the price.

! REMINDER In earlier chapters the indifference curves bowed in toward the origin (like the green isoquant in Figure 6.19) represented the *objectives* of the *person*, that is the thing that she wished to *maximize* based on her preferences, subject to some constraint, for example a limit on how much she could spend. Here the blue isocost lines represent the objective of the firm's owners, that is the thing they wish to *minimize*, while the curved isoquant is the *constraint* based on the feasible set of production techniques that produce at least the given amount of output.

Then the cost of using any particular combination of labor and capital goods depends on:

- Wages (w) paid per hour for the hours of labor hired (l), for a total cost of labor of wl.
- The rental cost of the capital goods (p_k) times the quantity of capital goods used (k) for a total cost of capital goods of $p_k k$.

Then the cost, $c(l,k)$, of a bundle of inputs is:

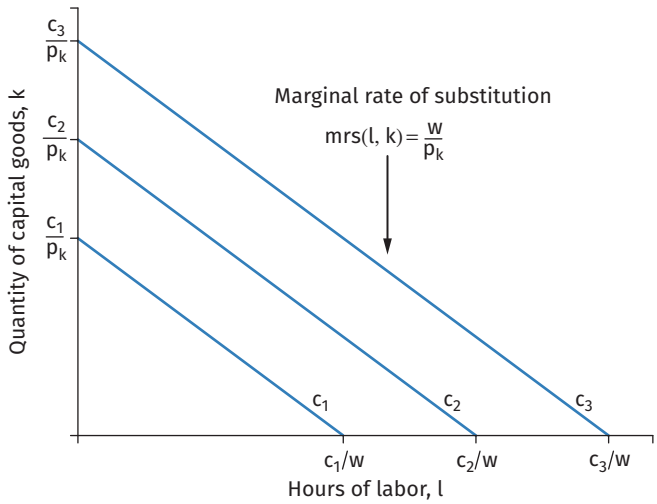
$$c(l,k) = wl + p_k k \tag{6.17}$$

Using Equation 6.17, we know the cost of every input bundle. So we can construct an **isocost line**, showing all the possible combinations of amounts of labor and amounts of the capital good that result in a constant or equal ("iso") level of costs. Rearranging Equation 6.17, we can find the equation for an *isocost line* associated with total costs of c:

$$\text{Isocost line} \qquad k = \frac{c}{p_k} - \left(\frac{w}{p_k}\right) l \tag{6.18}$$

ISOCOST LINE A line that represents all combinations of inputs that cost a given total amount.

Figure 6.18 Three isocost lines are presented: c_1, c_2, and c_3. Isocost lines closer to the origin are made up of less costly input bundles. The equation for an isocost line is given by $c = p_k k + wl$, where p_k is the cost per unit of renting capital goods, k is capital goods input, w is the wage, and l is the quantity of labor input. We can rearrange this equation in terms of the capital goods input, $k = \frac{c}{p_k} - \left(\frac{w}{p_k}\right)l$. The slope of the isocost line is determined by the marginal rate of substitution of capital goods into labor, $mrs(l,k) = \frac{w}{p_k}$, which is the tradeoff of using more labor in terms of the lesser quantity of capital goods that can be used, in order to hold constant the cost of the resulting bundle.



A set of isocost lines—called an isocost map—is shown in Figure 6.18. Each line corresponds to a constant cost level, $c_1 < c_2 < c_3$. The isocost lines represent the *objectives* of the owners of the firm when they are seeking to find the combination of inputs that will minimize the cost of producing some particular output. The slope of an isocost line is:

$$\frac{\Delta k}{\Delta l} = -\frac{w}{p_k} = -\text{marginal rate of substitution} \tag{6.19}$$

! REMINDER We refer to the trade-offs associated with an actor's objectives as a marginal rate of substitution, while the trade-offs arising from the actor's constraints are represented by a marginal rate of substitution.

We call the negative of the slope of the isocost line the *mrs* because, for cost to remain constant, the given level of the negative of the slope is the amount of the capital good that must be increased to substitute for a reduction in the labor input.

The owners would like to find the technique of production that will put them on the *lowest* feasible isocost line, that is, the one closest to the origin. The *constraint* limiting the owners' decision is the available technology or technologies as described by the production isoquant for the given level of output.

Remember the owners do not know that they will produce this given amount of output, and most likely they will not. Instead they are doing a

thought experiment, reasoning as follows: if we were going to produce this given amount, what would be the least cost way to do it. They do the same minimum cost exercise for every level of output that they might produce. They need this information along with information about buyers' demand for their product to determine the level of output that will maximize their profits.

The cost-minimizing choice of technique

Could point **d** in Figure 6.19—the technique given by the input bundle (l_d, k_d)—be the least-cost way of producing output $\underline{x}$? The opportunity cost of using less labor—that is the additional amount of capital goods that would be required to sustain the production of $\underline{x}$ using less labor—is small, shown

Figure 6.19 The minimum cost of producing a given level of output. To produce the output given by the isoquant $\underline{x}$, the least-cost input bundle is indicated by point **b** where $mrts = mrs$.

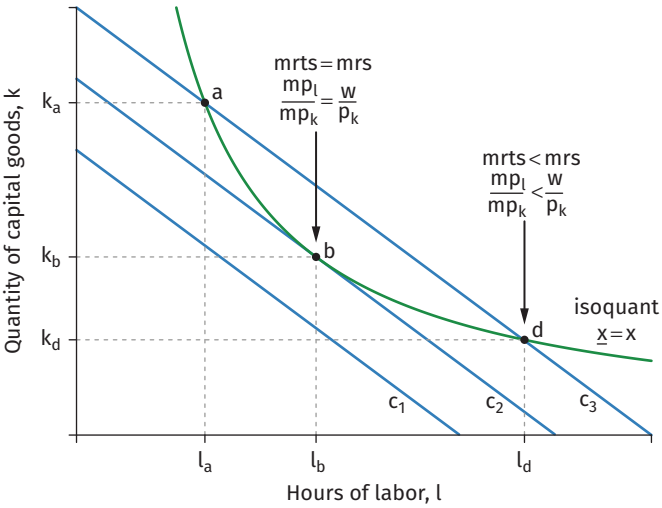


Table 6.3 Four rules: individual constrained optimization, societal Pareto efficiency, and firm cost minimization.

Tangency rules	Tangency of	Rule for what
$mrs = mrt$	An individual's indifference curve and feasible frontier	Individual constrained optimization
$mrs = mrts$	The firm's owners' production isoquant and isocost lines	Individual cost minimization
$mb = mc$	Restatement of $mrs = mrt$ using marginal benefits and costs	Individual constrained optimization
$mrs^A = mrs^B$	Two or more people's indifference curves	Societal (multi-person) Pareto efficiency

by the relatively flat isoquant at point **d**. Remember, this is because the marginal product of labor is low (due to the substantial amount of labor being used) and the marginal product of capital goods is high (because few are in use). This quantity–the *mrts*–falls short of the added use of capital goods that would substitute for the reduced labor hired, resulting in no change in the cost incurred by the firm (indicated by the steepness of the isocost curve). So we have, using the technique (l_d, k_d):

$$mrts = \frac{x_l}{x_k} \quad < \quad \frac{w}{p_k} \equiv mrs \tag{6.20}$$

This means that the given level of output could be produced at lower cost by substituting more capital goods for labor hours. The firm would continue to substitute capital goods for labor hours until the point where the ratio of marginal products equals the ratio of prices for the inputs at **b**, where:

$$mrts = \frac{x_l}{x_k} = \frac{w}{p_k} \equiv mrs \tag{6.21}$$

M-CHECK As with the $mrs = mrt$ rule, there are cases where the $mrts = mrs$ rule does not apply. For example, the cost-minimizing technique could be a corner solution in which no capital goods at all were used.

The *minimum cost technique of production* to produce output $\underline{x}$ is therefore the input bundle (l_b, k_b), where the ratio of marginal products equals the ratio of input prices, $\frac{x_l}{x_k} = \frac{w}{p_k}$, or what is the same thing, the slope of the $\underline{x}$ isoquant is equal to the slope of an isocost line.

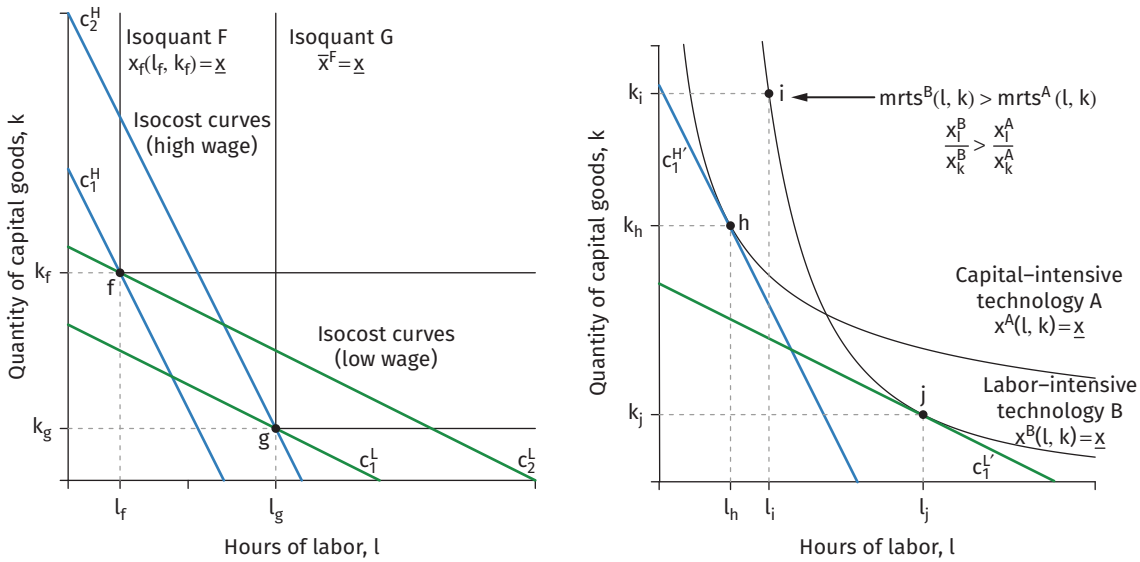
CHECKPOINT 6.11 **Choices of capital goods and hours of labor** Make sure you understand Figure 6.19 by explaining why, if the firm were producing at point **a**, it could *reduce* costs of producing the given amount of output by using *fewer* capital goods and *more* hours of labor.

Input prices and the choice of a labor- or a capital-intensive technique

Now, think about a firm that sells some product and is considering which of two technologies to use producing the output. One uses some powerful machinery (the capital good) and little labor while the other technology uses lots of labor and a smaller machine. For concreteness, think of the two technologies as similar to plowing a field using a powerful tractor or with a small garden-type roto-tiller. We refer to the contrast of these two technologies as a difference in **factor intensity**. The roto-tiller technology is called labor-intensive and the tractor technology is called capital-goods-intensive.

FACTOR INTENSITY A production function A is more labor-intensive than production function B if for any given ratio of wages to the price of capital goods, the cost-minimizing choice of inputs will be to hire more labor hours when using A than when using B.

Figure 6.20 Choosing a capital-intensive or labor-intensive technology to minimize costs. The cost-minimizing choice of technology depends on the wage and the cost of capital goods. In panel (b) the coefficients for the labor-intensive Cobb–Douglas technology B are $\alpha = 2/3, \beta = 1/3$, and for the more capital-goods-intensive Cobb–Douglas technology A, $\alpha = 1/3, \beta = 2/3$. Higher wages (a steeper blue isocost line) will lead the owner to implement the more capital-goods-intensive technology.



(a) Two Leontief technologies (b) Two Cobb-Douglas technologies

If these two alternative ways of producing the output were described by two Leontief technologies, then we could say that the one using the roto-tiller is the *more* labor-intensive, or what is the same thing (because there are just two inputs) the *less* capital-goods-intensive. In the Leontief technology, the ratio of inputs of labor to the inputs of the capital good required to produce a unit of output is a measure of the labor intensity of the technology. While the more accurate expression is to refer to capital-*goods*-intensive technologies, to save words we sometimes refer to technologies as "capital-intensive."

Figure 6.20 (a) illustrates the case with Leontief technologies and Figure 6.20 (b) illustrates the case with Cobb–Douglas technologies. In Figure 6.20 (a) the Leontief technology indicated by point **g** is more labor-intensive than the technology at point **f**. In Figure 6.20 (b), the Cobb–Douglas technology indicated by point **j** is more labor-intensive than the technology at point **h**.

Where substitution between inputs is possible–as with the Cobb–Douglas technology–the distinction between labor-intensive and capital-intensive technology is not so simple. The basic idea, however, is the same:

the labor-intensive technology is the one that the owners of a firm would choose to minimize costs *if wages were low* relative to the cost of capital goods. A capital-goods-intensive technology, likewise, is one that would be used by a cost-minimizing firm *if wages were high* relative to the costs of capital goods.

EXAMPLE In this video (tinyurl.com/y3vf276o) Economic historian Bob Allen addresses the question of why Britain industrialized when others did not (from the CORE project. www.core-econ.org).

Point **f** in Figure 6.20 (a). shows the inputs required to produce a single unit of output using the capital-intensive technology. Point **g** shows the same information for the labor-intensive technology. Which technology the firm will adopt in order to produce its product at the lowest cost depends on the *relative* cost of labor and capital goods, as indicated by the isocost lines in green and blue. If wages are low, then the isocost lines are flatter, as shown in the figure with the green isocost lines. If the firm uses the labor-intensive technology it will incur costs of c_1^L which is less than the cost it would incur if it used the capital-intensive technology when there are low wages (along c_2^L).

Higher wages (for the same rental cost of the capital good) are indicated by the steeper isocost lines in blue. Using the labor-intensive technology with higher wages (along c_2^H) would incur higher costs than using the capital-intensive technology (along c_1^H).

Figure 6.20 (b) shows an analogous situation with substitutability between the two factors of production with Cobb–Douglas technologies. Once again, the relative costs are shown by two isocost lines, $c_1^{L'}$ and $c_1^{H'}$. The unit isoquants show the different combinations of capital goods and labor that would produce the same output, $\underline{x}$. The owners of the firm would choose point **h** if wages were high and point **j** if wages were low. At points **h** and **i**, the marginal rate of technical substitution differs between the two isoquants, which we can see with the labor-intensive technology B having a much steeper isoquant at point **i** than the capital-intensive technology has at point **h**.

CHECKPOINT 6.12 Capital-intensive and labor-intensive technologies

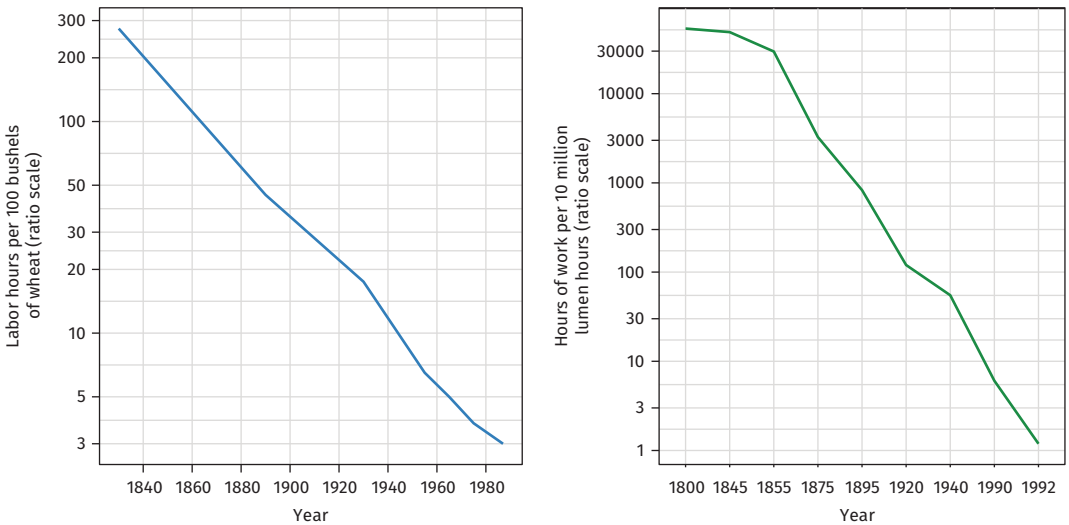
a. Using Figure 6.20 (a), show that there is one ratio of wages to the cost of capital goods such that the lowest cost of producing x will be the same using the two technologies.

b. Show that if a firm had just two technologies to choose from, the Leontief technology F from panel (a) and the Cobb–Douglas technology A from panel (b), it would choose the Cobb–Douglas technology if wages were either very high relative to the cost of capital or very low. But for some input price ratio in between, it would choose the Leontief technology.

c. Explain why this means that it is not always possible to designate a technology as more labor-intensive or more capital-intensive.

Figure 6.21 Improvements in farming and lighting technology over time. In both panels, improvements in technology show the reduced number of hours of labor required to obtain the indicated output (wheat and light). The vertical axis measures what is called a ratio scale so that, for example, the distance between 20 and 100 is the same as the distance between 10 and 50 (the ratio of the first to the second number is the same in both cases). This is equivalent to a logarithmic scale, so the rate of change of the measure is the slope of the lines shown. Notice that the vertical scales are different; the improvements in lighting shown are much greater on a percent per year basis than in agriculture for US farmers.

Sources: Nordhaus (1996); Spielmaker (2018).



(a) Improvements in agricultural technology

(b) Improvements in light technology over time

6.13 APPLICATION: TECHNICAL CHANGE AND INNOVATION RENTS

Production technologies shape how we live, and ongoing changes in technologies are revolutionizing the world. The Industrial Revolution and changes in technology since then have transformed the economies of Europe and North America from largely agriculture to manufacturing and later to service-based livelihoods.

Included in these changes were the shift of most work out of the home and into the factory or office, the enormous increase in the scale of production of typical firms, the widespread replacement of human labor by machines, and vast increases in the quantity of goods and services available along with a decline in the amount of time in one's lifetime spent working.

Figure 6.21 shows the scale of these productive improvements for two technologies: agricultural output and light. Panel (a) shows the change in

! REMINDER A *production technique* is one particular way of producing an output, (x, l, k). A *production function* (for example, the Cobb–Douglas) describes a *technology*, that is, a set of techniques.

› EXAMPLE *Forbes* magazine produces a list of the most innovative firms in the world (https://www.forbes.com/innovative-companies/list/). In 2018 Netflix, Tesla (electric vehicles), Facebook, and Amazon were in the top ten as was Hindustan Unilever (a consumer goods producer and marketer in India), and Naver (selling computer and web services based in South Korea).

the number of hours required to produce 100 bushels of wheat. A bushel of wheat is approximately 60 pounds or about 27.5 kilograms of wheat.

In 1830, farmers on average devoted 275 hours to produce 100 bushels of wheat or over 34 eight-hour workdays. This contrasts with merely three hours required to get the same amount of wheat in 1987 (and even less time today). This is a more than 90-fold improvement in productivity.

In panel (b), we show the amount of labor to obtain 1,000 lumen hours. A lumen is a standard measure of light intensity equivalent to the light of one candle. The increasingly steep slope of the line in panel (b) indicates an acceleration of the rate of decline in the amount of labor required to produce a given amount of light.

Figure 6.21 illustrates the process called **technical progress**, that is, the reduction over time in the quantity of inputs required to produce some given level of output. In the two cases shown the measure is output per unit of labor time. Ideally, however, technical progress is measured by an increase in overall productivity—for example an increase in q in the Cobb–Douglas function—that reduces the amount of both labor and capital goods required to produce the output.

Innovation rents

The main reason for the astounding reduction in the labor time it takes to produce the things we need—whether wheat or light—is that people are ingenious and reducing the cost of producing things can be very profitable. The owners of firms seek to reduce costs by their choice from existing ways of producing goods a least-cost technique; but they also seek to *innovate*, by finding new lower-cost techniques of production.

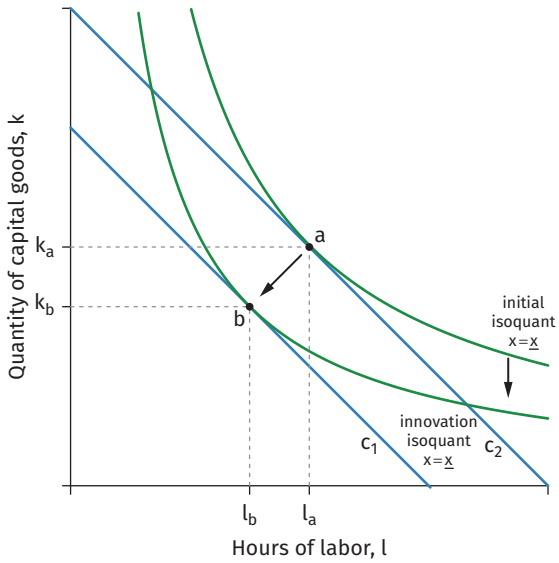
We present isoquants representing an existing and a new technology in Figure 6.22. The initial technology is shown with the cost-minimizing point **a** where the firm employs the combination of labor and capital goods (l_a, k_a).

A new technique of production will be of interest to a firm's owners only if it *lowers* costs of production *given current input prices*, the wage, w, and the rental price of capital goods, p_k. Because the innovation isoquant is closer to the origin the firm is able to produce the *same* amount of output with *fewer* inputs. The cost of producing $\underline{x}$ is now less. Notice that the new technology has made both labor and capital goods more productive; less of both is now being used to produce $\underline{x}$.

Other things equal—importantly the prices of the inputs and its output—if the firm produces the same quantity of x with lesser amounts of inputs per unit of x, it would necessarily increase its profit. The firm would therefore

TECHNICAL PROGRESS A reduction over time in the quantity inputs required to produce some given quantity of output.

Figure 6.22 Isocosts and technical progress. A firm innovates to reduce the cost of producing any given amount of output $\underline{x}$ which initially is given by point **a** with its initial labor and capital goods combination (l_a, k_a) on isocost c_2. With innovation the firm's constraint—the inputs required to produce that given amount—is eased, resulting in a new production isoquant with an expanded feasible set of production. As a result, the firm can, at going prices of labor and capital goods, w and p_k, employ a lower quantity of capital goods and less labor to produce output $\underline{x}$ at a lower total cost. Following the $mrts = mrs$ rule, the firm chooses the point at which its new production isoquant is tangent to the lowest possible isocost at point **b**, employing (l_b, k_b).



obtain an *innovation rent*. This is a rent because the firm's next best alternative–its fallback position–would be *not* to innovate. The innovating firm could, for example, lower its prices and capture a larger share of the market. To survive, other firms will have to innovate as well.

The innovating firm will continue to obtain higher profits until its competitors adopt the same or equivalent cost-reducing technical improvements. Once firms producing identical or similar products have matched the innovator's lower costs, if competition among firms is sufficient, the initial innovation rents will disappear.

CHECKPOINT 6.13 Innovation rent

a. Explain what is meant by the term innovation rent.

b. How will competition tend to eliminate the rent associated with a particular innovation?

EXAMPLE Innovation rents play the key role in determining the profitability and survival of firms. Apple, for example, keeps ahead of its competitors by being the first to introduce important innovations like the iPad or the iPhone X (with facial recognition). Business history also provides dramatic examples, such as IBM in the 1980s where a firm that had managed to maintain innovation rents for many product cycles loses its position by misjudging the next turn of the technological revolution.[7]

M-CHECK It is not always possible to say that one technology is more labor-intensive (or capital-intensive) than the other. To see why, notice that in Figure 6.23, you cannot say if the Cobb–Douglas technology in panel (b) is more labor- or capital-intensive than the Leontief technology shown in panel (a).

6.14 APPLICATION: CHARACTERIZING TECHNOLOGIES AND TECHNICAL CHANGE

Five characteristics of new technologies are important in determining the impact of this process of innovation on the economy. You are already familiar with three: the extent of *economies (or diseconomies) of scale*, the *overall productivity* of the technology, and its *labor (or capital-goods) intensity*. Two further characteristics also matter:

- *Complements and substitutes*: In some technologies a new input replaces some existing input. For example, computer-driven welding robots replace the work of manual (human) welders, reducing the demand for these workers. The same technology increases the demand for the engineers who design and program the robots. The robots and the manual welders are called substitutes—one replacing the other. The robots and the engineers are complements—the robots making the engineers a more important part of the production process.
- *Input substitutability*: Technologies differ in the degree to which inputs can be combined in different ways to produce an output. In some, inputs must be used in a *fixed* proportion: a truck needs a driver, adding a second driver to the one truck or a second truck to the one driver does not add much to the transportation services delivered. In this case we say that the ability to substitute capital goods (trucks) for labor (drivers) is limited. But other technologies allow much greater scope for substituting one input for another: it takes less of a researcher's time to do computations with a more powerful computer than with a small laptop, so using the larger computer is substituting capital (the larger computer) goods for labor (the researcher's time).

Figure 6.23 Substitutability between labor and capital goods with different unit isoquants.

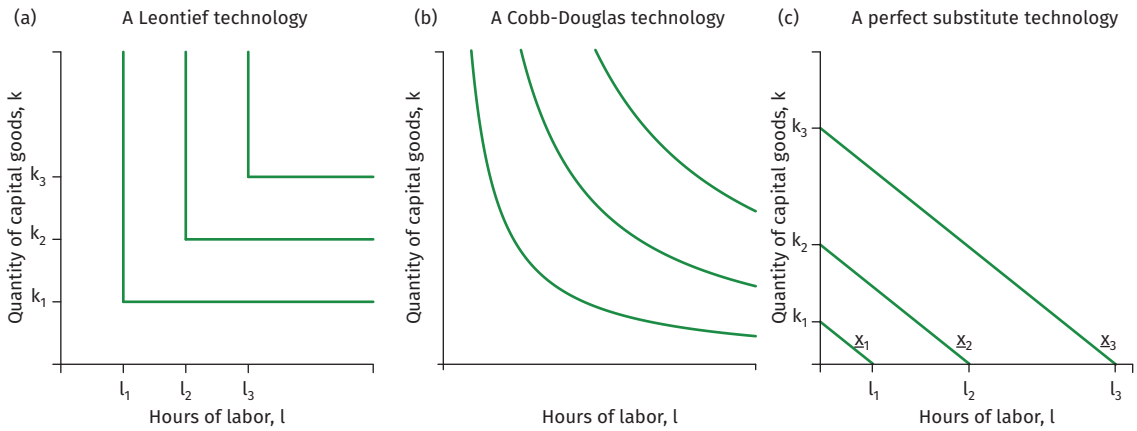


Table 6.4 The ways in which technologies differ and why it matters.

Characteristic	Example: Cobb–Douglas $x = ql^{\alpha}k^{\beta}$	Example from Industrial Revolution	Examples from today and future
Economies of scale	$\alpha + \beta > 1$ economies of scale; $\alpha + \beta < 1$ diseconomies of scale	Industrial Revolution increased economies of scale leading to larger firms and in many cases limited competition	New technologies (e.g. 3D printers) may reduce economies of scale but large first-copy costs ("prototyping") imply economies of scale (e.g. R&D for producing a drug)
Overall productivity	q	Increases in productivity allowed for improved living standards including less work	Is a long-term slowdown in productivity growth in our future?
Labor intensity	α	α fell and the ratio of capital goods to labor input rose	Labor with engineering and networking skills may be replacing both capital goods and other labor
Substitutes or complements	Inputs are complements	Capital goods were substitutes for some kinds of labor (manual, routine) and complements for others (engineering, design)	Artificial intelligence (AI) may become a substitute for even highly trained engineering and other labor
Possible substitution among inputs	Intermediate (possible but limited)	For many nineteenth- and early twentieth-century machining and production line processes, substitution was very limited	If opportunities for substitution are limited, then the continuing increase in the quantity of capital goods per worker could allow wages to rise relative to profits

Figure 6.23 illustrates three cases ranging from *no* opportunities for substitution at all in the Leontief production function in panel (a), to the case called perfect substitutes in panel (c).

Examples of how these dimensions of technologies altered ways of life in the past, and continue to do so today are given in Table 6.4. Key questions about technology today include:

- Are capital goods in the form of robots *substitutes* for workers doing routine tasks and at the same time *complements* of engineers who operate them? If so, as automation of production progresses the demand for the two kinds of workers will diverge, raising the income of engineers, and putting routine task workers out of work.
- As the amount of capital goods available for production continues to rise, will the supply of capital goods outstrip the demand? The answer depends in part on how readily machines and other capital goods can be substituted for workers. If the opportunities for substitution among inputs is similar to a Leontief production function–very limited substitution–then this "glut of capital goods" scenario would be likely. Were this to occur, the owners of the capital goods–wealthy people–would see their incomes *fall* relative to wages and salaries of workers. If, however, substitution opportunities are more like the "perfect substitutes" case in Figure 6.23 (c) then the capital goods owners are

EXAMPLE In 1983, Leontief wrote, "Any worker who now performs his task by following specific instructions can, in principle, be replaced by a machine. This means that the role of humans as the most important factor of production is bound to diminish—in the same way that the role of horses in agricultural production was first diminished and then eliminated by the introduction of tractors."[8] The disappearance of work that he feared did not materialize over the half century that followed. Leontief was referring to automation in manufacturing, but the same could be said now of artificial intelligence, which Anna Salomons and David Autor say could herald a "robopocalypse."[9] Economists Daron Acemoglu and Pascual Restrepo estimated in 2020 that the introduction of a robot replaces between three and six workers.[10]

likely to be able to sustain high incomes even as the ratio of capital goods to labor *increases*.

- As communication technology advances, the labor market will become globalized, not by intercountry migration but by remote working diminishing the importance of face-to-face in-place production. The effect will be to vastly increase the labor supply available to employers in the high-income countries.
- In high-income countries, sectors of the economy with labor-intensive production functions and where remote working is limited–such as education, security services, entertainment, child and elder care, and health services–are *increasing* their share of the economy. Declining fractions of jobs are now in capital-goods-intensive sectors such as manufacturing and agriculture. In the US for example, only one worker in seven is now employed in manufacturing or farming. A result of these sectoral shifts is an *increase* in the demand for labor. Will this result in *greater* scarcity of labor relative to capital goods and an *increase* in workers' bargaining power, resulting in *higher* wages and incomes?
- Will devices and algorithms associated with artificial intelligence eventually become a *substitute* for the work of engineers, lawyers, and other professionals, driving *down* the demand for their labor and *diminishing* their bargaining power, resulting in *lower* wages and incomes?

M-NOTE 6.8 **Substitutes, complements, and the elasticity of substitution**

These terms sound similar (because two include the stem *substitute*) but they refer to different aspects of a technology. The distinction between complements and substitutes is about whether or not there are *positive* or *negative* feedbacks between two inputs, whether they are synergistic or *not*. The elasticity of substitution captures whether changing the ratio of two inputs has a large or small effect on the ratio of the two marginal products. For a similar distinction, applied to consumption rather than production, see M-Note 7.8.

Complements and substitutes. If two inputs are complements then increasing one of them *raises* the marginal product of the other; they are substitutes if the reverse is true. So if we have a production function $x = x(l, k)$ then:

$$\text{Complements:} \quad \frac{\partial^2 x(l,k)}{\partial l \partial k} = x_{lk} > 0$$

$$\text{Substitutes:} \quad \frac{\partial^2 x(l,k)}{\partial l \partial k} = x_{lk} < 0$$

Example: In the Cobb-Douglas production function capital goods and labor are complements because the partial cross derivative of the production function with respect to capital goods and labor is positive.

From M-Note 6.7, the marginal product of labor in the Cobb-Douglas production function $x(l, k) = ql^{\alpha}k^{\beta}$ is:

$$\frac{\partial x}{\partial l} = x_l = \alpha q l^{(\alpha-1)} k^{\beta} \tag{6.22}$$

continued

And the effect of an increase in capital goods on the marginal product of labor, x_l, is:

$$\frac{\partial^2 x}{\partial l \partial k} = x_{lk} = \alpha\beta l^{(\alpha-1)} k^{(\beta-1)} > 0 \tag{6.23}$$

This is also the effect of an increase in labor on the marginal product of capital goods. So labor and capital goods are complements because this cross-partial derivative (x_{lk}) is greater than zero.

Elasticity of substitution. The extent to which one input can be substituted for another in production is called the *elasticity of substitution* defined as the percentage change in the input proportions chosen by a cost-minimizing producer that would result from a percentage change in the ratio of the wage rate to the price of capital goods, or:

$$\text{Elasticity of substitution} = \frac{\%\text{ change in } (k/l)}{\%\text{ change in } (w/p_k)} \tag{6.24}$$

The elasticity of substitution, η, is defined as follows:

$$\eta_{lk} = \frac{\%\Delta(k/p)}{\%\Delta w/p_k} = \frac{d\ln(k/l)}{d\ln(w/p_k)} \tag{6.25}$$

To find η_{lk}, we use the fact that the use of labor and capital goods by the firm has been determined by the $mrs(l,k) = mrts(l,k)$ rule, so we have:

$$mrs = \frac{w}{p_k} = \frac{\alpha}{\beta} \cdot \frac{k}{l} = mrts$$

Then,

$$\text{Rearranging} \qquad \frac{k}{l} = \frac{\beta}{\alpha} \cdot \frac{w}{p_k}$$

$$\text{Take the natural log of both sides} \qquad \ln\left(\frac{k}{l}\right) = \ln\left(\frac{\beta}{\alpha}\right) + \ln\left(\frac{w}{p_k}\right) \tag{6.26}$$

We take the natural log of both sides as that will allow us to find the ratio of percentage changes. We differentiate Equation 6.26 to find the percentage changes as stipulated in Equation 6.24:

$$\frac{d\ln(k/l)}{d\ln(w/p_k)} = \eta_{lk} = 1$$

The result is that the elasticity of substitution with a Cobb-Douglas production function is equal to 1.

CHECKPOINT 6.14 **Technologies** Think of your learning economics as a technology with two inputs: labor (your study time and attention) and capital goods (your studying space, internet, and computer capacities).

a. Are there economies of scale?

b. Can you substitute capital goods for labor?

c. Has learning become more capital-good-intensive or more labor-intensive as digital knowledge has replaced books?

6.15 APPLICATION: WHAT DOES THE MODEL OF INNOVATION MISS?

Our model of innovation captures essential parts of the process by which technical change revolutionizes an economy. But as the following example shows, it misses important aspects too.

A cluster of small firms in Sialkot, Pakistan produce about 40 percent of the world's soccer balls—30 million soccer balls per year—including the match balls for the 2014 World Cup. The industry is highly competitive not only among the hundred or so firms in Sialkot, but also on a world scale, with Chinese firms recently challenging the Pakistani dominance in the field. Firm owners are constantly interested in finding ways to cut costs. As the artificial leather that the balls are made from constitutes almost half the cost of a soccer ball, firm owners are particularly on the lookout for waste-saving methods of cutting the pentagons and hexagons that make up the balls.

Figure 6.24 A new technology. Using redesigned cutting dies, the white hexagons and black pentagons making up the ball could be cut with less leather wasted.

An Italian architect and her husband, an American economist, discovered a way to cut the pentagons and hexagons that would allow a considerable saving. (Unwittingly they had "discovered" what is called a "packing" principle already known by mathematicians.) They found a tool and die-maker in Sialkot to make some test dies (a cutting tool) using the new technique, expecting that it would quickly be adopted by the cost-conscious firms.

In May 2012 they gave 35 firms the new technology. They calculated that the new technology would increase profits of the companies adopting it by 10 percent. Fifteen months later only five of the firms had made any substantial use of the new cutting dies.[11] While the new design was easily copied, only one of the firms not given the new technology had copied it.

The reason, it seems, is that the employees who would have used the new dies (cutters and printers) were paid *piece rates*, that is, the employees were paid per panel they cut. The payment method mattered because the new technology did not speed up the process of cutting, which would have increased the pay of the cutters. Instead the cost reduction came from saving leather, which would enhance the profits of the owners, but would *not* have benefited the workers.

Because the cutters and printers did not stand to do better by saving leather, they had no interest in adopting the new technology. This was especially the case given the initial learning period in which the number of panels cut would actually be *lower* than before, meaning the workers would, for a short period, make less money. So they complained to their employers that the new dies did not work very well. Owners, lacking any independent way of verifying the competing claims of the Italo-American couple and their own cutters showed little interest in the new technology.

Except one firm. One of the larger firms had a different pay system—the cutters were paid a fixed monthly salary rather than per panel that they

cut. This firm purchased (and used) 32 of the new dies, apparently without resistance from the cutters. As long as none of the other firms adopted the new technology, this firm would then have been making substantial innovation rents due to the reduced cost of materials.

If the competitive process worked in Sialkot the way economists think that it should, then this firm should have expanded its share of soccer ball production, eventually forcing other firms to either adopt the new technology, or to drop out. We do not know if that is what happened.

This case makes it clear that firms are made up of people, and the sometimes incomplete information and conflicting interests among them constitute barriers to improvements that, in principle at least, would allow for mutual gains to be shared among workers and firm owners.

Figure 6.25 A worker at the firm that adopted the new technology.

Photo by David Atkin.

CHECKPOINT 6.15 Impediments to innovation

a. Why was the new technology not taken up by most firms?

b. Why was the one firm that did take up the new technology different?

6.16 CONCLUSION

Economics (as you read at the beginning) is the study of how people interact with each other and with our natural surroundings in producing and acquiring our livelihoods. The technologies studied in this chapter—summarized mathematically by production functions—describe how we can produce our livelihood by transforming nature—crops mineral resources and energy—in order to provide the goods and services that make up our standard of living.

The available technologies and the ways owners of firms seek to maximize their profits by choosing techniques of production that minimize their costs of production have important effects on how we interact with each other in this process (the other part of the definition of economics). If technologies are highly productive, and if exchanges on markets allow us to take advantage of economies of scale and learning-by-doing by specializing, then people will have the opportunity for high living standards—and adequate level of goods and services, and ample free time.

But economies of scale in production and learning-by-doing also make it likely that the economy will be dominated by a limited number of large firms whose owners will receive a large share of the income made possible by the high levels of production.

Technologies and the division of labor that results when people specialize according to comparative advantage is just one part of economic knowledge. Equally important are the wants and needs of people and how these are expressed in our willingness to pay for goods when they are supplied in markets. We turn next to market demand.

MAKING CONNECTIONS

Economies of scale and learning-by-doing: are among the main reasons for the division of labor and specialization, specialization makes important contributions to human well-being; but we will see in Chapter 8 that economies of scale and learning-by-doing may also limit the degree of competition in markets.

Markets as a means of coordination: The opportunity to exchange goods expands the set of feasible outcomes available to people and nations by facilitating specialization and the division of labor.

External effects, coordination failures, and poverty traps: The positive external effects associated with economies of agglomeration result in many possible Nash equilibrium patterns of specialization; countries may specialize in goods that keep them poorer than had they specialized in some other way.

Constrained optimization: the choice of technology using the $mrts = mrs$ rule: Minimizing cost subject to a constraint on the level of output produced represented by an isoquant and maximizing utility subject to a budget constraint have many features in common. They are both examples of maximization (or minimization) under constraints.

Innovation rents: A firm that succeeds in introducing a new technology that lowers costs of production at existing input prices can make substantial economic profits, called innovation rents, until others adopt the same or similar innovations.

IMPORTANT IDEAS

specialization	production function	marginal product
technique of production	average product	relative price
division of labor	marginal rate of transformation	marginal rate of
economies of agglomeration	mrt (production)	technical substitution
cost minimization	constant returns to scale	diseconomies of scale
economies of scale	isocost line	rental price of capital goods
wages	absolute advantage	technical efficiency
diminishing marginal productivity	mrt (exchange)	complements/substitutes
poverty trap	comparative advantage	(in production)
$mrts = mrs$ rule	marginal rate of substitution	innovation rents
elasticity of substitution	diversification	technical progress
production possibilities frontier	production isoquant	

MATHEMATICAL NOTATION

Notation	Interpretation
x, y	goods produced
$\bar{x}, \bar{y}$	maximum feasible production of x, y, given available technology and inputs
p	price
l	labor hours
k	quantity of capital goods
$f(\ )$	production function
a_l	minimum amount of labor hours to produce one unit of output
a_k	minimum amount of capital goods to produce one unit of output
α	exponent labor in Cobb–Douglas production
β	exponent of capital in Cobb–Douglas production
q	parameter of productivity, Leontief and Cobb–Douglas production
$\underline{x}$	a given amount of good x that can be produced by the inputs shown on an isoquant
w	wage
p_k	cost of renting capital goods
c	cost, or cost function

Note on superscripts: l: related to labor; k: related to capital goods.

CHAPTER

7

DEMAND

WILLINGNESS TO PAY AND PRICES

This division of labour . . . is the necessary . . . consequence of . . . the propensity to truck, barter, and exchange one thing for another. It is common to all men, and to be found in no other race of animals . . . Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog.

Adam Smith,
The Wealth of Nations (1776)[1]

DOING ECONOMICS

This chapter will enable you to:

- Apply constrained utility maximization to the problem of demand, relating a person's purchases of goods to their willingness to pay.
- Derive a person's demand curve from a utility function describing the person's preferences.
- Understand that consumption is often a social activity, so our preferences, for example, about how many hours we work, will depend on what others are consuming.
- Explain how people change their purchases when prices or income change.
- Understand how these responses reflect both income and substitution effects and use these concepts to explain the effects of a proposed carbon tax and citizen dividend.
- Use the concept of consumer surplus and understand the conditions under which it makes sense to sum the consumer surplus of many people.
- Explain how market demand curves can be derived from individual demand curves.
- Use the price elasticity of demand to explain the effects of price increases, for example, resulting from policies such as placing a tax on sugary drinks.

7.1 INTRODUCTION: MARKETS, UP CLOSE

Ancona is a town on the Adriatic coast of Italy southeast of Venice. It hosts one of the many daily fish markets that sell to European restaurants and fish dealers. Because fish (notoriously) spoil rapidly even with refrigeration, the price of fish on any one day depends largely on the amount of fish brought to the market that day (since none can be carried over from previous days). Economists view fish markets as a kind of ideal experiment for studying how *supply* and *demand* determine the prices at which goods are bought and sold.

Figure 7.1 Mosquito nets save lives; how widely used they are depends on the price. See Figure 7.3.

Photo by Ida Pap/Alamy Stock Photo.

Figure 7.2 shows the average daily prices and the average daily quantities of fish sold in the Ancona market:

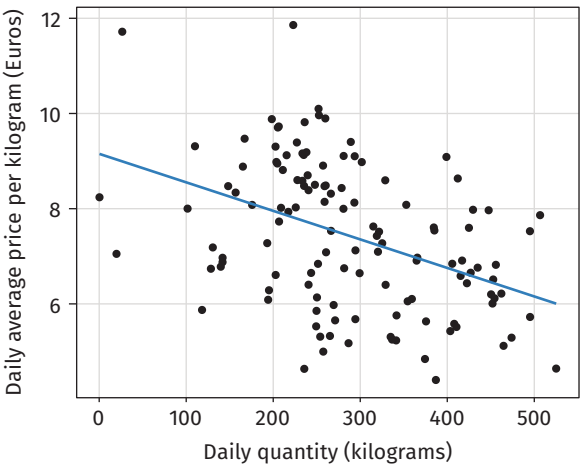
- if the price per kilogram of fish is *high*, the quantity of fish bought and sold is *less*; and
- if the price per kilogram of fish is *low*, *more* kilograms of fish are transacted.

One explanation for the downward-sloping line in the figure summarizing this relationship is that typical buyers in the Ancona fish market will buy more fish if the price is lower. Another is that the greater quantity of fish brought to market on any given day, the lower will be the average price per kilogram of fish.

To understand how the price and quantity of fish purchased is determined, an essential concept is demand, as measured by the amount a person is willing to purchase at any given price. For fish and other goods, knowing how the amount purchased depends on the price is also an important piece of information in the design of economic policies.

Figure 7.2 Prices and quantities of fish bought and sold in the Ancona market.

The plot of daily average prices of fish in the Ancona market against the same day's quantity of fish sold can be summarized by a downward-sloping curve.



✓ **FACT CHECK** Adam Smith's claim in the head quote that humans are unique among all animals in our division of labor and exchange of goods is probably right about dogs. But Smith is certainly wrong about the many other species such as ants and other social insects that practice a very advanced division of labor and specialization. Different species of fish exchange services in what are termed "biological markets."

Here is an example. There are now many low-cost and life-saving preventative health products such as insecticide-treated mosquito nets, tablets to eradicate parasitic stomach worms, and water purification products. In many countries in Africa, Asia, and Latin America, these products prevent illness and death of their users, and also limit the spread of infectious diseases to others, but they are used sparingly if at all. Some policymakers think these products should be provided free of charge to low-income families to encourage their use.

Other policymakers disagree, suggesting that there should be a cost to acquiring these products to discourage wasteful use through better targeting of who gets the products. The question then arises: How will the take-up of the products depend on the price? Will charging even a small price significantly discourage use?

Economists have conducted experiments in eight countries to find answers to these questions. In the experiments, potential users are randomly selected to be offered the goods free or at one or more different prices. The average use of the products at each price (including zero) is then recorded. Some of their results are shown in Figure 7.3.

Figure 7.3 shows that the effect of charging higher prices was to reduce the amount of the product used, in some cases by a substantial amount.

- In Zambia, for example, increasing the price of a disinfectant tablet from the equivalent of 9 to 25 US cents reduced the fraction of the population using the product from 76 to 43 percent.
- Only 43 percent of a group of pregnant Kenyan women purchased insecticide-treated mosquito nets when the price was 60 cents; virtually all used the nets when they were provided without charge.
- A program in Kenya that had initially given away deworming tablets for children, but later introduced a charge of 30 cents per child found that usage of the tablets fell from 75 percent of the affected population to just 18 percent.

! REMINDER Economists have researched preferences based on observing people's behavior in real situations and in experiments. We reviewed some of the key findings from this work in Chapter 2 and we will review more experimental data in later chapters.

On the basis of this information, the Poverty Action Lab at MIT, led by economics Nobel Laureates Abhijit Banerjee and Esther Duflo, suggested that there are good reasons to make these products available either without charge or highly subsidized to ensure very low prices.

We begin our analysis of how people spend their money—whether on fish or mosquito nets—with a basic fact: there are *limits* to how much a family or person can spend.

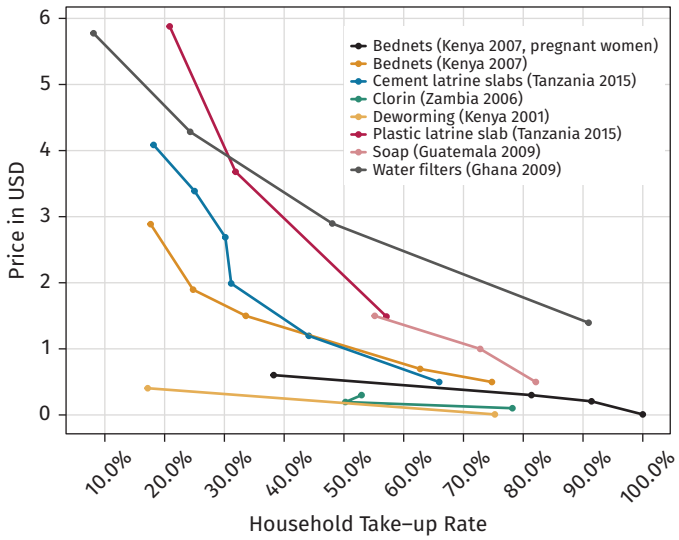
! REMINDER We saw budget constraints in Chapter 3: the slope was a price ratio that corresponded to the opportunity costs of spending money on one good in terms of how much less of the other good you can afford depending on your budget (the negative of the slope was the *mrt*).

7.2 THE BUDGET SET, INDIFFERENCE CURVES, AND THE RULES OF THE GAME

To understand how prices influence the take-up of one of the life-saving health products in Figure 7.3 or the amount of some good that we will

Figure 7.3 The demand for preventative health products: Take-up rates at various prices and when available for free. Our measure of demand is the take-up rate, that is, the fraction of the population that acquires the product (whether free or for a price). For most products the quantity demanded is substantially less when even a small price is charged compared with when the good is available for free; and higher prices are also associated with substantially lower quantity demanded than lower prices. To put the prices in perspective, at the time of the studies shown in the figure well over a third of the population of Kenya lived on less than $2 a day (adjusted for the purchasing power of the local currency).

Source: Dupas and Miguel (2017).



consume, think about someone who has a total amount of money to spend, m, that she has in cash, savings, or available credit.

The budget constraint

We shall consider a person, Harriet, and the decisions she needs to make. A person's budget set states what bundles (x, y) are *feasible* for her to consume given her budget and market prices of the goods:

$$m \geq p_x x + p_y y \tag{7.1}$$

$$m \geq \text{Prices of Goods} \times \text{Quantities of Goods Purchased}$$

Expressing this inequality as an equality—assuming that Harriet would not consume less than her budget allowed—we have the budget constraint:

$$m = p_x x + p_y y \tag{7.2}$$

$$\text{Budget} = \text{Prices of goods} \times \text{Quantities of goods purchased} \tag{7.3}$$

! REMINDER The *budget set* is the set of all feasible purchases of the bundles of x and y with current budget m, such that $m \geq p_x x + p_y y$. The budget constraint is the border of the budget set showing all combinations that *exhaust* the budget, i.e. for which the constraint holds with equality, $m = p_x x + p_y y$.

Equation 7.3 is a statement about prices and budgets. But it is also a statement about the rules of the game and preferences. They are as follows:

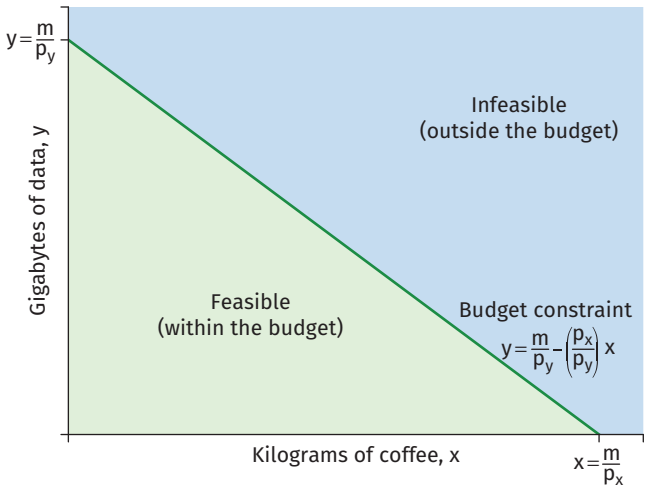
- *No gifts, thefts, or consumption as a matter of citizen rights:* You can consume only what you pay for; so no gifts or goods provided by government, or acquired by theft.
- *No altruism or concerns about environmental sustainability*: You consume the most that your budget allows and you consume it yourself, rather than giving it to or sharing it with others or consuming less than you could to reduce your carbon footprint.

We can rearrange the budget constraint, Equation 7.2, to obtain a line we can draw on the x and y-axes we use for indifference curves:

$$y = \frac{m}{p_y} - \frac{p_x}{p_y}x \tag{7.4}$$

We plot the budget constraint in Figure 7.4. Examining the two terms on the right-hand side of Equation 7.4, we can see that if Harriet were to consume only good y and no good x, then she would consume $\frac{m}{p_y}$ units of good y which is the intercept of the budget constraint with the y-axis. As Harriet buys

Figure 7.4 Budget constraint for coffee and data. The budget set is shaded in green and the budget constraint is the dark-green line on the border of the budget set. Consumption bundles (x, y) in the budget set and on the budget constraint can feasibly be obtained with the current budget (m) at going market prices for x and y, p_x and p_y. Outside the budget constraint, in the shaded blue area, the bundles of x and y cannot feasibly be obtained at going market prices with the existing budget. The budget set is another example of a feasible set.



more of good x, she moves along the budget constraint with the slope $-\frac{p_x}{p_y}$ indicating the rate at which she sacrifices good y for good x given the prices. If she were to buy only good x, she could afford $x = \frac{m}{p_x}$ units of good x.

The negative of the slope of the budget constraint is a **relative price** and measures the *opportunity cost* of obtaining good x in terms of the amount of good y that Harriet must sacrifice because her funds are limited. The (negative of the) slope of the budget constraint is another marginal rate of transformation; it tells us the terms on which a reduced amount of good y can be "transformed into" additional amounts of good x while just satisfying the budget constraint. So we have:

$$\text{(Negative of) the slope of the budget constraint} = \frac{p_x}{p_y} = mrt \tag{7.5}$$

M-CHECK We can take the first derivative of Equation 7.4, the budget constraint, to see that the slope of the budget constraint is:

$$\frac{dy}{dx} = -\frac{p_x}{p_y}$$

The $mrt(x,y)$ is the negative of the slope of the budget constraint or $-\frac{dy}{dx}$, that is, the opportunity cost of x in terms of y.

M-NOTE 7.1 Budget for coffee and data

For particular values of m, p_x, and p_y we can graph the budget constraint. Consider the following example:

- Harriet has a budget ($m = \$50$) to spend on kilograms of coffee, x, and gigabytes of data, y.
- The price of a kilogram of coffee, p_x, is \$10.
- The price of a gigabyte of data, p_y, is \$5.

Putting these pieces of data together, therefore, the budget constraint is $50 = 10x + 5y$. We can rearrange the budget constraint as we did in Equation 7.4:

$$y = \frac{50}{5} - \frac{10}{5}x$$
$$y = 10 - 2x \tag{7.6}$$

Equation 7.6 is a line with an intercept at $\frac{m}{p_y} = 10$ on the y-axis, an intercept of 5 on the x-axis and a slope of $p = -2$. Such a curve would look like bc_1 in Figure 7.5.

CHECKPOINT 7.1 Sketching a budget constraint

a. There are two goods: vegetables (x), which have a price of 4 euros per kilogram, and meat (y), which has a price of 10 euros per kilogram. You have a budget of 50 euros a week for meat and vegetables for your family. Sketch your budget constraint.

b. The price of vegetables increases to 5 euros per kilogram. What happens to your budget constraint? Sketch and explain.

RELATIVE PRICE A relative price is a ratio of one price to another.

Budget constraints, indifference curves, and the amount demanded

! REMINDER We reason here in the same way we did in Chapter 3 that the utility-maximizing choice is the point of tangency between the highest attainable indifference curve and the feasible frontier or budget constraint, thereby following the $mrs = mrt$ rule.

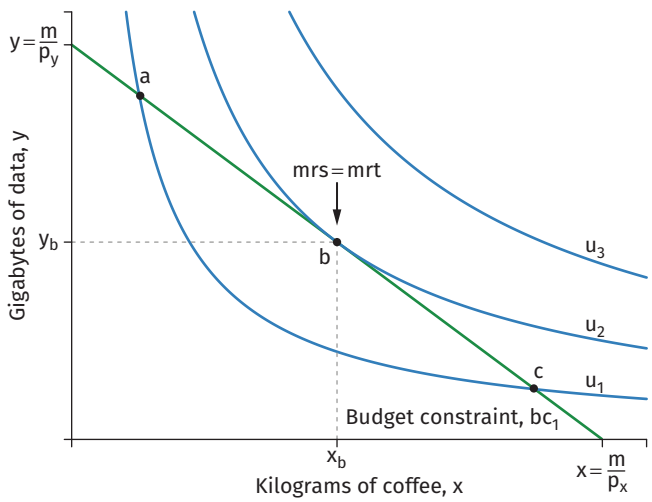
In Figure 7.5 we show three of Harriet's indifference curves. Remember at any given point on the indifference curve, the negative of its slope tells how much Harriet values the good on the x-axis compared to her valuation of the good on the y-axis, that is, her marginal rate of substitution ($mrs(x,y)$). Her marginal rate of substitution is her *willingness to pay* to get *more* of good x, namely how much of good y she would be willing to part with, in order to get one more unit of good x. So,

$$\text{(Negative of) the slope of an indifference curve} = \frac{\text{marginal utility of x}}{\text{marginal utility of y}}$$

$$\text{Willingness to pay data to get coffee} = mrs = \frac{u_x}{u_y} \tag{7.7}$$

Harriet wants to get to the highest indifference curve that she can, given her budget. This is the point at which the budget constraint is *tangent* to her highest attainable indifference curve. For the two curves to be tangent,

Figure 7.5 Utility-maximizing consumption bundle. Harriet maximizes her utility subject to her budget constraint bc_1. At point **a** the amount of data she is willing to pay for more coffee (the $mrs(x,y)$) is greater than the opportunity cost of getting more coffee (the $mrt(x,y)$). Or, what is the same thing: the indifference curve is steeper than the budget constraint, or $mrs > mrt$. Conversely, at point **c**, the amount of data she is willing to pay for more coffee (the $mrs(x,y)$) is less than the opportunity cost of getting more coffee (the $mrt(x,y)$). Or, what is the same thing: the indifference curve is flatter than the budget constraint, or $mrs < mrt$. She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x,y) = u_x/u_y$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x,y) = p_x/p_y$, following the $mrs = mrt$ rule introduced in Chapter 3.



the slope of her indifference curve must equal the slope of the budget constraint. Or, the marginal rate of substitution must equal the marginal rate of transformation.

$$mrs(x,y) = \frac{u_x}{u_y} = \frac{p_x}{p_y} = mrt(x,y) \tag{7.8}$$

This is the same rule that you learned in Chapter 3 because it determines how much of a good people will demand or want to buy at given prices. When Harriet uses the $mrs = mrt$, her trade-offs of one good for another in terms of utility ($mrs(x,y)$) equal the opportunity costs of the two goods in terms of each other ($mrt(x,y)$), where the opportunity costs are given by their prices. Remember that money she spends on one good means money she *cannot* spend on another good: capturing the essential idea of an opportunity cost.

The point where Equation 7.8 holds is Harriet's *utility-maximizing consumption bundle* or the quantity of goods x and y that Harriet will buy at prices p_x and p_y when her budget is m. We can now use her budget constraint (Equation 7.2) and her $mrs = mrt$ rule (Equation 7.8) to study how Harriet will react to changes in prices or her budget.

M-NOTE 7.2 **Using Lagrangian optimization to derive the *mrs = mrt* rule**

A common method that is used to find the optimal (maximum or minimum) outcome with an objective that has a constraint is to use a Lagrange multiplier (see the Mathematics Appendix for a fuller exposition).

$$\mathcal{L} = \text{Objective} + \lambda\,[\text{Constraint}] \tag{7.9}$$

In this case, the objective is to maximize utility, $u(x,y)$. The constraint is the budget constraint, $m = p_x x + p_y y$, which, expressed as a condition that must equal zero is $m - p_x x - p_y y = 0$, which is the constraint. Therefore, the Lagrangian equation becomes:

$$\mathcal{L}(x,y,\lambda) = u(x,y) + \lambda\left[m - p_x x - p_y y\right] \tag{7.10}$$

To find the utility-maximizing choices of x and y, we would differentiate the Lagrangian (Equation 7.10) with respect to x, y, and λ. For now, we shall simply differentiate it with respect to x and y and see that this process gives us the result that $mrs(x,y) = mrt(x,y)$.

We start by partially differentiating Equation 7.10 with respect to x and y and setting those partial derivatives equal to zero (imposing the first-order conditions for a maximum):

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= u_x - \lambda p_x = 0 \\ \Rightarrow u_x &= \lambda p_x \end{aligned} \tag{7.11}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial y} &= u_y - \lambda p_y = 0 \\ \Rightarrow u_y &= \lambda p_y \end{aligned} \tag{7.12}$$

continued

As both Equation 7.11 and 7.12 are equalities, we can divide the one equation by the other.

$$\frac{\text{Equation 7.11}}{\text{Equation 7.12}} = \frac{u_x}{u_y} = \frac{\lambda p_x}{\lambda p_y}$$

$$\lambda \text{ cancels} \qquad \Rightarrow mrs(x,y) = \frac{u_x}{u_y} = \frac{p_x}{p_y} = mrt(x,y) \tag{7.13}$$

Equation 7.13 is the $mrs = mrt$ rule we have used since Chapter 3. It shows that the constrained utility maximum is a set of purchases such that the marginal rate of substitution is equal to the relative prices or the marginal rate of transformation.

7.3 INCOME AND DEMAND: DIFFERENCES IN THE BUDGET

To study how prices and incomes (budgets) affect the demand for goods we ask a hypothetical "what-if" question: how much of good x would someone purchase if her budget were m and the price of good x were p_x and the price of good y were p_y.

A **demand function** shows the quantity purchased of x that results for the various values of the prices of both goods and the budget, p_x, p_y, and m. So $x(m, p_x, p_y)$ is the demand for x as income (m), or the price of x (p_x), or the price of $y(p_y)$ change. We use the term *demand curve* when we refer to the simpler two-dimensional graphical relationship $x(p_x)$ where we see how the amount purchased of the good varies with its price (p_x) holding constant all of the other influences on the demand for x.

We sometimes use a demand curve in which, instead of quantity sold depending on the price $x = x(p)$, *price* depends on the *quantity* sold, $p = f(x)$. This is called an *inverse demand curve* based on the **inverse demand function**, because it is the mathematical inverse of the conventional demand function.

The inverse demand function contains exactly the same information as the demand function and the inverse demand curve looks identical to the conventional demand curve (it is downward-sloping). What differs is the hypothetical question for which the inverse demand function provides an answer. Instead of asking how much of a good will be purchased at a

DEMAND FUNCTION A demand function shows how the amount of a good purchased by an individual varies with the prices of all goods and the individual's budget.

INVERSE DEMAND FUNCTION The inverse demand function (curve) answers the hypothetical question: what is the highest price at which a given amount of some good could be sold?

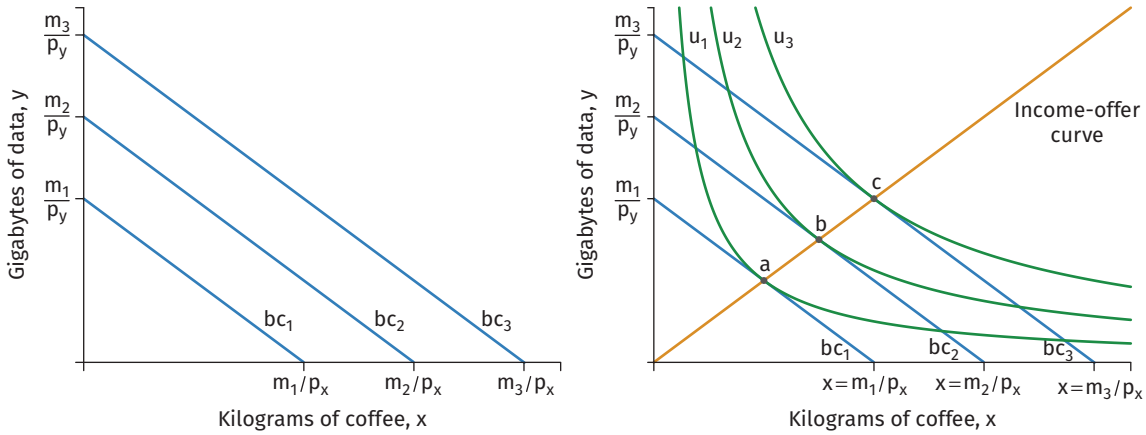
given set of prices and a budget, the inverse demand function answers the question: If the budget and the price of the other good are m and p_y what is the maximum price p_x that the buyer would be willing to pay to purchase an amount x of the good?

A change in income: The income-offer curve

To understand these changes, therefore, we examine Figures 7.6 (a) and 7.6 (b). As Figure 7.6 (a) shows, as Harriet's income changes, her budget constraint shifts. That is, the intercept with both axes, $\frac{m}{p_y}$ and $\frac{m}{p_x}$, shifts *up* as income (m) goes *up* and shifts *down* as her income goes *down*. Consider the three budget constraints in Figure 7.6 (a) where only income changes, but the prices of the two goods do not change.

- *Status quo*: She starts with an income of m_2 with intercepts $\frac{m_2}{p_y}$ and $\frac{m_2}{p_x}$.
- *Income decrease*: If her income *decreases* to m_1, then the intercepts of her budget constraint shift downward and to the left to $\frac{m_1}{p_y}$ and $\frac{m_1}{p_x}$, so she can buy *less* of both goods.

Figure 7.6 Harriet's budget constraint with shifts in income & her income-offer curve. In panel (a) Harriet's budget constraints with three levels of income are shown (m_1, m_2, and m_3) with the corresponding budget constraints bc_1, bc_2, and bc_3 shifting outwards as income increases. In panel (b) Harriet's budget constraints are shown tangent to three indifference curves, u_1, u_2, and u_3. The points where they are tangent are where $mrs(x,y) = mrt(x,y)$. The curve joining all the points at which Harriet maximizes her utility as her income changes illustrate her income-offer curve. To draw this figure we have set $\alpha = 0.5$.



(a) Budget constraint & shifts in Income

(b) The income-offer curve

- *Income increase*: If her income *increases* to m_3, then the intercepts of her budget constraint shift upwards and to the right to $\frac{m_3}{p_y}$ and $\frac{m_3}{p_x}$, so she can buy *more* of both goods.

Considering different levels in Harriet's income we can superimpose Harriet's indifference curves to find the consumption bundle for each income level that would maximize Harriet's utility using the $mrs = mrt$ rule. The path traced out by the points (x,y) as m increases is called her **income-offer curve**. Her income-offer curve is also called her *expansion path* because it shows the effect of *expanding* her feasible set (by increasing her budget). In Figure 7.6 (b), her income-offer curve is upward-sloping, showing the effect of an increase on her income on her consumption of the goods, x and y. As she gets more income, she would consume more of both goods.

The income-offer curve allows us to understand the *groups* of goods people consume.

- *Normal goods*: normal goods are goods like coffee and data where people buy *more* as their income *increases*, or *less* of them as their income *decreases*.
- *Inferior goods*: inferior goods are goods like cheap staples, such as white sandwich bread, basic rice, or instant noodles: people tend to consume *less* inferior goods as their income *increases* and *more* of them as their income *decreases*.

Figure 7.7 shows a situation in which Harriet's income increases, but her consumption responses for the two goods differ. For good y, on the vertical axis, Harriet consumes more of it as her income *increases* from m_1 to m_2: she *increases* her consumption from y_1 to y_2.

For good x, on the horizontal axis, on the contrary, Harriet consumes *less* as her income *increases* from m_1 to m_2: she *decreases* her consumption from x_2 to x_1 as her income increases. As a result, she has a *downward-sloping* income-offer curve.

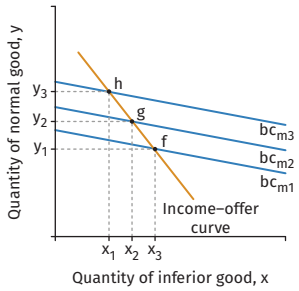


Figure 7.7 Inferior goods and an increase in income. The downward-sloping income-offer curve for an inferior good (x) and a normal good (y). As the person's budget increases, they consume *less* of good x, showing that x is an inferior good (decreasing from x_3 to x_2 to x_1). Consumption of good y increases as income increases, showing that y is a normal good (increasing from y_1 to y_2 to y_3).

CHECKPOINT 7.2 Inferior indifference curves

a. On your own set of axes, redraw Figure 7.7.

b. Add the relevant indifference curves to your figure. What do you think they look like? Explain.

c. What condition must be true at each of points **f**, **g**, and **h**?

INCOME–OFFER CURVE An income-offer curve describes consumption or other choices made by an individual for varying levels of income.

7.4 COBB–DOUGLAS UTILITY AND DEMAND

You already encountered Cobb–Douglas utility in Chapter 3. We build on that base and explore a person's choice of her utility-maximizing consumption bundle using indifference curves based on a Cobb–Douglas utility function and budget constraints. The Cobb–Douglas utility function has the following general form:

$$u(x,y) = x^{\alpha}y^{1-\alpha} \tag{7.14}$$

with α and $1-\alpha$ indicating the relative strength of preference for x and y respectively.

Using Cobb–Douglas utility, we can illustrate indifference curves for each good as the price of the good changes and we can derive a demand curve for each good. In Figure 7.5, we showed the indifference curves for two goods: kilograms of coffee (x) and gigabytes of data (y). In M-Note 3.4 (see also M-Note 7.3) we showed that the marginal rate of substitution is:

$$mrs(x,y) = \frac{u_x}{u_y} = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{y}{x}\right) \tag{7.15}$$

The $mrs = mrt$ rule, then requires that:

$$mrs(x,y) = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{y}{x}\right) = \frac{p_x}{p_y} = mrt(x,y) \tag{7.16}$$

From this relationship and the budget constraint as we show in M-Note 7.4, we can derive a demand curve for each good. The demand function for kilograms of coffee good x , shown in Figure 7.8 is:

$$\text{Demand function} \qquad x(m,p_x) = \frac{\alpha m}{p_x} \tag{7.17}$$

Equation 7.17 shows a relationship between quantity demanded (x) and price (p_x) such that the quantity demanded *decreases* as the price increases, or the quantity demanded *increases* as the price *decreases*.

Rearranging Equation 7.17 as $\frac{x \cdot p_x}{m} = \alpha$ meaning that with a Cobb–Douglas utility function, a person implementing the $mrs = mrt$ rule will spend a *fraction* of their total budget on x, that is:

- equal to the *exponent*, α, of x in the utility function, which is a constant, and therefore is,
- *independent* of the price of x and the price of y.

The fraction of the budget spent on the *other* good is *also* independent of changes in the price of x (it is equal to $1-\alpha$). So, because the budget m has *not* changed, the amount spent on y will also remain the same. Because the price of the other good has *not* changed, the amount of good y purchased is also unchanged.

! REMINDER The marginal rate of substitution indicates, for a given ratio of the quantities of the two goods x and y making up a bundle, how much the person values a small increment in the amount of x compared to how much they value a small increment of y is given by the ratio of the exponents. For example, if $\alpha = 0.75$, and $x = y$, then the marginal utility of x is three times the marginal utility of y (because $\alpha/(1-\alpha) = 3$).

With Cobb–Douglas utility for a given price of y, p_y and income (m), the only thing that differs with different prices of x, is the quantity demanded of good x.

Similarly, the amount spent on x does not depend on the price of y, which you can confirm from the fact that p_y does not appear in the demand function for x. These are not general features of demand functions; they are specific to the Cobb–Douglas utility function.

M-NOTE 7.3 Marginal rate of substitution, Cobb–Douglas utility function

Consider a Cobb–Douglas utility function:

$$u(x,y) = x^{\alpha}y^{1-\alpha}$$

The marginal utility with respect to each good is:

$$u_x = \frac{\partial u}{\partial x} = \alpha x^{\alpha-1}y^{1-\alpha}$$
$$u_y = \frac{\partial u}{\partial y} = (1-\alpha)x^{\alpha}y^{-\alpha}$$

Therefore, the marginal rate of substitution of x with respect to y is:

$$mrs(x,y) = \frac{u_x}{u_y} = \frac{\alpha x^{\alpha-1}y^{1-\alpha}}{(1-\alpha)x^{\alpha}y^{-\alpha}} \tag{7.18}$$

Note that:

$$\frac{x^{\alpha-1}}{x^{\alpha}} = \frac{1}{x}$$
and
$$\frac{y^{1-\alpha}}{y^{-\alpha}} = y$$

The marginal rate of substitution (Equation 7.18) becomes:

$$mrs(x,y) = \frac{u_x}{u_y} = \left(\frac{\alpha}{1-\alpha}\right)\frac{y}{x}$$

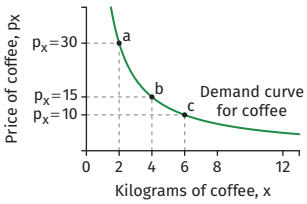
CHECKPOINT 7.3 Budget shares and exponents

Use the data in Figure 7.8 to confirm that at points **a**, **b**, and **c**:

a. the fraction of the budget spent on coffee is equal to the exponent of coffee and does not change as the price of coffee changes; and

b. do the same for the fraction of the budget spent on data.

Figure 7.8 A Cobb–Douglas demand curve for coffee. The demand function is $x(m,p_x) = \frac{\alpha m}{p_x}$. The budget is $m = \$120$ and the coefficient of coffee in the Cobb–Douglas utility function is $\alpha = 0.5$.



The inverse demand function

We can rearrange the function and find the **inverse demand curve**. The inverse demand curve is:

$$\text{Inverse demand} = p_x(x,m) = \frac{\alpha m}{x} = \frac{\text{amount spent on x}}{\text{amount of x purchased}} = \text{price}$$

Here, we have a downward-sloping demand curve where price *decreases* as the quantity demanded *increases*.

CHECKPOINT 7.4 The Cobb-Douglas utility function Harriet buys coffee and cookies to fuel herself while running her business. Her utility function for cookies (x) and cups of coffee (y) is given by the following utility function:

$$u(x,y) = x^{0.6}y^{0.4} \tag{7.19}$$

a. If Harriet already has one cup of coffee and two cookies, and a friend offers her either another two cookies or another cup of coffee, which would she take?

b. We assume that Harriet has a daily budget of \$10 to spend on coffee and cookies, where the price of a cup of coffee is \$3 and the price of a cookie is \$0.50. Suppose Harriet maximizes her utility given her budget constraint (i.e., she implements the $mrs = mrt$ rule). What fraction of her total budget will she spend on coffee?

M-NOTE 7.4 Cobb–Douglas demand functions

The Cobb–Douglas utility function is:

$$u(x,y) = x^{\alpha}y^{1-\alpha}$$

where $0 < \alpha < 1$. The individual maximizes this function subject to a budget constraint:

$$m = p_x x + p_y y \tag{7.20}$$

Remember that the negative of the slope of the indifference curve is the marginal rate of substitution and the negative of the slope of the budget constraint (which is also the ratio of the prices of the two goods) is the marginal rate of transformation. We found the following in M-Note 7.3:

$$mrs(x,y) = \left(\frac{\alpha}{1-\alpha}\right)\frac{y}{x} \tag{7.21}$$

So the utility-maximizing bundle that implements the $mrs = mrt$ rule must satisfy the following equation:

$$mrs(x,y) = \left(\frac{\alpha}{1-\alpha}\right)\frac{y}{x} \quad = \frac{p_x}{p_y}$$

$$\therefore p_y y = p_x x\left(\frac{1-\alpha}{\alpha}\right) \tag{7.22}$$

To find the demand function, substitute Equation 7.22 into the budget constraint, Equation 7.20, to isolate a value for x, which we then use to find y:

$$\begin{aligned} m &= p_x x + p_x x\left(\frac{1-\alpha}{\alpha}\right) \\ m &= \frac{(\alpha+1-\alpha)p_x x}{\alpha} \\ \therefore x(m,p_x,p_y) &= \frac{\alpha m}{p_x} \end{aligned}$$

continued

Substitute $p_x x$ into Equation 7.22 to find $p_y y$ and y:

$$p_y y = (1-\alpha)m \tag{7.23}$$

$$y(m, p_x, p_y) = \frac{(1-\alpha)m}{p_y} \tag{7.24}$$

We have therefore found the demands $\big(x(m, p_x, p_y), y(m, p_x, p_y)\big)$ as functions of the budget, m, and the prices of the goods, p_x and p_y, given the preferences for the goods, α. Notice that the demand for each good is *independent* of the price of the other good. The demand curve for each good in terms of its own price is a hyperbola.

CHECKPOINT 7.5 Demand with Cobb–Douglas utility

a. In Figure 7.6 we set $\alpha = 0.5$. Draw three Cobb–Douglas indifference curves with $\alpha = 0.7$ (just show how they would be different from the ones shown in the figure) and the budget constraint with three income levels.

b. Sketch the corresponding income-offer curve.

7.5 APPLICATION: DOING THE BEST YOU CAN DIVIDING YOUR TIME

We can apply the Cobb–Douglas utility function to a problem we all face: how to divide up the limited number of hours in our day between all of the things we would like to do, or must do to make a living. We simplify the problem by limiting our objectives to only two things: free time and consumption (similar to the problem involving Living and Learning in Chapter 3). Because we pay for our consumption with the wages we receive for working, and working means *not* having free time, we face a trade-off: more free time means less consumption and more consumption means less free time.

A trade-off between free time and consumption

Consider a worker, Scott, deciding how much leisure and consumption he would like. We assume (unrealistically) that his employer pays him an hourly wage and lets him choose how many hours a day he will work.

We define h as the fraction of the day that Scott spends *working* for wages, with $f = 1 - h$ the fraction of that day that is *free time* (f). Scott consumes his entire income, so we can express his daily consumption, x, as the total income he would receive if he worked 24 hours, w, times the fraction of the day that he works.

In Figure 7.9 we show Scott's feasible set of choices concerning consumption (wh) and free time ($1-h$), along with three indifference curves illustrating his preferences for the two goods. The feasible frontier is his budget constraint. The maximum that Scott could spend on consumption

is to have *no* free time (and no sleep) and to set working time at 1 allowing a total expenditure of w on consumption. So w is analogous to m–the maximum possible expenditure–in the previous budget constraints.

This limits his expenditure to the sum of consumption (x) plus his "expenditure" of free time (the fraction of time not working $1-h$), valued at the wage Scott would have received had he worked the entire day (w):

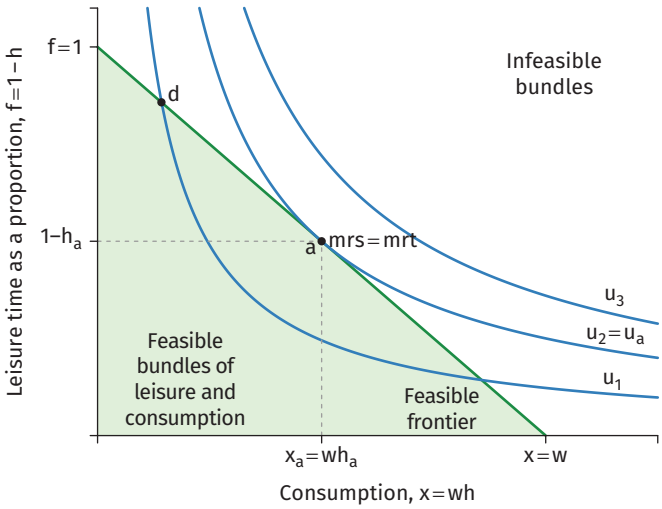
$$\text{Budget constraint:} \qquad x+(1-h)w \leq w \tag{7.25}$$

We let Equation 7.25 hold as an equality (he is not going to throw away money). Then rearranging the equality, we show how the *maximum* feasible amount of free time ($1-h$) depends on the level of consumption (x):

$$\text{Budget constraint} \qquad 1-h=1-\frac{x}{w} \tag{7.26}$$

The negative of the slope of the budget constraint is the marginal rate of transformation of reduced consumption into more free time. This is how much additional free time Scott is able to have by *giving up* one unit of consumption. It is also the opportunity cost of consumption, that is how much free time he must give up in order to have one more unit of consumption. The equation for the budget constraint (Equation 7.26) and Figure 7.9 show that the marginal rate of transformation is as follows:

Figure 7.9 Feasible set and indifference curves for working time and consumption. The constrained utility-maximizing choice, implementing the $mrs = mrt$ rule is at point **a**. The negative of the slope of the budget constraint is the opportunity cost of consumption in terms of free time given up. The negative of the slope of the indifference curves is the willingness to pay (giving up free time) for additional consumption. We used $\alpha = 0.5$ to draw this graph. So, using the equation for the *mrs* (equation 7.15), the slope of the indifference curves at some bundle $(x, 1-h)$ is $\frac{-(1-h)}{x}$.



$$-\frac{\Delta(1-h)}{\Delta x} = \frac{1}{w} = mrt = \text{opp. cost of consumption} \tag{7.27}$$

Scott's preferences are represented by the following Cobb–Douglas utility function that expresses how he values consumption (x) and free time ($f = 1-h$):

$$u(x,h) = x^{\alpha}(1-h)^{1-\alpha} \tag{7.28}$$

where, as before, $0 < \alpha < 1$ and the size of α indicates the relative preferences for the two goods. The negative of the slope of Scott's indifference curves is the marginal rate of substitution between consumption and free time. This is his willingness to pay (WTP) in reduced free time to obtain an additional unit of consumption. From M-Note 7.3 we know that the mrs for the Cobb–Douglas utility function is the ratio of the exponents times the ratio of the quantities, so, letting y be defined as $(1-h)$, we have:

$$-\text{indiff curve slope} = \frac{u_x}{u_y} = \left(\frac{\alpha}{1-\alpha}\right)\frac{\overbrace{1-h}^{y}}{x} = mrs = \text{WTP for consumption}$$

You can see from Figure 7.9 that the bundle of free time and consumption at point **d** could not be the best Scott can do. A point **d** the indifference curve is steep because $\frac{(1-h)}{x}$ is large, Scott has lots of free time and little consumption. So his willingness to pay for consumption is much greater than the opportunity cost of consumption (the flatter slope of the budget constraint). He should work longer hours.

The best Scott can do in this constrained optimization problem, maximizing his utility subject to his budget constraint and implementing the $mrs = mrt$ rule, is to select the bundle $(x, 1-h)$ such that the marginal rate of substitution is equal to the marginal rate of transformation, or

$$mrs = \frac{u_x}{u_y} = \frac{1}{w} = mrt \tag{7.29}$$

In M-Note 7.5 we show that at his utility-maximizing bundle, the fraction of the day Scott will work, h, is equal to α, the exponent of consumption in his utility function. As a result, the fraction of Scott's day that is free time will be $1-\alpha$, the exponent of "free time" in his utility function, a measure of how important free time is to him.

! REMINDER In the previous section you learned that for a person maximizing a Cobb–Douglas utility function with α the exponent of the x-good, the share of the budget spent on the x-good is α itself.

M-NOTE 7.5 Consumption, free time, and work hours

As explained in the previous section, we have:

Utility function	$u(x,h)$	$= x^{\alpha}(1-h)^{1-\alpha}$
Budget constraint from Equation 7.25	$x + (1-h)w$	$\leq w$
rearranged and as an equality	$(1-h)$	$= 1 - \frac{x}{w}$

continued

Below, we use the fact that this budget constraint can be rearranged to:

$$\text{Budget constraint: consumption} \leq \text{income} \qquad x \leq wh$$

$$\text{(using Equation 7.21)} \qquad mrs = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{y}{x}\right)$$

$$\text{using } y = 1-h \qquad mrs = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x}\right)$$

$$\text{(using Equation 7.26)} \qquad -mrt = \frac{d(1-h)}{dx} = -\frac{1}{w}$$

To determine Scott's utility-maximizing time worked, we use the $mrs = mrt$ rule:

$$mrs = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x}\right) = \frac{1}{w} = mrt$$

$$\text{Using } x = wh \qquad \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{wh}\right) = \frac{1}{w}$$

$$w\alpha(1-h) = (1-\alpha)wh$$

$$\alpha - \alpha h = h - \alpha h$$

$$\alpha = h$$

Scott's hours worked as a fraction of his available time is equal to α, showing that α is a measure of how much Scott values consumption relative to free time.

CHECKPOINT 7.6 Preferences for free time and consumption of goods

a. In many religions a high value is placed on leading a simple life with limited consumption, especially among the spiritual leaders of the religion. Write down a Cobb–Douglas utility function for a person with these values and contrast an indifference curve based on this utility function with the one shown in Figure 7.11 (based on $\alpha = 0.5$).

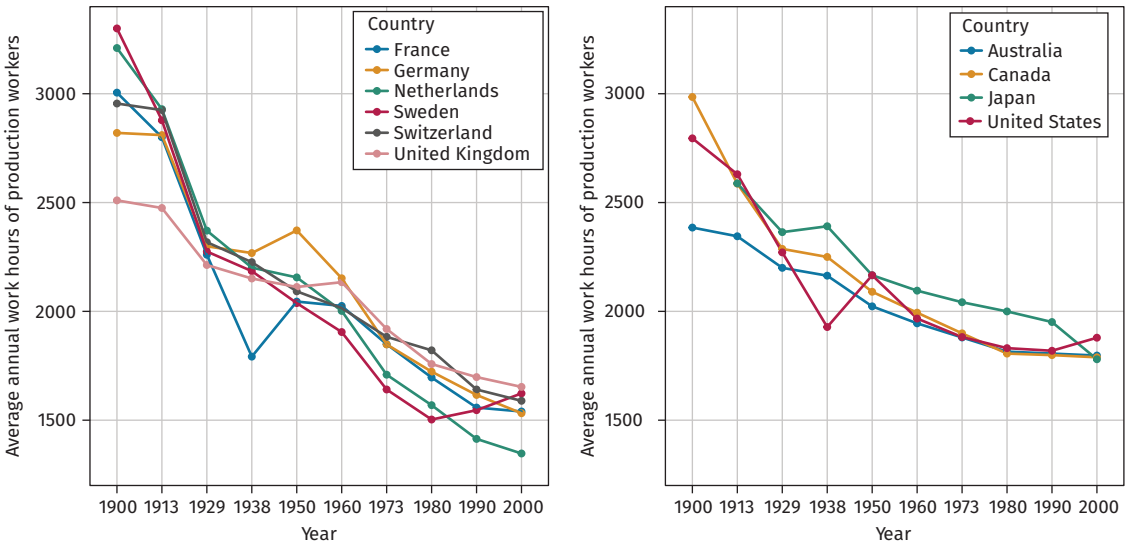
b. What would the indifference curves of a "workaholic" look like? What value of α might they be based on?

7.6 APPLICATION: SOCIAL COMPARISONS, WORK HOURS, AND CONSUMPTION AS A SOCIAL ACTIVITY

In Chapters 3 and 6 we looked at data on how men and women spend their time, and the increase in the amount of time women spend working for pay (the female labor force participation rate). Here we use the constrained optimization model to help understand another dramatic change in time use over the last century.

Figure 7.10 Work hours over time for a variety of countries. The data refer to annual average work hours for full-time production workers (meaning, excluding supervisory personnel)

Source: Huberman (2004).



(a) Work hours over time in European Countries

(b) Work hours over time in countries outside of Europe

Figure 7.10 shows that, in every country on which we have data, people have been working less on average. But there are important differences among the countries:

- In the Netherlands work hours fell from the equivalent of 62 hours, 52 weeks of the year to fewer than 27 hours per week.
- In Sweden, where work hours also declined dramatically, there was a small increase in work hours from 1980 to 2000.
- Work hours declined much less in the US than in most other countries–a decline of 33 percent compared to a decline of 58 percent in the Netherlands.
- In the US, as in Sweden, there was a slight increase in work hours at the end of the last century.

How can our model help explain these differences? We modify the model of choice of work hours to help us understand the differences among the countries and the changes over time (shown in Figure 7.10). The new idea that we will use to modify the model is that what people consume–the quality, quantity, and expense of what someone wears, or drives, or eats–is a signal to others and to themselves about where they stand in society *relative* to other people. That is, people judge the adequacy of their own level of consumption by comparisons with people's consumption.

Veblen effects, conspicuous consumption, and working time

The things we purchase in order to impress ourselves or others are referred to as *conspicuous consumption*. One way to model this is to say that we compare our consumption to that of the very rich, and the closer our consumption is to theirs, the better we feel.

To do this we now define "effective consumption" as how adequate our consumption feels to us given what others are consuming. To capture this idea, we define effective consumption as follows:

$$\begin{aligned}\text{Effective consumption} &= \text{Consumption} - \text{Veblen effect} \\ &\quad \times \text{Consumption of the rich} \\ x^v &= x - v\underline{x} \end{aligned} \tag{7.30}$$

Where x is Scott's consumption, $\underline{x}$ is the consumption of the rich, and v is a positive constant representing the Veblen effect.

The negative effect of the consumption of the rich on our utility is captured in the term v (named after Thorstein Veblen). Effective consumption defined in this way expresses the idea that the consumption of the rich has the effect of diminishing the adequacy that we feel for any particular level of consumption. This raises the marginal utility of consumption to compensate. Scott's utility now includes this idea:

$$\begin{aligned}\text{Utility} &= (\text{Effective consumption})^{\alpha}(\text{Free time})^{1-\alpha} \\ u &= (x - v\underline{x})^{\alpha}(1-h)^{1-\alpha} = (wh - v\underline{x})^{\alpha}(1-h)^{1-\alpha} \end{aligned} \tag{7.31}$$

Using Equation 7.21 for the $mrs(x, 1-h)$ with a Cobb–Douglas utility function, and recalling that $x^v = x - v\underline{x}$, Scott's marginal rate of substitution is now:

$$\begin{aligned} mrs(x, 1-h) &= \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x^v}\right) \\ &= \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x - v\underline{x}}\right) \end{aligned} \tag{7.32}$$

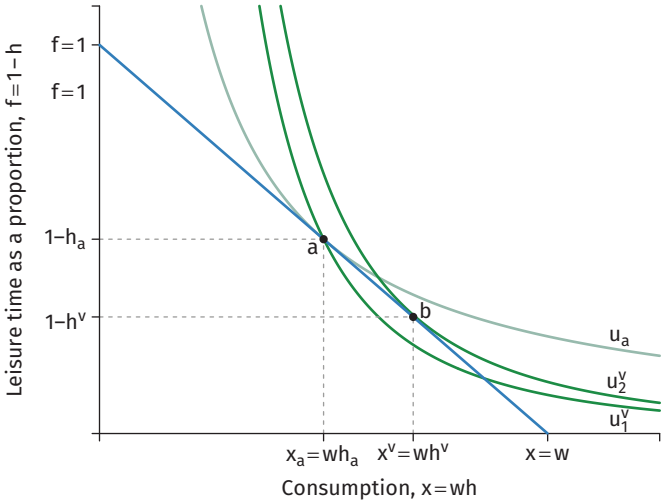
Increased consumption by the rich ($\underline{x}$) will diminish Scott's level of *effective* consumption, raising the $mrs(x, 1-h)$ (increased $\underline{x}$ reduces the denominator in the second term in Equation 7.32). So how much Scott values consumption relative to how much he values free time—his willingness to give up free time to consume more—is now *greater* than before. This means that in Figure 7.11 it has made the indifference curves *steeper*.

Point **a** in Figure 7.11, showed an initial choice that the worker, Scott, made *without* the Veblen effect, that is, when he did *not* worry about what others consumed. Point **b**, as a contrast, shows how the "Veblen effect" of the consumption of the rich affects the worker's choice.

The Veblen effect does *not* alter the feasible frontier, but it changes the marginal rate of substitution between consumption and free time. The indifference curves are steeper because at any level of *actual* (*not* effective)

HISTORY The term "conspicuous consumption" comes from the American economist and sociologist Thorstein Veblen (1857–1929). Well over a century ago he described exactly the trade-off we are modeling here. "The means of communication and the mobility of the population now expose the person to the observation of many persons who have no other means of judging his reputability than the display of goods . . . the present trend of the development is in the direction of heightening the utility of conspicuous consumption as compared with leisure."[2]

Figure 7.11 Feasible set and indifference curves for working time and consumption: Veblen effect. In the absence of the Veblen effect, the worker's maximum feasible utility is achieved at point **a** on indifference curve labeled u_a. The Veblen effect increases the person's willingness to pay for more consumption (by giving up more free time) and therefore steepens the indifference curve. At the new utility maximum, point **b**, he consumes more goods, works more, and consumes less leisure.



✓ **FACT CHECK** In 2001 the tax authorities in Norway began posting income tax records online, so that anyone could find out the income of their neighbors, friends, and coworkers. Huge numbers of people accessed the site. Ricardo Perez-Truglia studied the statistical relationship between Norwegians' income and measures of their "subjective well-being"—happiness and life satisfaction. Higher-income people were happier and had greater life satisfaction. But after incomes became public the differences between rich and poor in subjective well-being became much greater.[3]

HISTORY The people who set consumption standards, according to Veblen, are the *rich*. He wrote: "all canons and reputability and decency and all standards of consumption are traced back . . . to the usages and thoughts of the highest social and pecuniary class, the wealthy leisure class."[4]

consumption and free time, the marginal utility of effective consumption has *risen*. Why? Because when Scott compares himself to the rich it makes him feel as though he has *less* effective consumption than he actually has.

Figure 7.11 shows that when Scott maximizes utility by setting the $mrt(x,1-h)$ to the $mrs(x,1-h)$ Scott works longer hours, enjoys less free time, and increases the amount of consumption he purchases (see M-Note 7.6). The time Scott works now is greater than before, as you can see comparing points **a** and **b** in Figure 7.11 (see the M-Note 7.6):

$$h^v = \alpha + \frac{(1-\alpha)v\underline{x}}{w} > \alpha = h_a \tag{7.33}$$

Equation 7.33 shows the working time *without* the Veblen effect (namely, α) *plus* more time that Scott is now motivated to work due to the Veblen effect (namely, the fraction $\frac{(1-\alpha)v\underline{x}}{w}$). The inequality shows that the hours worked with the Veblen effect, h^v, are *greater* than those worked in the *absence* of the Veblen effect, h_a.

Veblen effects and falling inequality: An explanation of declining work hours in the twentieth century?

How does this model help us understand how working hours have changed over time and how work hours differ across countries? The model predicts

that the more rich people consume, the longer other people will work. So we would expect people to work longer hours in countries in which the rich are especially rich, and people to work less where the rich are only modestly richer than the rest.

M-NOTE 7.6 The choice of work hours with and without a Veblen effect

Using Equation 7.31, Scott's utility function when deciding between leisure and consumption may be represented as:

$$u = (x - v\underline{x})^{\alpha}(1-h)^{1-\alpha}, \quad \text{where } v \geq 0.$$

When $v = 0$, there is *no* Veblen effect, and the utility function is the same as in Equation 7.28 that we used in the previous section. When $v > 0$, there is a Veblen effect.

To calculate the time worked, we need to equate the $mrs(x, 1-h)$ to the $mrt(x, 1-h)$. Following 7.32 and 7.27:

$$\text{mrs} = \frac{\alpha}{1-\alpha}\frac{(1-h)}{(x-v\underline{x})} = \frac{1}{w} = \text{mrt}$$

Using $x = wh$

$$\begin{aligned}
\frac{\alpha}{1-\alpha}\frac{(1-h)}{(wh-v\underline{x})} &= \frac{1}{w} \\
\alpha w(1-h) &= (1-\alpha)(wh - v\underline{x}) \\
\alpha w - \alpha wh &= wh - \alpha wh - (1-\alpha)v\underline{x} \\
wh &= \alpha w + (1-\alpha)v\underline{x} \\
h^{v} &= \alpha + \frac{(1-\alpha)v\underline{x}}{w}
\end{aligned}$$

In the absence of Veblen effect ($v = 0$), the hours worked h depends only on the importance of consumption relative to free time in the utility function, α. A positive Veblen effect ($v > 0$) reduces effective consumption. Because there are diminishing returns to effective consumption (marginal utility of effective consumption is higher the less of it you have) the effect of there being less effective consumption is to increase the marginal utility of it. As a consequence, Scott *increases* his working hours (and consumption) and *reduces* his leisure.

You can also see that the higher the consumption of the rich, $\underline{x}$, the lower is *effective* consumption, and therefore the higher the hours worked (see Figure 7.11).

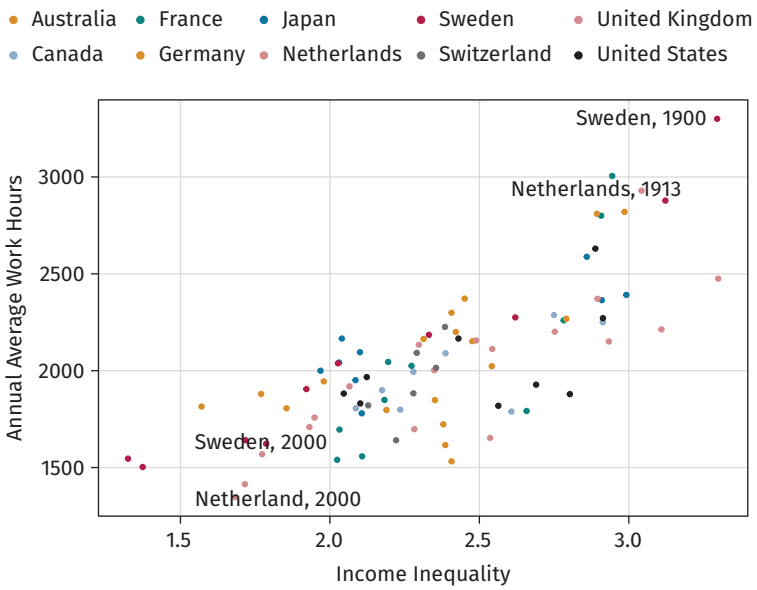
Figure 7.12 presents the average annual working hours and a measure of the fraction of all income received by the richest 1 percent of households, in ten countries over the twentieth century. The fraction of income received by the very rich is a key variable in the model because the consumption of the very rich divided by the wages of typical workers or $\underline{x}/w$ influences the size of the Veblen effect as equation 7.33 shows.

The figure shows that this prediction of the model with Veblen effects is borne out by the data: a larger share of income going to the very rich is associated with longer working time.

But it shows more: the decline in the relative incomes of the very rich is closely associated with the decline in work hours. Notice that Sweden is both the most unequal and "longest working" nation (in the early years of the data set) and also the most equal country with the most free time (in the

Figure 7.12 Inequality and work hours 1900 to 2000. The hours data are annual average work hours for full-time production workers. The income data are based on the share of total income received by the top 1 percent of households.

Figure 5a from Oh, S-Y., Park, Y., and Bowles, S., (2012), "Veblen effects, political representation, and the reduction in working time over the 20th century," *Journal of Economic Behavior & Organization*, 83(2). pp. 218–242. With permission from Elsevier.



more recent years). The countries that became *more* equal over this period also saw the greatest drop in work hours. The increase in work hours in both Sweden and the US at the end of the last century was associated with an increase in inequality in both countries.

In this model, conspicuous consumption by the very rich is a kind of "public bad." It is experienced equally by all members of society or at least can be by anyone with access to TV and social media (it is *non-excludable*). It is a "bad" and not a good because it reduces the utility of those it affects (it is *non-rival* in the disutility it creates). And the people affected then respond in ways that generate further negative external effects because the Veblen effect induces them to work and consume more, increasing the use of our limited environmental resources.

The Veblen effects model suggests some of the reasons for differing working hours among countries. But by itself it misses some important parts of the story. The most important missing element for the decline in work hours in the twentieth century is that voting rights were extended to include most adults early in the century. When overworked employees got the right to vote, in virtually all countries their trade unions and political parties demanded reductions in working hours.

The model with Veblen effects and the data provide an illustration of a more general point: consumption is not just a biological activity. Eating is not just nutrition. Clothing is *not* just keeping warm. Your home is more than four walls to keep out the weather. Consumption is a *social* activity. As Veblen said, our consumption is a *signal* to others and to ourselves about who we are. It is also a social activity in which we engage, for example, for the pleasure of the company of our friends.

CHECKPOINT 7.7 **Veblen effects** Use Equation 7.32 to show why the indifference curves in Figure 7.11 are steeper when there is a Veblen effect (as shown).

7.7 QUASI-LINEAR UTILITY, WILLINGNESS TO PAY, AND DEMAND

A useful interpretation of the marginal rate of substitution is possible when good y is *not* data, but instead the amount of money left over from the budget after purchasing good x. We used this when we first introduced budget constraints in Chapter 3. In the case where y is money for other goods, the marginal rate of substitution is literally the person's **willingness to pay** in money for the good x, the maximum amount of money she will pay for an additional unit of x when she has already bought the quantity x.

! REMINDER In previous chapters and earlier in this chapter we have considered x and y as two goods that changed hands by barter: so much code for some much coffee. In these models "willingness to pay" does not involve money, it is how much code you are willing to give up to get more coffee.

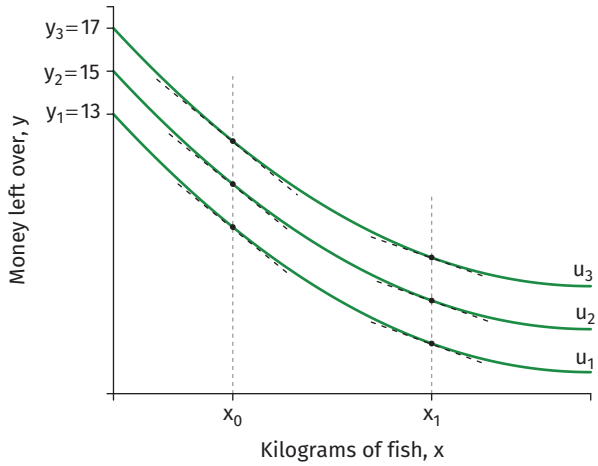
Quasi-linear utility and willingness to pay

In this case, the analysis of demand is greatly simplified if the marginal rate of substitution depends only on the amount of the good someone purchases, and not on the amount of money she has left over. Here are the simplifications:

- *Prices*: When y is money left over for other purchases, then $p_y = 1$ (the price of a dollar is a dollar), and we can simplify $mrt(x,y) = \frac{p_x}{p_y} = p_x$. So we can drop the subscript and just denote the price of the x-good as p. This is the opportunity cost of x or how much of the y good you have to give up to get a unit of x. Because the y good is money this opportunity cost is the price of x.
- *Marginal utility*: when utility is linear in y, then $u_y = 1$, that is, the marginal utility of money is constant and equal to 1. Therefore the marginal rate of substitution $mrs(x,y) = \frac{u_x}{u_y} = u_x$, the marginal utility of x.
- *Willingness to pay*: Because with the quasi-linear utility function the mrs is simply $u_x(x,y)$, that is, the marginal utility of x when consuming the

WILLINGNESS TO PAY The maximum amount a person would pay to acquire a unit of a good.

Figure 7.13 Harriet's indifference curves: quasi-linear utility. With quasi-linear utility, marginal rates of substitution depend only on the amount of the good x, and not at all on the amount of money left over to buy other goods, y. As a result, indifference curves with different levels of utility are vertical displacements of a single curve. This means that the slopes of the indifference curves when consuming x_0 amount of fish are the same independently of the amount of money she has for other purchases.



bundle (x,y), the $mrs(x,y)$ is also the maximum amount (in money units) that the person would be willing to pay to have a small increase in x. This means that the willingness to pay is measured in monetary units.

!REMINDER When a function is quasi-linear it depends *linearly* on one variable, e.g. y, and *non-linearly* on another variable, e.g. x, and has the form $u(x,y) = y + g(x)$. Hence it is *quasi* or "partly" linear.

We explore these two new ideas—the y-good is money left over and its marginal utility is a constant equal to 1—using the quasi-linear (QL) utility functions that you studied in Chapter 4 (sections 4.9 and 4.10):

$$u(x,y) = y + g(x) \tag{7.34}$$

This function illustrates the above bulleted properties:

- The marginal utility of the y-good is a constant equal to 1.
- As a result, the marginal rate of substitution is $u_x(x,y)$, which is the willingness to pay for an additional unit of x measured in the same monetary units as the y-good.
- Utility is measured in monetary units.

To see the last point, suppose the person spent nothing on the other good, then Equation 7.34 tells us that $u = y$: in this case, utility is the amount of money the person has to spend. If the amount of money the person has to spend increases by \$1, then utility increases by 1.

This does not mean that the *only* thing the person cares about is money. The amount of x the person consumes may matter a lot. It just means that how much it matters will be measured in money equivalents.

These properties make y more like wealth measured in monetary units (e.g. euros) than like a particular good such as data or coffee. So, we will often refer to y as *money*, understanding that it is really generalized purchasing power that can be spent on many other things possibly in many periods.

The feature of this utility function that the marginal utility of money available for purchasing goods is a constant is a reasonable approximation if purchases of x constitute a small fraction of a person's budget. If the purchase under consideration will use up 1 percent of the buyer's annual budget, then it is plausible to assume that this will *not* affect the marginal utility of the person's available money.

But if we consider the entire range of a person's possible income or wealth from destitute poverty to great affluence, then it would be far more realistic to let the marginal utility of money be greater for the poor person than for the very wealthy as would be the case with the Cobb–Douglas utility function for example.

Quadratic, quasi-linear utility

Many of the examples in this book use the particular class of **quadratic quasi-linear utilities**, where the function $g(x) = \bar{p}x - \frac{1}{2}\left(\frac{\bar{p}}{\bar{x}}\right)x^2$ is a quadratic function of the good x, and as a result the utility function is:

$$\text{Quadratic, quasi-linear utility} \qquad u(x,y) = y + \bar{p}x - \frac{1}{2}\left(\frac{\bar{p}}{\bar{x}}\right)x^2 \tag{7.35}$$

In Equation 7.35:

- **Satiation:** The parameter $\bar{x}$ represents the level of x at which the buyer is *satiated* with x and would consume no more even if the price were zero.
- **Maximum willingness to pay:** The parameter $\bar{p} > 0$ represents the buyer's maximum willingness to pay for the first unit of x when they do not have any x, i.e. when $x = 0$.

The marginal utility of x with the QQL utility is:

$$\frac{\Delta u}{\Delta x} = u_x(x,y) = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.36}$$

Equation 7.36 tells us the following:

M-CHECK

$\bar{p}$ is the person's *maximum willingness to pay* for good x. She won't pay more than $\bar{p}$ to get a unit of x.

$\bar{x}$ is the person's *satiation point* for x, beyond which her marginal utility of x is negative. She would prefer *not* to consume $x > \bar{x}$.

The point at which you are *sated* (verb) is where you reach *satiation* (noun) from consuming a good, like x. The intuition is easily seen with food: you reach satiation at that point where you do not want to eat another mouthful (the marginal utility hits zero) or, if you do, you know you'll regret it (the marginal utility will be negative). Or, it is the point at which you have reached bliss, which is perfect happiness or great joy, and at which, if you consumed or did any more, it would detract from that bliss, joy, or wonder.

QUADRATIC, QUASI-LINEAR UTILITY FUNCTION A quadratic, quasi-linear utility function is quadratic in one variable and linear in another variable.

- When $x < \bar{x}$, the buyer's marginal utility of x is positive, and she regards x as a *good*.
- When $x > \bar{x}$, the buyer's marginal utility of x is negative, and she regards x as a *bad*.

If y is budget left over to buy other goods, then the marginal utility of y is always 1, regardless of the levels of x and y. As a result, the marginal rate of substitution is equal to the marginal utility of x:

$$mrs(x,y) = \frac{u_x}{u_y} = u_x = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.37}$$

We can think about Equation 7.37 in the following way:

- Equation 7.37 is the equation for a line.
- Equation 7.37 has vertical intercept $\bar{p}$, which is the buyer's maximum willingness to pay.
- Equation 7.37 has a horizontal intercept $\bar{x}$, which is the point beyond which the buyer does not want to pay for good x (at $x = \bar{x}$, the buyer's willingness to pay is zero). $\bar{x}$ is the buyer's bliss point. To get the buyer to consume more than $\bar{x}$ of the good, you would have to *pay her*, rather than expecting her to pay you.
- Equation 7.37 shows that the demand curve has a slope of $-\frac{\bar{p}}{\bar{x}}$, which is the negative of the ratio of the maximum willingness to pay to the satiation point.

Equation 7.35 shows that Harriet's utility for the good x depends on how much she has relative to a level $\bar{x}$. When $0 \leq x \leq \bar{x}$, Harriet's utility increases if she has more x. But when $x \geq \bar{x}$, Harriet's utility decreases as she gets more x. We can plot quadratic, quasi-linear marginal rate of substitution or willingness to pay as a function of x as in Figure 7.14.

The *inverse demand function*, which gives the highest price Harriet will pay for each given total amount of the good is:

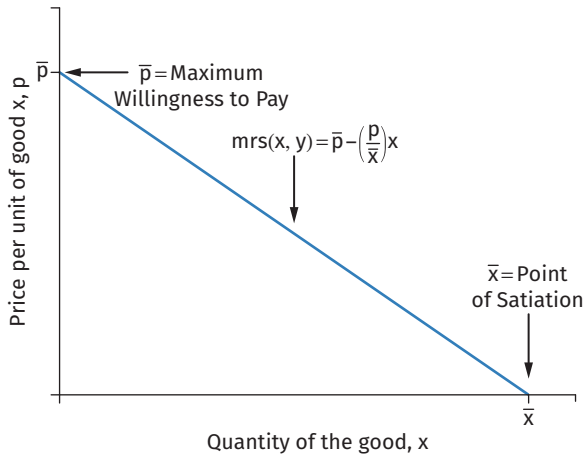
$$\text{Inverse demand curve:} \qquad p(x) = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.38}$$

The slope of the curve, $\frac{\Delta p}{\Delta x} = -\frac{\bar{p}}{\bar{x}}$ can be represented by single coefficient, $-\beta$, giving us:

$$\text{Inverse demand curve:} \qquad p(x) = \bar{p} - \beta x \tag{7.39}$$

Suppose Harriet has quadratic quasi-linear preferences between a good x and money left over to buy other things, with the willingness to pay $p(x) = \text{mrs}(x,y) = \bar{p} - \frac{\bar{p}}{\bar{x}}x$. She starts out with budget m and has the opportunity to buy any amount of the good x at the price p. If $p(0) = \bar{p} > p$, Harriet will buy at least some of the good. If she buys x_0 units of the good, Harriet's

Figure 7.14 Harriet's marginal rate of substitution (demand): quadratic quasi-linear preferences. The figure shows Harriet's $mrs(x, y)$ for a good, x, and money for other goods, y. That is, the downward-sloping line is Harriet's *willingness to pay* in money for an additional unit of good x for different levels of the quantity she has of good x and is therefore also her *demand curve* because it shows the relationship between the price of the good and how much of it she will buy at different prices. The vertical intercept, $\bar{p}$, is Harriet's maximum *willingness to pay* when she currently consumes zero units of good x. The horizontal intercept, $\bar{x}$, is her satiation point or bliss point, beyond which her marginal rate of substitution is negative so that x changes from being a good to a bad.



willingness to pay will have fallen to $p(x_0)$. If $p(x_0) > p$, then Harriet will buy more of the good, and so on, until $p(x) = p$.

At this point Harriet doesn't buy any more of the good, and spends the rest of her budget on other things. So from Harriet's willingness to pay (or marginal rate of substitution), we can derive the *quantity demanded* at a market price p, or:

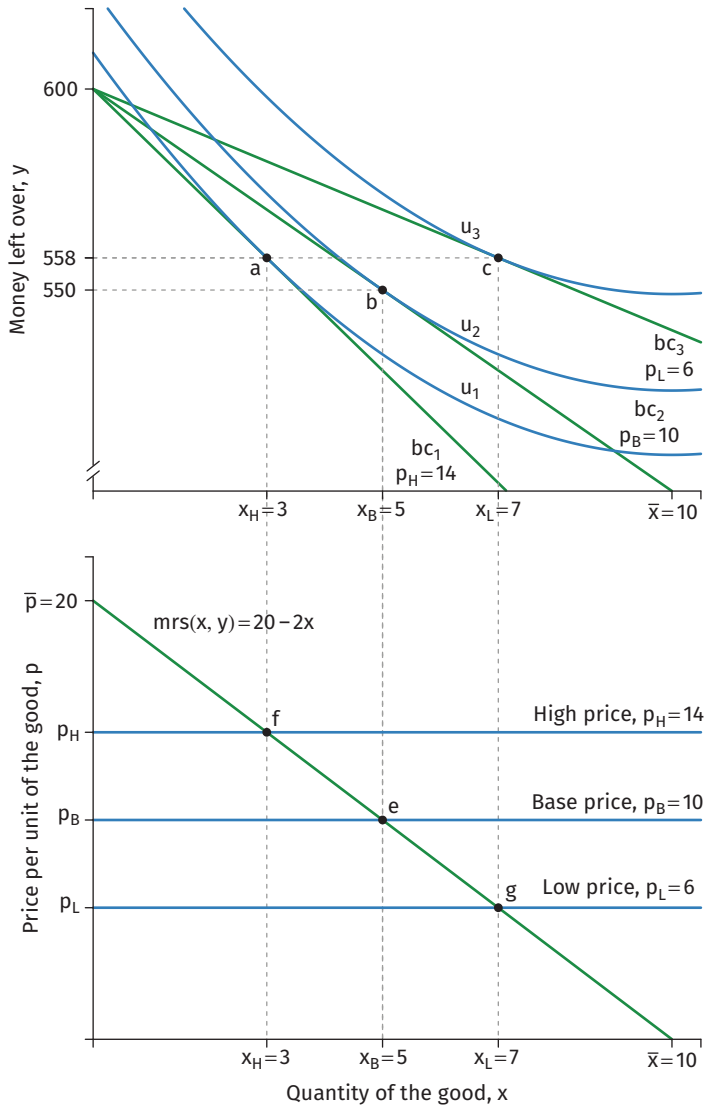
$$\text{Demand curve} \qquad x(p) = \bar{x} - \frac{\bar{x}}{\bar{p}}p = \bar{x} - \frac{1}{\beta}p = \bar{x}\left(1 - \frac{p}{\bar{p}}\right)$$

This equation says that for $\bar{p} > p > 0$ Harriet will consume an amount equal to a fraction less than one of her point of satiation ($\bar{x}$) given by the ratio the price of the good p to her maximum willingness to pay $\bar{p}$.

Figure 7.15 demonstrates this relationship by showing Harriet's utility-maximizing choices between x and y with her indifference curves and budget constraints for three prices of x in the top panel, while also showing her marginal rate of substitution of money for the good in the lower panel. The lower panel also shows how Harriet's marginal rate of substitution corresponds to a demand curve, by showing three different price levels and how the given price determines the quantity demanded at that price.

M-CHECK For example, when the market price is: $p = \$10$, $\bar{p} = \$20$, and $\bar{x} = 10$, then Harriet would like to buy five apples since her willingness to pay is the following: $mrs(5, y) = \$10$ for any y.

Figure 7.15 Harriet's utility-maximizing choice and marginal rate of substitution (demand): quasi-linear preferences. The top figure shows Harriet's indifference curves for kilograms of fish (x) and money left over for other goods (y). The figure shows her utility-maximizing choices at three prices for a kilogram of fish. Harriet's utility function is $u(x,y) = y + 20x - \frac{1}{2}\left(\frac{20}{10}\right)x^2$. Her budget constraint is $y = 600 - px$. The lower panel shows Harriet's $mrs(x,y)$ for a good, x, and money for other goods, y. Harriet's marginal rate of substitution is therefore $mrs(x,y) = 20 - 2x$. Her marginal rate of substitution, as her willingness to pay in money (y) for goods (x), is her *demand function* for x. She has a y-intercept of $y = 20 = \bar{p}$ (her maximum willingness to pay) and her x-intercept is $x = \bar{x} = 10$ (the amount of fish that satiates her appetite for fish, which is also the *maximum* quantity of fish she would consume were the price of fish zero). The slope of her marginal rate of substitution suggests she will exchange money (y) for fish (x) until her $mrs(x,y) = p$, i.e. when $20 - 2x = 10$, which implies $x = 5$ when $p = 10$.



M-NOTE 7.7 The demand for x and y with quadratic quasi-linear preferences

We begin with the $mrs = mrt$ rule, and then rearrange it to give us the demand for good x:

$$\begin{aligned} mrs(x,y) &= \bar{p} - \frac{\bar{p}}{\bar{x}}x = p = mrt(x,y) \\ \frac{\bar{p}}{\bar{x}}x &= \bar{p} - p \\ x(m,p) &= \bar{x} - \frac{\bar{x}}{\bar{p}}p \end{aligned} \tag{7.40}$$

Harriet will then use whatever remains of her budget (m) as money to spend on other goods, y, given what she spent on x at its price, p:

$$\begin{aligned} y &= m - px \\ y &= m - p\left(\bar{x} - \frac{\bar{x}}{\bar{p}}p\right) \\ y(m,p) &= m - p\bar{x} + p^2\frac{\bar{x}}{\bar{p}} \end{aligned} \tag{7.41}$$

Equation 7.41 shows that once we determined the demand for x, we can derive the demand for the other good as well.

CHECKPOINT 7.8 Satiation and willingness to pay Use Equation 7.37 to show that if $x = \bar{x}$ the willingness to pay is zero and at $x = 0$ the willingness to pay is $\bar{p}$.

7.8 PRICE CHANGES: INCOME AND SUBSTITUTION EFFECTS

When the price of a good changes, the consumption of the goods changes as we saw when deriving the demand curve in Figure 7.15. The total amount of the change is made up of two components:

- The **income effect**. A change in the price of a good alters people's real income, expanding or shrinking the feasible set of purchases. The effect of this change in real income on the goods purchased is the income effect.
- The **substitution effect**. When the price of a good changes, the change in the consumption of the good that is due to the change in relative prices (holding constant the buyer's real income) is the substitution effect.

INCOME EFFECT When the price of a good changes, this alters people's real income, expanding or shrinking the feasible set of purchases. The effect of this change in real income (with no change in price) on the goods purchased is the income effect.

SUBSTITUTION EFFECT When the price of a good changes, the change in the consumption of the good that is due to the change in relative prices (holding constant the buyer's real income) is the substitution effect.

Income and substitution effects for normal goods

Consider an increase in the price of good x, p_x. Harriet will change the amount of good x that she buys. We can *decompose* or separate this change into two effects. The first, the *substitution effect*, is how Harriet would change her purchases of x if she could hypothetically respond by choosing some point on the same indifference curve of her purchased bundle prior to the price change. The second is the *income effect*, which is the total change in her purchases minus the price effect, capturing the results of the reduced real value of her budget (and real income) caused by the price increase.

In Figure 7.16, Harriet starts at point **a** with bundle (x_a, y_a) on her initial budget constraint bc_a *before* the price increases. When the price increases,

Figure 7.16 Income and substitution effects: Cobb–Douglas utility. The total effect of a price change is the change in quantity demanded. The different effects shows how the total effect is broken up (or what economists call "decomposed") into the two parts of the substitution effect (a movement along the indifference curve) and the income effect (a movement to a new indifference curve). The substitution effect is shown by the hypothetical movement along u_2 from **a** to **c**. The income effect is shown by the movement to another indifference curve. Comparing points **b** and **c**, the income effect is measured by the difference in purchases of coffee between the two points $(x_c - x_b)$. The fact that the change in the price of coffee does *not* affect the level of purchases of data $y_a = y_b$ is an attribute of the specific Cobb–Douglas utility function we have used here, it is *not* a general result.

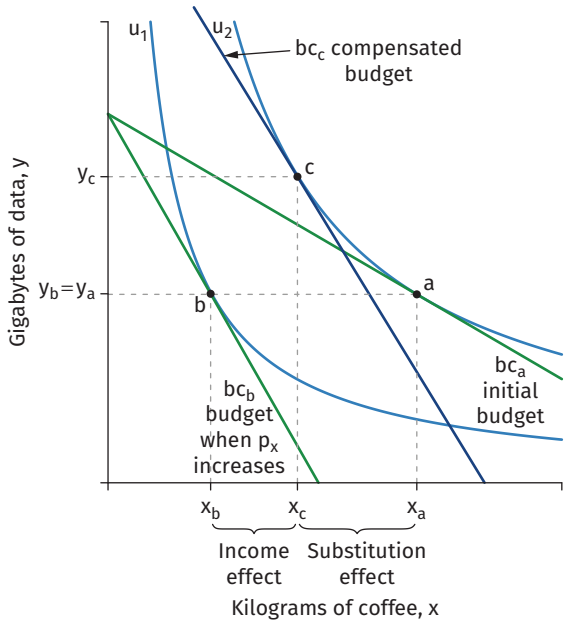


Table 7.1 Utility functions and their income and substitution effects. Remember that the substitution effect is captured by a movement along an indifference curve as prices or real incomes change, whereas the income effect is captured by a movement to a new indifference curve.

Utility function	Income effect	Substitution effect
Cobb–Douglas	Yes	Yes
Quadratic, quasi-linear	No	Yes
Perfect complements	Yes	No

the budget constraint pivots inwards to bc_b and creates a new utility-maximizing bundle at **b** with bundle (x_b, y_b).

To break down the two effects, we use the hypothetical idea of a **compensated budget constraint**. The compensated budget constraint takes the new prices of goods as given (it is parallel to the budget constraint after the price change), but it assumes the person has just sufficient income to purchase a bundle on their original indifference curve, at a new point of tangency with the original indifference curve u_2. The new tangency is at point **c** at bundle (x_c, y_c). The compensated budget constraint is entirely hypothetical: it is a thought experiment allowing us to look at the two different effects of a price change.

The difference between x_a and x_b is the *total effect* of the price change or the substitution effect plus the income effect. By construction the substitution effect causes a movement along the original indifference curve to point **c**. The difference between x_c and x_a is the *substitution effect*. The difference between x_b and x_c is the *income effect*.

! REMINDER In Chapter 6 we introduced the idea that inputs to a production process may be either substitutes (computer-driven machine tools and skilled machine operators) or complements (computer-driven machine tools and engineer-programmers).

Complements and substitutes in consumption

The size of the income and substitution effect will depend on whether the goods "go together" or have an "either/or" quality. If two goods are more enjoyable consumed *together*, then they are **complements**

BUDGET CONSTRAINT (COMPENSATED) A compensated budget constraint after a price change takes the new prices of goods as given (so it is parallel to the budget constraint after the price change), but gives the person just sufficient income to purchase a bundle on their original indifference curve, at a new point of tangency.

COMPLEMENTS IN CONSUMPTION Goods are complements in consumption if an increase in the quantity consumed of one raises the marginal utility of the other.

(coffee and cookies). "Either/or" goods are called **substitutes**: they are consumed *instead* of each other (tea and coffee). Perfect substitutes are goods between which the marginal rate of substitution does not depend on the ratio of quantity of each consumed (this means that the indifference curves are linear; see M-Note 7.8). Perfect complements are goods that are valuable only when consumed in some given proportion to each other (right shoes and left shoes, nuts and bolts). Indifference curves for perfect complements are L-shaped.

M-NOTE 7.8 **Complements and substitutes in consumption**

Remember, similar to the case of inputs into production (see M-Note 6.8) we define complements and substitutes in consumption by the effect that having more of one good has on the marginal utility of the other. The elasticity of substitution, by contrast, is about whether the goods are consumed in some fixed proportion—like right shoes and left shoes, or knives and forks, examples of a low elasticity of substitution—or can be varied in their proportions—like how much milk you put in your coffee, an example of a high elasticity of substitution.

Complements: Cookies and coffee. In the Cobb–Douglas utility function $u = x^{\alpha}y^{1-\alpha}$, x (coffee) and y (tea) are complements because:

$$\frac{\partial u}{\partial x} = \alpha x^{(\alpha-1)}y^{(1-\alpha)}$$

which, because $0 < \alpha < 1$, means that:

$$\frac{\partial^2 u}{\partial x \partial y} = (1-\alpha)\alpha x^{(\alpha-1)}y^{(-\alpha)} > 0$$

so the greater is the consumption of y, the higher is the marginal utility of x. By the same reasoning, the greater is the consumption of x, the higher is the marginal utility of y.

Substitutes: Coffee and tea. Here is a utility function in which x (coffee) and y (tea) are substitutes:

$$u = (x + \varepsilon y)^{\alpha} \tag{7.42}$$

where ε is a positive constant measuring how much the person prefers tea to coffee and $0 < \alpha < 1$. Therefore, finding the marginal utility of x by taking partial derivatives:

$$\frac{\partial u}{\partial x} = u_x = \alpha(x + \varepsilon y)^{\alpha-1} > 0$$

The marginal utility of x is positive. But how does the marginal utility of x change as consumption of y changes? We can work that out by taking the partial derivative of the marginal utility of x with respect to y:

$$\frac{\partial^2 u}{\partial x \partial y} = \varepsilon(\alpha-1)\alpha(x + \varepsilon y)^{\alpha-2} < 0$$

continued

SUBSTITUTES IN CONSUMPTION Goods are substitutes in consumption if an increase in the quantity consumed of one reduces the marginal utility of the other.

It is negative because $\alpha < 1$. This shows that the marginal utility of coffee is less the more tea the person consumes. The same reasoning shows that the marginal utility of tea is less, the more coffee the person consumes. For this particular utility function, tea and coffee are what is called perfect substitutes. In this case,

$$\frac{\partial u}{\partial y} = \varepsilon\alpha(x + \varepsilon y)^{\alpha-1}$$

so the individual's marginal rate of substitution, the ratio of the two marginal utilities, is as follows:

$$\frac{\partial u/\partial x}{\partial u/\partial y} = \frac{\alpha(x + \varepsilon y)^{\alpha-1}}{\varepsilon\alpha(x + \varepsilon y)^{\alpha-1}} = \frac{1}{\varepsilon} = mrs$$

Because the $mrs(x, y)$ is the negative of the slope of an indifference curve, the fact that it is constant means that the indifference curves are linear. This is what the fact that x and y are perfect substitutes means.

CHECKPOINT 7.9 **Income and substitution effects** Explain why the substitution effect of a price increase must be negative (using Figure 7.16 may help).

7.9 APPLICATION: INCOME AND SUBSTITUTION EFFECTS OF A CARBON TAX AND CITIZEN'S DIVIDEND

The decomposition of the results of price changes into income and substitution effects can be illustrated by a proposed carbon tax to reduce emissions of carbon dioxide and other greenhouse gases that contribute to climate change.

The prices of petroleum, coal, natural gas, and other fossil fuels do not include the costs of the environmental and climate-change external effects of their use. This means that people pay a *private cost* of using fossil fuels that is *lower* than the social (private plus external) cost of using them. The result—as in the case of overfishing in Chapters 1 and 5—is *overuse* of fossil fuels.

Now consider a tax imposed on the sale of fossil fuels, in conjunction with a transfer of the resulting tax revenues in equal amounts back to the members of the population, called a citizen's dividend. We ask: How would this so-called carbon tax and citizen's dividend policy reduce consumption of fossil fuels and affect citizens' consumption of other goods?

We consider two steps in our policy process:

- The substitution and income effects of the increased price of fossil fuels.
- The income effect of the citizen's dividend.

! REMINDER In Chapter 5 we explained that the social cost equals the private cost plus the (negative) external cost imposed by a person's action. In that case, a person's marginal private cost of fishing included their disutility of fishing, but the social cost included not only the private costs but also the negative external effect the fishermen imposed on others.

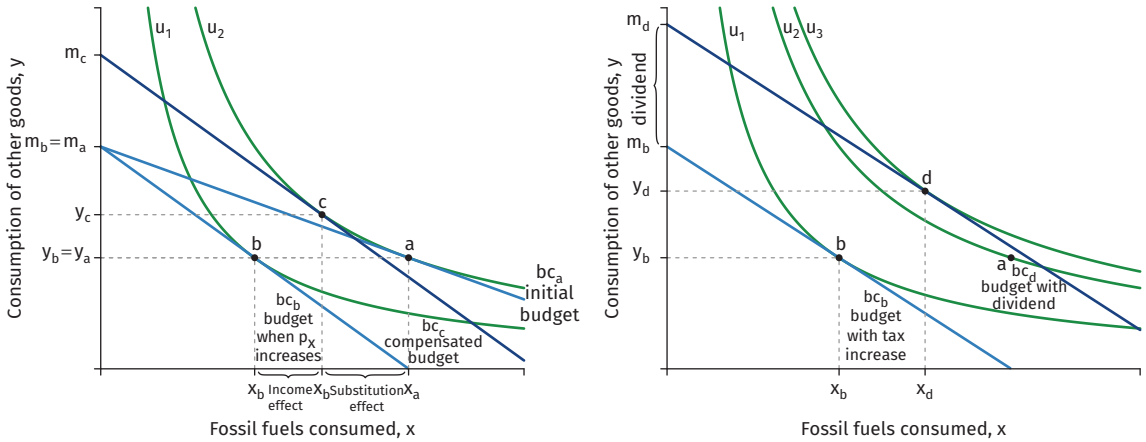
EXAMPLE In their 2017 op-ed (opinion editorial) in *the New York Times* three economists made "A conservative case for climate action," proposing the carbon tax and citizen's dividend that we analyze here. They included Martin Feldstein and Gregory Mankiw, chief economic advisers to US Presidents Ronald Reagan and George W. Bush, respectively.[5]

Reducing carbon emissions by imposing the tax

In Figure 7.17, the fossil fuel consumption (x) of a citizen is plotted on the horizontal axis, and the consumption of other goods measured in some currency is plotted on the vertical axis (y). Before the tax, a citizen is at point **a** in Figure 7.17 where the marginal rate of substitution equals the marginal rate of transformation, which is the existing price of fossil fuels, p_a. At **a**, the citizen has a utility of u_2 on the corresponding indifference curve.

The government then imposes a tax on fossil fuels. The tax increases the price of fossil fuels. With the increase in the price from p_a to p_b due to the tax, the citizen's budget constraint becomes *steeper*. It pivots inward round its y-axis intercept because the amount of the budget itself is unaffected, but the price has *increased*. So bc_b is the *new* budget constraint. As before, the citizen now maximizes her utility now consuming at point **b** where her marginal rate of substitution equals the new marginal rate of transformation, p_b. At **b** the constrained utility maximum, the citizen has decreased her consumption of fossil fuels from x_a to x_b, consistent with the policy goals of the tax to *decrease* consumption of fossil fuels.

Figure 7.17 Carbon tax with dividend. A citizen decides on consumption of fossil fuels (x) and other goods (y). Prior to the introduction of the carbon tax the citizen's budget is m_a and the budget constraint is the line labeled bc_a, with slope $-p_a$. The utility-maximizing bundle is on the budget constraint at point **a** where the citizen's indifference curve u_2 is tangent to the budget constraint. A carbon tax increases the price from p_a to p_b steepening the budget constraint. The result, shown in panel (a), is that the budget constraint pivots inward. The effect of the price change—a reduction in fossil fuel consumption from x_a to x_b—is the sum of the income and substitution effects. In panel (b) the citizen's dividend increases the household's budget so the budget constraint shifts outward and has a vertical intercept (the budget itself) $m_d > m_b$. The indifference curves shown here are based on a Cobb–Douglas utility function with $\alpha = 0.5$.



(a) Substitution and income effects of the price increase

(b) Second income effect due to citizen dividend

Setting aside the value that the citizen places on the overall mitigation of the climate-change crisis, the policy has lowered her utility because she is on a lower indifference curve, $u_1 < u_2$.

Is the tax fair?

All citizens will have lower utility, but the effect will differ by levels of income. In the U.S., the percentage reduction in *real income* imposed by the carbon tax will be *larger* among poorer households. This is because, as Figure 7.18 shows:

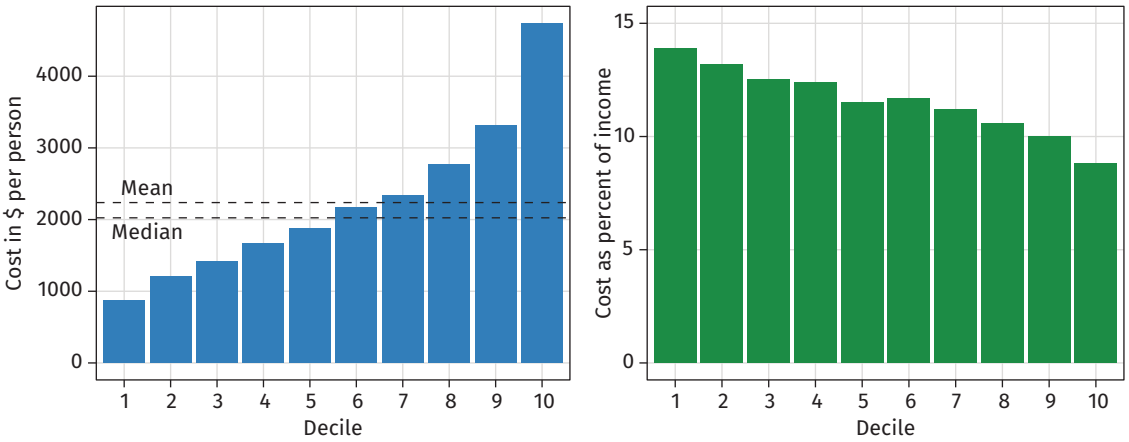
- while (panel (a)) higher-income people spend much more than lower-income households on carbon costs (think about air travel, heating, and air-conditioning large houses);
- expenditure on carbon costs *as a fraction of their total expenditure* is greater for lower-income households (panel (b)).

As a result, high-income people will pay more of the tax than low-income people, but the tax will lower the real income of poor people by a larger percentage than will be the case for higher-income people. In the US the carbon tax is regressive, meaning that the amount paid as a fraction of a household's income is greater for lower-income households. This is where the citizen's dividend comes in.

✓ **FACT CHECK** Using detailed data on household expenditures economists Anders Fremstad and Mark Paul calculated that a tax of \$50 per ton of CO_2, the revenues of which were distributed equally to all citizen households would raise the real income of 56 percent of US households and 84 percent of households with incomes less than the average income.[6]

Figure 7.18 Dividend distribution and carbon costs. The left-hand figure shows the absolute amount spent on carbon for each of the ten household income deciles (1 is poorest, 10 is richest). The right-hand figure shows the proportion of household expenditure on carbon by for the ten income deciles.

Source: Fremstad and Paul (2019) using data from the Energy Information Agency, the Bureau of Economic Analysis, and the Bureau of Labor Statistics.



(a) Average carbon costs per person by income decile

(b) Average proportion of expenditure on carbon by income decile

Increasing income and ensuring fairness through a citizen's dividend

To see how the citizen's dividend alters the result, return to Figure 7.17 (b). As in panel (a), the citizen is at point **b** with lower utility than before the tax. But now suppose the total carbon tax revenues collected are divided equally and distributed to each household, raising the budget of each from m_b to m_d. This is the citizen's dividend. As you can see, the effect of the dividend is to shift upwards the budget constraint by the same amount.

With the higher income, the citizen maximizes her utility at point **d**. For the citizen we have modeled, the dividend ($m^d - m^b$) is large relative to her previous budget.

The result is an increase in consumption of other goods, so that her level of utility is higher than it was before the tax, that is, comparing points **d** and **a**, $u_3 > u_2$. At **d**, they have higher consumption of other goods ($y_d > y_b$) and they consume a lower level of fossil fuels than they did previously, but greater than before the dividend ($x_a > x_d > x_b$). With the greater consumption of other goods, the citizen has higher utility and they obtain utility of $u_3 > u_1$.

All citizens receive *the same* dividend but higher-income households will pay a larger share of the tax from which the dividend is funded (because their carbon consumption is greater). As a result, lower-income households would pay *less* in taxes (which are proportional to the cost of the carbon they consume) than they receive in the citizen's dividend which is proportional to the mean carbon costs consumed.

Thus, in the US (and other higher-income countries) the carbon tax alone is **regressive**, but the carbon tax and the citizen's dividend taken together is **progressive**. The case we modeled in Figure 7.17 illustrates the increase in disposable income and utility of a poorer-than-average citizen.

International differences

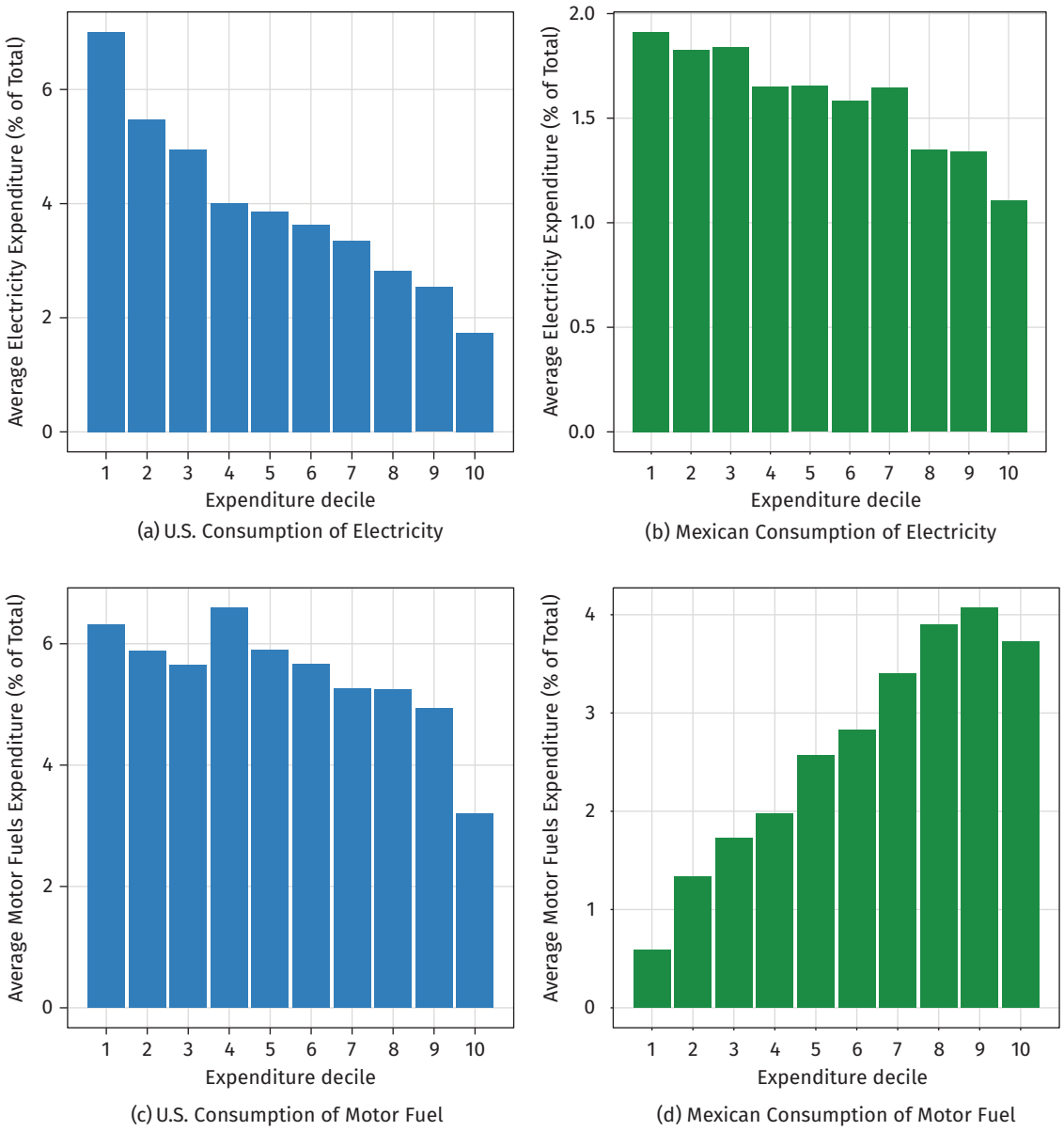
In some countries, however, the picture is reversed, with poor people spending a small fraction of their budget on carbon-related consumption and wealthy families spending a larger share. The data in Figure 7.19 for Mexico show exactly this, at least for motor fuels. As with carbon costs as a whole, the US data (in the left panels) show that the fraction of the

REGRESSIVE POLICIES A system of taxes and transfers or other policies that increase disposable income inequality is regressive.

PROGRESSIVE POLICIES A system of taxes and transfers or other policies that reduce disposable income inequality is called progressive.

Figure 7.19 US and Mexican consumption of motor fuel. Each county's income distribution is divided into deciles from poorest (1) to richest (10). The average consumption of motor fuel as a proportion of total family income for each decile is shown by the size of the bar for that decile. In the US, the consumption as a share of income is higher for lower deciles than for higher deciles. In Mexico, the consumption as a share of income is *lower* for lower deciles than for higher deciles.

Source: Pizer and Sexton (2019) using data from the US Consumer Expenditure Survey (Bureau of Labor Statistics) and Mexico's National Survey of Income and Expenditure (Instituto Nacional de Estadistica y Geografia).



household budget spent on electricity and motor fuels respectively falls dramatically as income rises.

But this is not the case for Mexico. The fraction of the household budget spent on electricity is only modestly lower for the upper-income households. And for motor fuels (car and truck use) the fraction of the budget spent on carbon consumption is much greater for high-income households than for those with lower incomes (among whom car and truck ownership is limited). In Mexico therefore, unlike in the US, a tax on motor fuels would be *progressive* rather than *regressive*.

CHECKPOINT 7.10 **Carbon tax and dividend** Using the data in Figures 7.18 and 7.19 explain why a carbon tax is regressive in the US but possibly not in Mexico.

7.10 APPLICATION: GIFFEN GOODS AND THE LAW OF DEMAND

The demand curves you have seen all slope downward: a lower price is associated with *more* purchases. This is called the **law of demand**, and the movement of prices and quantities purchased in opposite directions that the law predicts is widely observed. But there is a special kind of good–called a **Giffen good**–for which the law of demand is violated. For Giffen goods, a *higher* price is associated with a *greater* amount of purchases.

You already know that for an inferior good the amount purchased will decline as income rises. This is not really surprising; some of the low-cost foods that people eat when they have very limited budgets will not be purchased at all when they have more income to spend. A Giffen good really is surprising because less is purchased when *its own price* decreases.

!REMINDER An inferior good is one for which purchases fall as income rises. A Giffen good must be inferior, but not all inferior goods are Giffen goods.

How could this be? Think about a poor family consuming a large amount of an inferior good. When the price decreases there is both a price effect motivating the family to purchase more and an income effect resulting from the decrease in price. Because the good is inferior, the higher real income of the family motivates them to purchase *less* of the good. If the negative income effect is *greater* than the positive substitution effect, purchases will *decline* in response to a decline in the price.

Here is an example of a Giffen good. In China, for very poor households, rice is the main staple food and if they have enough money, they add other

LAW OF DEMAND The law of demand holds that a decrease in the price of a good will result in an increase in the quantity of the good purchased.

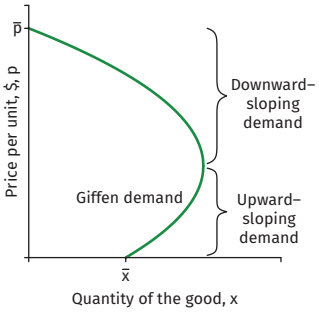
GIFFEN GOOD Over some range of prices, purchases of a Giffen good increase if the price rises, and fall if the price falls.

foods that make rice taste better, such as shrimp or beef. However, when the price of rice increases, this means the households have little money left over to buy beef or shrimp. Consequently, they will consume *more rice* even though its price has increased.

As a result, over some range of prices the demand curve for rice for these families is *upward-sloping*. Of course if the price of rice rose so high that the household purchased only rice, then further price increases would have to reduce the amount purchased, so the demand curve would then be downward sloping as the law of demand requires. Just such a demand curve is illustrated in Figure 7.20.

This is exactly what economists who studied subsidies of rice observed in Hunan, a region of China. They ran an experiment by subsidizing the price of rice, lowering the price the families actually paid. When they provided the subsidy, very poor households *reduced* their consumption of rice. That is, when rice was *cheap*, households consumed *less* of it. When they removed the rice subsidy so that prices *rose*, the households consumed *more* rice. For these households rice was a Giffen good.[7]

Figure 7.20 Demand for a Giffen good. The demand for a Giffen good is upward-sloping when the price is low. Over this region, a higher price is associated with more purchases. At a sufficiently high price, though, demand becomes downward-sloping.



CHECKPOINT 7.11 Inferior goods and Giffen goods What is the difference between an Inferior and a Giffen good? Can you think of examples of either kind of goods in your current consumption?

7.11 MARKET DEMAND AND PRICE ELASTICITY

The **market demand** for a good at any given price is the *sum* of the demands at that price of all the people making up the *demand side* of the market. We can compute the *market demand* by adding up the individual demand curves. If we plot all the demand curves with the quantity demanded of x on the horizontal axis and the price p on the vertical axis, this requires the horizontal summation of the individual demand curves. We use an upper-case X for market demand and a lower-case x for an individual demand.

Figure 7.21 shows how (on the left) an individual market demand curve is summed over ten people to produce the market demand curve (on the right), that is $X(p) = x_1(p) + x_2(p) + \ldots + x_{10}(p)$.

A linear market demand curve (quadratic quasi-linear utility)

If there are n identical buyers, each of whom has the same quadratic, quasi-linear utility for the good, with the same parameters $\bar{x}, \bar{p}$, each individual has the demand (from Equation 7.37):

DEMAND CURVE (MARKET) A demand curve provides the answer to the hypothetical question: what is the maximum amount of a good that can be sold at each price? The market demand curve is the horizontal summation of the individual demand curves of actual or potential buyers.

$$\text{Individual demand} \qquad x(m,p) = \bar{x} - \frac{\bar{x}}{\bar{p}}p \tag{7.43}$$

The market demand is then the sum of all the individual demands. But, since they are all equal for the identical people, this is the same as the number of people (n) multiplied by the individual demand curves. In the quadratic quasi-linear case, therefore, the market demand curve is given by the following:

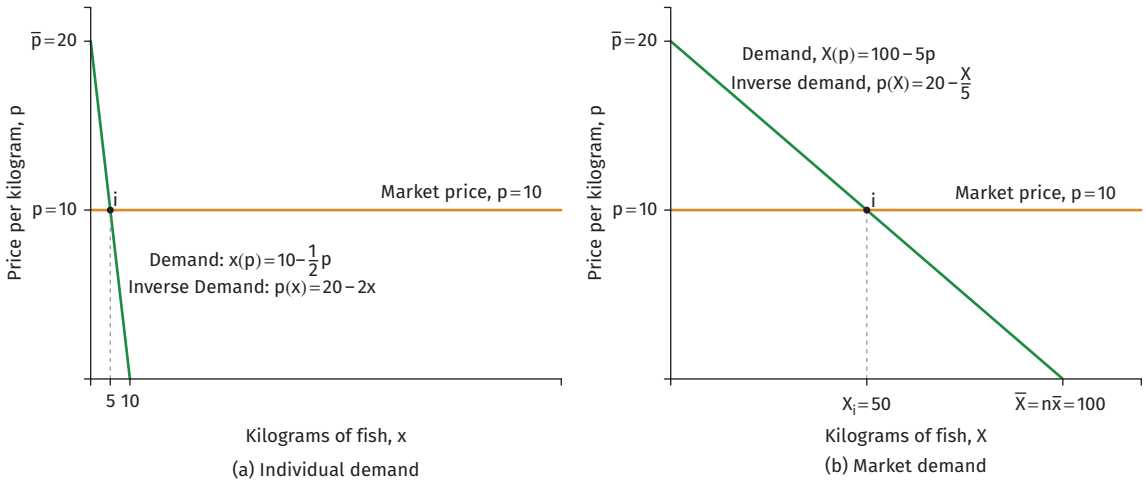
$$\begin{aligned}\text{Market demand} \qquad X(p,m,n) &= \text{Number of people} \times \text{Individual demands} \\ &= n\left(\bar{x} - \frac{\bar{x}}{\bar{p}}p\right) \\ &= \bar{X} - \frac{\bar{X}}{\bar{p}}p \end{aligned} \tag{7.44}$$

Because the market demand is the summation of individual demands, the market demand curve is also downward-sloping: quantity demanded falls as the market price increases.

Rearranging Equation 7.44, we can find the inverse market demand function:

$$\text{Inverse market demand} \qquad p(X) = \bar{p} - \frac{\bar{p}}{\bar{X}}X \tag{7.45}$$

Figure 7.21 Individual and market demand. In panel (a), we present Harriet's demand at different prices per kilogram of fish. On the right, is the *market demand* for fish, which is the sum of ten identical fish buyers' demands for fish (including Harriet). Notice that Harriet's individual demand curve is much steeper than the market demand curve. The change occurs because, for example, for every \$2 decrease in the price Harriet will buy one more unit of fish; for the market as a whole, each of ten people would buy one more unit of the good. As a result the slope of Harriet's demand curve is -2 while the slope of the market demand curve is $-\frac{1}{5}$, or just one-tenth as great.



The inverse market demand curve is linear with a vertical intercept of $\bar{p}$ (the maximum willingness to pay of buyers like Harriet), a horizontal intercept of $\bar{X}$, and a slope of $\frac{\Delta p}{\Delta x} = -\frac{\bar{p}}{\bar{X}}$.

M-NOTE 7.9 Market demand with ten buyers

Let us assume that the fish market is made up of Harriet and nine other buyers who are identical to her (a total of ten buyers). Harriet's quadratic, quasi-linear demand function was:

$$\text{Harriet's demand:} \quad x(p) = 10 - \frac{1}{2}p \tag{7.46}$$

If all the fish buyers are identical to Harriet, then we can sum their demand functions (quantity as a function of price), $x_i(p)$, to get the market demand, $X(p)$. This is the same as multiplying the demand function by the number of people, $n = 10$, to get the market demand function:

$$\begin{aligned} X(p) &= n\left(x_i(p)\right) \\ &= n(10 - \frac{1}{2}p) \\ &= 10 \times \left(10 - \frac{1}{2}p\right) \\ &= 100 - 5p \end{aligned} \tag{7.47}$$

Recall, though, that we typically graph price (p) as a function of quantity (X), or the *inverse* demand function, $p(X)$. We use the market demand curve to find the *inverse* market demand curve with price as a function of quantity by rearranging Equation 7.47 and similarly for Harriet with rearranging Equation 7.46:

$$\text{Harriet's inverse demand:} \quad p(x) = 20 - 2x \tag{7.48}$$

$$\text{Market inverse demand:} \quad p(X) = 20 - \frac{1}{5}X \tag{7.49}$$

Contrasting Equations 7.48 and 7.49 we can see that they have identical vertical intercepts equal to $\bar{p}$, but the slopes of the two functions differ.

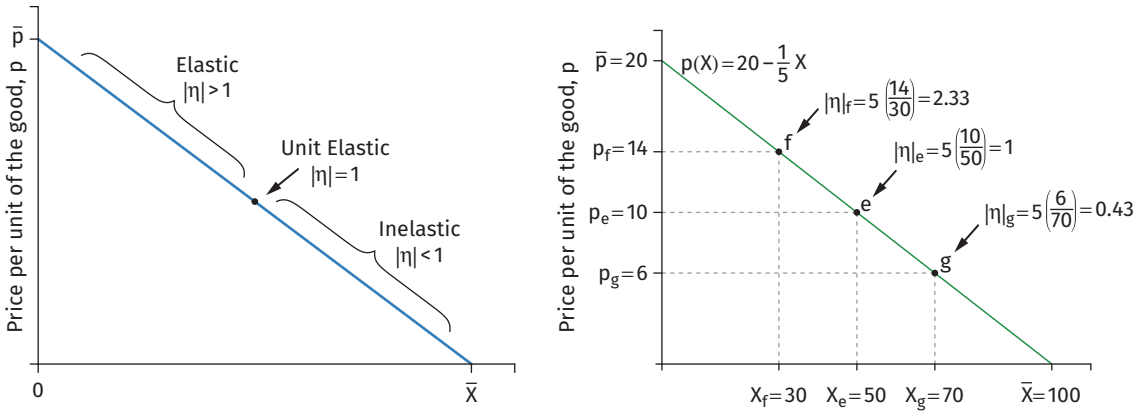
To see why, notice that $x = \frac{X}{n}$, and substituting this expression for x into Harriet's inverse demand (Equation 7.48) we get the equation for market inverse demand (Equation 7.49). This is why the market inverse demand curve has a slope equal to the slope of Harriet's inverse demand namely -2, divided by the number of total buyers (ten), for a slope of $-\frac{1}{5}$ for the market inverse demand curve. The market demand curve is therefore *flatter* than Harriet's relatively *steep* demand curve.

Price elasticity and the slope of the demand curve

For many issues of firm strategy and public policy an important question is: How much does quantity demanded change when there is a change in price or $\frac{\Delta X}{\Delta p}$? This is expressed in two different units: the units of the good (kilos of fish, X) and the monetary unit (dollars, p). But we often need to compare responsiveness *across* commodities–Is the demand for restaurant meals more or less responsive to differences in prices than the demand for motor vehicle fuel?

To allow for comparisons across commodities, we need a measure of responsiveness to price that does *not* depend on the units in which it

Figure 7.22 Price elasticity of demand: general and specific cases. In panel (a), we present the general relationship between the demand curve and the value of price elasticity of demand, η. The figure shows how price elasticity varies from a high value to a low value as you move left to right along the demand curve. In panel (b), we present the market demand curve for ten buyers like Harriet whose preferences are the horizontal sum of Harriet's resulting in a market demand curve of $p(X) = 20 - \frac{1}{5}X$. Consequently, we can calculate three values for price elasticity of demand using the formula of $\eta = \frac{\Delta X}{\Delta p}\frac{p}{X}$. The slope of the curve, $\frac{\Delta p}{\Delta X} = -\frac{1}{5}$; therefore inverting that value we see that $\frac{\Delta X}{\Delta P} = -5$. We can substitute in the values for p and X at each of the price quantity combination to find the value of price elasticity at each of the points **e**, **f**, and **g** as shown in the figure.



(a) General values of price elasticity

(b) Price elasticity for 10 people like Harriet in a market

is measured, whether the quantity demanded is in kilograms of fish or liters of Coca-Cola, whether the price is in yen or euros. We therefore describe the response of market demand to a change in price as the ratio of the *percentage* change in quantity demanded to the *percentage* change in price, $\frac{\Delta X/X}{\Delta p/p}$. This ratio is called the **price elasticity of demand** and is often represented by the Greek letter η (pronounced "ai-ta").

$$\text{Price elasticity of demand} \quad \eta_{Xp} = \frac{\%\text{ Change in quantity}}{\%\text{ Change in price}} = \frac{\Delta X/X}{\Delta p/p}$$

$$\eta_{Xp} = \frac{\Delta X}{\Delta p}\frac{p}{X} \tag{7.50}$$

PRICE ELASTICITY OF DEMAND The price elasticity of demand is the ratio of the percentage change in quantity demanded to the percentage change in price, $\eta_{Xp} = \frac{\Delta X/X}{\Delta p/p}$.

The price elasticity of demand at any point on the demand curve is equal to the slope of the demand curve multiplied by the ratio of price to quantity at that point. The price elasticity of demand falls into three categories:

$|\eta| > 1$ Demand is *price-elastic*, which means that the quantity demanded responds more than proportionally to a change in price.

$|\eta| = 1$ Demand is *unit price-elastic*, which means that the quantity demanded responds exactly proportionally to a change in price.

$|\eta| < 1$ Demand is *price-inelastic*, which means that the quantity demanded responds less than proportionally to a change in price.

M-CHECK The slope of the demand curve is negative, so the elasticity is also a negative number. We usually refer to the absolute value of the elasticity $|\eta|$; so a "larger" elasticity means at any given price and quantity a flatter demand curve.

Slope and price elasticity with a linear demand curve (the QQL case)

As you already know, the slope of the demand curve derived from a quadratic, quasi-linear utility function is constant, and hence a line. But its price elasticity changes as price and quantity change *along* the demand curve:

$$\eta = -\frac{\bar{X}}{\bar{p}}\frac{p}{X} \tag{7.51}$$

The term $\frac{\bar{X}}{\bar{p}}$ is constant, but the term $\frac{p}{X}$ is very large when p is close to $\bar{p}$ and X is close to zero, and goes to zero when p is close to 0 and X is close to $\bar{X}$. It is tempting to think of the price elasticity of demand for a particular good as a single constant number, but in general the price elasticity of demand *changes* with price and quantity demanded.

M-NOTE 7.10 Price elasticity along a linear demand curve

When there are ten people with demand functions as in M-Note 7.9, then we can evaluate the elasticity of demand as follows. Remember the following parameters: $\bar{p} = 20$; $\bar{x} = 10$; $n = 10$, therefore $\bar{X} = n\bar{x} = 100$.

We now evaluate price elasticity of demand at two different (X, p) points. Using Equation 7.51 when price is \$14 and quantity demanded is 30 units (refer to Figure 7.22):

$$\eta = \frac{\Delta X}{\Delta p}\frac{p}{X} = -\frac{\bar{X}}{\bar{p}}\frac{p}{X} = -5\left(\frac{14}{30}\right) = -2.33$$
$$\Rightarrow |\eta| = 2.33$$

Therefore, we would say that for ten people in the fish market, the price elasticity of demand is *elastic* because $|\eta| > 1$, which means that quantity demanded responds *more* than proportionately to a change in prices.

Now consider an alternative point with a lower market price, $p = 6$, and corresponding higher quantity demanded equal to $X(p) = 100 - (5)(6) = 70$. We then calculate the elasticity as follows:

$$\eta = \frac{\Delta X}{\Delta p}\frac{p}{X} = -\frac{\bar{X}}{\bar{p}}\frac{p}{X} = -5\left(\frac{6}{70}\right) = -0.43$$
$$\Rightarrow |\eta| = 0.43$$

continued

✓ **FACT CHECK** The demand for another sugary drink —Mountain Dew—is even *more* price-elastic than Coca-Cola, namely, $|\eta| = 4.39$. Coke and Mountain Dew are close substitutes. The price elasticity of demand for sugary drinks *as a whole* is estimated to be much lower, i.e. $|\eta| = 1.4$. This is because other drinks, like milk or tea, are not really close substitutes for sugary drinks. As a result a price increase for Coke might get you to switch to Mountain Dew, but not to tea or milk.

At the lower price, $p = 6$ (which is on the lower portion of the demand curve), the price elasticity of demand is *inelastic* because $|\eta| < 1$, which means that quantity demanded responds *less* than proportionately to a change in prices.

CHECKPOINT 7.12 The price elasticity of demand

Return to Figure 7.3 about preventative healthcare products and identify the products (and their price ranges) that are most price-*elastic* and most price-*inelastic*.

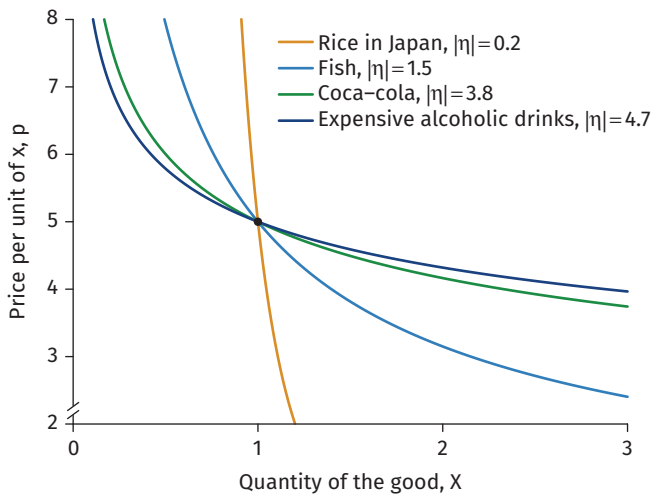
7.12 APPLICATION: EMPIRICAL ESTIMATES OF THE EFFECT OF PRICE ON DEMAND

Why are some goods more price-elastic than others?

Because we can observe price changes and how the quantity purchased changes as a response, we can estimate the price elasticity of demand for various goods. Some estimates are illustrated in Figure 7.23.

Figure 7.23 Comparison of elasticities for different goods. The demand functions shown are called iso-elastic, meaning that, unlike the case for linear demand functions, the price elasticity of demand is the same at every point on the curve. (Recall that iso means equal.) The functions have the form $X(p) = kp^{\eta}$. What the figure shows is that, for example, if the quantity of fish demanded when the price per kilo is 5 is 1 kilo, then if the price fell to 3 per kilo the demand would approximately double (increase from 1 to 2).

Sources: Akino and Hayami (1975); Chomo and Ferrantino (2000); Dhar et al. (2005); Miravete et al. (2020).



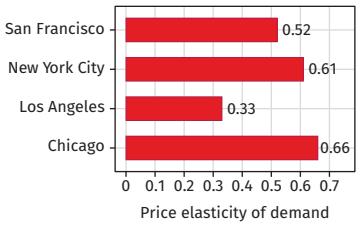
The demand for a good will be highly *inelastic* if it is "something that you cannot do without" and it also does not constitute a large fraction of your budget. Generalizing from this intuition we expect goods to be price-inelastic if:

- there are *few* substitutes for the good in question (e.g. brand loyalty, addiction, or prescription medication);
- it is considered to be a *necessity*, not a luxury (e.g. rice, not expensive liquor);
- it is not a large fraction of your total expenditures (e.g. fish);
- the person making the decision to buy is not the person paying for the purchase (e.g. when a doctor prescribes a drug that the patient will pay for).

Figure 7.24 presents a set of estimates for a single product, Uber rides, in different US cities.

Figure 7.24 Estimates of the price elasticity of demand for Uber rides in four US cities in 2015. The estimates are based on almost 50 million observations on Uber rides from the first 24 weeks of 2015 in Uber's four biggest US markets.

Source: Cohen et al. (2016).



The price elasticity of demand for sugary drinks and the effect of a tax

Obesity and its associated illnesses inflict extraordinary suffering and mounting healthcare costs around the world. Among high-income countries, the US and the UK have especially high rates of obesity, while obesity is relatively rare in Japan and Singapore. Among middle-income countries, Mexico has one of the highest obesity rates.

✓ **FACT CHECK** The fact that obesity rates (BMI > 30) in the US are ten times the rate in Japan suggests that these economic factors—which apply with about equal force in the two countries—cannot be the entire story. Differences in culture and public policies also matter.[8]

Among the contributors to the epidemic rise in obesity rates in recent years, economists have proposed, are two facts:

- as economies shift from farming and manufacturing to services the amount of calories we *use* in a day's work has *declined*, and
- the *cost* of calories, relative to other things we spend our money on, has *fallen*.

Governments around the world have addressed the second economic proposed cause of obesity–reduced cost of calories–by instituting so-called "fat taxes" either to tax the consumption of saturated fats or to tax the consumption of sugar. As of 2019, seven US cities and 34 countries had implemented such policies.

These taxes do *not* aim to increase government revenue. Instead, the government wishes to discourage citizens from consuming the goods because of concerns over the citizenry's health. Similar reasoning applies to "sin taxes," which are taxes on cigarettes and liquor, to discourage excessive consumption of those goods.

Sugary drinks that are commonly taxed in many countries include:

- fruit drinks (including sports and energy drinks);

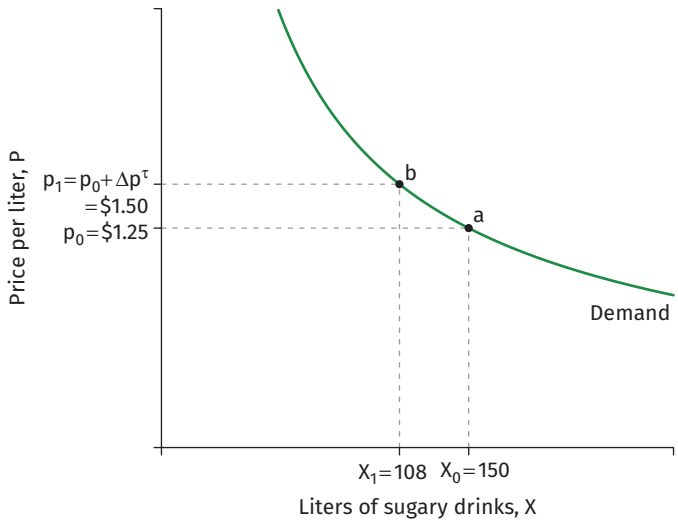
- pre-made coffee and tea (for example, bottled iced coffee and iced tea);
- carbonated and non-carbonated soft drinks (including cocktail mixes and powdered soft drinks).

American adults consumed an average of about 150 liters of these drinks per year during the years 2007-2016.

The demand curve and the price elasticity of demand derived from it provide essential pieces of information to assess the likely effects of the tax on sugary drinks. Because retailers frequently change prices of their drinks it is possible to estimate the price elasticity of demand. To do this, a team of economists recorded sugary drink sales, prices, and a long list of other possible influences on the individual's purchases (including health information and how much they "liked" sweet drinks).

On this basis they estimated that the price elasticity of demand for sugary drinks is about $|\eta| = 1.4$, meaning that a 10 percent increase in the price of the drinks would result in a 14 percent decrease in demand. This estimate–for sugary drinks as a whole–is much *less* than for one particular drink (Coca-Cola) because there are many *other* sugary drinks that are close substitutes for Coca-Cola. Figure 7.25 illustrates what an elasticity of this magnitude implies.

Figure 7.25 The effect of a price increase on the demand for sugary drinks when the price elasticity of demand is $|\eta| = 1.4$. The demand curve shown is iso-elastic with a price elasticity of demand of $|\eta| = 1.4$. There are two prices, the pre-tax price (p_0) and the post-tax price ($p_1 = p_0 + \Delta p$) where $\Delta p = 0.20 p_0$. At the higher price, the consumption of sugary drinks is less, comparing x_0 at point **a** to x_1 at point **b**.



In this example we do not ask what determines the price per liter p_0. Instead we ask the hypothetical question: How many liters would be demanded at various prices? We can study the effect of a sugary drinks tax on the amount consumed by comparing the price per liter without the tax to the amount consumed at the (higher) price when the tax is imposed.

Suppose the price of sugary drinks was initially $1.25 dollars per liter. At this price we can see that the typical person would have demanded about 150 liters. Figure 7.25 depicts this interaction. The starting price and quantity combination are shown by (X_0, p_0). If the effect of the tax was to raise the price of sugary drinks by 20 percent, the price after the imposition of the tax would be 1.50 per liter.

Recall that with a price elasticity of $|\eta| = 1.4$, this means a 10 percent increase in the price will result in a 14 percent decrease in quantity demanded. So, the effect of a 20 percent increase in the price is that the quantity demanded will decrease by 28 percent, down to 108 liters (x_1 in the figure) after the tax.

CHECKPOINT 7.13 Why is demand for sugary drinks price-elastic? Explain why the demand for sugary drinks as a *whole* in the US is *less* price-elastic than the demand for Coca-Cola or Mountain Dew.

7.13 CONSUMER SURPLUS AND INTERPERSONAL COMPARISONS OF UTILITY

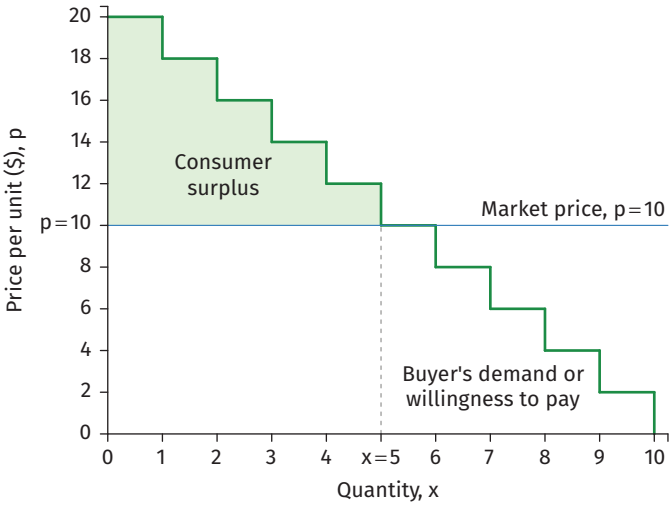
An individual's consumer surplus

When a person, call her Harriet as we did earlier, buys a good, she does so because she expects to derive a benefit that exceeds the price of the good. The difference between the most she would be willing to pay for the good and what she actually pays for it is called the **consumer surplus** that she received as a result of that purchase. Because it is measured as the *difference* between the maximum willingness to pay in money and the money that is paid, consumer surplus gives a measure in *monetary* terms of the benefits (or "*consumer welfare*") that a person derives from a purchase of a good. If we consider *not* buying the good as Harriet's fallback position, then her consumer surplus is an economic rent, namely a measure of what she gets above and beyond her next best alternative.

Figure 7.26 illustrates the consumer surplus available to Harriet by her purchases of five units of good x. The maximum she would pay for the first unit—think: access to a workout at the gym during a week, a film on Hulu—

CONSUMER SURPLUS The consumer's willingness to pay for a good minus the price at which the consumer bought the good, often summed across all units sold.

Figure 7.26 Harriet's willingness to pay and consumer surplus. The height of the steps in the step function is the maximum Harriet would pay to have the first, second, third, and so on unit of the good. For each unit she buys, she pays \$10. Her consumer surplus is the vertical distance between the price line at $p = 10$ and her willingness to pay for each additional unit, summed over the number of units she purchases. Her utility-maximizing consumption bundle is $x = 5$, when her willingness to pay equals the price: $mrs(x) = p = 10$. Summing over the bars, she receives consumer surplus, $CS = 10 + 8 + 6 + 4 + 2 = 30$.



HISTORY Jules Dupuit (du-PWEE) (1804–1866) was, according to Joseph Schumpeter, one of those "brilliant French engineers in the public service who contributed . . . so substantially to scientific economics." His studies of the maximum tolls that could be charged for travel over a bridge led him to discover not only the concept of marginal utility but also the idea of consumer surplus as the area under a demand curve and above a uniform price charged, exactly as is shown in Figure 7.26. He used these concepts to explore both price discrimination and the loss in social welfare associated with charging tolls on bridges. He has not to this day received much credit among economists for these discoveries (most think that consumer surplus was Alfred Marshall's idea), perhaps not surprisingly seeing that Dupuit's major paper on this subject was published in a journal whose title roughly translated is the *Chronicle of Bridges and Roadways*.[9]

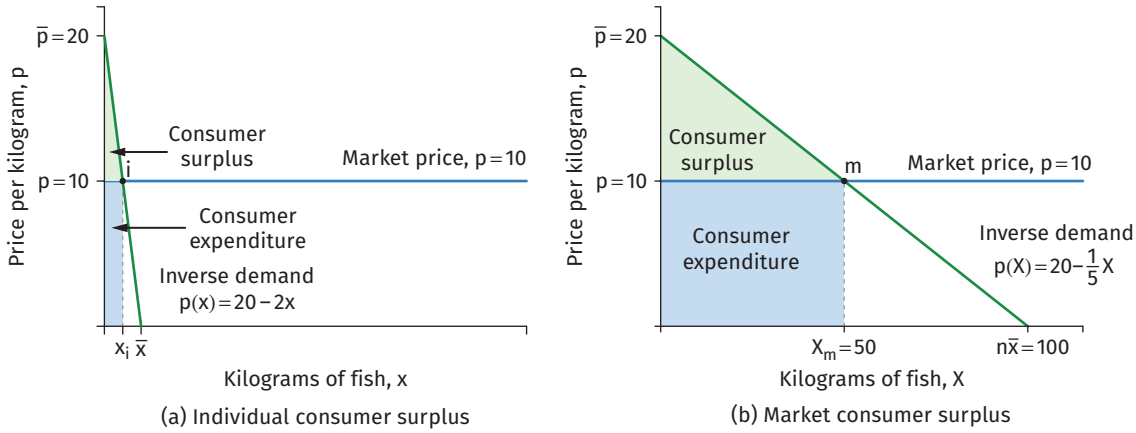
is \$20. But each successive unit is worth less to her, and if she already has four–workouts, films–the most she would pay for the fifth is \$12. The sixth would not be worth more than she'd pay for it. So she will purchase five units. Adding up her willingness to pay for each and subtracting what she actually paid–\$10 in each case she has a consumer surplus of \$30 (that's $10 + 8 + 6 + 4 + 2 = 30$).

The graph of her willingness to pay–called a step function–is her demand curve. When we think of people's purchase over a longer period of time, or the purchases of many people, we smooth out the "steps" and make a smooth curve (not necessarily a straight line).

Adding up consumer surplus for many people

We are often interested in a measure of how much people as a *group* benefit from the opportunity they have to purchase some good. A natural way to do this is to add up the consumer surplus enjoyed by each buyer. For this to make sense it must be that a dollar's worth of consumer surplus is as valuable to one person as to another. Unless we assumed this, we could *not* add the consumer surplus of one person–some dollar amount–to some other person's consumer surplus.

Figure 7.27 Individual and market consumer surplus. In panel (a) we show the utility-maximizing kilograms of fish Harriet (an individual buyer) buys and the consumer surplus she derives as a consequence. Her expenditure is the price she paid multiplied by the number of units she bought. In panel (b) we show the *market demand* for fish with the *market consumer surplus*.



There are really two parts of this key assumption:

- *We can make interpersonal comparisons of utility*: we can compare one person's well-being (or utility) with another. Recall from Chapter 3 that this means that we consider utility to be a cardinal measure that can be compared across people (like for height, *how much taller* is Simon than Harriet) rather than an ordinal measure (Simon is taller than Harriet).
- *The marginal utility of money left over for other purchases is the same to all people*: an additional dollar makes the same contribution to the well-being of one person as to another. This means that what Harriet would purchase with a dollar of money left over after spending on workouts contributes as much to her well-being as a dollar's worth of additional expenditure by a less fortunate person.

This would almost certainly *not* be true if one of them were very poor, so that a dollar would be worth a lot to them (it would be used to purchase food, or other essentials), and the other was very rich (the additional dollar would be spent on a luxury good).

To see why an individual's consumer surplus may not be something you can add up across many people of differing levels of income or wealth, consider something pleasurable, like a serving of excellent ice cream that is placed before you. The extent of enjoyment that eating it would give you—its taste and other reasons for enjoying it—would not be affected, if instead of your current income, you had half that amount.

But your willingness to pay for it would be less in the second case not because it became less tasty but because in the second case your marginal

utility of money left over would be higher. Remember the willingness to pay—the marginal rate of substitution—is the marginal utility of the ice cream divided by the marginal utility of money left over. If you were poorer, the marginal utility of money left over would be higher, and hence your willingness to pay would be lower.

Willingness to pay is therefore *not* a measure of how intrinsically "good" something is, even to a single person. So when we add up consumer surpluses across people we must be considering cases in which they are of approximately the same level of income.

In the Figure 7.27 (a), the *individual* consumer surplus (Equation 7.52) is the area of the light-green triangle above the actual price and below the demand (maximum willingness to pay) curve for Harriet. Consuming $x = 5$ units of fish provides Harriet with consumer surplus of \$25. This is exactly the same as if we had substituted Harriet's values for $\bar{x}$, $\bar{p}$, p and x into Equation 7.52.

The consumer surplus for *all buyers* is the area shaded in green in the Figure 7.27 (b) and because we have assumed all buyers are identical this is exactly $n = 10$ multiplied by the individual consumer surplus for Harriet.

M-NOTE 7.11 Consumer surplus with quadratic, quasi-linear utility

The demand curve based on quadratic, quasi-linear preferences is linear, as shown in Figure 7.27. Therefore, we can calculate the value of the consumer surplus—the green shaded triangle in panel (**a**) of that figure—using the following three data points:

- The person's maximum willingness to pay is $\bar{p} = 20$.
- The person actually pays the price $p = 10 < \bar{p}$.
- The total purchases, x_i = 5 for the individual and $10x_i = 50$ for the market.

Therefore, consumer surplus for the individual and the market is given by:

$$\text{Individual} \quad cs(x) = \frac{1}{2}(\bar{p} - p)x \tag{7.52}$$

$$\text{Market} \quad CS(X) = \left(\frac{1}{2}(\bar{p} - p)x\right) \times n$$

$$= \frac{1}{2}(\bar{p} - p)X \tag{7.53}$$

EXAMPLE To measure the utility gained by making a purchase and to sum this across individuals we need a measure of utility that is similar to money. If one person has \$1,000 and someone else \$100 we can say that the first person has ten times as much money as the second, irrespective of whether we measure their wealth in dollars, or in pennies.

CHECKPOINT 7.14 Individual and market consumer surplus Using the numerical examples in Figure 7.27, what is the consumer surplus:

a. for the individual?

b. for the market?

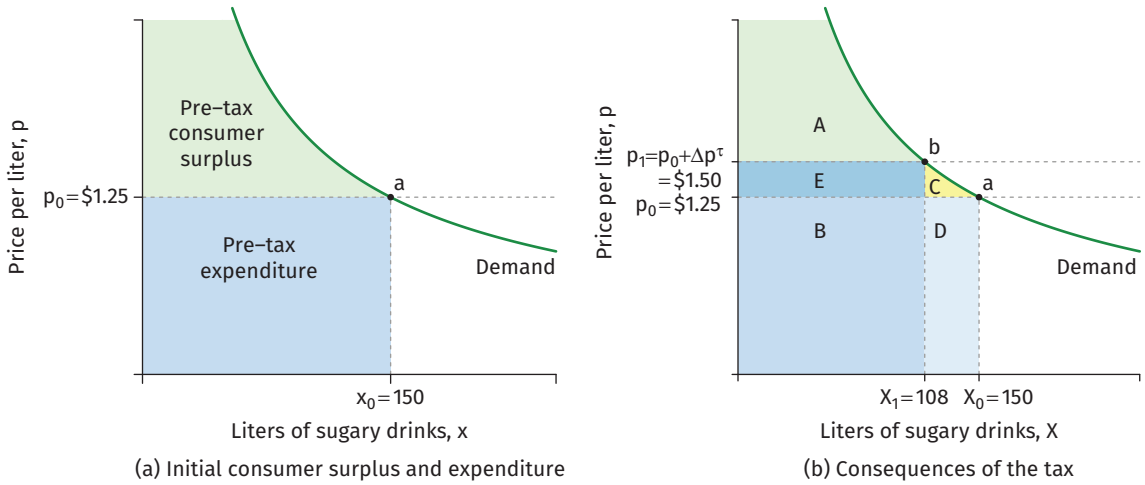
7.14 APPLICATION: THE EFFECT OF A SUGAR TAX ON CONSUMER SURPLUS

We showed earlier, in Figure 7.25 when discussing price elasticity of demand, that an *increase* in prices resulting from a tax will *decrease* quantity demanded by people. We now analyze the consequences of that tax for people's utility.

Figure 7.28 shows the consequences of the tax for consumer welfare (measured in terms of prices and quantities consumed). Figure 7.28 (a) shows how, at the initial price, consumer surplus is given by the combined area in green, shown by the total area A + C + E in Figure 7.28 (b).

After the tax at the new higher price, consumers will be left with consumer surplus equal to area A. Consumers will lose consumer surplus indicated by the area E + C. Area B is the portion of consumer expenditure that is unchanged by the tax. Consumer expenditure will decrease by the

Figure 7.28 Effects of a tax on consumer surplus and expenditure. There are two prices: the pretax price (p_0) and the posttax price (p_1). As the price increases, the consumption of sugary drinks decreases from X_0 at point **a** to X_1 at point **b**. In panel (a) the area shaded in green below the demand curve and above the price, p_0, is the pretax consumer surplus. The area shaded in blue is the pretax expenditure, equal to the price multiplied by the quantity people consumed, i.e. p_0X_0. In panel (b), area **A** is the consumer surplus that remains after the imposition of the tax. Area **B** is the expenditure that was common before and after the imposition of the tax. Area **D** is the decrease in expenditure by people who are unwilling to purchase sugary drinks at the higher price, p_1. Area **E** was consumer surplus before the tax, but is now part of expenditure. Area **C** plus area **E** is the *decrease* in consumer surplus as a consequence of the tax.



area D, but increase by the area E. That is, before the tax, consumers spent areas B + D. After the tax, consumers spend areas B + E.

In Chapter 8 we will return to the sugary drinks tax, looking at its impact on others, including firms' owners who will lose economic profits as a result.

Is it fair? Sugary drink taxes are regressive

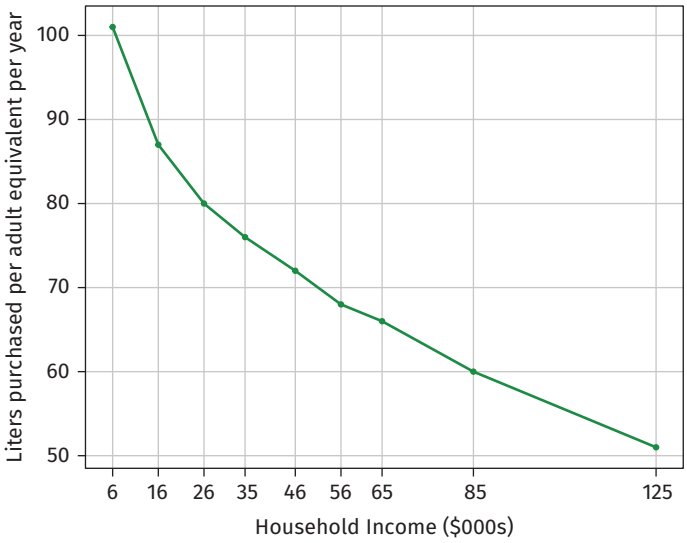
In 2017 voters in Santa Fe, the capital of the US state of New Mexico, voted overwhelmingly to reject a proposed tax on sugary drinks. The measure had been put forward by a popular mayor and would have directed the resulting revenue toward expanding preschool educational opportunities for the less well-off. It was opposed by the American Beverage Association.

Opponents of the measure held that the tax unfairly placed a burden on the less well-off. To address the potential unfairness of the tax the Santa Fe advocates of the sugar tax had linked the measure to the provision of a particular public service that was very much in demand among lower-income Santa Feans. But the very real substantial negative income effect apparently outweighed the promise of better educational opportunities.[10]

Figure 7.29 provides evidence about the consumption of sugary drinks in households of differing incomes based on matched data on purchases of

Figure 7.29 Consumption of sugary drinks in households of differing incomes. The figure shows the amount of sugary drinks purchased per 'adult equivalent' per year by income (measured in thousands of dollars in 2015 prices) from a panel data survey of 18,159 US households in 2017. The term 'adult equivalent' means that children in the households have been counted as some fraction of an adult.

Source: Allcott et al. (2019).



sugary drinks and household income. Notice the following pattern in the figure:

- households with *lower* incomes consume *larger* quantities of sugary drinks;
- households with *higher* incomes consume *smaller* quantities of sugary drinks.

As a result, a per-unit tax on sugary drinks is *regressive*: poorer households will pay more as a share of their household income. At the same time, however, decreasing the quantity of sugary drinks consumed by members of those households could be quite beneficial for health and for medical costs that those households incur.

We will return to the analysis of a sugary drinks tax in the next chapter, taking account of the consumer surplus that people lose, the profits owners lose, and the benefits made possible by the revenues the government raises.

CHECKPOINT 7.15 Policies to mitigate the income losses of less well-off people imposed by the regressive sugary drink tax The citizen's dividend—returning the tax revenues collected to citizens as an equal lump-sum payment to each family—is proposed as a way to counteract the regressive nature of the carbon tax. Explain why something similar would *not* accomplish this purpose in the case of the sugary drinks tax.

Experiences around the world of "fat taxes"

Denmark instituted a per-kilogram tax on saturated fats in 2011. Hungary introduced both sugar and fat taxes in 2011, where the percentage of the tax is proportional to the amount of sugar or fat in the good. In 2012 France introduced a tax on both added-sugar and artificially sweetened drinks of € 0.075 per liter (in 2015). Chile adopted a tax in 2015. In the US, several states and cities have implemented soda taxes, such as the tax implemented in San Francisco, CA in 2014. The aim of the tax is not primarily to obtain tax revenues but to *reduce* consumption of the offending foods so as to improve individuals' health and to reduce the cost burden of healthcare provision, including by the government. What has happened as a consequence of these taxes? Did the taxes achieve the government's aims?

Results from a study in Mexico showed that a 6 peso-per-liter tax on beverages with added sugar reduced the quantity demanded by between 6 percent and 12 percent over the year of the study (2014). Consumption decreased more among low-income families with the proportional decrease being between 9 percent and 17 percent over the year. The evidence also suggests that consumers switched to close–untaxed–substitutes that did not contain added sugar such as diet sodas, 100 percent fruit juices, and sparkling and plain water (with between 7 percent and 13 percent increases in these categories).[11]

In Hungary, the tax has had several effects. The tax has reduced consumption; it has also caused firms to change the recipes of their food items, a sensible response because the tax is proportional to the amount of the sugar or saturated fat the food item contains.[12]

Denmark's case is more complicated. The "fat tax" definitely reduced consumption of butter, margarine, and similar products, by 10 to 15 percent. People also changed their buying habits in terms of *where* they bought their butter and margarine: they switched to buying at discount stores. But, because these stores were aware of these buyer responses and the resulting positive shift in the demand curves they faced, they increased their prices on butter and fatty products more than high-end supermarkets did.[13]

The tax was unpopular in Denmark and was eventually repealed. Why? People had been crossing the nearby Swedish and German borders to do their shopping: one study showed almost half of Danish shoppers had gone across a border to avoid the tax.

These results illustrate the complexity of tax policy when the goal is to reduce consumption of a good. But, in places like Mexico, Denmark, and Hungary, we've seen significant and important decreases in the consumption of sugary drinks and fatty foods. In Mexico, particularly, this is important for many poor people who are disproportionately affected by health problems caused by high sugar consumption, especially when they cannot afford proper treatment of cardio-vascular diseases or obesity.

CHECKPOINT 7.16 Salt taxes and sin taxes: Putting the elasticity of demand to work To think through these issues, consider the following questions:

a. Centuries ago in China, France, and the British colony of India the salt tax was one of the major sources of government revenue. What is it about salt that made this tax a good way of raising revenue?

b. Explain why sin taxes levied on goods with price-elastic demand (e.g. alcoholic spirits) will be effective in changing peoples behavior, but not in raising revenue, while the opposite is true for goods with inelastic demand (e.g. cigarettes).

7.15 APPLICATION: WILLINGNESS TO PAY (FOR AN INTEGRATED NEIGHBORHOOD)

In Chapter 1 (section 1.15), we illustrated the idea of a Nash equilibrium and the process by which a group of people might arrive at such an outcome by the buying and selling of homes among members of two groups, "Blues" and "Greens." We showed the following:

- *Complete segregation:* The equilibrium composition of the neighborhood—one in which *none* of the residents wanted to switch their location and were able to do so—could be complete "segregation" of the Blues and Greens, even though everyone preferred an integrated outcome.
- *Multiple equilibria:* There was also an equilibrium in which the neighborhood was integrated.
- *Path dependence*: Which of the multiple equilibria—segregated or integrated—would be realized was *path dependent* (like whether the farmers in Palanpur planted early or late). Which equilibrium occurred depended on the recent history of the neighborhood.

These three characteristics—a Pareto-inefficient equilibrium, multiple equilibria, and path dependence—will also appear in the model we now introduce. But here we explicitly introduce a market in homes and people's willingness to pay.

In Milwaukee, Los Angeles, and Cincinnati towards the end of the last century over half of white residents, when asked, said they would prefer to live in a neighborhood in which 20 percent or more of their coresidents were African-American (one in five preferring equal numbers of each).[14] But few residents of these cities lived in integrated neighborhoods. They were asked about their preferences as part of a legal case concerning housing segregation in these and other cities. Most African Americans preferred fifty-fifty neighborhoods.

There are many reasons why members of a society might *not* want their residential communities to be highly segregated. Segregated living leads to racially segregated schools, friendships, and other social networks. Because group members would then be unlikely to have friends in the other group, segregated living could encourage group stereotypes and intolerance leading to conflicts between groups.

The respondents in the above surveys may have misrepresented their preferences, of course, but those sincerely seeking integrated neighborhoods would have been disappointed. The housing market in these cities produced few mixed white-African American neighborhoods even though these were apparently in substantial demand.

In Los Angeles, for example, virtually all white residents (more than 90 percent) lived in neighborhoods with fewer than 10 percent African-American residents, while 70 percent of African Americans lived in neighborhoods with fewer than 20 percent whites. Why was the result at the neighborhood level so seemingly at odds with the distribution of preferences of the individuals making up the neighborhoods? Imagine your surprise had we reported that one in five wanted a back-yard swimming pool and were prepared to pay the price for a pool, yet almost none had pools.

Figure 7.30 Segregation by income level in housing in the city of Santa Fe, near Mexico City.

Photo by Johnny Miller/Unequal Scenes.

✓ **FACT CHECK** If we were constructing a model to develop policies to ensure more integrated neighborhoods—rather than a teaching device to illustrate willingness to pay—we would introduce a great many elements that are essential to understanding segregation as a social and historical fact, and are not in our model. Included would be the role of policies of governments and banks that reinforce segregation. Also included would be a diversity in preferences including the desire on the part of some people to live in a homogeneous neighborhood of their own group, instead of our assumption here that people favor substantially integrated over totally segregated neighborhoods.

! **REMINDER** For segregation in New York City, have a look at Figure 1.18.

Why does willingness to pay get you a pool if you want one, but not an integrated neighborhood? To answer these questions we need an explanation of how highly segregated neighborhoods result, even if preferences were such that members of all groups would be better off with greater integration. We need to understand why the housing market produces a Pareto-inefficient level of segregation.

Residential segregation is the result of many aspects of how credit and housing markets work, and these differ across countries and even within the US among cities and states. Included are attempts by some banks, real estate sales people, and neighborhood residents to prevent integration. But there is another less obvious and perfectly legal way that segregated neighborhoods are sustained, even if most people preferred a more integrated community.

Willingness to pay for integration or segregation

HISTORY This way of thinking about segregation in residential neighborhoods was developed by Thomas Schelling, a Nobel Laureate in economics. You can run a computer simulation of how a population may segregate itself, even with very modest preferences for segregation, here (ncase.me/polygons/).

We will explain why this is true by modeling a single neighborhood (one of many in a large city) in which, when considered in isolation, all houses are equally desirable to all members of the population (they're identical). Peoples' preferences for living in this neighborhood depend solely on the racial composition of the neighborhood.

As before, Greens and Blues are two population groups that are equally numerous in the city. Greens prefer to live in a mixed neighborhood with slightly more Greens than Blues, and Blues correspondingly do not prefer segregation, but prefer a neighborhood with somewhat more Blues than Greens.

We have *normalized* the size of the neighborhood, setting it equal to 1, so we can refer to the fraction of Greens by g, which can vary from 0 to 1. For example, $g = 0.3$ means that the neighborhood is 30 percent Green and 70 percent Blue.

We will express the preferences of the Greens and the Blues by the maximum prices, p^G for Greens and p^B for Blues, they would be *willing to pay* for a house in the neighborhood, each depending on the *fraction* of homes in the neighborhood occupied by Greens g. The following willingness to pay (WTP) equations are a way to express the preferences described above:

$$\text{Blues' WTP} \qquad p^B(g) = \frac{1}{2}(g+\delta) - \frac{1}{2}(g+\delta)^2 + p \tag{7.54}$$

$$\text{Greens' WTP} \qquad p^G(g) = \frac{1}{2}(g-\delta) - \frac{1}{2}(g-\delta)^2 + p \tag{7.55}$$

where p is a positive constant reflecting the intrinsic value of the identical homes. Figure 7.31 shows the willingness to pay equation for the Greens as follows:

- a *low* willingness to pay for a house in an all-Blue neighborhood ($p^G(g = 0)$);

- a *higher* willingness to pay for a house in an all-Green neighborhood ($p^G(g = 1)$);
- the *greatest* willingness to pay in an *integrated* but Green-majority neighborhood (with 60 percent Greens)

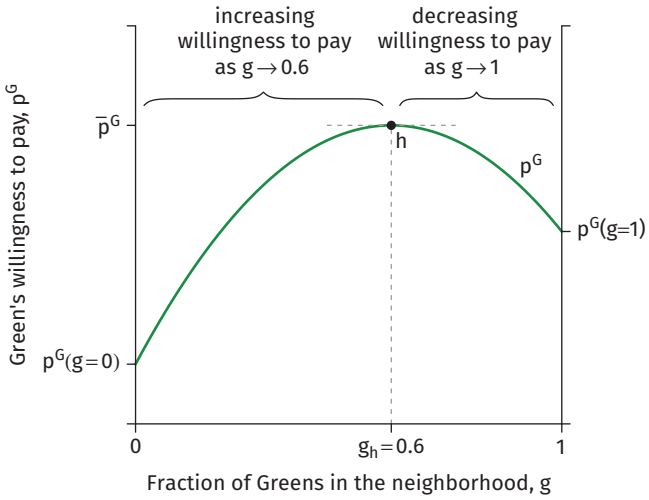
The term δ is a measure of the preference for segregation. We assume that Greens and Blues have similar preferences to live with their own group members, so δ is the same for the two groups.

To see how δ measures the degree of preferences for segregation, think about what the ideal neighborhood for a Green and for a Blue would be. If the ideal neighborhoods of members of the two groups are very *different*, then segregationist preferences are *strong*. Because the willingness to pay for a home—and therefore its value—depends on the composition of the neighborhood, the ideal neighborhood has a group composition that *maximizes* the value of owning a house in the neighborhood (or, what is the same thing, that *maximizes* willingness to pay for a home there).

What would each group of persons' ideal neighborhood look like? (M-Note 7.12 explains how these are derived.)

- **Greens:** The ideal neighborhood for greens (that which maximizes p^G) is composed of $g = \frac{1}{2} + \delta$ percent greens.
- **Blues**: Blues prefer an ideal neighborhood with $g = \frac{1}{2} - \delta$.

Figure 7.31 An illustration of the willingness to pay of Greens, p^G. Their willingness to pay reaches a maximum at point **h** where the proportion of greens $g_h = 0.6$. Between $g = 0$ and $g = 0.6$, Greens' willingness to pay is *increasing* as they move from being a minority to become a slight majority at 60 percent of the population. Between $g = 0.6$ and $g = 1$, Greens' willingness to pay is *decreasing* as they move from being a slight majority at 60 percent of the neighborhood to being 100 percent of the neighborhood. For this and the next figure we used $\delta = 0.1$.



As the difference between the ideal neighborhoods (that for which they would pay the highest price of a home) of the Greens and the Blues is 2δ we will refer to δ as the preference for segregation of the two groups (δ could differ between the two groups, or one group might not care about the racial composition at all, of course).

The willingness-to-pay curves and the degree of preferred segregation they express provide the essential building blocks for understanding how the housing market will work. But to put that information to work we need to turn to how the market will change or not depending on its composition.

This means we need to identify the Nash equilibria of the market (where there would be no forces changing the situation) and the points that are not Nash equilibria, in which people could do better by buying or selling a house in a way that changes the composition of the neighborhood. This is called an analysis of market *dynamics*, that is, how markets *change*.

! REMINDER We discussed *dynamics* in Chapter 5 when exploring how the fishermen reached the Nash equilibrium by comparing their marginal benefits and marginal costs.

M-NOTE 7.12 **Finding the preferred proportions**

We would like to find the proportion of Greens in the neighborhood that would maximize each group's willingness to pay. In Figure 7.31 this is 60 percent for the Greens at g_h. To see how we got this number we differentiate Equations 7.54 and 7.55, with respect to g.

$$\text{Greens:} \quad p_g^G \equiv \frac{dp^G}{dg} = \frac{1}{2} - (g - \delta) \tag{7.56}$$

$$\text{Blues:} \quad p_g^B \equiv \frac{dp^B}{dg} = \frac{1}{2} - (g + \delta) \tag{7.57}$$

Now, to find the g that maximizes the house value for the two groups, we set each of Equations 7.56 and 7.57 equal to zero and isolate g:

$$\text{Greens:} \quad g_{max}^G = \frac{1}{2} + \delta \tag{7.58}$$

$$\text{Blues:} \quad g_{max}^B = \frac{1}{2} - \delta \tag{7.59}$$

As can be seen, on either side of $g = 0.5$ lie the two preferred proportions of Greens for the two groups. They are separated by 2δ, that is, twice the degree of preferences for segregation.

Note: the specific functions shown in Equations 7.54 and 7.55 were chosen because they are easy to manipulate, they give an inverted U-shaped curve, and yield a transparent measure of segregation preferences, namely δ.

CHECKPOINT 7.17 How much difference does neighborhood composition make in the willingness to pay? In the willingness-to-pay Equations (7.54 and 7.55) let p = 100 and calculate

a. If the neighborhood is composed of 60 percent Greens (the Greens' ideal neighborhood) give the value of a home for a Green and a Blue.

b. The value of a home for a Green and a Blue in a neighborhood composed of an equal number of the two groups.

7.16 APPLICATION: MARKET DYNAMICS AND SEGREGATION

Remember, an equilibrium is defined by the *absence* of change. So to determine what level of integration or segregation we would expect to observe (the equilibrium) we need to better understand the *process* by which the neighborhood composition will change as a result of the dynamics of the market.

Home sales: A pathway to segregation

To do this, we now consider the conditions under which a house inhabited by a Green might be sold to a "Blue family," or vice versa. Imagine that prospective buyers from outside the neighborhood visit the neighborhood and just knock on the door of a randomly selected house. A sale takes place if the house is worth more to the visitor than it is to its current owner. If the current owner values it much more highly, no sale takes place. So houses never change hands among the same groups (because they value the houses identically). Therefore:

- *A Blue sells to a Green:* If a Green visits the house of a Blue, a sale will take place if $p^G > p^B$ and not if $p^G \leq p^B$.
- *A Green sells to a Blue:* If a Blue visits the house of a Green, a sale will take place if $p^B > p^G$ and not if $p^B \leq p^G$.

EXAMPLE This way of representing how the housing market works is an example of a market model in which people are paired and if there is a mutually beneficial trade that can be made, they do it. An equilibrium is a situation in which there are no more mutually beneficial trades to be done. Another example is our bargaining model of general equilibrium in Chapter 14.

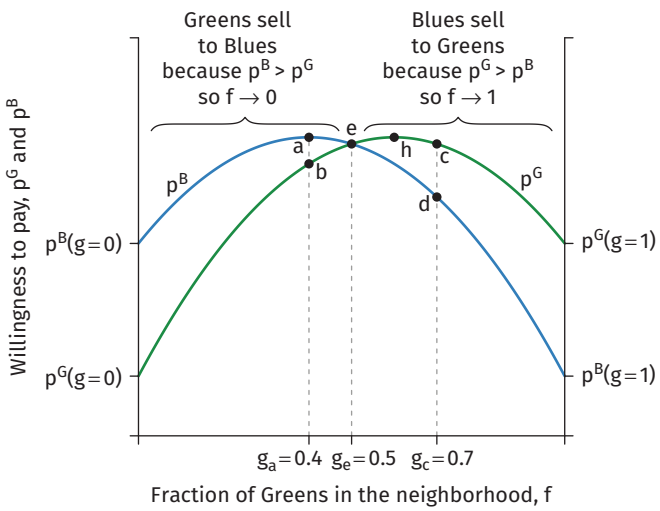
Remember that the homes are identical. While the residents in the neighborhood care about the composition of the neighborhood, they are "colorblind" when it comes to buying or selling houses: they sell if they are offered a price *above* what their home is worth to them, irrespective of the color of the buyer.

Figure 7.32 illustrates this. To check that you understand how it works, imagine that the fraction of Greens in the neighborhood at the moment is a bit greater than g_h (such as at $g_c = 0.7$). What would you anticipate happening? Check the values that Greens and Blues would then place on their homes: Greens would value a home in the neighborhood *more* than a Blue would value living in that home.

So a visiting Green would be willing to pay more for a Blue's home than the home was worth to the Blue. A transaction would take place—a home once occupied by a Blue would now be a "Green home." What would happen when the next Green showed up? The same thing. When would it stop? When there were no "Blue houses" left. The neighborhood would be entirely Green.

Of course had the initial value of g been a bit below g_h the process would have run in *reverse*, and we would have ended up with an entirely Blue neighborhood. Complete segregation would be the result in either case. The all-Green segregated neighborhood is a Nash equilibrium because there is

Figure 7.32 Dynamics of residential segregation. At **e**, the residents are at an unstable equilibrium. If the history of the neighborhood is such that it starts off at $g = 0.4$, corresponding to points **a** and **b** on the Green and Blue price curves, then, as a result of $p^B > p^G$, Greens will sell to Blues and the Greens will leave the neighborhood. This will continue until there are no Greens left. The neighborhood will then remain at $g = 0$, that is, no Greens in the neighborhood. With $g = 0$ a house is worth more to a Blue than to a Green, but there are no "Green houses" left for a Blue to buy. A similar process, but in the opposite direction works if in some period $g = 0.7$, then, as a result of $p^G > p^B$, Blues will sell to Greens and Blues will leave the neighborhood until $g = 1$, that is, no Blues in the neighborhood.



no incentive for a Blue to buy in the all-Green neighborhood given how much homes there are valued by the Greens living there and therefore the minimum price at which they would be willing to sell.

Think about it this way: individual neighborhoods in a city containing Green and Blue people will have neighborhoods of Greens and neighborhoods of Blues that are totally segregated. We could say that each neighborhood is locally homogeneous, but the city is composed of neighborhoods of the different groups and is therefore globally heterogeneous.

Which composition any particular neighborhood will exhibit will depend on history: if, in the recent past, g was less than $g = \frac{1}{2}$, we would expect to find $g = 0$. If g had been greater than $\frac{1}{2}$, we would expect to find $g = 1$. In other words we would expect to see a completely segregated neighborhood, one way or the other.

! REMINDER An unstable equilibrium at the boundary between two regions characterized by distinct movements in some variable is called a tipping point. Recall that there was a tipping point in the Palanpur game about about planting crops early or late (Chapter 2).

There is one composition of the neighborhood that is both integrated and a Nash equilibrium, namely $g = \frac{1}{2}$. At that composition the values of the homes to Greens and Blues do not differ, so no sales would take place. But we would not expect to see such an outcome. The reason is that if for *any* reason the composition moved a bit higher or lower, then the dynamic of buying and selling that propelled the neighborhood to complete segregation above would begin, with the same outcome of a completely segregated neighborhood. The composition $g = \frac{1}{2}$ is therefore a Nash equilibrium but it is not **stable**: a small movement above or below it will set in motion further moves *away* from the equilibrium.

M-CHECK The outcome $g = \frac{1}{2}$ need not even be possible, as would be the case if there were an odd number of homes!

Segregation: A coordination failure

A fully segregated neighborhood (either $g = 1$ or $g = 0$) is *not* the mutually preferred outcome. You can see from point **e** in Figure 7.32 that both Greens and Blues prefer a neighborhood with an equal number of each group to a neighborhood in which they are the only group living there. Greens and Blues alike would *prefer* to live in the *other* group's ideal neighborhood than to live in their own segregated neighborhood. All of the points **a** through **e** in the figure are Pareto superior to *either* of the complete segregation outcomes.

To understand why this Pareto-inefficient Nash equilibrium occurs, set aside the Greens and Blues for a moment and think about what kind of economic entity a neighborhood is. Take the age composition of a neighborhood, whether mostly young people live there, or mostly middle-aged or elderly, or mixed, and so on.

The age composition of the neighborhood has some of the aspects of a *public good*. It is *non-rival* because a person visiting the neighborhood and experiencing its composition does not subtract the experience of its composition from anyone else. It is *non-excludable* because the change applies to everyone in the neighborhood. The sale of a home—say from a middle-aged family to a young family—also has external effects on those not involved in the sale because it changes the age composition of the neighborhood for everyone.

At the beginning of this section we asked: Why does willingness to pay get you a swimming pool if you want one, but not an integrated neighborhood? The answer, referring back to the distinctions we made in Chapter 5, is that:

STABLE EQUILIBRIUM An equilibrium is stable if a sufficiently small displacement away from the equilibrium is self-correcting (leading to movement back toward the equilibrium).

- the pool is a *private good*, for which your willingness to pay provides sufficient motivation for some person or business to make the pool available;
- the composition of the neighborhood is a *public good*, resulting from the Nash equilibrium of a large number of people's actions; without coordination among the actors no person's willingness to pay for a more integrated neighborhood will result in the desired public good being realized.

The coordination failure represented by the completely segregated Nash equilibria does *not* occur because people do *not* want to live in integrated neighborhoods. To see this, imagine that the δ that we used in making the above figures had been 0.01 rather than 0.10. This means that the ideal composition of the neighborhood would have been virtually identical for the Blues and the Greens. But the result of the market dynamics would have been the same: total segregation.

This occurs because buying or selling a home has an *external effect* that buyers and sellers do not take into account. To see why, consider the neighborhood with composition g_c in Figure 7.32. With 70 percent Greens, both Greens and Blues in the neighborhood would prefer a "Bluer" composition. But, as you can see by comparing points **c** and **d**, homes in that neighborhood are worth *more* to Greens than to Blues. So Blues will sell to Greens, and the neighborhood will *not* become more Blue (as both Green and Blue residents prefer). It will move in the other direction, becoming Greener. As a consequence of exchanges, all of which are voluntarily entered into, the value of the housing in the entire neighborhood falls: *everyone* is worse off except for the two who make the sale and purchase.

Like the coordination problem of overfishing, or planting late in Palanpur this undesired outcome occurs because the residents are engaged in a non-cooperative game. They have no way of jointly agreeing on a neighborhood composition that they would all prefer to the segregated outcome.

Achieving a desirable and enduring neighborhood composition is a challenge that is not readily addressed by the measures we considered, for example, in the case of fishing.

CHECKPOINT 7.18 Self-correcting segregation?

a. Be sure you can explain why complete segregation is *stable*, that is, self-correcting.

b. Imagine all of the houses in the neighborhood were owned by a single person who could rent the houses for a monthly fee equal to some fixed fraction of the value of the house. What value of g would the owner implement, assuming that it was legal and otherwise acceptable for the owner to consider the group of the renter (Blue or Green) in offering a rental.

7.17 CONCLUSION

The ideas you have learned—especially the price elasticity of demand and the distinction between regressive and progressive taxation—are essential tools for understanding the effects of taxing sugary drinks, carbon emission, and other policies. We have learned, for example, that raising the price of a good with a large price elasticity of demand—such as sugary drinks—can have a substantial effect on how much people buy. Correspondingly raising the price of a good with a small price elasticity of demand—rice, medical care—will have more modest effects on how much people buy.

But if we are interested in public policy—such as sin taxes—then we need to know how the tax will affect the price. And this depends on the costs of producing the good, and how these vary with the amount of the good produced. In Chapter 6 we explained how the owners of a firm will select a mix of inputs of labor and capital goods or other inputs to minimize the cost of producing any given amount of their product. In the next chapter we turn to how costs of producing a good depend on how much of the good is produced, as described by cost curves.

MAKING CONNECTIONS

Feasible sets and indifference curves: These are the basis of our using constrained optimization to analyze the consumer's willingness to pay and resulting demand curves.

Social preferences: Our analysis of consumption as a social activity and Veblen effects takes account of the fact that people care about their own consumption and about the consumption of others to whom they are compared in assessing their social status.

Economics as an empirical social science: Models provide a lens for refining how we look at some aspect of a complex empirical reality whether it be the differing patterns of work hours over the twentieth century or the effect of a price change on the demand for fish.

Models: The map is not the territory: Our data on fish markets and declining work hours reveal a much more complex empirical reality than our models capture, which necessarily leave out aspects of the problem that might explain, for example, why Asian buyers paid less for their fish purchases, or why work hours fell so much more in Sweden than in the US.

Rents: As with other voluntary exchanges, people purchase goods because they expect to be better off by making the purchase compared to their fallback position (not making the purchase). Consumer surplus is a form of rent.

Ordinal and cardinal preferences: Our analysis of individual demand did not require utility to be measured cardinally; an ordinal ranking was sufficient. But when we aggregate the consumer surplus of the people making up a market, we are assuming that the marginal utility of money left over for other purchases is (a) the same for everyone involved and (b) is measured in the same units as money itself, a cardinal measure in which we can say, for example, that Harriet has twice as much utility as her brother.

Economics and public policy: The analytical tools for the analysis of demand—willingness to pay, the $mrs = mrt$ rule, demand curve, income, and substitution effects—provide the basis for better understanding public policies, for example sugar and fat taxes to address the rise in obesity and carbon taxes to mitigate global climate change.

Fairness: These policies may place greater burdens on the less well-off and thus raise questions of distributive justice. The dubious assumption on which consumer surplus is based—that an additional dollar is of the same value to a poor person as to a rich person—makes it difficult to apply this concept to cases in which people of differing wealth levels are involved.

IMPORTANT IDEAS

marginal rate of substitution	quadratic, quasi-linear utility	linear demand curve
quasi-linear utility	substitution effect	income effect
demand curve	compensated budget constraint	interpersonally comparable utility
market demand	price elasticity of demand	Cobb–Douglas utility function
slope of demand curve	work hours	regressive tax/progressive tax
Veblen effect	conspicuous consumption	complements/substitutes in consumption
sin tax	segregation	(un)stable equilibrium
dynamics	consumer surplus	consumer expenditure
mrs = *mrt* rule	marginal rate of transformation	

MATHEMATICAL NOTATION

Notation	Interpretation
x, y	individual quantities of a good
p_x, p_y	market prices of two goods x and y
m	budget available for buying goods
$u(\ )$	utility function
α	Cobb–Douglas exponent for good x
$g(\ )$	the nonlinear part of a quasi-linear utility function
h	fraction of the day spent working for wages
f	fraction of the day that is free time
w	wage
$\underline{x}$	consumption of the very rich
v	Veblen effect parameter
β	negative of the slope of the demand curve
ε	a measure of the preference for good y in utility function of substitutes
n	number of individuals in a market
X	market demand
η	elasticity of market demand with respect to price
g	fraction of Greens in a neighborhood
δ	degree of preferences for segregation

Note on superscripts: v: outcome with Veblen effect; QL: quasi-linear utility function; QQL: quadratic quasi-linear utility function.

CHAPTER

8

SUPPLY

FIRMS' COSTS, OUTPUT, AND PROFIT

We might as reasonably dispute whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper, as whether [prices are] governed by utility or cost of production . . . We have next to inquire what causes govern supply prices, that is prices which dealers are willing to accept for different amounts.

Alfred Marshall,
Principles of Economics (1890)[1]

DOING ECONOMICS

This chapter will enable you to:

- Understand the difference between accounting profits and costs and economic profits and costs.
- Summarize the facts about how the cost per unit produced varies with the level of output as a result of economies of scale and other influences.
- Explain why, if competition is limited, a firm's owners profit by restricting the firm's hiring and sales.
- Show why the price markup over costs is greater if there is less competition among firms.
- Explain how in some cases we can derive a firm-level and market supply curve based on firms' willingness to sell.
- See how production and sales of a good typically allow rents for both consumers and owners in the form of consumer surplus and economic profits.
- Understand why the Nash equilibrium price and quantity transacted in the model of perfect competition among price-taking buyers and sellers is Pareto efficient.
- Use the model of competitive supply and demand to show the effect of a tax on consumer surplus, economic profits, tax revenues, and deadweight loss.
- Contrast two benchmark models: a price-taking firm constrained by rising costs and a price-making firm constrained by demand.

8.1 INTRODUCTION: SOLAR PANELS AND ARMORED TRUCKS

In the summer of 1943 during World War II, Allied troops under General George Patton landed on the island of Sicily and began the battle to retake Italy from Italian fascist forces. Hitler soon ordered German units to the south of the peninsula to reinforce the retreating Italians. Palermo, the capital of Sicily, fell to American units on July 23.

At the same time, in support of the war effort, the Ford Motor Company had just launched the production of a new armored truck, the M20-GBK, in Chicago. A total of 3,791 M20-GBK armored trucks would be built over the next two years. General Patton (accompanied by his pet bull terrier Willie) would inspect one of the first to arrive on the European front of the war.

Because quickly getting the armored trucks built and transported to the war fronts was a top priority, Ford kept careful account of the labor time that was devoted to the production of the trucks, tabulating "man days of labor" (a great many of Ford's workers were women during the war) and total units produced each month. Although costs other than labor were of course involved—machinery and materials important among them—we lack data on these inputs. Ford's records nonetheless provide a glimpse into how the costs of producing goods vary with the *amounts* produced.

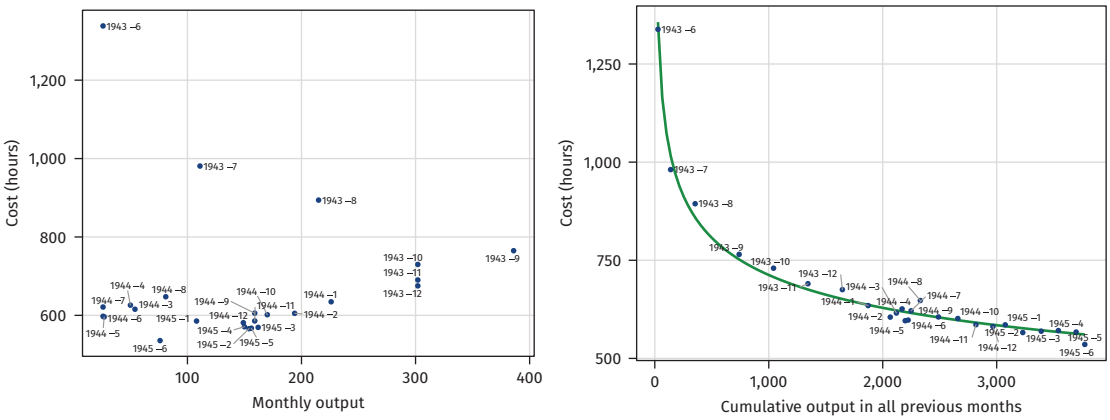
Figure 8.2 shows the evidence (photographed from records in Ford's historical archives). In panel (a), there is surprisingly little evidence that

Figure 8.1 The M20GBK armored truck built by Ford Motor Company 1943–1945.

Used by permission, Minnesota Historical Society.

Figure 8.2 Labor time to produce one M20-GBK armored truck. The horizontal axis in panel (a) is the usual dimension used for a cost curve: number of units produced in a given time period (in this case a month). The vertical axis is average labor hours devoted to the production of a unit in the given month. The horizontal axis in panel (b) is the cumulative total number of M20-GBKs produced since June 1943, when production began.

Source: Lafond et al. (2020).



(a) Costs were unrelated to number of units produced that month

(b) Costs declined as the total number of units ever produced increased

costs were lower for larger production runs. There is *no* relationship at all between the amount produced and the cost per unit. For example, we see that while the most costly month was June 1943, the beginning of the production run, when only 26 units were produced, the fourth most costly period was just four months later when the greatest number of units were produced.

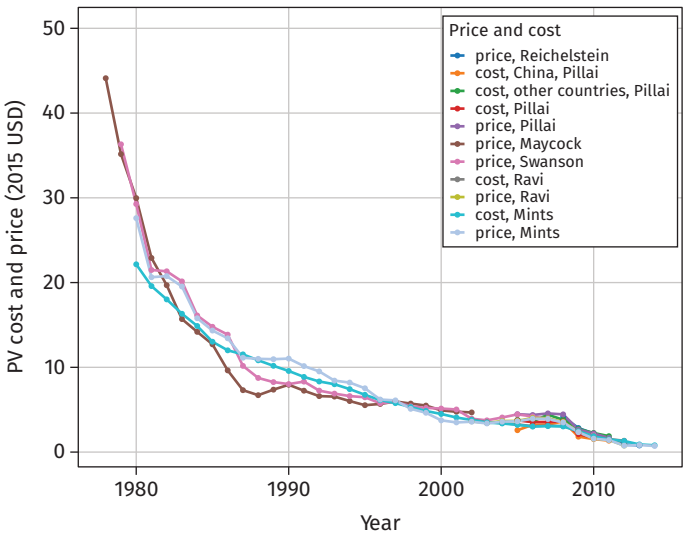
But panel (b) tells an entirely different story. Here we show the cost per unit in each month plotted *not* against how many were produced *that* month but instead, against how many units had been produced in all of the *previous* months. The downward-sloping line is evidence of what is called "learning-by-doing": costs were cut in half as the entire Ford team–engineers, managers, and workers alike–*learned* from their accumulating experiences. Remember from Chapter 6 that learning-by-doing is one of the sources of *decreasing* costs as output *increases*.

In panel (b), points farther to the right must be for later months (because they represent larger amounts of previous production) so it is possible that new technologies became available for the production process, accounting for the decline in costs. But this seems unlikely in light of the two-year time period and how long it takes to conceive, develop, and install new technologies.[2]

An even more remarkable fall in the costs can be seen in solar panel technology. Photovoltaic modules are the underlying components of solar panels. Figure 8.3 records both the cost and the price of photovoltaic

Figure 8.3 The declining cost of photovoltaic modules used to generate solar power. We show the results of the available studies of either prices or costs of the modules.

Source: Kavlak et al. (2018).



modules (in 2015 dollars to take account of inflation) using data from the US, China, and other countries over the period 1975-2015. The costs in 2015 were an astounding 1 percent of what they had been just 40 years earlier.

In contrast to Ford's armored truck, the major contributor to the decline in the costs of photovoltaic modules was research and development (both private and government funded in about equal measure). Learning-by-doing accounts for about 10 percent of the cost reduction. Despite a substantial increase in the scale of production of the PVs, economies of scale account for only about a quarter of the cost reduction, though this contribution has recently increased.

8.2 COSTS OF PRODUCTION: AN OWNER'S EYE VIEW

Reducing cost is an essential objective of owners of firms. They seek to expand the value of their ownership of the firm, which they do by making the firm be as profitable as possible. The means by which they obtain profit is by employing people and using machines to produce goods to sell to customers.

Figure 8.4 Solar photovoltaic modules at work.

Photo © Bildagentur Zoonar GmbH/Shutterstock.

In this they face the same kind of "doing the best you can" problem that you studied earlier as the individual's problem of constrained utility maximization. To solve the problem, the owners (or the managers they have hired) decide such things as how to produce—at the least possible cost—the good they wish to sell and how much to pay the employees they hire.

If we want to understand why a particular amount of some good is produced and put on the market, a critical piece of information is the cost of producing the good. To understand costs we take the "owner's eye view," a perspective that will always ask "If I did not invest my money producing this particular product, what else could I do with my funds, and what profits would I make in this alternative use?" The owner's eye view, in other words, takes account of the *opportunity costs* of how the owner's funds are used.

Keep in mind that economists sometimes say things like "firms decide how much to produce." Or the "firm" does this or that. But a firm is not a single person, it is a conglomeration of people. Most of the people do not have the authority to make decisions about such things. Among the people making up the firm, there may be conflicts of interest about the decisions being made. What we mean when we say "the firm does" something is that the *owners* of the firm or their *managers* decide on what is to be done and then exercise their *authority* over others in the firm to carry out their decisions.

For simplicity, we consider a firm that does not purchase any inputs other than labor and the capital goods (say, machines) owned by a single owner. We call such a firm "vertically integrated." And to simplify the analysis even

M-CHECK The characterization of costs here is *different* to the one we used in Chapter 6. Here, we explicitly recognize the role of the opportunity cost of capital. In Chapter 6 we avoided considering opportunity costs by assuming that capital goods are *rented*, not *owned*. The insights of Chapter 6 remain true as the same tools can be applied once the opportunity costs of capital are included.

further we assume that at the beginning of each period he purchases the machines and they run for a single period and then have to be replaced.

A firm's costs for a unit of output are then determined by the following:

- *wages* (w) paid at the end of the period for the amount of labor required to produce one unit (a_L);
- the *price of capital goods* (p_k) times the amount of capital goods required to produce one unit (a_k) called the accounting cost; and
- the *opportunity cost* (ρ) of the funds used to acquire the capital goods, termed the **opportunity cost of capital.**

Putting these ideas together, we obtain the cost to produce one unit of a good equals the labor cost plus the accounting cost of the capital goods plus the opportunity cost of devoting the value of the capital goods to the particular production process, or:

$$\text{cost of 1 unit} = \text{labor cost} + \text{accounting cost of capital} + \text{opportunity cost of capital}$$

$$c = wa_l + p_k a_k + \rho p_k a_k \tag{8.1}$$

$$= wa_l + p_k a_k (1+\rho) \tag{8.2}$$

The third term of Equation 8.1 is the opportunity cost of the firm's capital stock. The owner cares about their opportunity costs because had he not bought the machines, he could have made some other investment that would have yielded him some profits, perhaps purchasing a very safe financial instrument like a government bond.

Let the total cost of a single unit of the capital good be c_K.

$$c_k = p_k + \rho p_k = (1+\rho)p_k \tag{8.3}$$

c_k is therefore the *sum* of the price of one unit of the capital good (p_k) and the opportunity cost of devoting the owner's financial resources to this particular use (ρp_K). Now, including the cost of capital goods c_k into the cost function, we have:

$$c = wa_l + c_k a_k \tag{8.4}$$

Equation 8.4 tells us that the cost (c) to produce a single unit of output depends on the factors of production (labor and capital goods) required, the price of those factors, and the opportunity cost of capital goods.

OPPORTUNITY COST OF CAPITAL The opportunity cost of capital is the accounting rate of profit that a wealth holder would make on his next best alternative use of the funds used to acquire the capital goods used in production.

The difference between the owner's eye view of costs and the viewpoint of an accountant or an engineer is this: the engineer or the accountant would ignore the opportunity cost of the use of the owners funds and simply count the actual outlays–the payments to workers, and the cost of the machinery "used up" in producing the good. They would think about what is called the "accounting cost" or:

$$c^A = wa_l + p_k a_k \tag{8.5}$$

Because the owner and the accountant look at costs from a different perspective, they also have a different idea of how to count profits.

CHECKPOINT 8.1 Accounting and economic cost The "owner's-eye view" is called economic cost, to distinguish it from accounting cost. What is the difference between the two?

8.3 ACCOUNTING PROFITS AND ECONOMIC PROFITS

A firm's **economic profits** are given by its total revenues $r(x)$ which is just total output times the price at which it is sold, $p(x)$, minus total costs $c(x)$, which we can divide by the total output (x) to find the economic profit per unit:

$$\begin{aligned} \pi(x) &= r(x) - c(x) \\ &= px - cx \\ \therefore \text{Profit per unit} \quad \frac{\pi}{x} &= \frac{px - cx}{x} = p - c \end{aligned} \tag{8.6}$$

We can substitute c from Equation 8.2 into Equation 8.6 for the per-unit profit as follows:

$$\frac{\pi}{x} = p - wa_l - p_k a_k(1+\rho) \tag{8.7}$$

Equation 8.7 tells us that the per-unit profit is determined by the price at which goods are sold (itself determined by the demand curve and how many goods are on the market), the costs of the factors of production per unit of output, and the opportunity cost of capital.

But from the owner's perspective it is not the profits per unit, or even the total profits that matter: what he cares about is the total profits relative to how much money he invested in this firm that he could have invested elsewhere, making some profits in this alternative use of his funds.

ECONOMIC PROFIT Economic profit is accounting profit minus the opportunity cost of funds tied up in long-lived plant and equipment evaluated at the opportunity cost of capital, ρ. See also *accounting profit*. When we use the term "profit" without an adjective ("economic" or "accounting") we mean economic profit.

REMINDER In Chapter 6 we explained that the owners of a firm seek to maximize their profits in choosing how much to produce by:

- first determining the least-cost way to produce any level of output they might consider producing, and then
- using that information, along with information on the demand for their product and the degree of competition among firms selling similar products, determine the level of output that would yield the greatest profits.

We explained the first step in Chapter 6. Here we summarize the information from the first step in a cost function, and using this we explain the second.

What he cares about is the rate of economic profit, which is the ratio of the firm's economic profit per unit to the value of its capital stock. The value of the capital stock is the price of capital goods times the total capital goods used by the firm.

We can express the **rate of economic profit** (r^E) either as total profits divided by total value of the capital stock or as the profit per unit of output ($\frac{\pi}{x}$) divided by the value of capital per unit of output ($p_K a_K$), as shown in Equation 8.8:

$$r^E = \frac{\pi}{x p_k a_k} = \frac{\frac{\pi}{x}}{p_k a_k} \tag{8.8}$$

Accounting profits are given by total revenues minus the direct cost of production without taking account of the opportunity cost of capital. Or, what is the same thing (we show this in M-Note 8.1), the rate of accounting profits (r^A) equals the rate of economic profit (r^E) plus the opportunity cost of capital:

$$r^A = r^E + \rho \tag{8.9}$$

So, the rate of economic profit is zero when the rate of accounting profit equals the opportunity cost of capital:

$$\text{Zero (economic) profit condition} \qquad r^E = r^A - \rho = 0 \tag{8.10}$$

Equation 8.10 is called the *zero-profit condition*. (It could be more accurately named: the zero-*economic*-profit condition.) In the owner's eye view the condition is important. When the zero-profit condition is satisfied, the owner makes the same profits on his investment in the firm that he would make by investing in his next best alternative. If economic profits are positive, he is happy with his investment and will perhaps increase it. But if the opportunity cost of capital exceeds the accounting profits he is making, then economic profits are negative and he will be better off investing elsewhere.

CHECKPOINT 8.2 **Accounting and economic profits** Why would the owner of a firm consider economic profits (not accounting profits) as a measure of the success of his investment in the firm?

RATE OF ECONOMIC PROFIT Economic profits divided by the value of the capital stock.

ACCOUNTING PROFIT Accounting profit is the difference between sales revenue and the direct cost of the inputs used to produce output, excluding the opportunity cost of the funds tied up in financing long-lived assets such as buildings, intellectual property, and equipment. See also *economic profit*.

M-NOTE 8.1 Economic profits and accounting profits

To see the relationship between economic profits and accounting profits, we can substitute the value of $\frac{\pi}{x}$ given by Equation 8.7 into Equation 8.8 to have the following expression for the economic rate of profit:

$$r^{E} = \frac{p - (wa_l + p_k a_k(1+\rho))}{p_k a_k} \tag{8.11}$$

We can rearrange Equation 8.11 to separate out the opportunity cost of the capital goods used:

$$\begin{aligned} r^{E} &= \frac{p - wa_l - p_k a_k - \rho p_k a_k}{p_k a_k} \\ &= \frac{p - wa_l - p_k a_k}{p_k a_k} - \frac{\rho p_k a_k}{p_k a_k} \\ &= \frac{p - wa_l - p_k a_k}{p_k a_k} - \rho \end{aligned} \tag{8.12}$$

We know that the rate of accounting profit is:

$$r^{A} = \frac{p - wa_l - p_k a_k}{p_k a_k}$$

Therefore:

$$r^{E} = r^{A} - \rho \tag{8.13}$$

So the rate of economic profit is equal to the rate of accounting profit minus the opportunity cost of capital.

8.4 COST FUNCTIONS: DECREASING AND INCREASING AVERAGE COSTS

In Chapter 6 we explained how the owners or managers of a firm will select a technology and a combination of inputs of labor (l) and capital goods (k) to *minimize* the cost of producing a particular level of output x given the wage rate w and the price to hire (or rent) capital p_k. The solutions of this constrained optimization problem, applied to every level of output the firm might produce, then constitute the firm's **total cost** function $c(x)$, which represents the minimum total cost of producing the quantity x at the given costs of inputs. Figure 8.5 shows two different total cost functions.

In both panels there is a *fixed cost*, c_0, the firm must pay even at a zero level of output. A cost is said to be *fixed* if the firm cannot avoid it even if it produces nothing. Examples are the cost of the firm's capital goods which would be very costly to deinstall and sell if the firm chooses to produce less.

TOTAL COST Total cost is the minimum cost of producing an output level at given prices of inputs and the opportunity cost of capital. When we use the term cost without an adjective ("economic" or "accounting") we mean economic cost.

Also included are any patents or other intellectual property required in the production process.

Intellectual property such as patents, trademarks, and copyrights, even if owned by the firm, represent a fixed cost because the opportunity cost of the firm using them exclusively is the sum they could collect in fees by selling or leasing them. Other fixed costs are the wages and salaries of high-level managers who will be employed even if the firm is producing nothing, and the costs of any licenses, lobbying, or advertising in which the firm invests independently of its output.

Some costs are fixed in the **short run** but not in the longer run. The term short run does not refer to a period of time, but instead to what is *exogenous*, meaning fixed, that may become endogenous, meaning changeable, in the long run. About a firm's costs, for example, we assume that its stock of capital goods and technology is exogenous (constant) in the short run, but may be varied in the long run.

M-CHECK Marginal cost is the change in total cost, Δc, associated with a small change in total output, Δx, expressed as the ratio of the former to the latter, or $mc = \frac{\Delta c}{\Delta x}$, where Δx is small. Marginal cost is not the cost of producing the last unit (the units produced are identical, and the cost of each unit is the average cost, or $\frac{c(x)}{x}$).

The total cost curves in the two top panels of Figure 8.5 therefore show the minimum cost required to produce each level of output x, for a given wage w and cost of capital goods c_k. Buildings or equipment not being used because a firm has downsized can eventually be sold or rented to others.

Marginal cost is the change in total costs associated with a small change in output or $\frac{\Delta c(x)}{\Delta x}$, which is the slope of the total cost curve at any given point such as x_1.

Average cost is the ratio of total cost to output, which is the slope of a line from the origin to the point on the total cost curve corresponding to a given output, x_1.

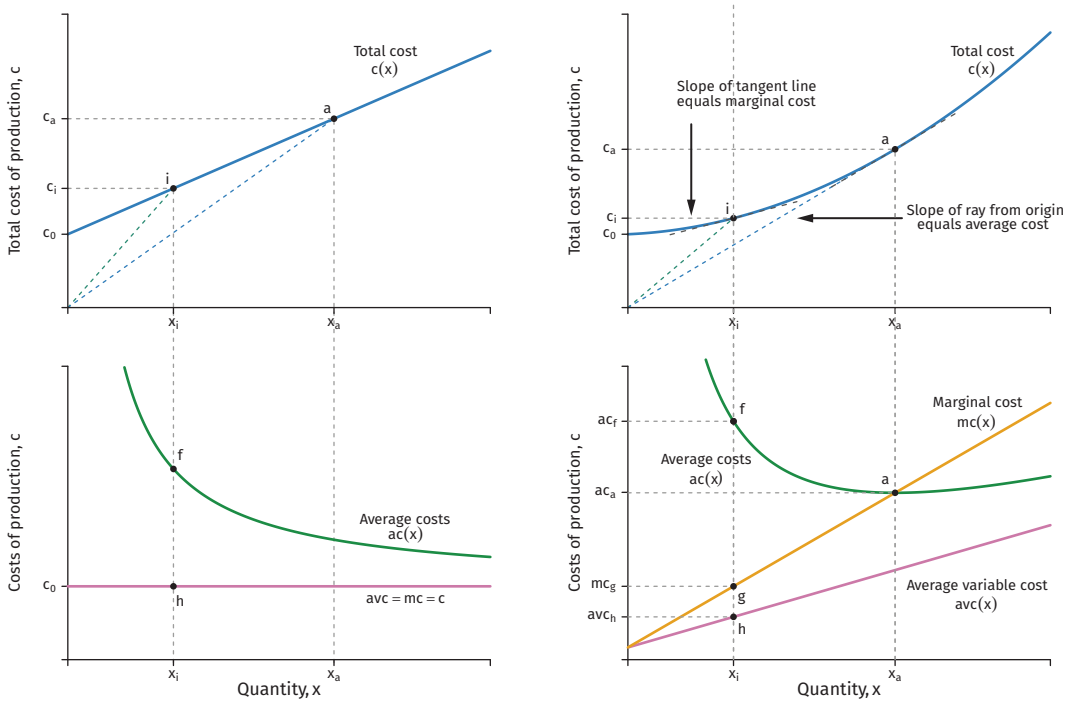
Average variable cost is the slope of a line from the y-intercept of the total cost curve to the point on the total cost curve. Examples of variable costs are the wages and salaries paid to employees engaged in production, the costs of inputs and energy used in the production process, and the wear and tear on the equipment used in production corresponding to the given output (not shown in the figure).

SHORT RUN The term does not refer to a period of time, but instead to what is exogenous, meaning fixed, that may become endogenous in the long run. About a firm's costs, for example, we assume that its stock of capital goods and technology is exogenous (constant) in the short run, but may be varied in the long run.

MARGINAL COST Marginal cost is the effect on total cost of producing a small amount more of output. It is the slope of the total cost function at each point.

AVERAGE COST Average cost is the cost per unit of output produced.

Figure 8.5 Total, average, marginal, and variable costs: increasing and constant marginal costs. Panel (a) and panel (b) represent different cases: *rising* marginal costs in panel (b) and *constant* marginal costs in panel (a). The upper and lower graphs show two ways of looking at costs for these two cases. In the figure we used the following cost function: Total Cost: $c(x) = c_0 + c_1 x + \frac{c_2}{2}x^2$ with $c_0 > 0$ in both panels, $c_1 > 0$ and $c_2 = 0$ in panel (b) and $c_2 > 0$ in panel (a). Based on that total cost function we have: marginal cost: $mc(x) = c_1 + c_2 x$, average cost: $ac(x) = \frac{c_0}{x} + c_1 + \frac{c_2}{2}x$, and average variable cost: $avc(x) = c_1 + \frac{c_2}{2}x$. All of the information used to produce the lower graphs is contained in the upper graphs (but in different visual form).



(a) Linear total cost with declining average costs and constant marginal costs

(b) Increasing total cost with U-shaped average costs and rising marginal costs

In Figure 8.5 (b) we show these average, marginal, and average variable costs. For a firm producing an amount x_1 units of its output:

- Point **f** shows the average cost, that is, the slope of the ray from the origin to point **i** on the total cost function in the top panel.
- Point **g** is the marginal cost of production, namely the slope of the total cost function at point **i** in the top panel.
- Point **h** is the average variable cost if the firm is producing x_1.

Table 8.1 Cost functions, average cost (ac), marginal cost (mc), and average variable cost (avc). The cost function is $c(x) = c_0 + c_1 x + \frac{c_2}{2}x^2$. Evidence on cost functions is presented below. See Table 8.2.

Fixed cost (c_0)	Linear cost (c_1)	Quadratic cost (c_2)	Description	Example
0	> 0	0	Constant ac = mc	Many firms in the long run (no fixed costs)
> 0	> 0	> 0	U-shaped ac curve; rising mc	Many firms in the short run, (Figure 8.5 (b))
> 0	0	0	Declining ac; mc=0	Digital production with "first copy costs"
> 0	> 0	0	Declining ac, constant mc	Many firms (Figure 8.5 (a))
0	> 0	< 0	Declining avc	Many firms

EXAMPLE Think about producing this book. Writing it has incurred large fixed costs (the research and writing, decades of experience in the classroom, the publisher's advertising and editorial advice) called first copy costs, namely the cost of producing just one copy of the book. But the marginal costs which are approximately equal to the average variable costs, that is the further costs of producing additional copies of the physical book or e-book, are very limited. For some texts, like *The Economy* by The CORE Team, an open-access introduction to the field, the fixed costs of producing the content was in excess of 2 million US dollars, but the marginal cost is literally zero and it is available at price = marginal cost, namely free (www.core-econ.org.).

Notice in the top-right graph that for a point like **a** where the ray from the origin has the same slope as a tangent to the curve, it will be the case that $mc = ac$, which, as you can see in the lower-left graph (also point **a**), the average cost is at a *minimum*.

CHECKPOINT 8.3 Draw cost curves For each of the cases (the rows) in Table 8.1 draw the average, average variable, and marginal costs curves. *Hint*: Whenever you draw average and marginal cost curves, if the marginal cost curve intersects the average cost curve, make sure this happens at the point where average cost is lowest! At zero output, the average variable cost is equal to the marginal cost, and is increasing when it lies *below* marginal cost.

8.5 APPLICATION: EVIDENCE ABOUT COST FUNCTIONS

As you will see later in this chapter and in Chapter 9, how the process of competition works and how we should model firms' owners' profitmaking strategies hinge on facts about cost curves, and on the following question in particular: Does the average cost curve of a firm slope *upward*, so that costs per unit *rise* with increased total output?

There are two sets of influences on the shape of the cost curve:

- *scale economies*: whether the production function used by the firm is characterized by economies of scale, diseconomies of scale, or constant returns to scale; and
- *input costs*: whether the cost of acquiring inputs–c_k and w in our example above, but including other inputs–is greater, less, or the same for differing levels of output.

Here are two examples of how input prices will depend on the amount the firm is producing. Hiring a large amount of labor in a small labor market may require higher wages than would be necessary if the firm were hiring fewer (we will model such a case in Chapter 11 (section 11.12)). Another example: a large buyer–like Walmart–may be able to bargain for lower wholesale prices of the merchandise it sells than a smaller firm.

We want to find out, then, what empirical research has been able to determine about the shape of the average cost curve in different industries. There are many ways to investigate this including:

- *Engineering data:* Engineering evidence about the relationship between physical inputs and output, to determine if larger outputs can be produced with less than proportional increases in inputs.
- *Statistics on costs*: Statistical studies of how costs vary with the amount produced.
- *Estimates of production functions allowing inferences about economies, diseconomies of scale, or constant returns to scale:* Statistical estimates of production functions, for example to determine if the sum of the exponents for the firms' inputs in a Cobb–Douglas production function are greater than one, less than one, or about one.
- *Surveys of managers and owners*: Direct evidence from firm managers on their understanding of the extent of economies of scale and the shape of their cost curves.

Engineering evidence: Physical inputs and outputs

Engineers design production processes and entire production systems to find the least-cost ways of producing given units of output. They also seek to determine the least-cost size of a production unit such as a plant, asking, for example: Would the costs of producing some amount be less using one large plant or two small plants? And, unlike economists, they start from input–output relationships measured in physical units.

When translated into costs using the prices of the physical inputs required the advantages of large firms can be substantial. A handbook for chemical engineers tells us that, to store dry chemicals in a fiber drum, the cost per square foot of storage would be 28 percent lower for a plant needing a 61-gallon drum than for a small plant needing only a 15-gallon drum.

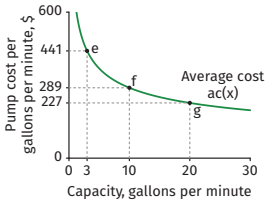
In Figure 8.6 we plot similar costs per unit of capacity from the same engineering handbook. The cost counted here is the price of the equipment in question, that is the capital good involved in the production process. The capacity is in physical units–gallons per minute for the pump. Economies of scale are quite common in this kind of manufacturing process.

! REMINDER If doubling all inputs *more* than doubles output then the production process exhibits *economies of scale*; if the increase in output is *less* than double, we have *diseconomies of scale*. And if output exactly *doubles* we have *constant* returns to scale.

✓ FACT CHECK Here is an example of how an engineer might use knowledge of the technology to estimate costs. The cost of piping for conveying liquids in a plant is proportional to the *circumference* of the pipe, but its capacity (in gallons or liters) is proportional to the area of a cross-section of the pipe. Because doubling the circumference of a circle (the cost of the pipe) *more* than doubles the area of the circle (the capacity of the pipe) the cost *per gallon or liter* conveyed by the pipe is *lower* the *larger* is the pipe's capacity.

Figure 8.6 Cost of equipment involved in chemical production processes. The cost *per gallon* per minute of capacity decreases as the capacity of the pump increases.

Source: Peters et al. (1991), with 2020 prices.



Econometric evidence on economies of scale and average costs

Table 8.2 summarizes some econometric studies, showing that economies of scale or constant returns to scale and declining or flat average cost curves are common. There is little evidence that average costs rise with higher levels of output.

Table 8.2 Evidence on economies of scale and decreasing average costs.

Source	Industry	Estimated slope of average cost (AC) curve and/or economies of scale
Nerlove (1963)	Electricity supply	"Marked" economies of scale especially for smaller firms
Griliches and Ringstad (1971)	Norwegian manufacturing	Economies of scale
Christensen and Greene (1976)	Electric power generation	Declining AC for 1955; flat for 1970
Bittlingmayer (1982)	Iron pipe	Declining AC (including declining mc)
Friedlaender et al. (1983)	Top three US auto firms	Declining AC for GM and Chrysler, rising for Ford
Caves et al. (1984)	Airline industry	Declining AC for a given market, constant AC for expanding market
Dawson and Hubbard (1987)	Dairy	Substantial declines in AC at modest size, slight increases at larger sizes
Hall (1988)	US manufacturing	Flat (17 of 21 industries); economies of scale (2) and diseconomies of scale (2)
Klette (1999)	Norwegian manufacturing	Flat AC in "most industries." "Moderate diseconomies" of scale in a few industries
Koshal and Koshal (1999); de Groot et al. (1991); Laband and Lentz (2003)	Higher education	Economies of scale
Cockburn and Henderson (2001); Henderson (2000)	Pharmaceutical research	Economies of scale for firms smaller than the largest, flat for the largest three firms
Ashton (2003)	Privately owned water services (UK)	"Slight" diseconomies of scale, rising average variable costs
Martín and Voltes-Dorta (2011)	Airports	Economies of scale
Ertan et al. (2018)	US firms	Flat AC curves
De Loecker et al. (2020)	Publicly traded US firms 1955–2016	Economies of scale (increased between 1980 and 2016)

Managers' assessments of their firm's cost curves

Alan Blinder and his collaborators adopted a research strategy rarely used by economists: they simply asked managers and CEOs what they thought their cost curves looked like. Specifically, they were asked if "their variable costs per unit are roughly *constant* when production rises" or if instead variable costs were *rising* or *falling* with increased production. Of the 190 firms that responded only 11 percent reported that average variable costs *increased* with additional output. The rest reported having either downward-sloping or flat average variable cost curves, meaning, even in the absence of fixed costs, *declining* or *constant* average costs.[3]

The conclusion most consistent with the engineering, statistical, and survey evidence is that average costs *do not* rise with increased output. This means that marginal cost is either *equal* to or *less* than average cost, which is what we will assume in the models to be developed.

For reasons of simplicity we frequently assume that marginal and average costs are equal. But this ignores the important role of *fixed costs* which, if they are substantial, will lead to declining average costs.

CHECKPOINT 8.4 Rising or falling costs of inputs Give examples (additional to those above) illustrating that the cost of a firm's input may either rise or fall the more of the input is purchased.

✓ **FACT CHECK** In the Blinder study, managers at ten of the firms were unable to answer the question even after the idea of average variable costs was explained in a number of different ways, suggesting that our models above about how firms choose output levels and prices may not apply to all firms.

! REMINDER The *price-setting power* that you studied in Chapter 4 describes the situation faced by the monopolistically competitive firm. They do not have "take-it-or-leave-it" (TIOLI) power.

8.6 A MONOPOLISTIC COMPETITOR SELECTS AN OUTPUT LEVEL

The firm's cost function is part of the information needed to answer the question: How do the firm's owners determine the level of output that the firm will produce and the price at which the output will be sold?

We model the firm's owners seeking an *output* level and a *price* of their product that will maximize profits:

- using a *given* technology;
- facing a *given* set of input prices; and
- subject to the demand curve for their product (the *constraint*).

We provide an answer based on a simple cost function in which average cost is a constant, so average cost = $ac = c = mc$ = marginal cost.

We model a single firm selling a unique product—for example, Honda selling the Accord model of the many cars they produce. As the only producer of this particular product—other businesses can sell other cars, but no other firm can sell a Honda Accord—the firm is a monopolist, that is the single seller of what is called a *differentiated* product.

HISTORY The seemingly contradictory term "monopolistic competition" is due to Edward Hastings Chamberlin whose book of that title along with Joan Robinson's *Imperfect Competition* (both published in 1933) provided an alternative to the model called "perfect competition." We use Chamberlin's term rather than Robinson's because we see competition among large firms with substantial market power as an alternative benchmark model, not as an "imperfection" of the "perfect competition" approach.

REMINDER In Chapter 7 you learned that the demand curve for a product will be steeper if there are fewer close substitutes for the product, so that the reduction in sales associated with higher prices is modest. As a result, firms seek to design and advertise products so that they will have or appear to have few close substitutes. Later, you will see that the maximum profit per unit of output sold is greater the steeper is the demand curve for the product.

M-CHECK Here we use linear demand functions, such as those we derived in Chapter 7 from quadratic, quasi-linear utility functions.

REMINDER A firm's economic profit is the difference between sales revenue and the total cost of producing the output, $\pi(x) = r(x) - c(x) = p(x)x - c(x)$. Remember that the firm's costs include the opportunity cost of the capital goods it uses. When we use the term "profit" without an adjective ("economic" or "accounting") we mean economic profit.

Monopolistic competition and product differentiation

But potential buyers have alternatives: for example, a Subaru Forester is a substitute for a Honda Accord. Even if there is no substitute product, being a monopolist does not mean that the firm can sell *any* amount that it produces at any price it chooses. It is constrained by a downward-sloping demand curve that effectively tells the firm: "if you produce more cars or you will have to lower your price to sell them all." Or: "if you charge a higher price you will not be able to sell as many goods."

These firms engage in what is called **monopolistic competition**: they are the *single seller of their own product*—Subaru cannot produce and sell an Accord—but they *compete* with other firms selling products that are close substitutes for their own product. A monopolistically competitive firm will face less competition and make higher profits the *more different* its product is from what other firms are selling.

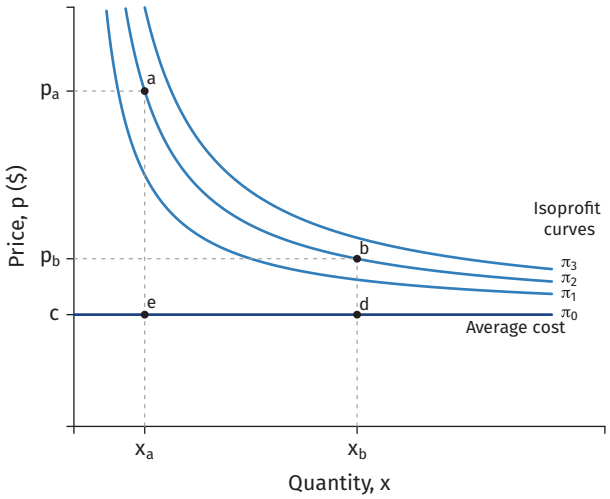
Economists say that these firms produce a *differentiated product*, meaning that a firm's product does not just "happen to be" different from its competitors' products but that firms actively try to make its product be or seem to be as different as possible from the products of other firms. **Product differentiation** strategies include distinctive design and advertising to promote brand loyalty. Another is the use of trademarks, which are privately owned (they are intellectual property) and cannot be used by firms other than by their owner. These prevent other firms from competing directly (only Nike can sell Nike shoes). Active product differentiation is just one of a great many strategies that owners of firms deploy to maximize their profits. Others include innovation to reduce costs (that you studied in Chapter 6), preventing other firms from producing identical or similar products, and lobbying government bodies for favorable tax or regulatory treatment. But here we will study just one dimension of this profit-maximizing process: deciding on a quantity of goods to produce and the price at which to sell them.

Monopolistically competitive firms set prices and levels of output to maximize profit. To understand a firm's choice of a particular combination of a price and a quantity, we will use the same architecture for constrained

MONOPOLISTIC COMPETITION Monopolistically competitive firms are the only sellers of the particular good they produce, but they compete with other firms that sell similar products.

PRODUCT DIFFERENTIATION Product differentiation is a business practice aimed at making the firm's product appear more distinct from or less similar to substitute products.

Figure 8.7 A firm's isoprofit curves. The labels to the right of the isoprofit curves indicate the level of profits associated with every point on the curve, and $\pi_3 > \pi_2 > \pi_1 > \pi_0 = 0$. Selling x_a units of output at price p_a (point **a**) yields the same profits as selling x_b units at price p_b (point **b**). If the price is equal to the average cost, no profits are made no matter how many units are sold (for example, points **d** and **e**).



maximization that we have used throughout the book with its two basic elements:

- The *objectives* of the decision maker represented as a new type of *indifference curves* called *isoprofit* ("equal-profit") *curves*, giving the price–quantity combinations that, if implemented, would yield a given level of profit.
- The *feasible set* of choices that the decision maker can implement, that is, all of the price–quantity combinations that the firm could actually implement given the demand curve for its product.

Isoprofit curves: The owner's objectives

The **isoprofit curves**, along which profit is equal to some constant, are just another kind of indifference curve because the owners of the firm are indifferent between making a particular level of profit by selling a lot of their product for a relatively low price, or alternatively, by selling a few units at a high price. Some of the firm's isoprofit curves are shown in Figure 8.7.

ISOPROFIT CURVE An isoprofit curve shows combinations of prices and quantities sold of a good yielding equal profits to the owners of a firm.

! REMINDER As we showed in Chapter 7, for any bundle of goods that she is currently consuming (x, y), a consumer has a marginal rate of substitution of money, y, for the good, x, which is her *willingness to pay* for the good. A buyer's willingness to pay for all possible levels of purchase is their individual *inverse demand function*. The market demand function for x is the horizontal sum of all buyers' demand functions for the good.

HISTORY In his 1948 introductory economics textbook, Nobel Laureate Paul Samuelson, called the "founder of modern economics," wrote that monopolistic competition "includes most firms and industries" while the perfectly competitive firm "includes a few agricultural industries."[4]

Figure 8.8 Joan Robinson (1903–1983) taught economics at the University of Cambridge and developed one of the first models of what she called "imperfect competition." The term "monopsony" is also due to her, as were many colorful turns of phrase. Concerning debates between critics and defenders of capitalism she wrote: "No one is conscious of his own ideology any more than he can tell the smell of his own breath." And "the misery of being exploited by capitalists is nothing compared to not being exploited at all" (that is, being without a job). Her heated "capital controversy" with American economists Paul Samuelson, Robert Solow, and others raised doubts about the general equilibrium model of perfect competition. (We will see in Chapter 14 that subsequently Robinson's critique was substantially vindicated.) A letter to a female student in 1970 from Paul Samuelson, perhaps the most influential economist of the twentieth century, concluded: "P.S. Do study economics. Perhaps the best economist in the world happens also to be a woman (Joan Robinson)."[5]

Photo by Album/Alamy Stock Photo.

Every point in the figure is some price–quantity combination that the firm might think about selecting. Not all of these are possible; what is feasible will depend on the demand curve. So think about the isoprofit curves as hypothetical statements. Point **a** and point **b**, for example, on the same isoprofit curve indicate that *if* the quantity indicated on the horizontal axis could be sold at the price indicated on the vertical axis for both of these points, then the total profits of the firm would be the same. Higher profits are above and to the right (indicating a higher price and/or higher quantity sold).

The negative of the slope of any one of the isoprofit curves is the marginal rate of substitution between selling more and charging a higher price. It is a measure of the owner's "willingness to pay"–meaning to cut prices–in order to increase sales by one unit.

In Figure 8.7 you can see that at point **a** the steep isoprofit means that the owner is willing to cut the price a substantial amount to sell an additional unit. The reason is that at the high price p_a selling the additional unit is worth a lot to the owner. At point **b** the opposite is the case: at the low price p_b selling an additional unit is not worth much to the owner. So the isocost curve is flatter.

Now think about the isoprofit curve given by the horizontal line coinciding with the firm's (constant) average cost curve. The information it conveys is that at a price equal to the average cost the economic profits per unit sold are zero, so selling a lot (point **d**) or a little (point **e**) yields the *same* amount of profit for the firm, namely zero. So the owner's willingness to pay to sell more–to reduce the price further–is zero. Now that the owners of the firm have ranked every possible price–quantity combination as more, or less, or equally as profitable we ask: Which of these combinations could the firm actually implement?

M-NOTE 8.2 Slope of an isoprofit curve: Two methods

To find the slope of the isoprofit curve we write the equation for profits as follows:

$$\pi = \pi(p,x) = px - cx \tag{8.14}$$

$$\pi = x(p-c) \tag{8.15}$$

For any two points on an isoprofit curve, the profit difference associated with the difference in price dp is exactly compensated by the (*opposite* signed) profit difference associated with the difference in quantity dx. Taking account of both effects, the *difference* in profit between the two points is zero.

We want to find the differences in x and p that are consistent with no difference in profits, that is, being on the same isoprofit curve. To do this we totally differentiate Equation 8.15 with respect to dx and dp and we set the result equal to zero.

$$d\pi = dx\frac{\partial \pi}{\partial x} + dp\frac{\partial \pi}{\partial p} = 0 \tag{8.16}$$

continued

We can use Equation 8.15 to find $\partial\pi/\partial x = p - c$ and $\partial\pi/\partial p = x$ and substitute them into Equation 8.16 to find the total derivative:

$$d\pi = dx(p-c) + dpx = 0 \tag{8.17}$$

We can rearrange Equation 8.17 to find dp/dx, which is the slope of the isoprofit curve:

$$\begin{aligned} \frac{dp}{dx} &= -\frac{p-c}{x} \\ &= -\frac{\partial\pi/\partial x}{\partial\pi/\partial p} \\ &= -\frac{\Delta\pi \text{ due to change in } x}{\Delta\pi \text{ due to change in } p} \end{aligned} \tag{8.18}$$

Alternatively, we can calculate the slope of the isoprofit curve directly. First, by rearranging Equation 8.14 we have the expression of the isoprofit curve

$$p(x,\pi) = c + \frac{\pi}{x} \tag{8.19}$$

where π is a constant and p is a function of x.

- As shown in Figure 8.7, when $\pi = \pi_0 = 0$ the isoprofit curve is just $p = c$, the horizontal line.
- When $\pi = \pi_1 > 0$, the graph of the isoprofit curve $p = c + \frac{\pi_1}{x}$ is the hyperbola $\frac{\pi_1}{x}$ shifted upward by c units.
- Given the constants $\pi_3 > \pi_2 > 0$, we have:

$$p(x,\pi_3) = c + \frac{\pi_3}{x} > c + \frac{\pi_2}{x} = p(x,\pi_2) \text{ for all } x$$

That is, for any given quantity x, to achieve a higher profit, the firm has to charge a higher price. Therefore, the isoprofit curve π_3 is above π_2.

We can calculate the slope of the isoprofit curve (Equation 8.18) by taking the first derivative of Equation 8.19 with respect to x:

$$\frac{dp}{dx} = -\frac{\pi}{x^2}$$

Recall that $\pi = (p-c)x$, so we can rewrite the above expression as:

$$\frac{dp}{dx} = -\frac{(p-c)x}{x^2} = -\frac{p-c}{x}$$

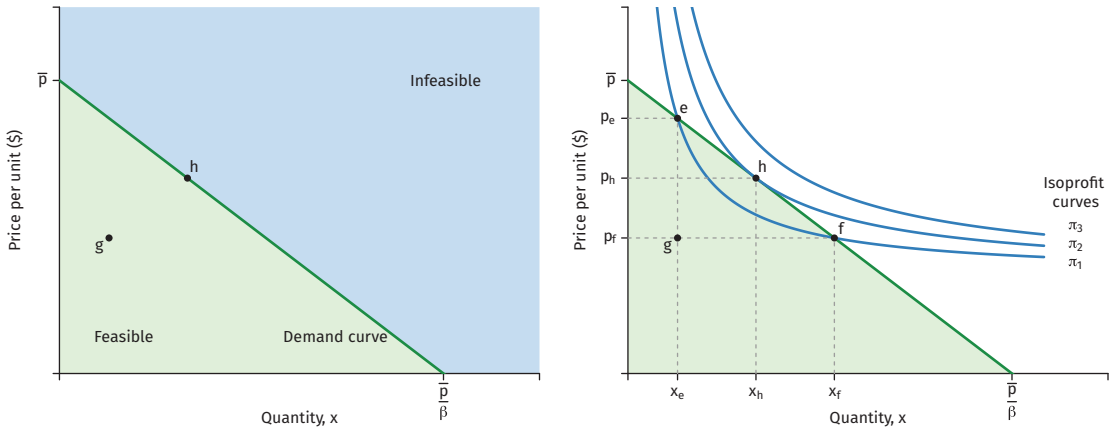
which is the same as Equation 8.18.

CHECKPOINT 8.5 Average costs and the zero economic profits isoprofit curve Explain why the average cost curve (the line **c**, **e**, **d** in Figure 8.7) is also the isoprofit curve for zero economic profits.

Demand: The constraints on profit maximization

To choose the profit-maximizing price and quantity combination, we introduce the demand curve facing the firm, shown in Figure 8.9 (a). The demand curve shown is:

Figure 8.9 Feasible price and quantity combinations and the isoprofit curve. A demand curve for the firm's products determines the feasible combinations of prices and quantity that it can choose in its maximization problem, as shown in panel (a). In panel (b) we include the firm's isoprofit curves. Along an isoprofit curve, profit is constant and the firm wants to choose the highest isoprofit curve given what is feasible. That is, the firm will choose the isoprofit curve that is tangent to the demand curve. At the point of tangency, the firm will choose its quantity of production and the corresponding price on the demand curve.



(a) Feasible combinations of price and quantity

(b) Isoprofits and demand

$$\text{Inverse demand} \qquad p(x) = \bar{p} - \beta x \tag{8.20}$$

and its slope $\frac{\Delta p}{\Delta x} = -\beta$.

The figure divides the entire space into the *feasible* set of price and quantity combinations and the set of *infeasible* combinations. The firm now has to decide among the feasible points.

You can see from Figure 8.7 (a) that while both **h** and **g** are feasible, point **h** has both higher prices and a larger quantity of sales, so the firm will surely not choose point **g**. So the choice is narrowed down to points on the boundary of the feasible set, which is the demand curve.

! REMINDER We used the $mrs = mrt$ rule in earlier chapters. The owners of the firm maximize their profits where the slopes of the two curves are tangent and $mrs = mrt$.

The negative of the slope of the demand curve, $-\Delta p/\Delta x = \beta$, is the marginal rate of transformation of a larger quantity sold into a lower price charged. As before, we interpret the mrt as an opportunity cost, in this case β is the opportunity cost in the necessarily lower price of selling one unit more.

Putting the owner's objectives (the isoprofit map) together with the constraints (the demand curve) we can see in Figure 8.9 (b) that the owner of the firm can choose among points **e**, **f**, and **h** on the demand curve (the feasible frontier); all three points are feasible.

- At point **e**, the owner will compare the slope of the isoprofit curve which is the willingness to pay (cut the price) to sell an additional unit (the mrs)

with the slope of the demand curve, that is, the opportunity cost or the reduction in the price that will be required to sell an additional unit (the mrt). The fact that the isoprofit curve is steeper than the demand curve (or $mrs(x_e, p_e) > mrt(x_e, p_e)$) means that the owner of the firm can make more profits by selling more output at a lower price.

- At point **f**, the opposite is true: the marginal rate of substitution is less than the marginal rate of transformation ($mrs(x_e, p_e) < mrt(x_e, p_e)$) and the owners of the firm can make more profits by selling less output at a higher price.
- Point **h** results in higher profits on isoprofit curve π_2 rather than points **e** and **f** on isoprofit π_1 ($\pi_2 > \pi_1$). At point **h** $mrs(x_h, p_h) = mrt(x_h, p_h)$ and the firm is on the highest isoprofit curve it can attain subject to the constraint of the demand curve.

Using the $mrs = mrt$ rule, the price–quantity combination that maximizes profits is shown in Figure 8.9 (b) and in M-Note 8.3:

slope of isoprofit curve = slope of demand curve

The negatives of these slopes are the marginal rate of substitution and the marginal rate of transformation.

marginal rate of substitution = marginal rate of transformation

$$\frac{p-c}{x} = \beta \tag{8.21}$$

M-NOTE 8.3 Isoprofit curves, the feasible set, and maximum profits

The slope of the isoprofit curves. In the previous M-Note, you saw that the slope of the isoprofit curve is

$$\frac{dp}{dx} = -\frac{p-c}{x} \tag{8.22}$$

The marginal rate of substitution is the negative of the slope of the isoprofit curve.

$$-\frac{dp}{dx} = \frac{p-c}{x} = mrs(x, p) \tag{8.23}$$

Equation 8.23, the marginal rate of substitution between the quantity sold and the profits per unit sold, is profits per unit (the numerator) *divided* by the number of units sold (the denominator).

The slope of the demand curve. The inverse demand curve is the frontier of the feasible set and using the demand curve shown in Equation 8.20:

$$p(x) = \bar{p} - \beta x$$

Differentiating the price with respect to the total amount sold, its slope is $-\beta$. The negative of the slope of the demand curve is the marginal rate of transformation of x into p, or in this case more goods into lower prices at which they sell, that is, $mrt(x, p) = \beta$.

continued

The $mrs = mrt$ *rule for maximum profits.* The price and quantity that maximize the firm's profits are the combination of p and x such that the slope of the isoprofit curve is equal to the slope of the demand curve, or $mrs(x, p) = mrt(x, p)$:

$$mrs(x, p) = \frac{p - c}{x} = \beta = mrt(x, p)$$

8.7 PROFIT MAXIMIZATION: MARGINAL REVENUES AND MARGINAL COSTS

There is a second, equivalent way to determine the price and output level that will maximize a firm's profits. Recall that the firm's revenue is the product of the price at which it sells the output, $p(x)$, and the output sold (x):

$$\text{Total revenue} \qquad r(x) = p(x)x \tag{8.24}$$

M-CHECK In this chapter, we assume that for any given amount sold, all customers pay *the same* price. In Chapter 9 we consider the case of price discrimination, when a firm charges different customers different prices.

Like the cost function, the **total revenue** function has corresponding average and marginal values:

- **Average revenue** is total revenue divided by total output (or the ratio of revenue to output): $ar(x) = \frac{r(x)}{x}$. Notice that because $r(x) = p(x)x$, average revenue is also the price the firm can charge when it is selling output x: $ar(x) = p(x)$. The amount of revenue a firm gains *per unit sold* is simply the *price* of the good.
- **Marginal revenue** is the *change* in total revenue associated with a small change in sales, that is, $\frac{\Delta r}{\Delta x}$. Marginal revenue is therefore the *slope* of the total revenue function at a given output x.

The firm's profit, π, is the difference between its *revenue* from sales of its output, $r(x)$, and the cost of producing output, $c(x)$:

$$\text{Profit} \qquad \pi(x) = r(x) - c(x) \tag{8.25}$$

The owners of the firm would like to find a level of production x that maximizes $\pi(x)$ given the inverse demand function $p(x)$ and the cost function $c(x)$. We have already seen how this can be done using the firm's isoprofit curves and demand curve. An equivalent method is to plot the total revenue curve $r(x) = p(x)x$ together with the total cost curve $c(x)$ as in Figure 8.10.

AVERAGE REVENUE Average revenue is the revenue per unit of output, which is the price.

TOTAL REVENUE Total revenue is total sales times price per unit sold.

MARGINAL REVENUE Marginal revenue is the change in total revenue associated with a small change in sales.

The vertical distance between the two curves at any point, is the total amount of profit. Figure 8.10 shows how, starting at $x = 0$, the vertical distance first *increases* up to a maximum, then *decreases* as revenues begin to increase less than costs. Eventually, at high levels of output, costs exceed revenues.

The maximum profit of the firm occurs at a level of output, x^m, where the slope of the total cost curve is equal to the slope of the revenue curve at x, as shown in Figure 8.10:

$$\frac{\Delta r(x)}{\Delta x} = \frac{\Delta c(x)}{\Delta x}$$
$$\text{marginal revenue} = \text{marginal cost} \tag{8.26}$$

The firm will then find the corresponding price for its good by substituting the profit-maximizing quantity back into its demand curve. Figure 8.10 illustrates this way of finding the output level (and price that maximizes the firm's profits). Setting marginal revenue equal to marginal cost ($mr = mc$) is another rule for constrained maximization and is therefore added to the table of rules in Table 8.3.

M-NOTE 8.4 **The monopolistic competitor chooses a price and output level: general case**

The information relevant to the problem is:

- The firm's total cost $= c(x) = cx$
- The *inverse demand curve* for the firm's product is $p = p(x)$, and
- The firm's *total revenue* is $r(x) = p(x)x$
- so, the firm's total *profit*, $\pi = r(x) - c(x) = p(x)x - cx$

To find the rule that determines the profit-maximizing level of output, we differentiate the profit function with respect to x, and set the result equal to zero. So we have:

$$\pi(x) = p(x)x - c(x)$$

$$\text{Product rule} \quad \frac{d\pi}{dx} = \underbrace{\overbrace{\frac{dp}{dx}x}^{\text{Lost revenue due to fall in price}} + \overbrace{p}^{\text{Revenue gained through increased sales}}}_{\text{Marginal revenue}} - \underbrace{c}_{\text{Marginal cost}} = 0 \tag{8.27}$$

Equation 8.27 tells the owner of the firm to expand production as long as the additional revenues resulting from a small increment in sales—the marginal revenue—exceed the addition to total cost associated with the increment in output. The first term on the right-hand side is the revenue lost due to the *negative* effect of selling more on the price at a given level of output (because the demand curve is downward-sloping). The second term on the right-hand side is the *positive* effect of selling more on revenues (at a given price). So the rule (or first-order condition) determining the profit-maximizing level of output of the firm is to choose the level of x that equates marginal revenue to marginal cost.

Figure 8.10 The level of output that maximizes economic profit. Economic profit at output x is the difference between total revenue at x, $r(x) = p(x)x$, and the total cost of producing x, $c(x)$. Economic profit is maximized when marginal revenue, the slope of the tangent to the revenue curve, is equal to marginal cost, the slope of the total cost curve. The maximum profit is the difference between revenue and total cost where the slopes are equal, that is, at points **a** and **b**. The distance **ab** in the upper panel is shown as the area of the green shaded rectangle in the lower panel. The *slopes* of the total cost and total revenue curves in the top panel are the *levels* of the marginal cost and marginal revenue curves in the lower panel. The price p^m in the lower panel corresponds to the *slope* of the ray from the origin in the top panel to point **a**. The slope of the ray from the origin in the top panel is the average revenue, that is, the price p.

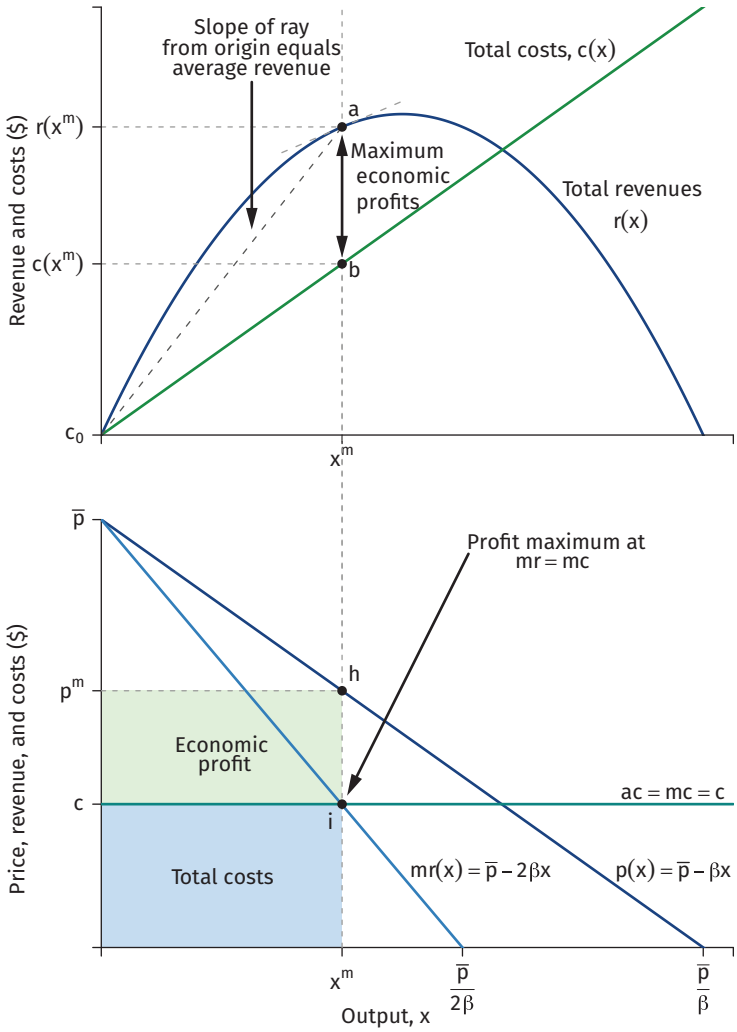


Table 8.3 Five rules: individual constrained optimization, societal Pareto efficiency, and firm cost minimization.

Tangency rules	Tangency of	Rule for what
$mrs = mrt$	An individual's feasible frontier and indifference curve	Individual constrained optimization
$mrs = mrts$	The firm's owners' production isoquant and isocost lines	The firm's owners' constrained optimization
$mb = mc$	Restatement of $mrs = mrt$ using marginal costs and benefits	Individual constrained optimization
$mr = mc$	Restatement of $mrs = mrt$ using marginal costs and revenues (M-Note 8.6)	The firm's owners' constrained optimization
$mrs^A = mrs^B$	Two or more people's indifference curves	Societal (multi-person) Pareto efficiency

M-NOTE 8.5 Output, price, and profit of a monopolistic competitor: the example of a linear demand curve

The relevant information for this case is:

$$\text{Inverse demand} \quad p(x) = \bar{p} - \beta x \tag{8.28}$$

$$\text{Total revenue} \quad r(x) = p(x)x = (\bar{p} - \beta x)x = \bar{p}x - \beta x^2 \tag{8.29}$$

$$\text{Economic profit} \quad \pi^m(x) = r(x) - c(x) = \bar{p}x - \beta x^2 - cx \tag{8.30}$$

The marginal revenue is the derivative of total revenue with respect to output (the slope of its revenue function):

$$\text{Marginal revenue} \quad mr(x) = \frac{dr(x)}{dx} = \frac{dp}{dx}x + p = \bar{p} - 2\beta x \tag{8.31}$$

You can see from the first expression in Equation 8.31 that there are two effects on revenue of a small increase in sales. The first effect is the revenue lost on all of the firm's sales due to the fall in price $\left(\frac{dp}{dx}x\right)$. The second effect is the revenue gained by selling more, which is just the price itself.

The owners maximize profit by choosing the output level that equalizes marginal revenue and marginal cost and then selling that output at the highest price possible, as given by the inverse demand curve:

$$\bar{p} - 2\beta x = c \tag{8.32}$$

Solving Equation 8.32 for x, we have:

$$x^m = \frac{\bar{p} - c}{2\beta} \tag{8.33}$$

To find the price, we substitute the output (Equation 8.33) into the inverse demand curve (Equation 8.28):

$$p^m = \bar{p} - \beta x^m$$

$$\text{Substitute in } x^m \quad = \bar{p} - \beta\left(\frac{\bar{p} - c}{2\beta}\right)$$

$$= \frac{1}{2}\bar{p} + \frac{1}{2}c$$

$$\text{Add and subtract } \frac{1}{2}c \quad p^m = c + \frac{1}{2}(\bar{p} - c) \tag{8.34}$$

The firm charges a price that is *greater* than its marginal costs by one-half the difference of the maximum price and its marginal costs. Price exceeds marginal cost. As a result, the firm makes economic profit equal to π^m.

continued

To find economic profit we substitute the monopolist's quantity into Equation 8.30.

$$\pi^m = \left(p(x^m) - c\right)x^m = \frac{(\bar{p} - c)^2}{4\beta} \tag{8.35}$$

CHECKPOINT 8.6 **Numerical example** Assume new values for the parameters in which we are interested: let $\bar{p} = 100$ and $\beta = 1$ and the firm has constant marginal costs $c = 1$.

a. Find the formula for its profit.

b. Find the profit-maximizing quantity.

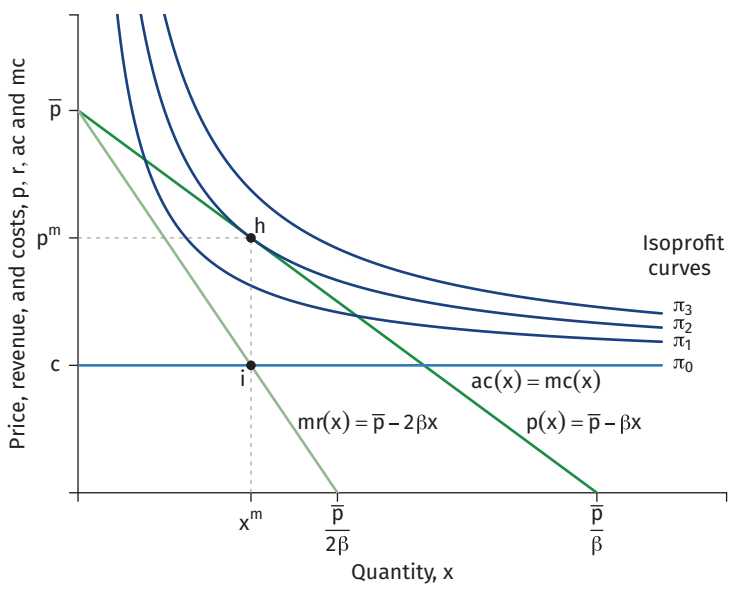
c. How much profit do the owners make? What is the price and how much greater is the price than marginal costs?

Two methods of choosing a profit-maximizing price and quantity compared

We have introduced two methods of determining the profit-maximizing price and quantity defined by:

- the tangency of the feasible frontier given by the demand curve and the isoprofit curves; and
- equating marginal revenue and marginal costs.

Figure 8.11 **Isoprofits, demand, marginal revenue, and marginal costs.** The firm chooses the quantity at which marginal revenues equal marginal costs. This quantity coincides with the quantity the firm would have chosen when its isoprofit curves were tangent to its demand curve. Along π_0 economic profits are zero as the price equals marginal cost.



The two methods are illustrated in a single graph in Figure 8.11 where point **h** is the point of tangency between the demand curve and the firm's highest feasible isoprofit curve, and point **i** shows the marginal revenue equated to the marginal cost. Whether determined by the owner selecting point **h** (by the $mrs = mrt$ rule), or point **i** (by the $mr = mc$ rule) the price (p^m) and quantity sold (x^m) will be the same. M-Note 8.6 confirms that the methods give the same result.

M-NOTE 8.6 The two methods give the same result

To determine the combination of price and quantity sold that maximizes the profits of the firm we used two methods. This M-Note shows that the methods are identical (also shown in Figure 8.11).

The first was based on the demand curve as the frontier of the feasible set and the family of isoprofit curves as the basis for choosing the most profitable point on the demand curve. Using the inverse demand curve in Equation 8.20 gave us the following condition:

$$\text{marginal rate of substitution} = \text{marginal rate of transformation}$$

$$\frac{p-c}{x} = \beta \tag{8.36}$$

The second method showed that the maximum profit is the level of output such that

$$\text{marginal revenue} = \text{marginal cost}$$

$$\overline{p} - 2\beta x = c \tag{8.37}$$

We can rearrange Equation 8.36 to show that it is the same as Equation 8.37. We first replace p in Equation 8.36 by the equation for the inverse demand curve which gives the value of p for each value of x:

$$\text{marginal rate of substitution} = \text{marginal rate of transformation}$$

$$\frac{(\overline{p} - \beta x) - c}{x} = \beta$$

$$\frac{\overline{p} - c}{x} - \beta = \beta$$

$$\frac{\overline{p} - c}{x} = 2\beta$$

$$\overline{p} - c = 2\beta x$$

$$\overline{p} - 2\beta x = c$$

$$\text{marginal revenue} = \text{marginal cost}$$

8.8 THE MARKUP, THE PRICE ELASTICITY OF DEMAND, AND ENTRY BARRIERS

An important determinant of the amount of profit that a firm's owner will make at the profit-maximizing level of output and price is the extent of competition that the firm faces. The reason is that competition from other firms producing similar products limits the extent to which it will be profitable to raise prices substantially above costs.

The markup and the markup ratio

We can summarize the relationship between a firm's price and costs by the markup and the markup ratio. The **markup** is the firm's profit per unit or the price it charges per unit minus its costs per unit, or $p - c$. Remember that the firm's profits are given as follows:

$$\text{Economic profit} \quad \pi(x) = p(x)x - cx$$

$$\text{Markup or economic profit per unit} \quad \frac{\pi}{x} = p(x) - c \tag{8.38}$$

! REMINDER Recall from Chapter 7, that the price elasticity of demand is a measure of the extent to which the amount sold varies with the price. The price elasticity of demand, η is the percentage change in quantity, x, associated with a percent change in price, p:

$$\eta = \frac{\frac{\Delta x}{x}}{\frac{\Delta p}{p}} = \frac{\Delta x}{\Delta p}\frac{p}{x} = \frac{dx}{dp}\frac{p}{x}$$

Because the numerator and denominator will always be of opposite sign, η is negative. But when we say that the elasticity is "smaller" or "larger" we are referring to the absolute value, $|\eta|$.

Equation 8.38 is called the markup, because it measures the degree to which the owners of the firm "mark up" prices above the cost to produce the good. The **markup ratio** (μ) is the ratio of the markup to the cost:

$$\text{Markup ratio} = \frac{\text{Profit per unit}}{\text{Cost per unit}}$$

$$\mu = \frac{p - c}{c} \tag{8.39}$$

The markup ratio measures the degree to which the price exceeds cost. If prices were equal to costs, i.e. $p = c$, then the markup ratio would be zero ($\mu = 0$). As price becomes greater than cost, the markup ratio measures the extent of that difference as μ gets larger and larger.

In Figure 8.12 we illustrate the same monopolistically competitive firm as we showed in Figure 8.11. We now provide information on the profit-maximizing price (8) and quantity (4) when demand is described by the initial demand curve. Demand is quite elastic, meaning (remember) that an increase in prices would be associated with a substantial reduction in sales. This could be the case because there is a second firm selling a very similar product. An increase in the firm's price would motivate some of its customers to switch to buying from the other firm.

But if the firm could devise a way to make the demand curve less elastic, like the steeper demand curve in the figure, it would then have a different—and more profitable—constrained optimization problem to solve. This is why a firm's profit-maximizing strategies include product differentiation: advertising, trademarking, and design can make the firm's own product seem "more different" from other firms' products, so that fewer customers will switch in response to a price increase. As a result there are fewer close substitutes for the firm's product, which as you know from Chapter 7 makes the firm's demand curve less elastic.

MARKUP The markup is the difference between the price at which a good sells and its cost (including the opportunity cost of the capital goods used).

MARKUP RATIO The firm's markup ratio is the profit per unit of output (the markup) divided by unit costs.

The markup ratio and the elasticity of demand

Think about the owner of the firm facing this new demand curve. The (negative of the) slope of the demand curve has risen from 1 to 4, meaning that the expression $\frac{\Delta x}{\Delta p}$–the inverse of that slope is now 0.25. At point **h**–the price (8) and quantity (4) he is currently implementing–the price elasticity of the new demand curve has fallen from 2 prior to the change to 0.5 (using Equation 8.40 in the margin note).

If he raised the price 10 percent, his sales would fall by 5 percent, so his total revenue would increase. And if he produces less, then his costs would be lower. This means that raising prices is a win–win option: at the same time increasing revenues and cutting costs. So he will raise the price.

How much will he raise the price? In Figure 8.12 the firm facing the new steeper demand curve finds a new profit maximum at point **g** where both the markup and the owner's profit are higher than at point **h**, the initial profit maximum. In M-Note 8.7 we show that a constrained profit maximum such as point **h** with the old demand curve and point **g** with the new demand curve:

EXAMPLE Point **h** is on both the initial demand curve and the new demand curve. We use the equation for the definition of price elasticity of demand to calculate the elasticity in each case. Starting with the initial demand curve, we find:

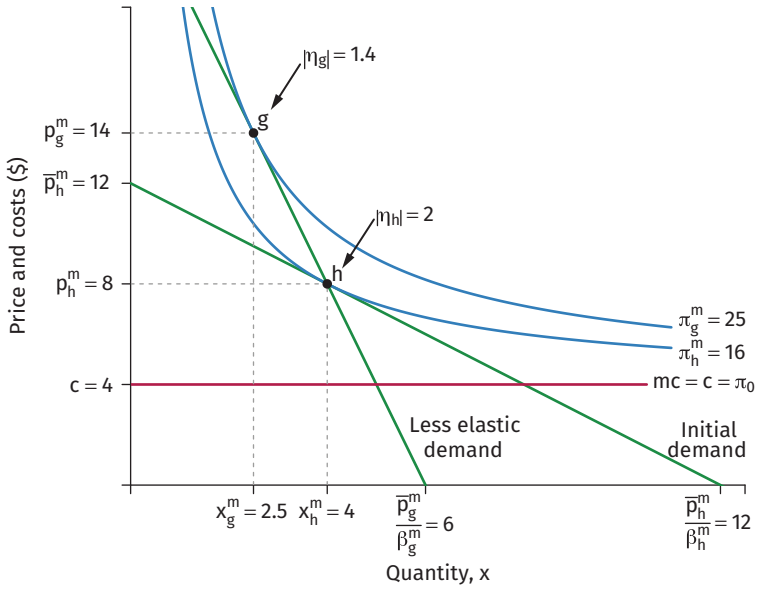
$$|\eta_h| = \frac{\Delta x}{\Delta p}\frac{p}{x}$$
$$= 1\left(\frac{8}{4}\right) = 2$$

On the new demand curve, we find $|\eta|$ as follows:

$$|\eta_h| = 0.25\left(\frac{8}{4}\right) = 0.5 \quad (8.40)$$

Because $|\eta_h| < 1$ on the new demand curve, the firm will maximize its profits by reducing the quantity produced and raising the price to p_g^m.

Figure 8.12 Increasing the markup ratio by making a demand curve less elastic. Initially (point **h**, as in the previous figure) the firm faces a steeper demand curve. So the $mrs = mrt$ tangency no longer holds. The new profit maximum is at point **g**, increasing both the markup ratio (10/4 vs 4/4) and the owner's total profits (25 vs 16).



M-CHECK When $|\eta| = 1$ a change in price has no effect on revenue which, using the demand curve, $x(p)$ to express the amount sold in terms of the price, we can write as $r = px(p)$. Then

$$\frac{dr}{dp} = x + \frac{dx}{dp}p = 0$$

$$\text{or,} \quad 1 = -\frac{dx}{dp}\frac{p}{x} \equiv |\eta|$$

The second line follows by dividing the first line by x and rearranging.

$$\mu = \frac{p-c}{c} = \frac{1}{|\eta|-1} \tag{8.41}$$

If $|\eta|$ is large, as will be the case when the firm has many competitors selling close substitutes for the firm's product, the markup ratio is close to zero (meaning that economic profits fall as competition increases). If, on the contrary, η gets closer to 1, then μ gets larger (μ tends toward infinity as $|\eta| \to 1$).

Does Equation 8.41 mean that if $|\eta| < 1$ the markup will be negative? No it does not. The reason is that the owner of the firm would never select a point on a demand curve at which $|\eta| < 1$ because it cannot be a profit maximum. You have seen above that when $|\eta| = 0.5$ raising prices increased total revenue and lowered total costs. So such a point could *not* be a profit maximum. This is a general rule. The profit maximum will always be at a point on the demand curve at which $|\eta| > 1$.

M-NOTE 8.7 **The markup ratio and the price elasticity of firm demand**

We derive the relationship that must hold at a profit maximum between the markup ratio (μ) and the price elasticity of demand (η). We begin with the total revenue of the firm $r(x, p(x))$ using the inverse demand curve $p(x)$. Then marginal revenue can be written:

$$\text{Product rule} \quad \frac{dr(x)}{dx} = p(x) + \frac{dp(x)}{dx}x$$

$$= p(x)\left(1 + \frac{dp(x)}{dx}\frac{x}{p}\right)$$

$$\text{so, marginal revenue:} \quad mr(x) = p\left(1 + \frac{1}{\eta}\right) \tag{8.42}$$

Recall that elasticity is a *negative* number because the demand curve is downward-sloping, and therefore $\frac{dp}{dx} < 0$. Consequently, we can rewrite the equation above using the absolute value of η, that is, $|\eta|$:

$$mr(x) = p\left(1 - \frac{1}{|\eta|}\right) \tag{8.43}$$

At the profit maximum we must have that the expression for marginal revenue just derived is equal to marginal costs, c.

$$mr(x) = p\left(1 - \frac{1}{|\eta|}\right) = c = mc(x)$$

$$\frac{p}{c}\left(1 - \frac{1}{|\eta|}\right) = 1$$

$$\frac{p}{c} = \frac{1}{1 - \frac{1}{|\eta|}}$$

$$\Rightarrow \frac{p}{c} = \frac{1}{\frac{|\eta|-1}{|\eta|}}$$

$$\Rightarrow \frac{p}{c} = \frac{|\eta|}{|\eta|-1} \tag{8.44}$$

continued

Now, subtracting $\frac{c}{c} = 1$ from both sides to get an expression for the markup ratio we have:

$$\mu \equiv \frac{p-c}{c} = \frac{|\eta|}{|\eta|-1} - 1 \tag{8.45}$$

Then we can substitute $1 = \frac{|\eta|-1}{|\eta|-1}$ for the 1 in the previous equation :

$$\mu = \frac{|\eta|}{|\eta|-1} - \frac{|\eta|-1}{|\eta|-1}$$

$$\mu = \frac{|\eta|-|\eta|+1}{|\eta|-1}$$

$$\mu = \frac{1}{|\eta|-1}$$

which is Equation 8.41.

Barriers to entry that raise the markup

Because the markup ratio and profits will be greater if the demand curve is less elastic, and this will be the case if there are fewer competing firms in a market, profits will be greater if it is difficult for new firms to enter a market to compete with **incumbent firms**. A **barrier to entry** is anything that makes it difficult for a new firm to enter a market to compete when the owners of incumbent firms are making economic profits. Sometimes barriers to entry are called "moats" because, like the protective water barriers around old castles, they protect the incumbent firms from intrusion by "outsiders."

Barriers to entry include:

- **Predatory pricing**: Incumbent firms may temporarily charge a price less than their average variable costs so as to force price reduction and to inflict losses on a new firm attempting to enter a market. Predatory pricing is illegal in many countries, but it is hard to secure convictions in these cases because it is difficult to prove that the strategies were specifically targeted at competitors.

INCUMBENT FIRMS Firms already selling in a market.

BARRIERS TO ENTRY Anything making it difficult for new firms to enter a market, including intellectual property rights, economies of scale in production, and predatory pricing.

PREDATORY PRICING Predatory pricing occurs when an incumbent firm charges a price lower than its (marginal) costs, seeking to drive its competitors out of business.

- *Economies of scale in production and learning-by-doing*: A new firm with initially limited output and experience will have a disadvantage until it can grow larger or accumulate experience in producing the good.
- *Intellectual property and licensing*: Exclusive use of trademarks, production processes, or knowledge protected by *intellectual property rights* and other government-enforced monopolies (e.g. licenses).
- *Naturally occurring barriers to entry*: Such barriers include limited access to natural resources or advantageous spatial locations owned by incumbent firms.
- *Network economies of scale in demand*: For many goods or services, the value to the consumer depends on the number of others purchasing the good. An entering ride–hail service could not compete with Uber and Lyft unless it already had a large number of drivers, which it could not get unless it had a large number of consumers.

Some barriers to entry are not the result of deliberate strategies by firms; they just exist by the nature of the product being produced. Economies of scale and naturally occurring barriers are examples.

But some barriers are constructed or at least heightened by the one or more incumbent firms. Predatory pricing, establishing monopolies on essential knowledge through intellectual property rights and licensing, and advertising a company-owned trademark are examples.

CHECKPOINT 8.7 Price elasticity of demand and profits Use the numerical values in Figure 8.12 to do the following:

a. Confirm that the profits at point **g** are 25 as shown.

b. Show that the elasticity of demand at point **h** is 2 as indicated, and is 1.4 at point **g** using Equation 8.40.

c. Show numerically that under the new, less elastic demand curve at point **h**, the initial profit maximum Equation 8.41 does not hold, and explain why.

8.9 APPLICATION: EVIDENCE ON THE MARKUP IN DRUG PRICES

Because the profit-maximizing markup ratio depends on the extent of competition that a firm faces, it differs substantially among firms, across different industries, and over time. In Chapter 9 (Figure 9.27) we provide evidence on the markup ratio for firms in the US economy as a whole, showing a sharp increase from around 0.2 during the mid-1980s to almost 0.6 three decades later.

For some firms the markup ratio is much greater than these economy-wide averages. A reason is the limited competition facing firms selling prod-

ucts in which they hold intellectual property rights such as trademarks, copyrights, or patents.

An example is the pharmaceutical sector, in which patents on drugs or ingredients of drugs constitute barriers to entry by competing firms. Only Bristol Myers Squibb (BMS) or those companies that it licenses can sell the hepatitis-C drug declatasvir. But BMS still has competition from Gilead Sciences, the only company that can sell another hepatitis-C drug, sofosbuvir.

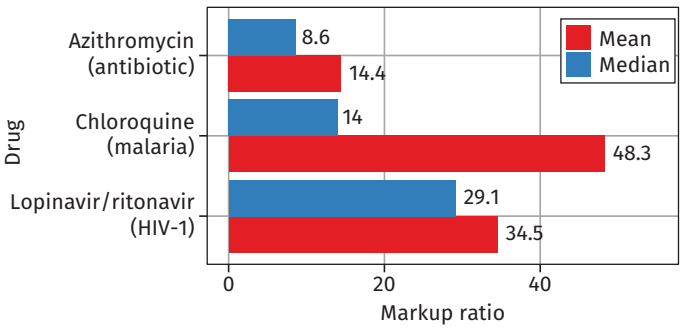
One set of estimates of markups on drugs is in Figure 8.13. Seeking some guidance on what a vaccine for COVID-19 might cost to produce as the pandemic swept the world in early 2020, a team of pharmaceutical scientists devised a measure of the "minimum cost" of producing what they considered as possibly similar drugs. (There was no vaccine for COVID-19 at the time.) For example they used an online database for tracking actual shipments of the necessary chemicals ("active pharmaceutical ingredients") and their prices to find the least-cost inputs.

The cost estimates include these ingredients (the major part of the total cost), and the costs of production (called "formulation and tabletting") and packaging. They also include a 10 percent markup over accounting costs that represents the opportunity costs of the capital goods used, and the cost of a profit tax. They assumed that production was at a level to exploit the possible economies of scale in production.

Their estimate is a measure of the marginal cost (and the average variable cost). It does *not* include costs unrelated to the production of the particular drug such as advertising, lobbying, legal, and other expenditures attempting to promote a favorable legal and regulatory environment for

Figure 8.13 Estimated markup ratios for drugs, 2020. The data shown are measures of $\mu = (p - c)/c$ based on a single minimum cost estimate (c) along with prices charged for the drugs shown in the seven to 11 countries for which this information was available. The blue bar for chloroquine, for example, indicates that the profits per treatment sold (that is $p - c$) were 14 times the minimum cost of the treatment (c). We show both the mean and the median because extraordinarily high prices in the US contribute to a higher mean markup than is representative of the countries taken as a whole.

Source: Authors' calculations based on data from Hill et al. (2020).



the companies concerned, research, and the costs of trials required for regulatory approval.

The price data that the researchers used to calculate the markup ratios in Figure 8.13 came from national drug price databases or online pharmacy sites. Where more than one price was available, the lowest price was used.

The middle set of bars in Figure 8.13, for example, is for a treatment of malaria, which the pharmaceutical scientists estimated has a cost of $0.30 for a 14-day course of the drug. The lowest price at which the drug was being sold was less than that, in Bangladesh, and considerably more than that in India, Malaysia, Sweden, South Africa, China, and the UK, where the same 14-week course of the drug was sold for $8. In the US the same course of the drug sold for $93. The estimated markups are based on these prices, which for the other drugs shown also include data from France, Brazil, Malaysia, Sweden, Turkey, and other countries.

We do not know if the firms in question actually achieved markup ratios of the amounts shown. For example they may have been using production methods that are *not* the least-cost solution to their cost-minimization exercise. Remember, the cost curve indicates the *minimum* cost of producing a particular amount, and firms may not have implemented this in combining the "active pharmaceutical ingredients" to produce their "final finished product."

Limited competition provides firms' owners with opportunities to increase their profits not only in the prices at which they sell their products, but also in the prices which they pay of inputs into their production.

CHECKPOINT 8.8 **Markups** Why do you think the drugs shown in Figure 8.13 are sold at prices so much higher than the minimum cost of production?

8.10 WILLINGNESS TO SELL: CAPACITY CONSTRAINTS AND MARKET SUPPLY

Recall that for a particular good x, the demand curves introduced in Chapter 7 provided answers to a hypothetical "what–if" question: Given the prices of all other goods and an individual's budget, how much of good x would she purchase if she could buy any amount she wished consistent with her budget when offered various prices for the good?

The **supply curve for a firm** provides an answer to a similar hypothetical question: If the price at which the firm could sell any amount it wished were p, what amount (x) of output would the owners of the firm choose to produce and sell? Answers to this question for all possible values of p give

SUPPLY CURVE (INDIVIDUAL FIRM) A supply curve provides the answer to the hypothetical question: What amount will be supplied for each given price? The individual firm's supply curve is the portion of the firm's marginal cost curve that is not less than the average variable cost.

us the supply curve that expresses the amount produced as a function of the hypothetical price, or $x(p)$.

The hypothetical question is a thought experiment by which we construct the supply curve as an economic concept. It is not something that owners of firms ask themselves. How much a firm places on the market will typically affect the price at which it can sell the goods. So firms do not take the price at which they will sell the good as a given. They choose both the price to charge and the quantity to market based on their costs and the demand for the product (using the constrained optimization methods described in sections 8.6 and 8.7).

With declining costs there is no firm supply curve

There is another reason why the owners of firms do not ask themselves the "what–if" question above about how much the firm would produce: in many cases the answer would be either "nothing" or "an infinite amount." To see why, consider the firm in Figure 8.15.

The firm's average costs are *declining* with increased output, while marginal costs are constant. Suppose we asked the owner of this firm the question: "How much would you produce if you could sell any amount you wished at a some price p?" The owner would think the question odd,

Figure 8.14 Supply and demand according to Bill Gates. Bill Gates of Microsoft tweeted his hand-drawn sketch of a figure similar to this one in 2018 with the heading "At the time I was in college this was basically how the global economy worked" adding "cost . . . increases as supply increases." He went on "But software doesn't work like this. Microsoft might spend a lot of money to develop the first unit of a new program, but every unit after that is virtually free to produce. . . . And it's not the only example: data, ebooks, even movies work in similar ways."[6]

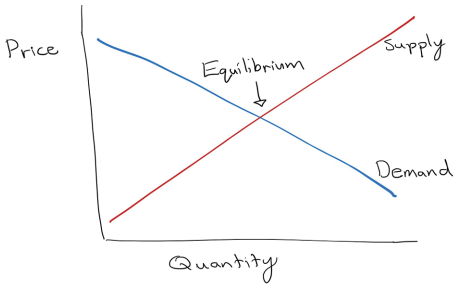
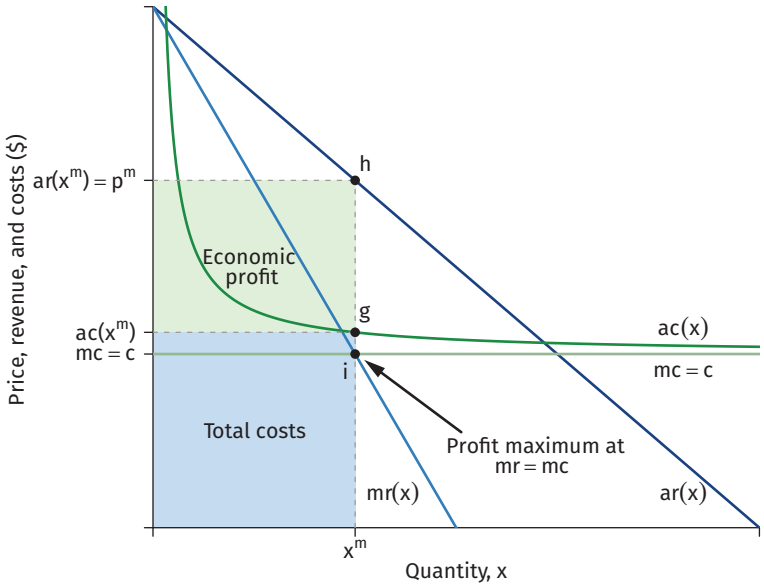


Figure 8.15 Profit maximization with declining average costs. Like the firm shown in Figure 8.10, this firm has fixed costs and linear variable costs and therefore its total costs are $c(x) = c_0 + c_1x$, its average costs are $ac(x) = \frac{c_0}{x} + c_1$ and the firm's marginal costs are $mc(x) = c_1$.



because she would of course know that she would not be able to sell any amount she wanted at any positive price.

But if she was willing to answer the question anyway, she would notice that as long as she could produce at a level such that average costs were less than the price specified, she would be making economic profits. The more she produced, the more profits she would make.

So, if she took the question literally, she would say that she would produce an infinite amount of goods. Or, more practically, as much as she possibly could, that is, until her firm ran into some constraint on its capacity to produce (what is called a *supply constraint* not shown in the cost function). If she were asked how much she would direct her firm to produce if the price were $p < mc$ her answer would be "nothing" because no matter how much she produced her economic profits would be negative.

From this example we see that for a firm whose average costs fall as more output is produced unless there is something that eventually limits how much the firm can produce, the firm supply curve does not exist. The same is true for a firm with constant average costs (equal to marginal costs): for a $p > ac$ the owners would want to produce an infinite amount, and for $p < ac$ they would in the long run produce nothing; they would go out of business.

Firms' supply curves when they face capacity constraints

Practically speaking, firms *do* face supply constraints (also called capacity constraints) that limit their production. A supply constraint is some level of production $\underline{x}$ such that it is not possible (at any cost) to produce more than the constraint. Sources of supply constraints include:

- *Natural (long-run) limits on expanding one or more input*: if arable farmland is limited in supply, for example, or the density or zoning of urban locations prevents expansion.
- *Short-run limits on increasing capital goods* or some other input.
- *Intrinsic limits on administrative capacity*: it may be simply impossible to coordinate the activities of a very large organization.

These supply constraints may be particularly binding in the short run.

To derive a market supply curve we will simplify by assuming the following:

- Firms have constant average and marginal costs up to some supply constraint, that is, $mc = ac = c$.
- Firms *cannot* produce beyond the supply constraint, $\underline{x}$.
- Though firms produce an identical product (such as a sugary drink), they differ in their marginal and average costs.

Differences in their costs may arise, for example, from their machinery being more or less up to date, or their management being more or less competent.

This means that an individual firm's supply function is L-shaped like those shown in Figure 8.16. Firms I, J, and K are all firms in the sugary drinks industry, producing liters of sugary drinks that they would like to sell in the sugary drinks market.

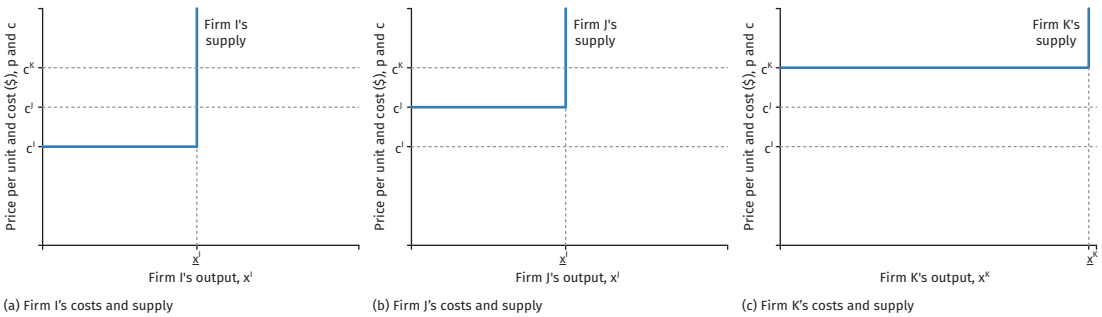
Their costs curves are flat and equal to their marginal and average costs (c^I, c^J, and c^K) up to their capacity constraints ($\underline{x}^I$, $\underline{x}^J$, and $\underline{x}^K$), and then *vertical*. In this instance, the firms' capacity constraints could be determined by sizes of their plants, the availability of ingredients, or other natural limitations that limit the liters of sugary drinks they can produce in the short run.

M-CHECK The marginal cost curves are vertical at the capacity constraints because (in the short run) the firm simply cannot produce more than that amount (implying that marginal costs are infinite at that point).

Though they appear very different, the capacity-constrained L-shaped supply curves of the individual firms have an interpretation similar to the individual demand curve. Remember, the demand curve indicates the buyer's *maximum* willingness to *pay* to acquire an *additional* unit of the good. Similarly, the height of the supply curve is the *least* amount that one could *pay* to a seller—as a take-it-or-leave-it offer—to purchase an additional unit.

Therefore you can think of the supply curve as a minimum **willingness-to-sell** curve. The minimum level at which Firm I is willing to sell is a price, p, that is at least as great as its costs, c^I. For any price greater than c^I, Firm I is willing to sell a quantity up to its capacity constraint, $\underline{x}^I$. The other

Figure 8.16 Costs, capacity constraints, and supply curves of three firms. The marginal (and average) costs and capacity constraints of three firms are shown: Firm I, Firm J, and Firm K. They are arranged from lowest cost (Firm I) to highest cost (Firm K). Firm I and Firm J have the same capacity constraint ($\underline{x}^I = \underline{x}^J$), but I's marginal and average costs are lower than J's ($c^I < c^J$). Firm K has a larger capacity than Firms I and J, that is, $\underline{x}^K > \underline{x}^I = \underline{x}^J$, but Firm K's costs are greater than I's and J's, $c^K > c^J > c^I$.



(a) Firm I's costs and supply (b) Firm J's costs and supply (c) Firm K's costs and supply

WILLINGNESS TO SELL The lowest price at which a seller would be willing to sell a good is the willingness to sell.

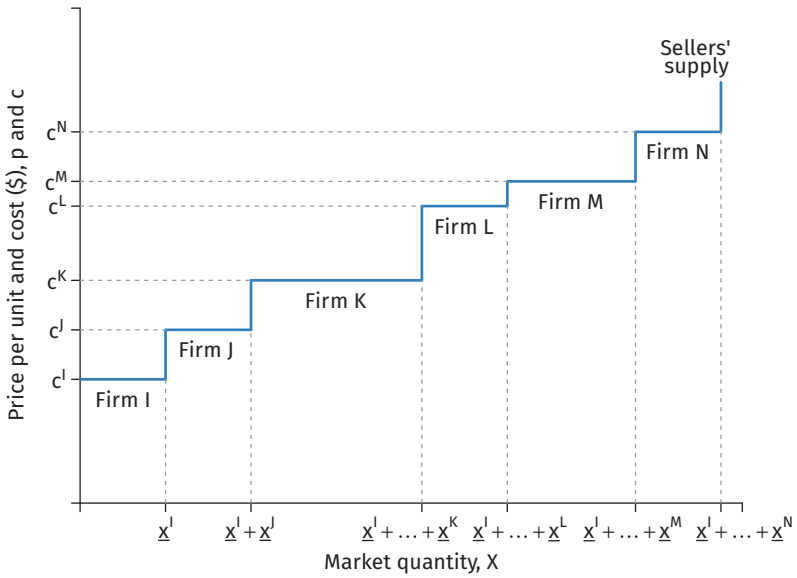
firms behave similarly: they are willing to sell goods for any price greater than their minimum willingness to sell or willingness to accept up to their capacity constraints.

CHECKPOINT 8.9 Willingness to sell Explain how the supply curves in Figure 8.16 indicate the firm's minimum willingness to sell. How is this similar to the demand curve as a measure of buyers' maximum willingness to pay? Why is the first a minimum and the second a maximum?

8.11 ECONOMIC PROFITS AND THE MARKET SUPPLY CURVE

Just as we constructed a market demand curve by horizontally summing individual willingness to pay demand curves in Chapter 7, the willingness-to-sell step function in Figure 8.17 is the horizontal sum of the cost curves for firms I, J, and K shown in Figure 8.16, along with the costs of three additional higher-cost firms L, M, and N with their corresponding cost curves and capacity constraints.

Figure 8.17 A stepwise willingness to sell function. The step function is the horizontal sum of the supply curves of the firms: I, J, K, L, M, and N over the ranges of their given capacity constraints. The firms are ordered from least cost (Firm I with c^I) to highest cost (Firm N with c^N). The cost curves for firms I, J, and K are shown in Figure 8.16. The costs of firms L, M, and N are all higher than I, J, and K. The firms' capacity constraints differ, as shown by the differing lengths of the outputs they produce on the horizontal axis.

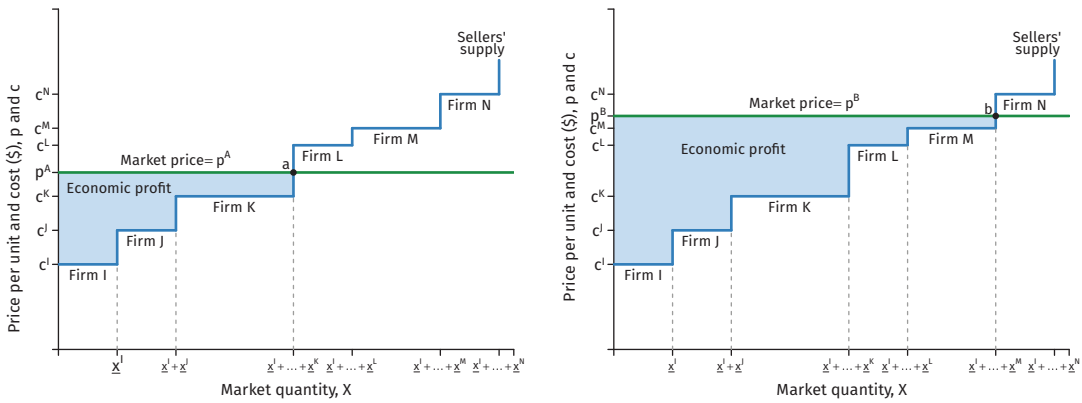


Remember, the supply curve provides answers to the hypothetical question: At some particular price, p, how much output will be supplied? Figure 8.18 shows two different hypothetical prices: p^A (a low price) and p^B (a higher price). The figure shows that at the lower price, p^A, only three firms have a willingness to sell (considering their costs) that is less than the given market price.

Presuming there is sufficient market demand at this price (remember, the price is just a "what–if" thought experiment), each firm will sell a quantity equal to its capacity constraint and will sell a total number of units equal to $x^I + x^J + x^K = X$. This makes the price and quantity sold at **a** in panel (a) one point on the market supply curve: $X = X(p)$. The other firms L, M, and N will not produce anything.

The higher price p^B, exceeds the minimum willingness to sell of two additional firms: Firm L and Firm M with costs c^L and c^M. So they enter the market, and along with other lower-cost firms they produce at their capacity constraints ($\underline{x}^L$ and $\underline{x}^M$). This means that in panel (b), point **b** is another point on the market supply curve. The supply curve for the entire market is constructed by repeating the same exercise with other

Figure 8.18 Economic profits with the step-wise supply function. In panel (a) the price is p^A and three of the firms—I, J, and K—produce at that price because the price is greater than each of their costs, $p^A > c^I$, $p^A > c^J$, and $p^A > c^K$ meaning that each of them makes economic profits per unit equal to the price minus their costs. Therefore, $\pi^I = x^I(p^A - c^I)$, $\pi^J = x^J(p^A - c^J)$ and $\pi^K = x^K(p^A - c^K)$. These profits are shown in the shaded blue area above the costs and below the price. In panel (b) the price is p^B and, following similar reasoning to the above, Firm L and Firm M enter the market as the price is now greater than their marginal costs. They make corresponding economic profits and the profits of the incumbent firms' increase. Firm N's cost remains above the price and therefore it does not produce any output.



(a) A low price and low economic profits

(b) A higher price and high economic profits

hypothetical prices. So the step function giving the willingness to supply is itself the **market supply curve**.

Economic profits, market supply, and firm entry or exit

For each unit that a firm sells at a price above its cost per unit, it receives an *economic profit*, namely the difference between the price at which it sells the good and its costs. Therefore for Firm I, when it sells at price p the largest output it can produce is $x^I = \underline{x}^I$, so its profits are $\pi^I(x^I) = x^I(p - c^I)$. In similar fashion, the other firms choose an output level and make economic profits based on their costs, supply constraints, and the price at which they sell their output.

The total profits at the two different prices are shown by the shaded blue areas in Figure 8.18. Firm N does not enter the market even at the higher price, because its economic profits would be negative, i.e. accounting profits less than the opportunity cost of the capital goods used. Similarly, if the price were at p^B with all but Firm N supplying to the market, but then it fell to p^A, Firms M and L would cease production, as they would be making negative economic profits.

Economic profits and losses therefore regulate whether firms enter the market (or not, like Firm N) or cease production (like Firms M and L at the lower price), a topic to which we return in Chapter 9.

In most cases we have many more than five firms supplying the market so that showing each of them separately as in Figure 8.18 would be impractical. In Figure 8.19 we present a smoothed version of the same market supply curve. The smoothed supply curve we show based on the previous stepwise functions is linear, but it can have any shape as long as it is upward-rising.

CHECKPOINT 8.10 **The market supply curve** In Figure 8.19, why does the market supply curve slope upwards?

8.12 PERFECT COMPETITION AMONG PRICE-TAKING BUYERS AND SELLERS: SHARED GAINS FROM EXCHANGE

Even before Adam Smith's *Wealth of Nations*, economists developed models of how market competition works. Today, the most widely taught of these in introductory economics—the model of perfect competition—is based on the intersection of the supply and demand curves as illustrated in Figure 8.20,

SUPPLY CURVE (MARKET) A supply curve provides the answer to the hypothetical question: What amount will be supplied for each given price? A market supply curve is the horizontal sum of the firms' supply curves.

Figure 8.19 The smoothed market supply curve. The figure shows a smoothed version of the stepwise function in Figure 8.17. Points **a** and **b** here correspond to points **a** and **b** in that figure.

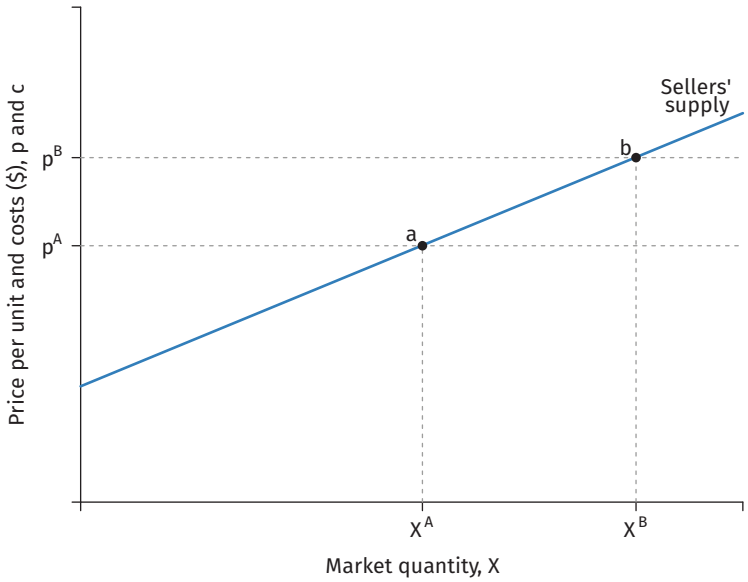
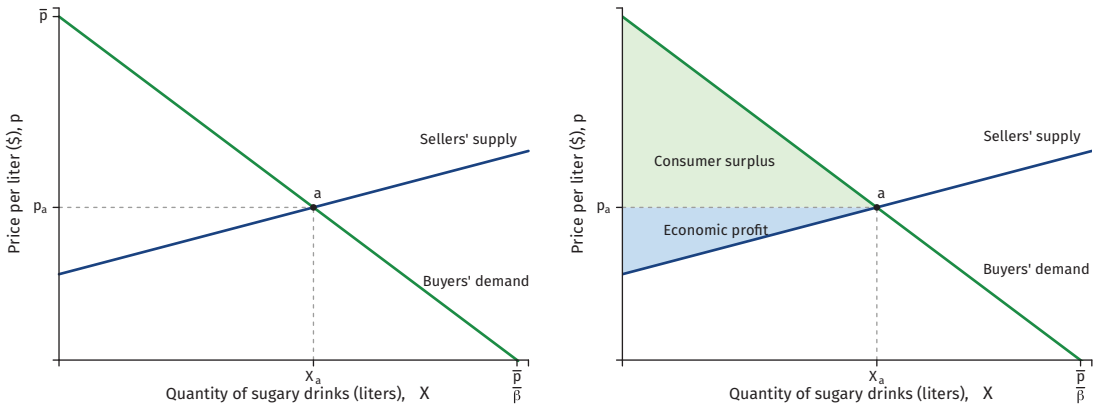


Figure 8.20 Competitive equilibrium in the market for sugary drinks. In panel (a), the intersection of the buyers' demand curve and the sellers' supply curve gives the competitive equilibrium in the sugary drinks market. The intersection of the curve provides the market-clearing price: the price at which quantity supplied by the sellers equals quantity demanded by the buyers. The intersection is shown by point **a** with quantity demanded X_a and market price p_a. In panel (b), the consumer surplus is the area shaded in green beneath the demand curve and above the market price, p_a. The economic profits are shown by the shaded area in blue above the supply curve (the minimum willingness to sell) and the market price, p_a.

! REMINDER Consumer surplus is the difference between a buyer's maximum willingness to pay and the price they actually pay for the good.



(a) The demand and supply market equilibrium

(b) Consumer surplus and economic profit at market equilibrium

using the sugary drinks market as an example. In panel (a) we show an upward-rising supply curve (like the smoothed variant of the one just derived) and a downward-sloping demand curve consistent with those derived in Chapter 7.

A Nash equilibrium of price-taking buyers and sellers

The price and quantity given by the intersection of the two curves is the equilibrium of what is called the **perfectly competitive equilibrium model**. This is because the hypothetical questions that we used to construct the supply and demand curves correspond to a way that economists have commonly represented as perfect competition. In both cases we asked the buyer or seller to hypothetically take the prices as given, and then say how much they would want to buy or sell assuming they could choose any such amount that they pleased.

EXAMPLE In this video (https://tinyurl.com/y4tcdrcn) Richard Freeman explains why you can't outsource responsibility (from the CORE project. www.core-econ.org).

Taking a price as *given* is called **price-taking**. It means not considering different prices at which one could buy or sell.

Three things to note about price-taking:

- *Price-taking is a choice of about strategies*: Price-taking does *not* mean that a buyer or seller *cannot* change the price (of course they can, a buyer or seller can post any price they wish at which they are ready to buy or sell); price-taking is just an assumption about the strategies that a player in a game—a buyer or seller—considers when deciding what to do.
- *Price-taking is a behavior*: Price-taking may make sense in some situations but not in others. Price-taking is *not* an attribute of a firm or other seller or buyer, it is a behavior that buyers and sellers may adopt as a best response to a particular situation. So when you use the term price-taking you should be thinking about the situation, not the person or the firm.
- *Price-taking cannot be an assumption*, unless it has been demonstrated as a result, showing that taking prices as given is the best that people can do given the strategies adopted by others.

Here is how price-taking can be a result of a model not a presupposed assumption. From the way we constructed the market supply curve we

PERFECTLY COMPETITIVE EQUILIBRIUM In a perfectly competitive equilibrium, supply equals demand, and neither buyers nor sellers can benefit by altering their price or quantity.

PRICE-TAKING Price-taking is a strategy that an economic actor may follow, taking as given the prices at which one might buy or sell.

know that *if* the price is p_a and firms are acting as price-takers, *then* firms will supply a total of X_a. From the way we constructed the market demand curve, we know that *if* the price is p_a and buyers are price takers, *then* the total quantity demanded will be X_a.

If this is the case, then the intersection of the supply and demand curves (point **a** in Figure 8.20) is a Nash equilibrium: both buyers and sellers are best responding to the price p_a that is being charged or offered by the *other* side of the market.

Because supply equals demand, there are no sellers wishing to sell at that price and unable to find buyers, and there are no buyers wishing to buy at that price and not finding a seller. The only relevant actors, therefore, are those transacting at the price p_a.

They are all best responding to the strategies of everyone else, given the limited strategy sets available to them in this model. Remember, firms' only choices concern prices and quantities; they do not have the option of choosing other strategies such as product differentiation. This shows that if the buyers and sellers are acting as price-takers, then the intersection of the supply and demand curves—both constructed on the basis of a hypothetical price-taking question—is a Nash equilibrium: no actor could do better by adopting a different strategy. In section 8.15 we will ask whether being a price-taker is indeed the "best they can do" if the strategy sets of the buyers and sellers are expanded to include such things as product differentiation. But first let's see how the model works.

! REMINDER We know that owners of firms will find it profitable to differentiate their product so as to become monopolistic competitors, and that monopolistically competitive firms will not act as price-takers: they will affect the prices at which they sell their outputs by selecting a level of production.

HISTORY In his 1890 *Principles of Economics* Alfred Marshall labeled the area between the market-clearing price and the supply curve as a "producer's surplus" analogous to his consumer's surplus.[7] Others have called this the "seller's surplus." Like Marshall starting from the fact that firms differ in their costs, we have derived a "willingness to sell" based market supply curve for which at either the firm level or for the market, the area between the price and the supply curve is a rent. This is because it is a payment above the seller's next-best alternative which is to not produce and sell the good at all, or to sell it at cost. But this is just economic profit, so we do not introduce Marshall's term.

Shared gains from trade and Pareto efficiency

In Figure 8.20 (b), we show the consumer surplus and economic profits—the gains from exchange or rents—which result if the price is p_a.

Consumer surplus and economic profits are similar in two ways:

- They are both rents, that is utility (for consumers) and income (for firms' owners) in excess of their next-best alternative (not buying and not producing and selling, respectively).
- They are similar in the dimensions in which they are measured, namely, money. This means that as long as we assume that an extra euro is worth the same to all consumers and all owners, we can do more than add up the surpluses received by consumers or owners. We can add the two classes of rent together. So the area between the supply and demand curves from the origin to however many goods are transacted is the total rent made possible by the transactions.

There are three important characteristics about the price and quantity at the intersection of the supply and demand curves:

- **Market clearing**: Because the outcome (p_a, X_a) is a point on *both* curves, the amount supplied is equal to the amount demanded, so there is no excess demand (demand greater than supply) or excess supply (supply greater than demand).
- *Price is (approximately) equal to marginal cost*: To see this, remember that costs differ among firms, but for a given firm costs are constant up to some supply constraint. So over this range of the firm's output, average and marginal costs are equal. Figure 8.18 shows that at any point on the supply function–like **a** or **b**–the price is either *equal* to the average and marginal cost of the highest-cost firm producing for the market, or (as is the case with **a** and **b**) it is *between* that cost and the marginal and average cost of the lowest-cost firm that is not producing for the market.
- *Total rents (consumer surplus plus economic profits) are maximized*: A consequence of the above result–that price is approximately equal to the marginal cost of the highest-cost producer in the market–is that there is no other price and quantity that could feasibly be transacted for which the sum of consumer surplus and economic profits would be larger.

The last point means that the Nash equilibrium of this interaction of price-taking buyers and sellers is Pareto efficient. Starting from the price–quantity combination (p_a, X_a) there is no alternative technically feasible price and quantity transacted under which consumers could be made better off without making owners worse off and conversely. The buyers could be made better off, of course, if the price were lower, but this would make the sellers worse off. And vice versa for a price increase.

CHECKPOINT 8.11 A price-taking Nash equilibrium Explain why at quantity X_a in Figure 8.20, all buyers and sellers acting as a price taker and transacting at the price p_a is a Nash equilibrium.

8.13 THE EFFECTS OF A TAX: CONSUMER SURPLUS, PROFITS, TAX REVENUES, AND DEADWEIGHT LOSS

We can use the perfectly competitive model to analyze the effects that changes in the supply curve or the demand curve have on prices and amounts transacted, and the resulting changes in consumer surplus and economic profits. We will illustrate this application of the model by returning to the tax on sugary drinks introduced in Chapter 7.

MARKET CLEARING A market clears when the amount supplied is equal to the amount demanded.

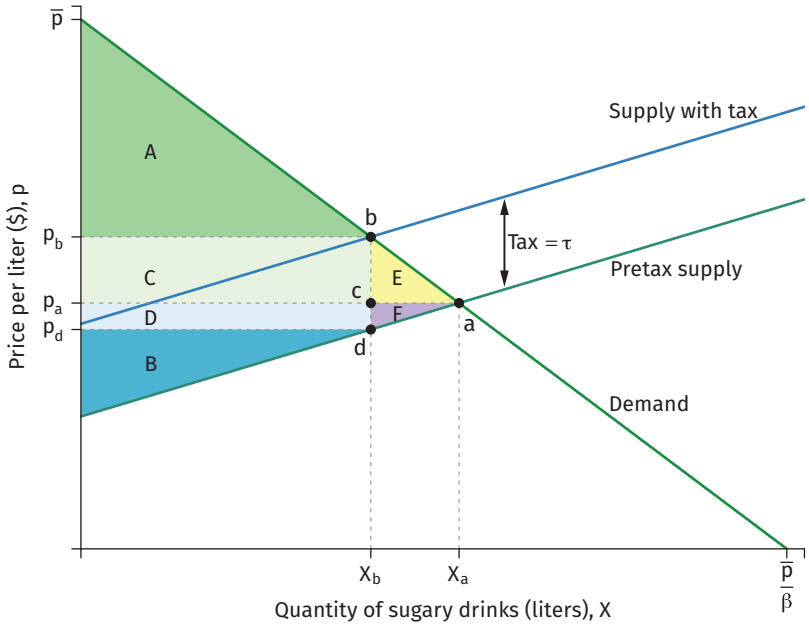
As we did in Chapters 5 and 7, we will use comparative statics: compare the Nash equilibrium after the shift in demand or supply with the status quo Nash equilibrium or the Nash equilibrium *before* the change.

- The word *static* refers to the Nash equilibrium because at a Nash equilibrium there are no reasons for the actors to change what they are doing.
- The process is *comparative* because we compare two or more states before and after a change.

We consider a tax that is imposed on firms producing sugary drinks at the rate τ per liter of the drink. Figure 8.21 shows the demand curve for sugary drinks and the supply curve prior to the tax with the Nash equilibrium price p_a and the quantity sold X_a indicated by point **a**, along with the respective consumer surplus and economic profits.

When a tax is imposed, it increases the costs for each firm by the amount of the tax per unit, τ. So, if Firm F's cost per unit before the tax was c^F, then the cost after the imposition of the tax is $c^F + \tau$. Because the supply curve

Figure 8.21 The effects of a tax on sugary drinks. Before the tax, the total quantity of sugary drinks sold is X_a with price p_a. After the tax, the supply curve shifts upward because of the increased cost imposed by the tax, with a corresponding decrease in quantity demanded to X_b and a higher price p_b. Consumers and firm owners both bear some cost of the tax. With the tax, only those firms with lower marginal costs, indicated by point **d** will continue to produce and sell the sugary drinks.



is the horizontal summation of the cost curves of each of the firms and the cost for every firm has risen by the amount of the tax, the market supply curve shifts upward by that amount too, as is shown in Figure 8.21. Imposing the tax therefore reduces the amount of sugary drinks that will be produced at any given price.

The new Nash equilibrium is point **b** with price p_b and quantity sold X_b. The intended effect of the tax–a reduction in the quantity consumed–is achieved: consumption falls by the amount $X_a - X_b$.

From the figure we can identify the following changes:

- *Consumer surplus lost*. The consumer surplus which before the tax was the areas $A + C + E$ is now just A because *fewer* units are purchased and at a higher price. In Chapter 7 we noted that the consumers of sugary drinks tend to be poorer than average (in the US) and raised the concern that the tax might be thought to unfairly burden them.
- *Economic profits lost*. But the consumers are not the only ones who have suffered losses. The total economic profit, too, has been reduced. Before the tax it was the area $D + F + B$ but because the amount sold is *less* and the price minus the tax paid per unit is now lower, economic profits are just B after the tax.
- *Tax revenues and government services gained*. There is now a substantial amount of tax revenue equal to the tax times the number of units sold, or areas $C + D$ in the figure. These revenues support governmental programs–education, public safety, income transfer programs for the less well-off–that provide benefits for both consumers and owners.
- *Deadweight loss*. The area of the triangle $E + F$ is the *deadweight loss* associated with the tax. The top part of it (E) is consumer surplus lost by consumers (you already know this from Chapter 7); the lower part (F) is the firm's owner's lost economic profits.

To understand why the tax creates a "deadweight loss," think about two contrasting effects on the total rents enjoyed in the form of consumer surplus and economic profit:

- *Redistribution of the rents*: From consumers and owners on the one hand to the government (and to those who benefit from government-financed services); here, it is the case that what the consumers or producers *lost* the government *gained*.
- *Reduction of total rents*: This is the deadweight loss triangle; and it is not a transfer from one group to another, it is a quantity of benefits that existed prior to the tax that is lost.

We cannot say that the owners or consumers are worse off as a result of the tax. The tax revenues resulting from the policy may finance public policies that confer benefits on both groups that more than offset the rents that

they lost either in transfers to the government or as deadweight losses. And all of those affected by the tax may benefit from the reduction in illnesses related to the overconsumption of sugar and the costs that these inflict on family members, taxpayers and others, including the elevated cost of insurance for the entire population associated with these illnesses.

CHECKPOINT 8.12 **Consumer surplus** Explain the meaning of each of the colored and lettered areas in Figure 8.21, starting with the situation before the tax and then looking at the post-tax situation.

8.14 APPLICATION: THE DISTRIBUTIONAL IMPACT OF PUBLIC POLICIES—RENT CONTROL

We have illustrated the supply and demand model by the tax on sugary drinks that would internalize the external effects of the actions that people take, thereby mitigating a coordination failure. But the same model can also be used to study a policy that is motivated by concerns about the unfairness of the distribution of wealth or income.

Rent control is a legally binding limit on the rents that landlords can charge tenants. Landlords (owners of housing that is rented out) are typically much wealthier than the people they rent to. Rent control is advocated as a way to redistribute income from the landlords to the tenants. Adequate housing is also considered by many to be a **merit good**, one like access to healthcare, education, voting, and fair trials that on moral grounds should be available to all irrespective of their income. This is a second reason commonly proposed in support of rent control. Rent control laws are common in some major US cities, including Los Angeles, San Francisco, New York, and Washington DC. Rent control is typically bundled with restrictions on the conditions under which a landlord can evict a tenant. In other counties residential rents are also regulated by governments. In Germany, for example, there are legal restrictions on rent increases in excess of the rate of inflation (so that the real value of rents varies little over time).

Rent control has similar economic logic to the minimum wage: it seeks to improve the economic conditions of less well-off people (renters, low-wage workers) by imposing a price (a lower rent, a higher wage) that favors their

RENT CONTROL A policy regulating the rent that a landlord can charge, most commonly limiting the size of a rent increase that is permitted.

MERIT GOOD A merit good is one that it is thought on moral grounds should be available to all irrespective of their income.

! REMINDER The term "rent" in economics has two uses. It commonly refers to the payment for the temporary use of some piece of property such as a tenant pays to her landlord. But it is also used more generally as the amount of utility, profit, or other benefit that a person receives in excess of their next-best alternative. The two uses are bound to be confusing when thinking about the housing market, so we will use the expression "economic rent" for the second use of the term.

interests. As with the minimum wage example we use in Chapter 11, a policy change like rent control produces winners and losers.

In 1994 San Francisco voters approved rent control on rental housing built before 1980. Landlords responded by demolishing older rental structures and building new rental housing that was exempt from controls. Many offered their tenants large sums of money to vacate their apartments. The vacated apartments could then be converted to owner-occupied condominiums, which were also not covered by the new law.

By comparing what happened in otherwise similar rental units built before and after this key date, researchers were able to assess the impact of the rent control policy. The gains to tenants in the apartments affected by rent control were substantial, averaging between $2,300 and $6,600 per year, and totaling $214 million annually.[8]

But the reduction in the supply of rental housing resulted in a city-wide average *increase* in rent of 5 percent. So, among the losers, in addition to landlords, were those who were renting units *not* covered by the new law.

The rental housing market: Renters' surplus and landlords' economic profit

We can use the model of supply and demand to study the impact of rent control. Recall that in Chapter 7 the supply and demand model allowed us to identify two components of the gains from trade:

- *Consumer surplus*: because for most buyers their willingness to pay exceeded the price they paid.
- *Economic profit*: because the price at which a good is sold exceeds the marginal cost of its production, contributing to the profits of the owners of the firms producing the good.

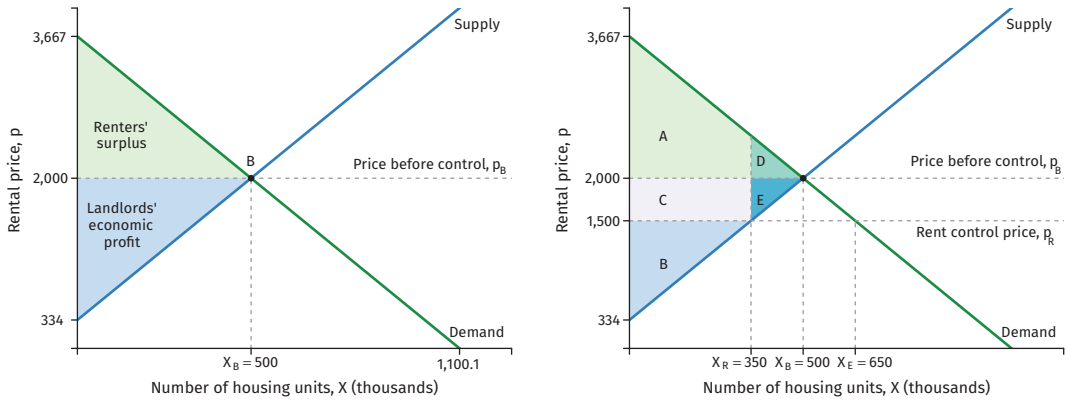
Here we adapt those concepts to the rental market, so that we now have:

- *renters' surplus* arising from the fact that for most renters their willingness to pay for their apartment exceeds the rent they actually pay (similar to consumers' surplus); and
- *landlords' economic profit*, which exists because the rent that most landlords receive exceeds the marginal cost of providing a unit of housing to the market.

Figure 8.22 illustrates these concepts at the equilibrium of a hypothetical rental market. To make sense of the model assume that there are two classes of people in a city: landlords and renters. Renters considerably outnumber landlords, which in a democracy gives them the possibility of passing legislation limiting the rents that landlords can charge.

The horizontal axis is the number of units of housing. To simplify we assume all housing units are identical in quality and that landlords are

Figure 8.22 A model of supply and demand for housing with rent control. A price control *lowers* the price of housing from p_B to p_R and results in excess demand $(X_E - X_R)$. Renters obtain renters' surplus (consumer surplus) that is *greater* than they previously had at a higher price. Landlords have a *lower* economic profit than before the rent control. As a consequence of the rent control there is lost consumer surplus and economic profit, that is, deadweight loss. The deadweight loss is similar to the deadweight loss of taxes that we considered in Chapter 7 and earlier in this chapter in section 8.13. To make this figure we used the following data: Demand: $p(X) = 3667 - \frac{10}{3}X$. Supply: $p(X) = 334 + \frac{10}{3}X$.



(a) Housing market clears before rent control

(b) Housing market with rent control

unable to charge different rents to different people. So there will be just a single rental price, which is measured on the vertical axis. The supply curve tells you, for any given rent, how many units of housing will be offered. A higher price will bring more units onto the market even in the short run as landlords find ways of converting unused space into apartments. In the long run, higher rentals will raise the profitability of owning rental apartments and stimulate new construction.

The demand curve provides the answer to the question: If the rental price is p, how many units of housing will be demanded? At a lower rent, more units are demanded, as more people choose to live in the city, or not to live with their parents or with roommates.

Rent control reduces total benefits from exchange and rearranges who gets it

The introduction of the rent control reduces the rental price to p_r. The landlords respond by supplying fewer units, reducing the number available to X_R. With fewer units being rented, notice that the willingness to pay of the "least willing" renter (the height of the demand curve at X_R) exceeds the marginal cost of putting additional units on the market. Therefore, there are people who would have been willing to rent units beyond the X_R, being

offered at a price exceeding the marginal cost. So the demand for rent-controlled housing exceeds the supply of rent-controlled housing. Rent control has two effects:

- *Redistribution to renters*: A portion of what was before the landlords' economic profit, is now part of the renters' surplus. This was the intended effect of the policy.
- *Reduction of the total of the renters' surplus plus landlords' economic profit*: The deadweight loss (foregone gains from trade) resulting from the reduced supply of rental housing under rent control is partly lost by renters (the top triangle of the deadweight loss space) and partly by landlords (the bottom triangle).

The net effect of these two changes is that landlords definitely lost. Their economic profit is *less* than before for two reasons: first, they experienced some of the deadweight loss (Area E); second, they lost some of what was before their share of the total gains from trade to the renters (Area C).

Evaluating the effect on the renters is more complicated. Like the landlords, renters experienced some deadweight loss, but they also gained some of what was previously landlords' surplus. Their net gain is area C minus area D.

In Figure 8.22 the surplus gained at the expense of the landlords is greater than the deadweight loss experienced by the renters. So, the policy benefited them, as intended, even though it reduced the supply of housing. Rent control is a way of dividing up a smaller pie, with a larger slice going to the renters. Because the pie is smaller as a result of the deadweight loss (Areas D plus E), the loss by the landlords (Areas E plus C) must be greater than the surplus gained by the renters (Areas C plus D).

Not only is the pie smaller, but a different supply curve (a flatter one) would have shown that rent control hurt the less well-off, the renters as a group, rather than helping them as it did in this example. The costs inflicted on the renters in the form of deadweight loss could have exceeded the gains they made by capturing some of what previously had been the landlords' surplus.

CHECKPOINT 8.13 Distribution of surplus under rent control Using a diagram similar to Figure 8.22, sketch supply and demand curves such that the costs experienced by renters (their "share" of the deadweight loss) will *exceed* the gains made possible by the fact that renters get a larger share of the total gains from trade.

Is there a better way to help the less well-off? A bargain between renters and landlords

Is there a better way to help the less well-off? Suppose the renters have joined an organization to promote their interests, and the president of the

Renters' Association sits down for a talk with the head of the Landlords' Association.

Because it was she who had asked for the meeting, the renter's representative starts off: "We are willing to vote to rescind the rent control if you and your landlord pals will simply transfer some money to us so that we are as well-off or better than we are under rent control."

"How much would you need?" is his reply. She simply hands him Table 8.4.

From Table 8.4 you can see that if the landlords transferred $137.5 million per month to the renters, the renters would be as well-off as they would be under rent control, and the landlords would be much better off (paying $137.5 million directly to the renters is better than losing a total of $212.5 million in lower rents and deadweight losses). Buying off the renters by paying them directly (rather than enduring the losses imposed by rent control) would strike the landlords' representative as a bargain. Of course, the president of the renters group would be quick to point out that were the landlords to transfer $212.5 million to the renters, then the landlords would be no worse off than they were under the rent control, and the renters much better off (getting a transfer of $212.5 million beats the net benefits to the renters from lower rents but on a reduced number of apartments rented).

The two might then bargain and agree on some intermediate amount, under which both landlords and renters would be better off.

This episode of course is fanciful. It is difficult of think of how the transfers from landlords to renters could take place in any practical way. But it underlines an important objective: if possible, policies to grant a larger slice of the pie to the less well-off should be designed to make the pie larger or at least not make it smaller.

HISTORY This kind of bargain is an example of what is called a "Coasean bargain" named for Ronald Coase (1910–2013) who proposed exactly these kinds of bargains between people in cases where one imposes a negative external effect on the other. He showed that decentralized bargains between people can achieve Pareto improvements over the status quo outcome. Coase won a Nobel Prize in Economics in 1991 for his work on external effects, transaction cost, and organizational structure. We explore Coasean bargaining in great depth in Chapter 14.

Weighing the gains and losses to different groups in society

The fact that in the case of rent control, the landlords' losses must exceed the surplus gained by renters is *not* a reason to oppose the policy: remember, the rent control policy was intended to help the renters, and it did.

Table 8.4 Monthly gains and losses compared to the no rent control market equilibrium. Entries in the table are based on Figure 8.22, with p_B = $2,000, p_R = $1,500, $X_B = 500$, $X_R = 350$, and $X_E = 650$.

Landlord and renter economic surplus gains or losses	Area in Figure 8.22	Amount ($ m.)
Previously landlord economic rent, now renter surplus	C	175
Landlords' share of deadweight loss	E	37.5
Renters' share of deadweight loss	D	37.5
Landlord loss	C + E	212.5
Renter net gains	C – D	137.5
Net loss	D + E	75

✓ **FACT CHECK** The data in Figure 15.14 in Chapter 15 shows that many countries have found ways to give a larger slice of the pie to the less well-off while also growing the pie. Compare the US, France, and Italy to Germany, Norway, and Finland. We can see that Germany, Norway, and Finland have enjoyed both faster growth in average incomes and a larger share of income going to the less well-off (they are less unequal than the US, France, and Italy).

Policies to redistribute income are often advocated on the grounds that providing additional income to one group (typically less well-off) is more highly valued by the policymaker or the electorate than the incomes lost by some other (typically higher-income) group. The basic idea here was introduced in Chapter 3 when we considered the difference between cardinal and ordinal utility.

Placing a greater weight on the gains of the renters than on the losses of the landlords can be done if the utility of people can be compared. In this case, the policymaker might conclude that the needs that will be more likely to be met by the renter family—more adequate housing, for example—are more important than the reduction in spending—perhaps on a vacation home—that the better off landlord will experience.

Going back to Table 8.4, you can see that if we placed a value on the gains by the renters that is four times the value placed on the *costs* to the landlords then the gains to the renters are (4 × 137.5 million) 550 million. These gains greatly outweigh the costs to the landlords of 212.5 million.

When we compare the gains and losses made to particular people or members of groups, we necessarily make value judgments, that is, we make judgments based on moral or ethical values. This is clear when we are evaluating the dollar gains of the renters as ethically 'more important' than the dollar losses of the landlords. But the point applies with equal force when we simply treat all of the dollar gains and losses as equivalent: that is also an ethical judgment.

M-NOTE 8.8 **Valuing gains and losses differently across groups**

One way to value the gains or losses of different groups is to express these values as statements about the marginal utility of income to each of these groups. A higher value placed on income gained by the poor would be to say that their marginal utility of income is higher.

To do this you might specify some cardinal utility function relating the well-being of person i to their income as follows:

$$\text{Utility function:} \qquad u_i = u(y_i) = \alpha \ln(y_i)$$

where α is a positive constant and y_i is the individual's income. Then (differentiating this utility function with respect to y and recalling that the derivative of $\ln y$ with respect to y is just $1/y$) we see that:

$$\text{Marginal utility of person i:} \qquad = \frac{\partial u}{\partial y_i} = \frac{\alpha}{y_i} \tag{8.46}$$

If all members of the population have the same utility function (though their incomes differ) then Equation 8.46 means that the marginal utility in income of someone with an income of $50,000 is four times as great as the person with $200,000 in income.

CHECKPOINT 8.14 Valuing the losses and benefits of rent control Imagine you are a policymaker considering imposing the rent control whose distributional effects are shown in Table 8.4. You place a higher value on the gains to the renters than on the losses to the landlords, because you wish to raise the living standards of the less well-off (the renters) even at a cost to those who are better off (the landlords). You just saw that if your value on the gains to the renters is four times your value on the losses to the landlords, then the benefits of the policy exceed the cost.

What is the *smallest* value placed on the gains of the renters that would make the benefits of the policy exceed the costs?

8.15 PERFECT COMPETITION AMONG PRICE-TAKERS: AN ASSESSMENT

As is the case with any model, the usefulness of the perfectly competitive Nash equilibrium of price-taking buyers and sellers depends on the insights it can generate despite not being an exact representation of the empirical problem at hand. For example, while consumers of sugary drinks are price-takers, that assumption does not apply to some of the producers. Just two firms—Coke and Pepsi—in 2020 accounted for well over two-thirds of the carbonated soft drinks sold in the US. They are effectively **price-making** duopolists or at best oligopolists, not price-takers.

But, to take our example, maybe the model captured *enough* of the reality of the market for sugary drinks to be a useful tool of analysis. Contrasting two Nash equilibria of a market with price-taking buyers and sellers using the comparative static method provided a simple way of identifying and adding up some of the gains and losses associated with the tax on sugary drinks, even if it was far from a complete picture.

Stepping back from the sugary drinks market to the analysis of market competition in general, there are four conditions under which the Nash equilibrium of price-taking buyers and sellers would be a plausible model:

- *Standardized products*: firms produce a product that is *indistinguishable* or barely so from that of its competitors.
- *Limited barriers to entry*: there are a sufficient number of competitors so that no firm or other actor can affect the price in its favor by altering the amount that it buys or sells.
- *Rising or supply constrained average cost curves*: remember that if the average cost curve is decreasing, the supply curve does not exist.

HISTORY Today, when people think about economics, the first thing that comes to mind is "supply and demand." But in the middle of the last century, in the most famous economics textbook ever, the topic "Determination of price by supply and demand" was put off until p. 447. Exactly ten pages later, the author, Paul Samuelson, wrote: "This is all there is to the doctrine of supply and demand. All that is left to do is to point out some of the cases to which it can be applied and some to which it cannot."[9]

PRICE-MAKING Price-making is a strategy that an economic actor may follow, altering the price at which they offer to buy or sell, or altering the level of output in ways that change the price at which they can transact.

- *Market equilibrium with market clearing*: so that the intersection of the supply and demand curves approximates the real situation we are studying.

The first two conditions—standardized products and price-taking by buyers and sellers—define what competition means in the "perfect" sense. But the third and fourth conditions—rising cost curves and market-clearing equilibria—are equally essential to make the model work.

To evaluate this theory we can use two types of standards:

- Is it *coherent*: does it make sense internally? Or, for example, are its assumptions contradictory?
- Does the model and the predictions based on it *correspond to the reality* it is designed to describe?

Is the perfectly competitive model coherent? We have shown that if buyers and sellers act as price-takers then the intersection of the supply and demand curves is a Nash equilibrium. But we now have to ask: Is acting as a price-taker a Nash equilibrium? Does the perfectly competitive model arbitrarily limit the strategies that buyers and sellers can follow? Is price-taking the best a buyer or seller can do?

We have already seen that to reduce the price elasticity of demand and thereby to raise profits, firms will seek to differentiate their products and to limit entry of firms to the markets in which they sell. If they are successful in raising barriers and differentiating their products they will be able to profit from being price-*makers*. In this case the assumption of profit maximization and price-taking are contradictory. So price-taking is not "doing the best they can" and *cannot* be part of the Nash equilibrium. In cases where this is true, the perfectly competitive model is incoherent.

There is a second source of incoherence in the model: the price-taking assumption makes it impossible to explain *how* a market would reach an equilibrium if it were not initially at the intersection of the supply and demand curves. If at the current price supply exceeds demand, for example, then some of the actors must *lower* the price at which they offer to transact. This is the common-sense explanation of how the market reaches an equilibrium. But if buyers and sellers are price-takers, then they *cannot* change their prices. We return to this problem in Chapters 9 and 14 where we explain how dropping the price-taking assumption can address this problem.

✓ **FACT CHECK** The ingredients and recipe by which Coke is made is a legally protected trade secret; and the company also has a government enforced monopoly on the trademark Coca Cola (and Coke).

Does the model and its predictions correspond to the reality that it claims to describe? Models are useful because they do *not* correspond to all of the complexities of real economic problems. So this is a question about where the model applies sufficiently to be useful and where it does not. While we gain important insights about the sugary drinks market from the model of perfect competition—despite the huge market shares of Coke and Pepsi—there are other markets in which the model will be substantially misleading.

We will see in Chapters 11 and 12 that neither the assumption of price-taking nor the condition that markets clear corresponds to the reality of some of the most important markets of a modern capitalist economy. Examples include the credit market (where banks set interest rates and the demand for loans typically exceeds the supply) and the labor market (where employers set wages and the demand for labor typically falls short of the supply, meaning that people are unemployed). In Chapter 9 we will see that barriers to entry can be sufficient so that price-making strategies for selling goods and services best characterize a substantial part of modern economies.

There are other limitations of the perfectly competitive model. We have shown that the Nash equilibrium is Pareto efficient by restricting our analysis to consumers and their surplus and owners and their profits. But there will be others affected by the production and consumption of the goods or services in question who are not considered in showing that the price-taking competitive equilibrium is Pareto efficient. These external effects concern:

- those who produced the goods;
- those affected by the external environmental effects of the goods production; and
- others affected by the external effects of the good's consumption (think of the health impacts of sugary drinks and their effects on family members and taxpayers).

So our statement that the perfectly competitive equilibrium is Pareto efficient is based on a useful but incomplete model. We will consider these and other external effects in Chapter 14.

None of these limitations means that the model is "wrong" or cannot be insightful in answering some questions. But they do suggest that the domain over which the model of perfect competition applies in the real economy is somewhat restricted.

CHECKPOINT 8.15 **Perfect competition: Pro and con** For each of the following products give reasons why the perfect competitive model would or would not be an appropriate analytical tool: Pharmaceuticals, streaming music, restaurants, barbers or hairdressers, private tutors (e.g., in math or language training).

8.16 TWO BENCHMARK MODELS OF THE PROFIT-MAXIMIZING FIRM: PRICE-TAKERS AND PRICE-MAKERS

The domains for which the price-taking model is of limited applicability—markets with substantial barriers to entry, labor markets, credit markets for example—are all interactions in which some form of price-making figures

prominently in the profit-maximization strategies of firms. Here we bring together what you have learned about price-taking and price-making firms.

The price-taking owners of firms take the market price as given in choosing an output level to maximize profits. The size of firms in this model is limited by the fact that marginal costs are assumed to *rise* with increasing levels of output and will exceed average costs because average costs are (after some level of output) rising with increased output. As a result, for any given price, firms will expand up to the point that price equals marginal cost, but no further. Higher levels of output would reduce profits.

As you have learned, the model does not apply to cases in which average costs decline with increased output or are even constant as output expands. In these cases, if instead the price exceeds average costs, then the firm will want to grow without limit.

A downward-sloping average cost curve (like the cost curves in Figure 8.15) raises two problems for the theory:

- If the firm grows without limit, then we have to question the assumption that all firms are small relative to the size of the market so that the decisions the owners of firms make will not affect the market price. A consequence is that the process of competition itself may destroy the conditions under which a large number of small firms could compete.
- Price cannot be equal to marginal cost because average cost (at any level of output) exceeds marginal cost, so firms that set $p = mc$ would eventually go out of business. A consequence, we will see in the next chapter, is that the resulting prices and quantities will not be Pareto efficient.

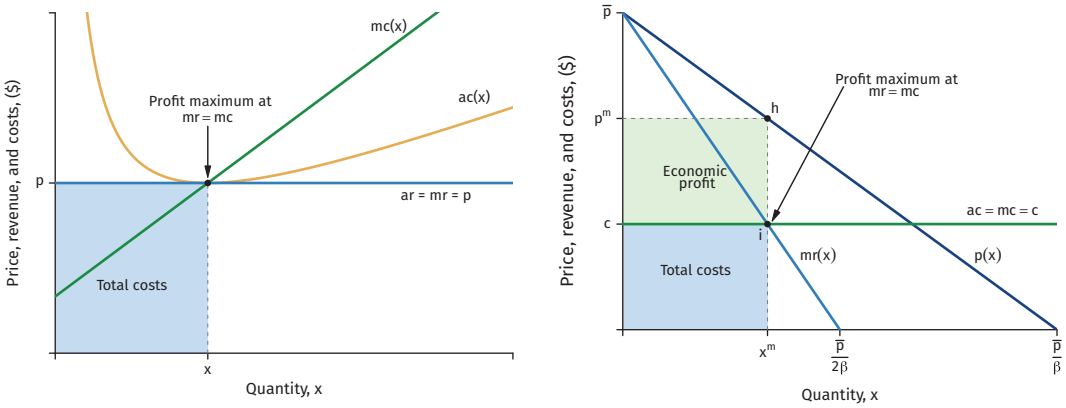
An alternative benchmark model accepts the evidence that long-run average cost curves are flat or even downward-sloping and explains why firms do *not* grow without limit by the fact that firm sales are constrained by a downward-sloping demand curve. Demand curves slope downward because, in practice, many firms have a *limited* number of competitors selling the same product that they are producing and the absence of close substitutes—a Honda is not a Ford.

The two benchmark models—perfect competition and monopolistic competition—are illustrated in Figure 8.23. For the monopolistically competitive firm, on the right, the downward-sloping demand curve replaces the upward-rising marginal cost curve as the limit on firm growth. The flat cost curve replaces the flat demand curve as a part of the model facilitating firm growth. Under these conditions, firms will not grow without limit.

So the downward-sloping demand curve allows us to reconcile flat or downward-sloping cost curves with a limit on firm size.

But by accepting the monopolistic competition model as a benchmark, one has to *give up* the result that, in equilibrium, the price will be equal

Figure 8.23 The perfectly competitive firm and the monopolistically competitive firm: two benchmark models. In the perfectly competitive model, the firm's demand (*not* the demand for the entire market) is flat and marginal costs are rising. Price equals marginal cost at the profit maximum and the firm makes no economic profit ($p = mr = mc$, therefore $\pi = 0$). In our monopolistically competitive alternative model, costs are flat, demand is downward-sloping, and the profit maximum is where $mrs = mrt$ (the isoprofit curve is tangent to the demand curve at point **h**) or, what is the same thing, $mr = mc$ at x^m with price $p^m > c$. The firm makes positive economic profits because $p^m - c > 0$.



(a) Perfectly competitive firm

(b) Monopolistically competitive firm

to marginal cost. This will be important when we consider the relationship between market competition and the efficiency of the resulting allocations.

Which of the two models is best to use depends on the question that you would like the model to help you answer.

For both reasons, the model of monopolistic competition would seem to be a better representation of the sugary drinks market. We will study the process of competition among large firms in the next chapter.

CHECKPOINT 8.16 Two benchmark models In Figure 8.23 explain what limits the growth of the firm in the two cases (panels (a) and (b)).

8.17 APPLICATION: THE DYNAMICS OF FIRM GROWTH AND THE SURVIVAL OF COMPETITION

The downward-sloping demand curve can place a limit on firm growth, but other factors may work in favor of continuing growth in size, especially of large firms. These include:

- *Radically declining average costs* as occur in the production and sale of knowledge-intensive products such as software and pharmaceuticals where "first-copy costs" (c_0 in our cost function) are substantial while

✓ **FACT CHECK** The last few decades have witnessed spectacular growth in some firms. In 2020 for example, Walmart employed over two million people worldwide, and in the US about half of all employees work in firms with more than 500 workers. Between 1988 and 2012, the average firm size in the US grew modestly from around 18 employees to a bit over 20.[10]

marginal costs are effectively zero. These costs are for the "first copy"–of the drug, of the app, or another piece of software–because every copy beyond the first has a marginal cost of close to zero.

- *Demand-side increasing returns* introduced in Chapter 7, in which the willingness to pay (or to endure advertising) is greater for the one-millionth member of a network than it is the first hundred and
- *Learning-by-doing* which gives advantages to incumbent firms with a larger amount of cumulative sales as the example in Chapter 6 of the M20 armored truck during World War II showed.

This brings us to the question: If successful firms grow, taking over a larger market share, then why doesn't the number of firms in a given market shrink, reducing the level of competition?

To provide a possible answer, a completely different set of models of firm size and competition has been proposed, inspired by biological models of competition for fitness in the natural world. In these "life history" models, firms are born, grow, and die, so the size of any firm in existence is typically growing. But they do not grow without limit (they die) due to influences such as bad luck or mismanagement that are not directly related to prices and average costs.

Think about a similar but simpler question. Why is the average age of many populations *constant* despite the fact that every member of the population is getting older? The answer is that those *leaving* the population are *older* than those being born into it.

A similar possible explanation for why the average firm size could stay constant even if some firms grow perpetually may be that when firms die, they are replaced on average by one or more smaller firms. So, constant growth of surviving firms can be consistent with a *constant average firm size*.

To see how this could work, let us say that there is a firm i whose size, by some measure, say sales or employment, is s_{it} at time t. Firm i grows in size at the percentage rate, $g^s > 0$. Firm i's size in the next period (time $t+1$) would therefore be:

$$s_{it+1} = (1+g^s)s_{it} \tag{8.47}$$

We consider a constant number, n, of firms in the industry and the average size of a firm at time t is $\bar{s}_t$. Next, suppose that in every time period, as a result of competition with other firms, each of the n firms dies with a probability f (it *fails*). Firm deaths may occur because owners or managers make strategic errors, or conditions change (new technologies, new competitors). Where there is substantial competition among firms, then these and other adverse events will result in the *failure* of the firm. Where competition is weak, firms can endure a great many of these events before finally dying.

Once dead, a firm is replaced by a firm of size $\underline{s} < \bar{s}$. In other words, the new firm is *smaller* than the average firm size in the industry.

We can then find values of the growth rate of firms g^s, the size of new firms $\underline{s}$ and the "mortality rate" of firms f such that there is some average firm size that *does not change* even if all of the surviving firms grow at the rate of g^s. In other words, we can find a set of values $(g^s, \bar{s}, f)$ for which $\bar{s}_{t+1} = \bar{s}_t$.

To see this, we say the *weighted average firm size* in $t+1$, that is, (s_{t+1}), is composed of the following:

- the increased size of all surviving firms, $(1+g^S)s_t$ multiplied by the probability that the firms survived, $(1-f)$;
- plus the probability that a firm did *not* survive (it failed), f multiplied by the size of the average *replacement* firm, $\underline{s}$

Putting these terms together we find the weighted average firm size:

$$\bar{s}_{t+1} = (1-f)(1+g^s)\bar{s}_t + f\underline{s} \tag{8.48}$$

To find the conditions under which firm size could be *unchanging* from period to period, we equate the average firm size in time t and $t+1$. That is, we find the common value of $\bar{s}_t = \bar{s}_{t+1} = \bar{s}$ satisfying the equations:

$$\bar{s}_{t+1} = (1-f)(1+g^s)\bar{s}_t + f\underline{s} = \bar{s}_t \tag{8.49}$$

Using this constant value, $\bar{s}$, and dropping the time subscripts, we can simplify the above to get:

$$\underbrace{g^s\bar{s}(1-f)}_{\text{firm growth effect}} = \underbrace{f(\bar{s}-\underline{s})}_{\text{firm death effect}} \tag{8.50}$$

The above equation gives us a condition such that the average firm size will *not* change despite the growth of surviving firms. The average firm size will be constant if its growth caused by firms surviving with probability $(1-f)$ and growing at the percentage rate g^s (the left side of the equation) is *offset* by the *shrinkage* in firm size caused by firms dying with probability f and being replaced by firms that are on average *smaller* by the amount $\bar{s} - \underline{s}$ (the right side of the equation).

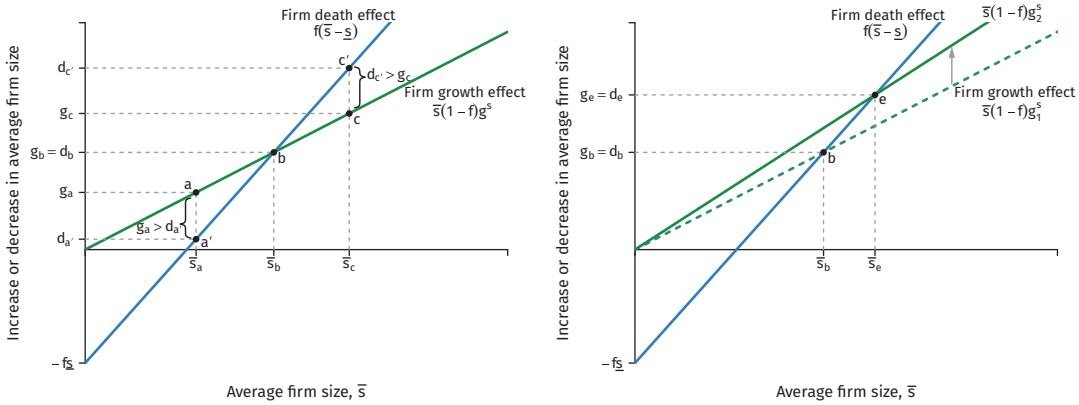
The resulting constant firm size can then be found by solving Equation 8.50 for $\bar{s}_t$ as is illustrated by point **b** in Figure 8.24 where the equations for the firm growth effect and firm death effect intersect:

$$\bar{s}_b = \frac{\underline{s}f}{f - g^s(1-f)} \tag{8.51}$$

At a lower level of the average firm size, $\bar{s}_a < \bar{s}_b$, the firm growth effect (g_a) *exceeds* the firm death effect $(d_{a'})$. The reason is that the average size of firms is small, so that when a failed firm is replaced by a new firm, the

M-CHECK The dynamic process of firm growth, death, and replacement will yield an equilibrium positive firm size as long as the rate of firm growth g^s is not too large relative to the rate of failure, that is as long as $g^s < \frac{f}{(1-f)}$. If $g^s > \frac{f}{(1-f)}$ or what is the same thing $g^s(1-f) > f$, then in Figure 8.24 the slope of the firm growth effect (the former expression) line exceeds the slope of the firm death effect line so the two lines do not intersect and firms will grow to infinite size.

Figure 8.24 Constant average firm size with every firm growing in size. In panel (a), the average firm size is s_b. This occurs when the firm death effect $f(\bar{s} - \underline{s})$ is just equal to the firm growth effect $\bar{s}(1-f)g^s$ at point **b**. The second panel shows that with a more rapid growth of existing firms (growth rate increasing from g_1^s to g_2^s), the constant firm size would increase from $\bar{s}_b$ to $\bar{s}_e$ at point **e**.



resulting decrease in average firm size (the firm death effect) is small. As a result firms will grow *more quickly* than they die or shrink, until point **b**.

At a *higher* average firm size $s_c > s_b$ the replacement of failed firms by smaller firms will have greater force because the difference between the average firm size and the size of the replacement firms is larger. As a result the firm death effect shrinking average size ($d_{c'}$) will exceed the growth effect (g_c). So average firm size will shrink down to s_b.

Recalling that we take f as one measure of the degree of competition it follows that the average size of firms will be greater when competition is less (corresponding to a lower f). The average size of the firm will also be greater if firm growth is greater, as is shown by the higher firm growth line in Figure 8.24 (b), or if the size of the replacement firm is greater. (The latter would be the case if barriers to entry are greater for smaller firms.)

The model could be extended to take account of the possibility (suggested by recent evidence from the US) that firm growth is greater for large firms or that smaller firms are more likely to fail. Additionally, if firms fail by merging with a larger firm, then there is no smaller "replacement firm." All of these possibilities make the puzzle of constant firm size more difficult to reconcile with surviving firms growing.

Economists will continue seeking models in which both competition and the advantages of large-scale production play an important role. The biologically inspired "life history" models of the firm allow for both economies of scale and the enduring importance of competition. Whether they will prove insightful for the analysis of firm behavior in other respects remains to be seen.

CHECKPOINT 8.17 The birth and death of firms Using Figure 8.24, explain the effect on the constant average firm size of

a. an increase in the size of the replacement firms $\underline{s}$, and

b. an increase in the degree of competition among firms (higher value of f)

8.18 CONCLUSION

We began Part II of this book in Chapter 6 with an explanation of how economies of scale (along with learning-by-doing) lead to specialization by task and product, and how the resulting division of labor requires a method of distributing goods from the specialist producers to the generalist end user. Markets play an essential role in coordinating this production and distribution process.

Economics has demonstrated that for those goods and services where conditions approximating perfect competition exist, and, where environmental or other external effects are absent or small, markets can perform this task reasonably well, at least by comparison to alternatives such as centralized allocation of goods and services by governments.

But the factors that make specialization so beneficial—economies of scale and learning-by-doing—may also be inconsistent with the price-taking model of competition that is the basis for economists thinking that markets do a good job. We turn, then, to models of how firms compete that are more consistent with what is known empirically about modern economies.

MAKING CONNECTIONS

Profit maximization using feasible sets and indifference curves: The owner's choice of price and output level is another constrained optimization problem, with the demand curve as the frontier of the feasible set and the owner's objectives summarized by a set of isoprofit curves.

Consumer surplus, economic profits, and mutual gains from exchange: Both consumers and firms' owners will typically receive rents arising from the production and sale of goods.

Restricting output and hiring: A larger slice of a smaller pie: The firm facing a downward-sloping demand curve will increase its profits by restricting hiring and sales so as to sustain higher prices; the firm's customers and would-be employees do less well as a result.

Economics as an empirical science: Our models are informed by evidence; for example that average costs are not greater (and may be less) at higher levels of output, and that intellectual property rights and other barriers to entry can result in substantial price markups over cost.

Two contrasting benchmark models: (a) firms that face a flat demand curve and are constrained by rising costs and (b) monopolistic competitors with flat cost curves and are constrained by a downward-sloping demand curve.

Public policy: The supply and demand model of price-taking buyers and sellers on competitive markets illuminates the effects on consumers, owners, and others of a tax on sugary drinks and rent control.

Learning-by-doing, the survival of competition, and dynamic analysis: The declining cost of the M20 armored car during World War II and of solar panels recently along with the model of how average firm size might remain constant even if all firms are growing are examples of a dynamic analysis, looking at how something changes over time rather than comparing the equilibrium before and after some exogenous change, as in comparative static analysis.

IMPORTANT IDEAS

total cost	average cost	merit good
fixed cost	falling average cost	$mrs = mrt$ rule
accounting cost and profit	economic cost and profit	monopolistic competition
revenue	average revenue	marginal revenue
profit maximization	marginal cost	$mr = mc$ rule
price markup over cost	barriers to entry	perfect competition
consumer surplus	economic profit	deadweight loss
price-takers/price-makers	market dynamics	tax
estimates of cost curves	firm size	birth and death of firms
interpersonally comparable utility	rent control	cardinal utility
landlords' economic profit	Coasean bargaining	renters' surplus
markup	markup ratio	price elasticity of demand

MATHEMATICAL NOTATION

Notation	Interpretation
x	output of a firm
X	market output of a good
p	price of a unit of output
a_l, a_k	minimum labor and capital goods to produce a unit of output
c	total costs of production
w	wage
p_k	price of capital good
ρ	opportunity cost of capital
c_k	total cost of a capital good
$r(x)$	total revenue function
π	economic profits
r^A, r^E	accounting profit rate, economic profit rate
$-\beta$	slope of the demand function
η	elasticity of market demand with respect to price
μ	markup ratio over costs of production
τ	per unit tax
s	firm size
g^s	growth rate of firm size
$\underline{s}$	average size of a firm entering to the market
f	probability of failure ("death") of a firm
$\bar{s}$	average firm size

Note on superscripts and subscripts: c: competitive market; t: time period; l: hours of labor, k: amount of capital goods.

CHAPTER

9

COMPETITION, RENT-SEEKING, AND MARKET EQUILIBRATION

[A]ccording to Dr. Johnson [an eighteenth-century dictionary compiler], competition is "the action of endeavouring to gain what another endeavours to gain at the same time." Now, how many of the devices adopted in ordinary life to that end would still be open to a seller in a market in which so-called "perfect competition" prevails? I believe that the answer is exactly none. Advertising, undercutting, and improving ("differentiating") the goods or services produced are all excluded by definition—"perfect" competition means indeed the absence of all competitive activities.

Friedrich Hayek,
"The Meaning of Competition" (1948)[1]

DOING ECONOMICS

This chapter will enable you to do the following:

- Understand competition as a strategic process among firms, actively rent-seeking through price-making, choice of output, advertising, innovation, product differentiation, and more.
- See why (except under special conditions) the "price equals marginal cost" condition for Pareto efficiency of the level of output will not be a Nash equilibrium.
- Explain how perfect price discrimination, like perfect competition, is an abstract ideal model illustrating conditions under which price would equal marginal cost.
- Show how the prices, quantities sold, the price markup over costs, economic profits, and consumer surplus vary with the extent of competition in a market.
- Explain how barriers to entry restrict competition among firms, increase deadweight inefficiency, and raise owners' profits while reducing customers' consumer surplus.
- Understand how rent-seeking by buyers and sellers may equilibrate supply and demand.
- See how the forces of supply and demand work by altering the fallback positions of buyers and sellers.

9.1 INTRODUCTION: "STAY HUNGRY, STAY FOOLISH"

The year 1890 was a big one for computers. It was the year that Herman Hollerith, frustrated with the eight years it took to tabulate the 1880 US census, decided to introduce punch cards into data collection and entry processes for the census. It took one year rather than eight to tabulate the 1890 census, a massive improvement in productive efficiency, that heralded a revolution that would echo through the twentieth and twenty-first centuries.

Charles Flint, recognizing the genius of Hollerith's system, bought out Hollerith's company and, merging it with several of his own companies, created the company that became, in 1924, the International Business Machines company (IBM). For decades, IBM was at the forefront of computing innovation, providing large machines to companies, governments, and universities; and raking in huge profits as a result.

But IBM, like its competitors, didn't see the coming wave of the personal computing **industry** and the dominance of software companies. It took other innovators—Bill Gates and Paul Allen of Microsoft, Steve Jobs and Stephen Wozniak of Apple, among others—to see the potential that a personal computing machine would have in the everyday lives of people around the world.

Even Steve Jobs—famous for advising Stanford University's graduating class in 2005 to "stay hungry, stay foolish"—could not keep up with the pace of things.[2] It took him and others a couple of decades after the invention of the personal computer to see the potential for personal musical devices—the iPod—and to design a smartphone—the iPhone. The iPhone would go on to dislodge the giants of the cell phone industry, Nokia and Research in Motion (the manufacturer of the Blackberry).

Big, near-monopolistic, profitable, and established firms did not foresee the coming changes to their industries. The story of IBM echoes within and across industries. Kodak, a once-dominant firm in the photography industry, failed to develop its own digital camera even though people within the company had invented one. Their invention had been discarded because of the damage it might do to their core business, selling film and film-based cameras.

Microsoft, which had won the personal computing battle of the early 1990s through its innovative software, failed with its smartphones (losing to Apple), search engine (losing to Google, now Alphabet), its web browser (losing to a variety of Chrome, Firefox, and other non-Internet Explorer

Figure 9.1 An Apple Macintosh Classic introduced in 1990.

Photo by Bombaert Patrick/Alamy Stock Photo.

INDUSTRY An industry is a set of firms producing similar products.

packages), or its networking systems (losing to the Linux-based Apache system).[3] Microsoft has subsequently innovated once more with cloud-based computing with Microsoft Azure.

When we think of sports, politics, the job market, or social status climbing, we usually think of competition as a process in which competitors actively seek to gain advantage over others. In the economy, too, competition is the equivalent of warfare where there are big winners, and, as IBM and Kodak found out, big losers. Depending on the nature of the competition, consumers too can be big winners, or big losers.

But as Friedrich Hayek pointed out in the head quote, what economists call perfect competition as taught in introductory economics courses differs from competition in politics and war. Buyers and sellers are passive "price-takers" (who take the price as given). In this chapter, we present a more empirically grounded view of competition. Firms are not price-takers, they are price-makers. Firms set prices and innovate to capture a larger market share, and even create and then dominate entirely new markets.

To do this we study how competition works, whether there are a substantial number of firms producing similar products or there is a single firm selling a unique product. We consider the perfect competition model as a special case, like we did in Chapter 8.

Figure 9.2 Antoine Augustin Cournot (1801–1877) developed the first mathematical models of the process of competition among firms ranging from two (duopoly), to a few (oligopoly), to many. He studied mathematics as an undergraduate and later received advanced degrees in the fields of astronomy, mechanics, and law. He is credited with having invented the supply and demand curve analysis, and having had a hand in persuading Léon Walras—a founder of the school of "neoclassical economics"—to take up the subject.

9.2 MODELING THE CONTINUUM OF COMPETITION: FROM ONE FIRM TO MANY

To understand how the number of firms in a market affects the outcomes for firms and consumers, we start with what is called the Cournot ("Coor-NO") model of competition.

To determine the outputs the two or more firms will produce and the price at which the goods will sell we cannot consider the firms in isolation as we did when we modeled monopolistic competition in Chapter 8. The firms are engaged in a strategic interaction: their owners know that their profits depend on not only their own firm's actions but the actions taken by the others. This is why we use game theory to understand the process of competition. Cournot competition is represented by a game with the following characteristics:

- *Players*: The owners of firms in an industry (if there is more than one firm) sell an *identical* or *standardized* product.
- *Strategies*: Each firm simultaneously selects a level of output to produce and sells that entire output at the highest price possible given by the other firms' sales and the industry inverse demand curve.
- *Payoffs*: For every set of outputs for each of the firms (called the *industry output profile*) there is a level of economic profit (possibly zero or negative) received by each firm's owners.

- A *Nash equilibrium*: is an *industry output profile* and a price such that each firm's output level is a (profit-maximizing) best response given the other firms' output levels when the *single price* at which all of the firms outputs are sold is determined by the industry inverse demand curve and the total output produced by the firms.

There are two important consequences of this setup.

First, firms in the Cournot model are not price-takers; they are price-makers. As in the models of monopolistic competition in Chapter 8 the owners of the firm know that the level of output they put on the market will affect the price at which they can sell it. They deliberately "make" prices by choosing *how much* to produce, taking account of the fact that they are constrained by a downward-sloping demand curve.

Second, competition in the Cournot model conforms to the **law of one price**. The law of one price states that in equilibrium identical goods or services will transact at the same price. The reason why prices do not differ among the firms is that if one firm's output were selling at a higher price than other firms' outputs, then buyers would switch to the lower priced firms.

Identical prices for identical goods may seem a truism too obvious to warrant a law of its own, but it is not always true: airlines, for example regularly charge different prices to different categories of customers–the elderly for example–for exactly the same seats on the same flights. We introduce this case–price discrimination–later in this chapter, and other violations of the law of one price in subsequent chapters.

Cournot's model allows us to consider a continuum of competition that we illustrate by three cases differentiated by the number of firms in the industry, n:

- *Monopoly*, $n = 1$: there is only one firm in an industry.
- *Duopoly*, $n = 2$: a second firm shares the industry demand.
- *Oligopoly* and "*unlimited competition*", $n > 2$: several firms (oligopoly) or many firms (unlimited competition) share the demand.

"Monopoly" and "monopolistic competition"

The outputs and prices that maximize profits in the monopoly case $n = 1$ are identical to the case of monopolistic competition introduced in Chapter 8 as long as the firm's product is differentiated in some way (by trademark for example) so that no competitor can sell an identical product.

REMINDER Remember from Chapter 7 that the inverse demand curve ($p(X)$) gets its name from the fact that it is the mathematical inverse of the conventional demand function, which makes quantity sold a function of the price ($X(p)$).

REMINDER The Nash equilibrium describes a situation in which, were it to occur, none of the firms' owners would have an incentive to alter their strategy. It does not describe the *process* by which the equilibrium would come about.

LAW OF ONE PRICE The law of one price states that in equilibrium identical goods or services will transact at the same price.

✓ **FACT CHECK** We use this cost function for simplicity, not for realism. As we saw in Chapter 8, firms typically have some fixed costs and large firms may enjoy either economies of scale in production or increased bargaining power in purchasing inputs from suppliers (meaning that larger firms pay less for inputs and have lower costs).

The term "monopoly" suggests a sole seller of a product with few substitutes. Examples would be a single local seller of electricity or the drug Daraprim (life-saving treatment of an HIV AIDS related illness) which in 2015 was priced at $750 a pill, having previously been sold (presumably at a profit) for $7 a pill.

By contrast, the term "monopolistic competition" stresses that many single sellers of a differentiated product (literally monopolies) do have to compete with firms selling close substitutes (unlike Turing Pharmaceuticals, the seller of Daraprim). Examples are the sugary drinks discussed in Chapter 8 with very price-elastic demand (Coca-Cola at $|\eta| = 3.79$ and Mountain Dew at $|\eta| = 4.39$). These highly price-elastic demand curves are an indication of the competitive nature of the sugary drinks industry even though just a few firms sell most of the drinks purchased. The price elasticity of demand for sugary drinks as a whole, however, is much lower at $|\eta| = 1.4$. So if that entire sugary drinks market were served by a single firm, the term "monopoly" would be appropriate.

The economic environment: Demand, revenue, costs, and profits

! REMINDER If costs are $c(x) = cx$, then average costs (the total costs divided by the total output) are $ac(x) = \frac{c(x)}{x} = c$ and marginal costs (the change in total costs associated with a small increase in output Δx) are $\frac{\Delta c(x)}{\Delta x} = c$.

For each of these cases (where the number of firms, $n = 1, 2$, few, many) we will use the *inverse demand curve* for which price depends on the quantity sold (we assume that firms sell everything they produce, so we use "sales" and "output" interchangeably). For simplicity we assume the demand curve is linear where p is price, X is total output and sales in the market, $\bar{p}$ is the maximum price (given by consumers' maximum willingness to pay) when output is zero, and $-\beta$ is the slope of the industry inverse demand curve, $\frac{\Delta p}{\Delta x}$:

$$\text{Inverse demand curve} \qquad p(\mathrm{X}) = \bar{p} - \beta \mathrm{X}$$
$$\text{where} \qquad \mathrm{X} = x^1 + x^2 + \ldots + x^n$$

In the model, all firms have an identical cost function with *no fixed costs* and a *constant marginal and average cost*, c.

$$\text{Cost function} \qquad c(x) = cx \tag{9.1}$$

! REMINDER When we refer to a "firm" seeking to maximize profits or making some other decision we mean the owners of the firm (who will receive the profits and directly or indirectly through management make the relevant decisions).

The opportunity costs of the machinery, intellectual property, and other capital goods used in production are included as a cost of production. In this case, the firm's marginal and average cost are equal and do not vary with the firm's output. When a firm sells its product at a price greater than its costs it makes economic profits, meaning that accounting profits exceed the opportunity cost of capital. To calculate the firm's economic profits, we use its revenue and costs as follows:

$$\text{Firm i's revenue} \qquad r^i = p(\mathrm{X})x^i$$
$$\text{Firm i's profit} = \text{Revenues} - \text{Costs}$$
$$\pi^i = p(\mathrm{X})x^i - cx^i$$

We can substitute the inverse demand curve $p(X)$ into Firm i's profit function to find their profits:

$$\begin{aligned}\pi^i &= p(X)x^i - cx^i \\ &= (\bar{p} - \beta X)x^i - cx^i \\ &= (\bar{p} - \beta(x^1 + \cdots + x^n))x^i - cx^i \end{aligned} \tag{9.2}$$

A firm's profits depend both on their own production and on the production of other firms (included in X). Each firm has a negative external effect on the revenues of other firms by producing more output.

A coordination problem: Overharvesting fish and crowding the market

With a downward-sloping demand curve and more than one firm, firms face a coordination problem like overharvesting fish from the lake in Chapter 5. What do firms competing on markets have in common with fishing people depleting the basis of their livelihood? The common idea is *over*harvesting—whether it is fish or customers—that could be prevented if the firms or fishermen coordinated their actions rather than acting singly.

Just as the Port Lincoln lobstermen in Chapter 5 discovered that they could benefit by making a common decision to *limit* the number of traps they set, so too will firms discover that they could make higher profits if they were able to agree to restrict sales and benefit from the resulting higher prices, rather than competing. We can think of the potential buyers of the product of the firms as analogous to the fish in the lake: the more customers that one firm "harvests" the fewer are left for the other firms.

This means that, like the lake, the market is a common property resource: like catching fish, selling to customers is *rival* because the customer one firm sells to will not buy goods from another firm. But the firms *cannot be excluded* from competing on the market. The analogy of market competition to the overfishing problem is summarized in Table 9.1.

Comparing the objectives of the players—the utility function of the fishermen and the profit function of the firms (in the table)—you can see the following:

- The *firm's level of output* x^i is analogous to the fisherman's fishing time h^i—it is necessary to the firm's objectives (profit) and it is also a cost.
- *Other* firms' level of output (contributing to the total output sold on the market, X) is similar to other fishermen's fishing hours contributing to H (total hours fished) because it has a *negative effect* on the firm's objective.
- This negative external effect in each case is represented by β.

We shall see that as long as firms do not coordinate (forming what is called a cartel) they end up producing more than would be Pareto efficient (when considering only the producers). The firms' owners make lower profits than

Figure 9.3 Jean Tirole (1953–) is a French economist who has transformed the way that economists think about problems of monopoly and competition policies, taking account of the limited information available to policymakers and representing a government regulator as a principal and the owners or managers of firms as agents. Like Jules Dupuit, the founder of cost–benefit analysis well over a century earlier, Tirole studied engineering at the National School of Bridges and Roadways before turning to economics. His 2017 book *Economics for the Common Good*, is a passionate statement that economics can and must help to address urgent problems facing humanity such as climate change and the future of work along with his thoughts on the appropriate role of markets and governments in rising to these challenges.[4]

Photo by REUTERS/Alamy Stock Photo.

Table 9.1 Comparison between overharvesting fish and "overharvesting" customers. The lake and the market are both common property resources: additional fishermen and additional firms *cannot be excluded*, and they compete with incumbent fishermen or firms for fish stocks and customers, so the resources are *rival*.

	Overharvesting fish	Overselling
Objective	utility $u^i = y^i - \frac{1}{2}(h^i)^2$	profit $\pi^i = p^i x^i - cx^i$
Production and demand	Fish caught (y^i) $y^i = h^i(\alpha - \beta H)$ $H := h^1 + h^2 + \ldots + h^n$	Revenue ($p^i x^i$) $p^i x^i = (\overline{p} - \beta X)x^i$ $X := x^1 + x^2 + \ldots + x^n$
External effect, $n = 2$	β: j's fishing reduces i's catch	β: j's sales reduce i's revenue

if each firm had produced less or if there were fewer firms. There is one difference between the overharvesting fish model and the overharvesting customers problem modeled below. In Chapter 5 the only players in the game were the fishermen themselves (they did not sell the fish they caught, they just ate them).

When we model markets, there are not only the owners of the firms but also customers. We will see that when firms fail to coordinate so as to limit their sales, then the firms' owners make lower profits, but customers benefit from lower prices. Correspondingly, when firm owners do cooperate, agreeing jointly on a level of output and a price, they can all improve their profits, but at a cost to the customers who then would face higher prices.

Before we study the coordination failure among firms we look at what happens with one firm in the industry. This constrained maximization problem and its solution are identical to that for the monopolistically competitive firm shown in Figure 8.10 in Chapter 8.

CHECKPOINT 9.1 Overharvesting

Thinking about Table 9.1:

a. Which is the parameter in both the overharvesting fish and overselling to consumers set up that represents the negative external effect of one actor's actions on the others?

b. Explain why α in the fishing problem is like $\overline{p}$ in the selling problem.

c. In the selling problem there are constant costs of producing more. Is this true in the case of fishing?

9.3 REVIEWING THE MONOPOLY CASE, n = 1

In the monopoly case the firm's output and sales x is the same thing as the industry output and sales X (i.e. $x = X$) so the firm's revenue is $r(x) = p(x)x$. The firm costs are $c(x) = cx$, and so its profit is:

$$\text{Profit} \qquad \pi(x) = p(x)x - cx$$

Which, using the inverse demand function to represent the price, is:

$$\text{Profit} \qquad \pi(x) = (\bar{p} - \beta x)x - cx$$

As was the case for the monopolistically competitive firm in Chapter 8, the firm chooses the profit-maximizing level of output and sales by finding the value of x that equates marginal revenue to marginal cost or

$$\text{Marginal revenue} = \text{Marginal cost}$$

$$\frac{\Delta r(x)}{\Delta x} = \frac{\Delta p}{\Delta x}x + p = \bar{p} - 2\beta x = c \tag{9.3}$$

Solving Equation 9.3 for x gives the profit-maximizing level of output and sales for the Cournot model with $n = 1$, that is, the monopoly case:

$$x = \frac{\bar{p} - c}{2\beta} \tag{9.4}$$

The top panel of Figure 9.4 depicts a monopoly firm, Firm A, choosing to produce output x^{Am} which it then sells at the price p^m given by the inverse demand function. The superscript m is for monopoly. The firm's total revenue is $x^{Am}p^m$ and its total profits are $x^{Am}(p^m - c)$, the green-shaded area. The area of the unshaded triangle $p^m h\bar{p}$ is the amount of consumer surplus that this firm's sales generate by selling x^{Am} and the price p^m.

! REMINDER If you are unsure how this expression for the marginal revenue is derived return to M-Note 8.5.

! REMINDER In Chapter 8 we introduced the model of monopolistic competition (due to Joan Robinson and Edward Chamberlin) that originated during the Great Depression, showing how the owners of a firm decide how much to produce when there are a small number of firms producing products that are differentiated, but are similar enough so that raising a price will lead some potential buyers to switch to a competitor. Here we study a model in which a single product is produced by *n* firms with *n* ranging from 1 (monopoly) to many (similar to perfect competition).

9.4 DUOPOLY: TWO FIRMS' BEST RESPONSES AND THE NASH EQUILIBRIUM

If there are two firms, they compete for a share of the market demand for the identical product they are selling. This is called a **duopoly**. The strategic analysis of a duopoly and competition among more than two firms is based on each firm's best response (meaning profit-maximizing level of output) to each of the other firms' choices of outputs. We derive each firm's best-response function by first considering the profit-maximizing decision of a single firm and how this depends on the output level of the other firms.

! REMINDER Consumer surplus is a measure of consumer welfare that is the difference between a buyer's willingness to pay and what they actually pay for each unit of the good that they consume, generally summed over all purchasers of the good (requiring the assumption that the marginal utility of the money left over in the person's budget after the purchase is identical for all consumers).

Duopoly profit, revenue, and costs

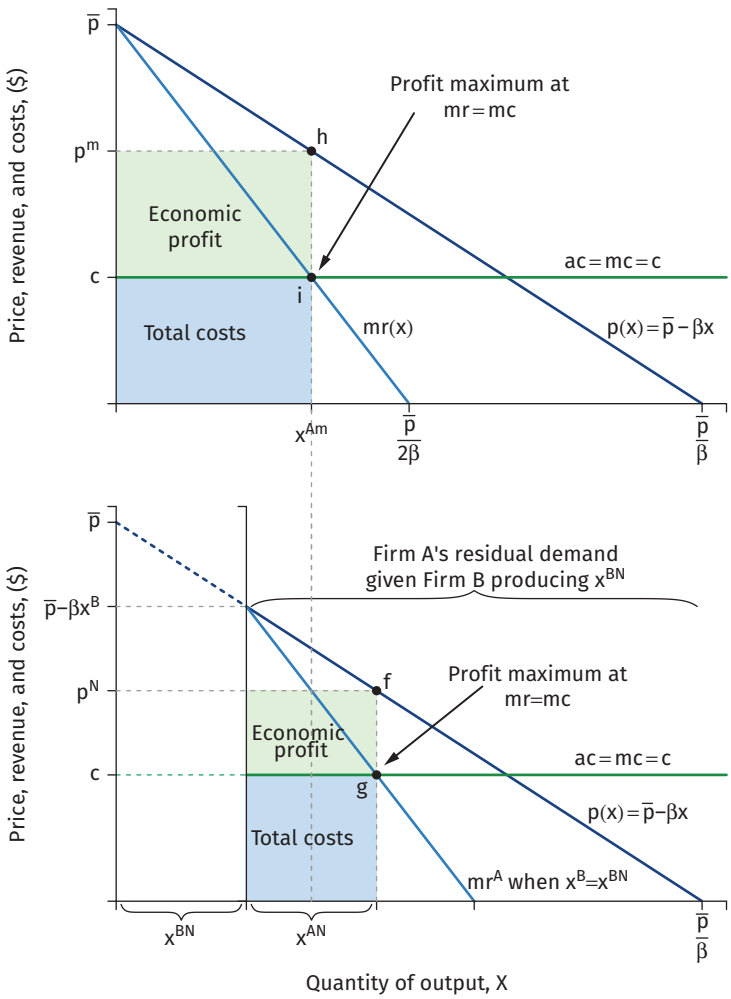
The industry or market output (X) is the sum of the firms' outputs. The market price will be affected by both firms' outputs. We assume that firms sell everything they produce, so sales and production are the same quantity, namely x. Therefore when there are two firms, A and B, the market output is as follows:

DUOPOLY When there are just two firms selling the same output, we call the industry a duopoly and we call each firm a duopolist.

$$\text{Duopoly market output} \qquad X = x^A + x^B \tag{9.5}$$

We show how the duopoly case differs from monopoly in the lower panel of Figure 9.4. If the new firm, B, produced nothing, Firm A would be a monopoly, as before. So, consider the case of interest when Firm B produces and will sell some amount x^{BN}. These are sales that Firm A will not be able to make: customers who have been removed from the market. The result is that Firm A now faces a reduced demand curve, shown in Figure 9.4

Figure 9.4 Comparison of monopoly and duopoly. The top panel shows a single firm, A, choosing the profit-maximizing output level. In the bottom panel a second Firm B has entered the market with sales of x^{BN}, so Firm A's residual demand curve is now given by the maximum price at which amounts greater than x^{BN} can be sold.



as a rightward shift of the vertical axis by the amount of sales that will be implemented by Firm B.

Firm A then faces the problem depicted in the lower panel. The solid portion of the demand curve is the portion of the market "left over" for Firm A. This is called the *residual demand curve* expressed as $x^A(p) = X(p) - x^B$.

The profit maximum for the firm will be to produce and sell an output such that the marginal cost equals the marginal revenue (now based on the residual demand curve), or x^{AN}, and then to sell this amount at the price p^N.

The consumers are identical and the firm's products are also identical so this price is also the price at which Firm B sells its output. We can represent the duopoly case mathematically, starting with each firm's revenue and profit function:

M-CHECK When we say that the functions are *mirror images*—sometimes also called *symmetrical*—it means that if we were to repeat the process for B, each time you encounter x^A, substitute in x^B, each time you encounter x^B, substitute in x^A. For example, B's profit function is $\pi^B = p(X)x^B - cx^B$.

$$\text{A's Revenue} \quad r(x^A, x^B) = p(X)x^A$$

$$\text{A's Costs} \quad c(x^A) = cx^A$$

$$\text{A's Profit} \quad \pi^A = r(x^A, x^B) - c(x^A) \tag{9.6}$$

$$= p(X)x^A - cx^A \tag{9.7}$$

Because the firms are identical, Firm B's revenue, costs and profit are *mirror images* of these. The market demand curve is given by the same function we used for the monopoly scenario.

$$\text{Market Demand} \quad p(X) = \bar{p} - \beta X$$

$$= \bar{p} - \beta(x^A + x^B) \tag{9.8}$$

We can substitute Equation 9.8 into Equation 9.7 to obtain Firm A's economic profit:

$$\pi^A = (\bar{p} - \beta x^A - \beta x^B)x^A - cx^A \tag{9.9}$$

Firm B's economic profit similarly is:

$$\pi^B = (\bar{p} - \beta x^A - \beta x^B)x^B - cx^B \tag{9.10}$$

Because each firm's profit depends on the output of the other firm, we can see that the profits of the firms are *interdependent*. In effect, B's sales shift the demand curve of Firm A to the left, meaning that for any given price, the amount that firm A can sell is less, the greater are the sales of Firm B. This is a type of *negative external effect* of one firm's production on the other firm's demand curve and profitability. As a result, like the Fishermen's Dilemma in Chapter 5, the two firms are engaged in a *strategic social interaction* and facing a coordination problem.

Best responses in a duopoly

Acting independently, each firm will choose a strategy, the quantity of output that it will produce, as a *best response* to the strategy—the quantity of output—of the other firm. Each firm will have a *best-response function*,

its profit-maximizing output for each potential level of output of the other firm. We know from the analysis of Figure 9.4 that for any given output and sales of the other firm, each firm will produce the amount that equates its marginal revenue and its marginal cost. For the duopoly case M-Note 9.1 shows that this gives us, for Firm A, the following rule to follow:

$$\text{A's best response} \qquad x^A(x^B) = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}x^B \tag{9.11}$$

Comparing Equation 9.11 to the profit-maximizing rule for the monopoly in Equation 9.4 we see that Firm A's best-response function says the following:

- *First term*: produce what you would have produced had you been a monopoly, $\frac{\bar{p} - c}{2\beta}$ *minus* the
- *Second term*: one half what the other firm produces, $\frac{1}{2}x^B$.

M-NOTE 9.1 Best responses in a duopoly

Firm A's profits are:

$$\pi^A = (\bar{p} - \beta x^A - \beta x^B)x^A - cx^A$$

Given the output of Firm B, we find the output of Firm A that maximizes A's profits by differentiating the above equation with respect to x^A and setting the result equal to zero.

$$\frac{\partial \pi^A}{\partial x^A} = (\bar{p} - \beta x^B) - 2\beta x^A - c = 0 \tag{9.12}$$

Equation 9.12 requires that Firm A maximizes profit by choosing the output level that makes marginal revenue equal to marginal cost:

$$mr^A(x^A, x^B) = (\bar{p} - \beta x^B) - 2\beta x^A = c = mc(x^A) \tag{9.13}$$

Rearranging Equation 9.13 to isolate x^A, we find:

$$x^A(x^B) = \frac{\bar{p} - c - \beta x^B}{2\beta}$$

And rearranging further to isolate the term involving x^B we get:

$$\text{A's best response} \qquad x^A(x^B) = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}x^B \tag{9.14}$$

We can repeat the derivation for Firm B, to find Firm B's best-response function:

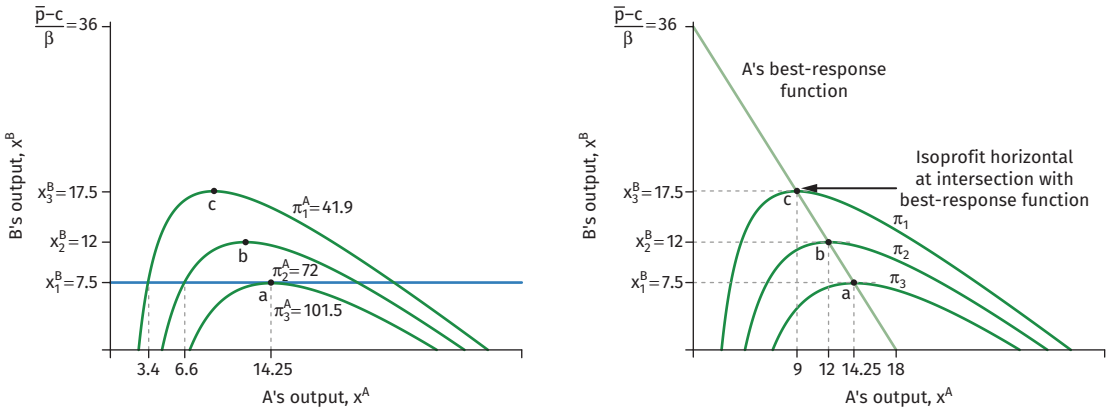
$$x^B(x^A) = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}x^A \tag{9.15}$$

The two best-response functions show that as Firm A produces more, Firm B produces less, and vice versa.

Isoprofit curves in a duopoly

To understand the strategic relationship between the two duopolists we introduce Firm A's **isoprofit curve** that shows those combinations of

Figure 9.5 Isoprofit and best-response curves for one duopolist. In panel (a), three of Firm A's isoprofit curves are shown along with a horizontal blue line indicating a hypothetical level of output that Firm B might choose, namely 7.5 units of output. Panel (b) shows the derivation of A's best-response function plotted in dark green. Firm A's isoprofit curves are the same as in panel (a). The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$. A's best-response function is given by Equation 9.11 (which is the same as 9.14 in M-Note 9.1).



(a) The feasible set and choosing a best-response isoprofit

(b) The duopolist's best-response function

outputs (x^A, x^B) that result in the given level of profit for the owners of Firm A. Figure 9.5 presents three isoprofit curves, each corresponds to a different level of profit: π_1^A, π_2^A and π_3^A where $\pi_3^A > \pi_2^A > \pi_1^A$. Firm A's profits are greater for isoprofit curves *closer* to the horizontal axis. The reason is that the less Firm B produces (that is the closer to the horizontal axis) the higher will be Firm A's profit for any level of output that it chooses.

Firm A's best response function is constructed exactly as were the best response functions of the two fishermen in Chapter 5. To see how this is done, suppose Firm B is producing and selling 7.5 units, and Firm A considers producing some small amount (3.4), placing it on the uppermost isoprofit curve π_1^A in Figure 9.5 (a) (which is the lowest level of profits shown). The owners of Firm A could do better if they proceeded to the right along the horizontal line (producing more) until they encountered the middle isoprofit curve π_2^A in Figure 9.5 (a) producing 6.6. But they could do still better if they increased output to 14.25.

Producing more than 14.25 would bring them back down to the middle isoprofit curve with lower profits. So, $x^A = 14.25$ is Firm A's best response to Firm B's producing 7.5. Remember, as was the case with the fishermen, this is an entirely hypothetical exercise: we have no reason to think that Firm B will in fact produce 7.5. We are simply doing a "what–if" thought experiment to map what would be the best response for Firm A if Firm B actually were to produce that output.

! REMINDER Remember that when Abdul and Bridget engaged in the Fishermen's Dilemma, the same was true of their indifference curves in (h^A, h^B) coordinates: they had higher utility the closer their indifference curve was to their own effort's axis. We also used isoprofit curves in Chapter 8, but in that chapter the isoprofit curve was determined by the price of the good and the quantity of output rather than by the two firms' quantities.

We can generalize this example. A horizontal line from Firm B's axis representing a given level of B's output is the feasible frontier for Firm A's constrained profit maximization, and Firm A's best response is the level of output where the isoprofit curve is tangent to the horizontal line. We show three dashed lines in Figure 9.5 (b) corresponding to three levels of output by Firm B: x_1^B, x_2^B and x_3^B. For each of these *given* levels of output by Firm B, Firm A's best response is on the highest isoprofit curve that is *tangent* to the horizontal line representing Firm B's output.

Firm A's best-response function is made up of points on the isoprofit curves like points **a**, **b**, and **c** in Figure 9.5 where the isoprofit curve is *horizontal*. The best response must be horizontal because at the profit maximum the isoprofit curve is tangent to a horizontal line.

An equivalent figure for Firm B would show the best-response function made up of the points on of B's isoprofit curves are tangent to a *vertical* line representing a given level of output by Firm A. (See Figure 9.15.)

M-NOTE 9.2 **The slope of an isoprofit curve**

Here we show why the best response of Firm A to any level of output of Firm B is where Firm A's isoprofit curve is horizontal.

The equation for a particular isoprofit curve of Firm A—one with $\pi^A = k$ where k is some constant—has the form:

$$\pi^A(x^A, x^B) = \left(\bar{p} - \beta(x^A + x^B)\right)x^A - cx^A = k$$

The isoprofit curve for this particular level of profit (k) is made up of points with differing levels of x^A and x^B but with the same level of profit, namely $\pi^A = k$. Because k is a constant these points necessarily satisfy the following equation:

$$d\pi^A = dx^A \cdot \frac{\partial \pi^A}{\partial x^A} + dx^B \cdot \frac{\partial \pi^A}{\partial x^B} = 0 \tag{9.16}$$

Which we can rearrange as follows:

$$dx^B \cdot \frac{\partial \pi^A}{\partial x^B} = -dx^A \cdot \frac{\partial \pi^A}{\partial x^A} \tag{9.17}$$

To find the slope of the isoprofit we need to find $\frac{dx^B}{dx^A}$:

$$\text{slope} = \frac{dx^B}{dx^A} = -\frac{\partial \pi^A / \partial x^A}{\partial \pi^A / \partial x^B} \tag{9.18}$$

The denominator is always negative because the more one firm sells, the lower will be the residual demand and hence the profits of the other firm.

Where:

$$\partial \pi^A / \partial x^A = 0 \tag{9.19}$$

the slope is zero so the isoprofit is flat. This point is also the best response of Firm A to the level of output of Firm B, because Equation 9.19 is the condition defining A's profit-maximizing output, as you can see from M-Note 9.1.

CHECKPOINT 9.2 Reinterpreted BRFs Redraw Figure 9.5, but instead draw Firm B's best-response function and isoprofit curves.

Nash equilibrium in Cournot duopoly

In Figure 9.6 we plot the duopolists' best responses together. Because the firms are identical, the best-response functions are symmetrical. So, for example the vertical axis intercept of A's best-response function (36) is the same as the horizontal axis intercept of B's best-response function. The symmetrical best-response functions together with the fact that neither firm has any particular bargaining advantage—such as being first mover or having the power to make a take-it-or-leave-it offer—also means that at the Nash equilibrium, the firms will produce the same output. So the Nash equilibrium will lie on the 45-degree line in the figure where the firms' outputs are equal.

Each best-response function is negatively sloped because the more one firm produces the less will be the profit-maximizing level of the other firm's output. The Nash equilibrium occurs at the intersection of the firms' best-response functions: where each firm best responds to the strategy of the other player. At the equilibrium, the firms produce x^{AN} and x^{BN} and both sell their output at the following price:

$$p^{N} = \bar{p} - \beta(x^{AN} + x^{BN}) \tag{9.20}$$

Are we certain that this is a Nash equilibrium? We constructed it as a mutual best response, so it should be. But in our construction we assumed that each firm took the other's output and sales as given, their customers already 'extracted' from the market, like fish taken from the lake by another fisherman and no longer "available" to be caught.

Have we overlooked any opportunity for profit that the actors might adopt? The owner of one of them, Firm A for example, might reason that seeing that the products of the two firms are identical, he could capture the entire market just by offering a price somewhat lower than p^{N}.

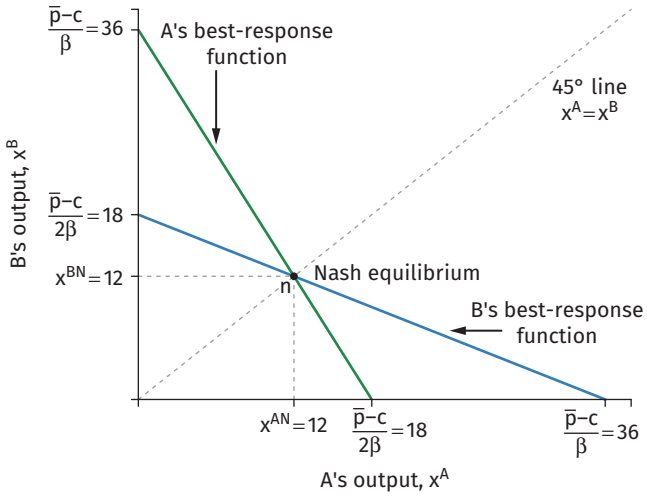
But suppose he tried this. Remember the other firm has already produced X^{BN} and is going to sell that amount at the highest possible price (whatever that price is). So Firm B would match or beat any price Firm A selected. As a result the price-cutting strategy would be self-defeating. This confirms that the intersection of the two best-response functions is indeed a Nash equilibrium.

How does the duopoly outcome contrast with the monopoly case?

- the presence of the second firm (B) dilutes the monopoly power of the first firm (A);
- we say that B *crowds* the market, leading to a larger total output than in the monopoly case;
- therefore, there is a lower market price closer to marginal cost; and
- there is lower total economic profit of the two firms compared to the single monopolist.

HISTORY Seeing that Antoine Cournot developed this model almost a century before John Nash (1928–2015) was born, it may seem strange that we do not refer to the Cournot equilibrium, and economists sometimes do. But to limit confusion we will stick to the more familiar term *Nash* equilibrium. The best-response functions that we use here (and in the case of the interacting fishermen earlier) are also Cournot's idea.

Figure 9.6 Best-response functions for the two duopolists. Because the Nash equilibrium is a mutual best response it must be a point on both best-response functions. There is only one such point, namely, the intersection. The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$. The best-response functions are given by Equations 9.14 and 9.15 in M-Note 9.1.



M-NOTE 9.3 Nash equilibrium output with two firms

The output levels given by the intersection of the duopolists' best-response functions is a Nash equilibrium. To find these equilibrium outputs, we substitute B's best-response function (Equation 9.15) into A's best-response function (Equation 9.14):

$$x^A = \frac{\bar{p}-c}{2\beta} - \frac{1}{2}\left(\frac{\bar{p}-c}{2\beta} - \frac{1}{2}x^A\right)$$

$$x^A = \frac{\bar{p}-c}{2\beta} - \frac{\bar{p}-c}{4\beta} + \frac{1}{4}x^A$$

Subtract $\frac{1}{4}x^A$ $\quad \frac{3}{4}x^A = \frac{\bar{p}-c}{4\beta}$

Multiply through by $\frac{4}{3}$ $\quad x^A = \frac{4}{3}\left(\frac{\bar{p}-c}{4\beta}\right)$

Cournot (Nash) equilibrium output $\quad \therefore x^{AN} = \frac{\bar{p}-c}{3\beta} = x^{BN}$ (9.21)

If you derive B's best-response function as an exercise, you will see that B's Nash equilibrium output is a mirror image of A's as shown in Equation 9.21.

There is a simpler way to solve for the Nash equilibrium output in this symmetric setting. At the symmetric Nash equilibrium, both firms will choose the same output $x^{AN} = x^{BN} = x^N$. Therefore, x^N should best respond to itself, that is:

continued

$$x^N = \frac{\bar{p}-c}{2\beta} - \frac{1}{2}x^N$$
$$\frac{3}{2}x^N = \frac{\bar{p}-c}{2\beta}$$
$$x^N = \frac{\bar{p}-c}{3\beta}$$

Graphically, the symmetric Nash equilibrium is the intersection of one of the best-response functions and the 45-degree line $x^B = x^A$, as shown in Figure 9.6.

M-NOTE 9.4 Prices, total output, and profits with two firms

To find the prices and profits associated with the Nash equilibrium levels of output, we sum x^{AN} and x^{BN} to get total market output and then use the inverse demand function to find the equilibrium price. From the equilibrium price, we can find each firm's profit and sum them to find the total economic profit in the duopoly. Remember, the superscript N denotes a Nash equilibrium.

Total output $\quad X^N = x^{AN} + x^{BN} = \dfrac{2(\bar{p}-c)}{3\beta}$

Using the inverse demand function
$$\begin{aligned} p^N &= \bar{p} - \beta(x^{AN} + x^{BN}) \\ &= \bar{p} - \beta\left(\frac{2(\bar{p}-c)}{3\beta}\right) \\ &= \frac{1}{3}\bar{p} + \frac{2}{3}c \end{aligned}$$

Add and subtract $\frac{1}{3}c$ $\quad p(X) = c + \frac{1}{3}(\bar{p}-c)$

Firm profits $\quad \pi^{AN} = (p(X)-c)x^N = \dfrac{(\bar{p}-c)^2}{9\beta} = \pi^{BN}$

Total profits $\quad \Pi^N = (p^N - c)(x^{AN} + x^{BN}) = \dfrac{2(\bar{p}-c)^2}{9\beta}$

Relative to the monopoly outcome, market output is higher, market price is lower, and total profits are lower in the duopoly.

CHECKPOINT 9.3 Willingness to pay, slope of the demand, and duopoly output Let us assume new values for the parameters in which we are interested: let $\bar{p} = 100$ and $\beta = 1$, and the firm has constant marginal costs $c = 1$.

a. Find each firm's best-response function.

b. Find the Nash equilibrium quantity for each firm.

c. How much profit does each firm make? What is the price and how much greater is the price than marginal costs (that is, what is the markup)?

9.5 OLIGOPOLY AND "UNLIMITED COMPETITION": FROM A FEW FIRMS TO MANY FIRMS

A feature of Cournot's approach is that it allows us to use a single general model to study the entire range of competition from monopoly to an industry with very many firms. This contrasts with a common approach, which is to present entirely distinct models of profit maximization by a single firm (monopoly or monopolistic competitor) on the one hand and perfect competition on the other. We include a summary of the general results that we derive in Table 9.2.

Oligopoly

We can represent the process of competition and its outcome graphically for n firms, taking account of two characteristics of the Nash equilibrium:

- Because the firms are identical, in equilibrium they produce the same level of output.
- For each firm, the Nash equilibrium output must be a best response to the *total output* being produced by the $n-1$ *other* firms.

To see how this works we arbitrarily pick one firm ("the firm") and study its choice of an output level given what all of the other firms are producing. It does not matter which firm we pick because they are identical. If "the firm" best responds to the outputs chosen by the other firms and those firms (like "the firm") also best respond to what the other firms do, then the result is a Nash equilibrium.

We can visualize the equilibrium by plotting the *total output* of the other firms, X^{-i}, on the vertical axis and the best response of "the firm," x^i, on the horizontal axis, as we do in Figure 9.7.

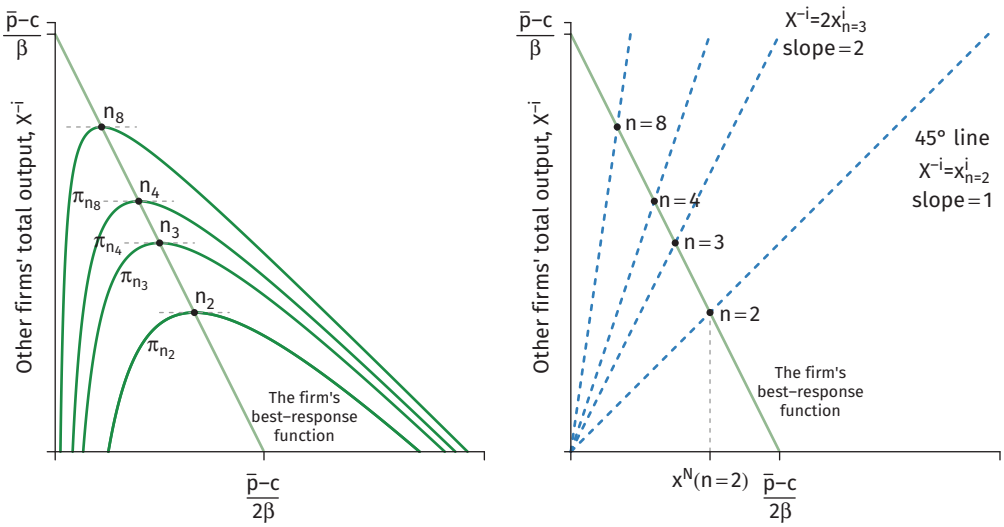
In Figure 9.7, the firm's best-response function is plotted against the total output of the other firms in the industry. The best response shows each point at which the firm's isoprofit curve is tangent to a horizontal line (shown by the dashed gray lines) indicating some hypothetical given output of the other firms, such as at points $\mathbf{n_2}$ and $\mathbf{n_3}$ (each of which have corresponding outputs for the firm).

As before, the best-response function is made up of points where the isoprofit curves are *horizontal*. In the Nash equilibrium, two conditions must be met:

- the output of the firm must be on its best-response function; and
- the output of each of the $n-1$ other firms must be equal to the best-response output of "the firm", so the total output of the other firms has to satisfy $X^{-i} = (n-1)x^i$.

M-CHECK The superscript $-i$ means "not Firm i" and uppercase letters refer to totals, so X^{-i} is the total output of the $n-1$ firms *other than* firm i.

Figure 9.7 Nash equilibrium output with *n* firms. In both panels the arbitrarily selected firm's output is on the horizontal axis, and the total output of the remaining $n-1$ firms is on the vertical axis. Panel (a) shows the construction of the best-response function of "the firm." Each of the other firms has an identical best-response function. In panel (b), each dashed ray from the origin has a slope of $n-1$ and shows for different values of n, the total output levels of the other firms that are $n-1$ times the output level of "the firm." We know that this condition must be true in equilibrium because the firms are identical.



(a) The firm's isoprofits along with their best-response function

(b) The firm's Nash equilibrium output for different numbers of the other firms

The second condition means, as you have already seen, that if there is just *one other* firm, then equilibrium lies on the 45-degree ray from the origin. The ray from the origin is a line $X^{-i} = x^i$ which has a slope of 1, namely the number of other firms.

As a result the Nash equilibrium is the point on both "the firm's" best-response function and the line $X^{-i} = (n-1)x^i$, namely their intersection at each of the Nash equilibrium points: $\mathbf{n_2}$ for $n = 2$ firms, $\mathbf{n_3}$ for $n = 3$ firms, and so on.

We show in M-Note 9.5 that the firm's best response when it faces $n-1$ competitors is as follows:

$$x^i(X^{-i}) = \frac{\bar{p}-c}{2\beta} - \frac{1}{2}X^{-i} \tag{9.22}$$

Equation 9.22 is the monopoly firm's level of output $\left(\frac{\bar{p}-c}{2\beta}\right)$ minus $\frac{1}{2}X^{-i}$, which is the same output that we found (in Equation 9.11) for Firm A in the duopoly case where $n = 2$ and $X^{-i} = x^B$. And Equation 9.22 replicates the output choice for the monopoly (given in equation 9.4), too, because in that case, $n = 1$ so the second term drops out (the output of the other firm is zero because there is no other firm).

! REMINDER M-Note 9.1 and Figure 9.5 explained why the best response of Firm A is at a point where A's isoprofit curves are horizontal.

! REMINDER If the index i has been used to identify firms (firm i, firm j etc.), then the superscript $-i$ means "all the other firms that are *not* i." The same notation is used for people, players in games, and so on.

Figure 9.7 shows that with eight firms in the industry each of them is producing less than the monopoly output; but the total output is greater because the inverse demand curve is downward-sloping. The market price will therefore be *lower* with many firms competing.

M-NOTE 9.5 Cournot competition with many firms: best responses

To study the case with many firms we focus on a single firm i (called "the firm") whose output is x^i and all of the $n-1$ other firms taken together, whose total output is X^{-i}. So total production and sales is $X = x^i + X^{-i}$ and the demand curve seen by the firm when there are $n-1$ other firms is:

$$\text{Demand faced by the firm} \qquad p(x^i, X^{-i}) = \bar{p} - \beta(x^i + X^{-i}) \tag{9.23}$$

From this we have the firm's revenue:

$$\text{The firm's revenue} \qquad r(x^i, X^{-i}) = p(x^i, X^{-i})x^i = (\bar{p} - \beta X^{-i})x^i - \beta(x^i)^2$$

And the firm's profit:

$$\text{The firm's profit} \qquad \pi(x^i, X^{-i}) = \left(\bar{p} - \beta(n-1)X^{-i}\right)x^i - \beta(x^i)^2 - cx^i \tag{9.24}$$

Therefore, given X^{-i}, we partially differentiate the total revenue function with respect to x^i:

$$\text{The firm's marginal revenue} \qquad mr(x^i, X^{-i}) = (\bar{p} - \beta X^{-i}) - 2\beta x^i \tag{9.25}$$

The firm maximizes profit by choosing the output level that equates marginal revenue to marginal cost:

$$mr(x^i, X^{-i}) = (\bar{p} - \beta X^{-i}) - 2\beta x^i = c$$

$$\text{Solving this equation for } x^i \qquad x^i(X^{-i}) = \frac{\bar{p} - c - \beta X^{-i}}{2\beta} \tag{9.26}$$

Rearranging Equation 9.26 we can find the firm's best-response function:

$$\text{The firm's best response} \qquad x^i(X^{-i}) = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}X^{-i} \tag{9.27}$$

M-NOTE 9.6 Nash equilibrium output with many firms

The intersection of the firm's best-response curve with the line $X^{-i} = (n-1)x^i$ is the Nash equilibrium, which we find by substituting it into the firm's best-response function, Equation 9.27:

$$x^i = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}\left[(n-1)x^i\right]$$

$$x^i + \frac{1}{2}(n-1)x^i = \frac{\bar{p} - c}{2\beta}$$

$$\frac{2x^i + (n-1)x^i}{2} = \frac{\bar{p} - c}{2\beta}$$

$$\frac{(n+1)x^i}{2} = \frac{\bar{p} - c}{2\beta}$$

$$\text{Multiplying by } \frac{2}{(n+1)} \qquad x^N(n) = \frac{\bar{p} - c}{(n+1)\beta} \tag{9.28}$$

M-NOTE 9.7 Equilibrium price n firms

We can substitute the equilibrium outputs from Equation 9.28 into the inverse demand curve to find the price and each firm's profit:

$$\begin{aligned} p^N &= \bar{p} - \beta n x^N \\ &= \bar{p} - \beta n \left(\frac{\bar{p} - c}{\beta(n+1)} \right) \qquad (9.29) \\ &= \frac{1}{(n+1)}\bar{p} + \frac{n}{(n+1)}c \end{aligned}$$

Add and subtract $\frac{1}{(n+1)}c$ $\quad \therefore p^N(n) = c + \frac{1}{(n+1)}(\bar{p} - c)$ (9.30)

As Equation 9.30 shows, the price in Cournot competition is equal to the marginal cost, c, plus $\frac{1}{n+1}$ times the difference between the maximum price, $\bar{p}$, and marginal cost.

Notice that for any given number of firms in the market, the price *does not* depend on β, the (negative of the) slope of the inverse demand curve. Phrased differently, we might ask why do the two βs cancel out in Equation 9.29? With the demand curve we are using, if β increases, the demand curve rotates downward with the vertical axis intercept unchanged, so two things happen:

- A larger β means a steeper (relatively more inelastic) demand curve, which would induce firms to *raise* prices.
- But, a larger β also corresponds to a *decrease* in demand, which means prices would *decrease*.

CHECKPOINT 9.4 Higher willingness to pay and competition Let us assume the following values: $\bar{p} = 100$, $\beta = 1$, and the firm has constant marginal costs $c = 1$. Let $n = 43$ firms.

a How much output will each firm produce with the parameters as described above?

b What will the market price be, how much profit will each firm make, and what is the markup over marginal costs?

9.6 UNLIMITED COMPETITION AND THE PRICE MARKUP OVER COSTS

Additional competing firms dilute the market power of the incumbent firms (those already in the industry). Adding firms to the market leads to each firm producing less but with a larger total output. As a result of the greater output the market price decreases and becomes closer to marginal cost and the firms make lower total profits approximating the results of the model of perfect competition. Like Cournot, we call this very large n case "unlimited competition." We will use Cournot's term, because, while these

results are similar, his model is quite different from the model associated with the term's perfect competition in which firms act as price-takers.

M-Note 9.7 shows that the price at which each of the firms sell their product is:

$$p^{\mathrm{N}} = c + \frac{1}{n+1}(\bar{p} - c) \tag{9.31}$$

Equation 9.31 says that as n becomes very large the following happens:

- the second term in Equation 9.31, involving $\bar{p} - c$, approaches zero;
- the price therefore falls to just above the cost per unit (c); and
- the firms' economic profits all but disappear because profit per unit, that is, the markup $(p - c)$ approaches zero.

Because price equal to marginal cost and economic profits equal zero are features of the equilibrium of the model of perfect competition among price-takers, we take "unlimited competition" in the Cournot model as the "competitive" pole of the continuum extending all the way to $n = 1$, namely monopoly. The Cournot model therefore includes results approximating "perfect competition" as a limiting case. We will also see that it can incorporate the effects of barriers to firms entering the market, so that "perfect competition" is never realized.

! REMINDER The markup ratio—here written as $\mu(n)$ to stress the markup ratio's dependence on the number of firms in the industry—is defined as: $\mu(n) = \frac{p(n)-c}{c}$.

The markup ratio introduced in Chapter 8—profits divided by costs—falls as the number of firms—that is the extent of competition—increases. In Figure 9.8 we plot $p^{\mathrm{N}}(n)$ along with the cost c per unit which does not vary with the number of firms. The difference between the two is the markup. Notice from Figure 9.8 and Equation 9.31 that the markup will be positive for any finite number of firms, so firms will make some economic profits.

At the end of Chapter 8 we asked: Why does the size of firms not grow forever, eventually eliminating competition? Here we ask the opposite question: Why do firms not continue to enter the market until there are so many firms competing that it approximates "perfect competition?"

M-NOTE 9.8 The markup ratio, $\mu(n)$ and the degree of competition

We know from M-Note 9.7 that in the Cournot model the price with n firms is given by Equation 9.30, which is:

$$p^{\mathrm{N}}(n) = c + \frac{1}{n+1}(\bar{p} - c) \tag{9.32}$$

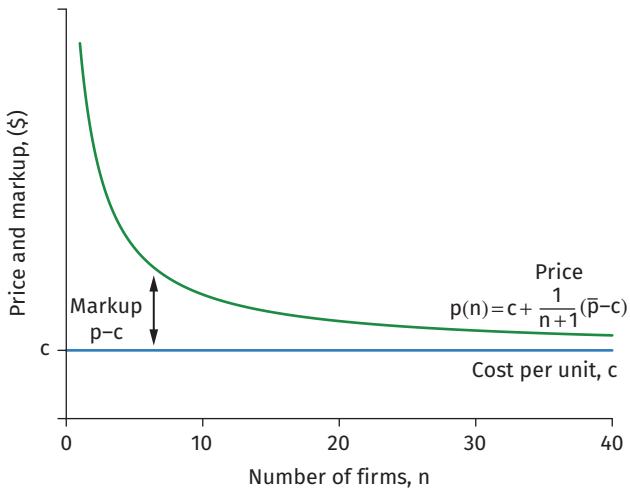
Let us rearrange the equation and find the markup ratio:

$$\text{Subtract } c: \quad p^{\mathrm{N}} - c = \frac{1}{n+1}(\bar{p} - c)$$

$$\text{Divide by } c: \quad \frac{p^{\mathrm{N}} - c}{c} = \frac{1}{n+1}\left(\frac{\bar{p} - c}{c}\right) = \mu(n) \tag{9.33}$$

Equation 9.33 shows that the markup ratio for an industry with a degree of competition given by n firms is the maximum possible markup ratio times the inverse of $n+1$. An increase in competition (increase in n) will *reduce* the markup ratio as can be seen from Equation 9.33.

Figure 9.8 Price and markup over costs. In M-Note 9.7 we showed that the Nash equilibrium price with n firms is given by Equation 9.31, which is the green curve in the figure. It slopes downward because while each firm produces less, the total effect of more firms is to *increase* industry output. The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$.



CHECKPOINT 9.5 The markup ratio Assume parameter values of $\bar{p} = 20$, slope $\beta = 0.5$, and marginal cost, $c = 2$.

a Calculate the markup ratio in each of the following situations: monopoly, duopoly, and oligopoly with $n = 5$ and $n = 43$ respectively.

b Draw a figure with the number of firms (n) on the horizontal axis and the markup ratio on the vertical axis. Label four points corresponding to your answers to a. and sketch the line connecting these points.

! REMINDER The *opportunity cost of capital* is the rate of return on an *alternative* use of a firm owner's funds.

9.7 MARKET DYNAMICS: BARRIERS TO ENTRY AND THE EQUILIBRIUM NUMBER OF FIRMS

We have so far illustrated our three cases by assuming particular values of n the number of firms: $n = 1, n = 2$, and $n > 2$ (either few, or many). We now consider what determines the number of firms, meaning: what determines the extent of competition.

To understand the factors affecting the extent of competition, we ask: What determines the *equilibrium* number of firms, that is, the number of firms that does not change over time? We asked a similar question in Chapter 5 when we studied the equilibrium number of people fishing on a lake. This was the largest number of people fishing such that the utility

! REMINDER In Chapter 5 the Nash equilibrium number of people fishing on the lake is the largest whole number n^N such that the utility of those fishing $u(n^N)$ is not less than the utility that they would receive at their fallback option, or $u(n^N) \geq u_z$.

M-CHECK The equilibrium number of firms must be a whole number because the entry of a "fractional firm" would be meaningless, just as with the case of the equilibrium number of people fishing on the lake it did not make sense to talk about a "fractional fisherman" taking up fishing.

from fishing is not less than the utility of the fallback option that people would experience if they did not fish.

The equilibrium number of firms in an industry is determined in a similar way. For a firm to attempt entry to the market it must be the case that the expected profits to be obtained by attempting to enter the market are *greater* than or equal to the *opportunity cost* of entering, or greater than the fallback option of the potential entrant. So the Nash equilibrium number of firms in the industry, n^N, will be the largest whole number such that the **expected profit** exceeds the opportunity cost of attempting to enter. We will see that the equilibrium number of firms will be smaller:

- the more the equilibrium price *falls* as the number of firms in the market *increases*;
- the *more profitable* are the alternative uses of the funds that a firm's owners might commit to entering the market (the opportunity cost of capital);
- the extent of *barriers to entry* of new firms.

! REMINDER The term barriers to entry refers to anything making it difficult for new firms to enter a market, including intellectual property rights that give incumbent firms a monopoly on particular technologies or, economies of scale in production or in demand, and predatory pricing.

! REMINDER An *incumbent firm* is a firm that is already producing in the industry.

The decision to enter an industry and barriers to entry

We model barriers to entry as a probability that a firm attempting to enter will fail. The greater are the barriers to entry, the higher is the probability of failure for a firm attempting to enter. The new firm attempting to enter a market is identical to the *incumbent* firms, so if it succeeds in entering, it becomes just one of $n+1$ identical firms.

To determine the conditions under which a new firm would enter, we have to return to the "owner's eye view of costs" introduced in Chapter 8. In deciding whether to enter a market, the owners of the firm consider the value of the capital goods they will devote to that project. They then compare the profits they expect to make if they attempt to enter the industry with the profits they could make if they devoted those same funds to an alternative use. But entering a new industry is *risky* because the owners commit some funds to the project (entering the industry) but they do not know what the outcome of their attempt will be. To consider just the extremes it could be either:

- *Failure*: There is a probability (b) that attempting to enter will fail resulting in no revenues to offset the costs of attempted entry.
- *Success*: Alternatively, there's a probability ($1-b$) that the firm succeeds in entering the market and is able to sell its products at the same price as the other firms, thereby offsetting the costs of the initial investment.

! REMINDER

Firm profits $\pi(x^i) = (p-c)x^i$

Profit per unit $\dfrac{\pi(x^i)}{x^i} = (p-c)$

EXPECTED PROFIT In a situation of risk, expected profit is the sum of profits occurring under each contingency multiplied by the probabilities that each contingency occurs.

The key idea for modeling the risk faced by an entering firm is that it incurs its costs with certainty regardless of whether it succeeds or fails, but earns profits only if it succeeds. The entering firm intending to produce an amount x with certainty pays the cost cx. The risk arises because with some probability b (for *barriers to entry*) the firm will fail to sell the product.

A firm considering entering a market with n firms, if successful, will be in a market with a total of $(n+1)$ firms selling its product at the price that results when there is an additional firm in the market.

With barriers to entry affecting the probability that the firm will successfully enter an industry, the firm therefore calculates its **expected profit** per unit of output produced as follows:

$$\text{Expected profit per unit} \quad \frac{\hat{\pi}}{x^i} = \underbrace{-cb}_{\text{Failed entry}} + \underbrace{(1-b)(p(n+1)-c)}_{\text{Successful entry}}$$

$$\frac{\hat{\pi}}{x^i} = \underbrace{(1-b)p(n+1)}_{\text{Expected price}} - \underbrace{c}_{\text{Cost}} \tag{9.34}$$

The firm will attempt entry if the expected profits of entry are positive. This requires that the expected price exceed its costs:

$$\text{Expected price} \quad \hat{p}(n+1) = (1-b)p(n+1) > c \quad \text{cost} \tag{9.35}$$

The rule governing firm entry and exit is the following:

- *Entry*: If with n firms already in the market, the prospective entering firm's expected price (taking account of the probability b that the price will be zero) *exceeds* the cost (including the opportunity cost of investing capital in this firm), then the owners will decide to attempt to *enter*.
- *Exit*: Some firms will exit the industry for reasons outside the model, and in addition if at the existing n some firms' profits *fall short* of what they could make in an alternative investment, then they will *exit* the market.

The equilibrium number of firms—the number such that no new firms will be attempting entry is given by the largest whole number value of n for which Equation 9.35 is positive.

! REMINDER We introduced barriers to entry and the extent to which firms are able to mark up prices over costs in Chapter 8 (sections 8.8 and 8.9).

M-CHECK The expected price is the weighted average of the price if entry fails ($p=0$) and the price if entry succeeds ($p(n+1)$) with the weights being the probability of these two events occurring, that is b for failure and $1-b$ for success.

M-NOTE 9.9 Barriers to entry and the equilibrium number of firms

Here we show how the level of barriers to entry determines the equilibrium number of firms in the industry. From Equation 9.35, we know that potential entrants consider an expected price, $\hat{p}$, when deciding whether to enter.

continued

We can now substitute $\hat{p}$ (Equation 9.30) into the condition the number of firms remaining constant given by Equation 9.35 to find n^{N} the equilibrium number of firms:

$$
\begin{aligned}
(1-b)p(n) &= c \\
(1-b)\left(c+\frac{1}{n+1}(\bar{p}-c)\right) &= c \\
(1-b)c+(1-b)\left(\frac{1}{n+1}(\bar{p}-c)\right) &= c \\
(1-b)\left(\frac{1}{n+1}(\bar{p}-c)\right) &= bc \\
\frac{1-b}{n+1}(\bar{p}-c) &= bc \\
\text{Multiply by } (n+1) \qquad (1-b)(\bar{p}-c) &= nbc+bc \\
\text{Subtract } bc \qquad nbc &= (1-b)(\bar{p}-c)-bc \\
\text{Simplify RHS} \qquad nbc &= \bar{p}(1-b)-c \\
\text{Divide by } bc \qquad n^{\mathrm{N}} &= \frac{\bar{p}(1-b)-c}{bc} \qquad (9.36)
\end{aligned}
$$

The Nash equilibrium number of firms n^{N} is greater therefore:

- the greater is the consumers' maximum willingness to pay ($\bar{p}$);
- the smaller are the barriers to entry (b); and
- the smaller is the firm's marginal costs (c), which include the opportunity cost of capital (ρ).

EXAMPLE Consider a person choosing whether to purchase capital goods worth \$1,000 in starting a firm, where the risk-free annual rate of return, ρ = 5% which is the rate of profit (profits per dollar invested) had he used his funds in some other way. Then the opportunity cost of that owner's capital (what they would earn if they did not purchase the capital goods for the startup firm) is \$50 (\$1,000 × 1.05 − \$1,000) per year. Economic profits are accounting profits in excess of this opportunity cost. For more information, review these ideas in Chapter 8.

Determinants of the equilibrium number of firms

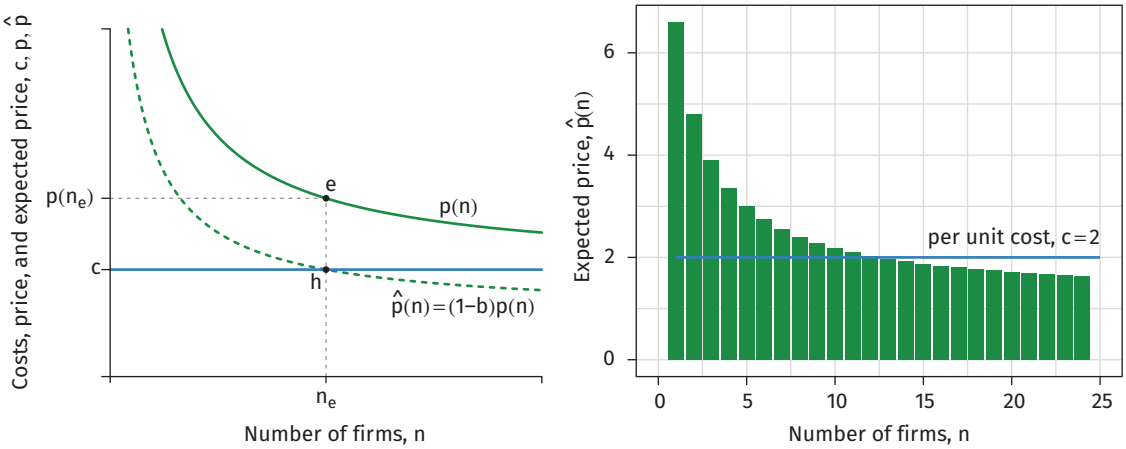
Figure 9.9 illustrates how the equilibrium number of firms in the industry is determined by Equation 9.35. The expected price ($\hat{p}(n)$) slopes downward because as more firms enter the total amount to be sold increases and the price falls (due to the downward-sloping market demand curve). The cost curve (including the opportunity cost of capital) is horizontal and does not depend on the barriers to entry, because the firm pays the cost whether or not it succeeds in selling its output.

In Figure 9.9 the price at the Nash equilibrium for each number of firms $p(n)$ is shown in green. All firms in the industry will sell at this price. With barriers to entry the probability that a firm attempting entry will fail and *not* be able to sell its output at all is b. A firm that fails to enter will face a price $p = 0$ (it can't sell its goods). So the expected price of the firm considering entering is the dashed line, $\hat{p} = (1-b)p(n)$. As $\hat{p}$ exceeds the cost as long as there are fewer than n^{N} firms in the industry, entry of new firms will occur until $n = n^{\mathrm{N}}$.

A summary of the factors affecting the degree of competition

We can now consider some of the economic factors that affect the level of competition in an industry by increasing or decreasing the equilibrium number of firms.

Figure 9.9 The equilibrium number of firms with barriers to entry. It is in a firm's interest to attempt to enter as long as $\hat{p} \geq c$, that is, as long as the expected price for entering the market is greater than the costs. In panel (b) the barriers to entry are $b = 0.4$ and costs $c = 2$, with corresponding equilibrium number of firms $n = 12$. With $n = 12$, the expected price is 2.03; with $n = 13$ the expected price falls to $\hat{p} = 1.97 < 2 = c =$ cost. So the 13th firm will *not* enter.



(a) Computing the equilibrium number of firms

(b) Equilibrium number of firms: Example (numerical example)

- An *increase in barriers to entry* (b): Increasing b will shift down the expected price function as shown in Figure 9.10, resulting in a smaller number of firms, less competition, and a *higher* price markup over costs.
- *An innovation that reduces the cost of production:* Decreasing c (not shown in the figure) shifts down the cost line, increasing the equilibrium number of firms and hence the degree of competition.
- *An increase in the opportunity cost of capital* ρ: This could occur if profitability in some other economy increased, improving the firm owner's next best alternative use of their funds, or if the central bank's monetary policy increased the cost of borrowing. Because an increase in the opportunity cost of capital increases c, the increase will *reduce* the equilibrium number of firms in the industry and support a less competitive environment.

In Figure 9.10 you can see that with *low* barriers to entry, the number of firms is given by the intersection of costs and $\hat{p}(n, b_L)$ (point **h**) with a relatively *higher* number of firms n_L^N. With *high* barriers to entry, the number of firms is given by the intersection of costs and $\hat{p}(n, b_H)$ (point **g**) with number of firms n_H^N. With greater barriers to entry, there are *fewer* firms and lower *expected profits* because of how hard it is for firms to enter

Figure 9.11 How barriers to entry affect the equilibrium number of firms. The extent of entry barriers affects the equilibrium number of firms. There are two expected price lines: with low barriers to entry (b_L) the expected price line is $\hat{p}_L(n, b_L)$ and with high barriers to entry (b_H) the expected price line is $\hat{p}_H(n, b_H)$. The high barriers to entry line is *lower* than the low barriers to entry line. This is because, with a substantial probability that the entering firm will fail (high barriers to entry), then taking account of the probability that the entrant will *not* be able to sell its goods (meaning a price of zero) the expected price is lower.

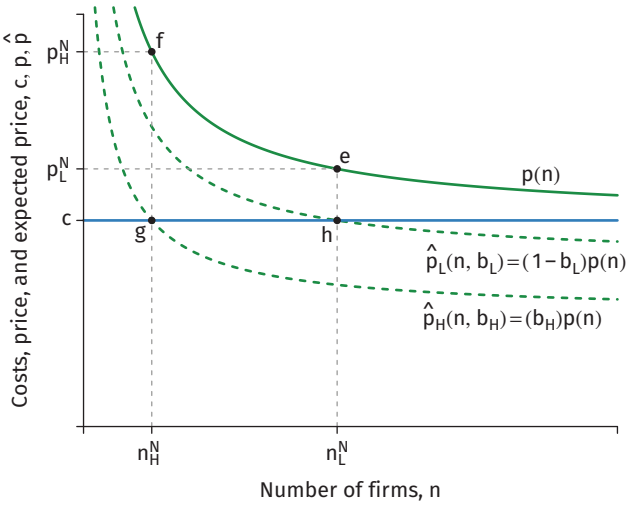
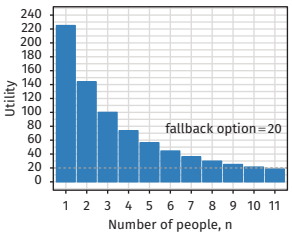


Figure 9.10 Reminder: Overexploitation of a common property resource. This is Figure 5.18 repeated here to recall a similar process: the overharvesting of fish. The height of the bars is the utility gained by the people fishing the lake and how this depends on their number. Each fishermen's fallback option (20) is similar to the firms' cost; it is the opportunity cost of entry. Utility is the benefit of entering, like the entering firm's expected price except there is no risk of failure with fishing. The equilibrium number of fishers is 10.



the market. We bring together all of the results on Cournot competition as n varies from 1 to many in M-Note 9.10.

Figure 9.10 and M-Note 9.10 show that if there are barriers to entry—whether high or low—then at the Nash equilibrium each of the firms will sell goods at a price *above* the marginal cost. With a positive price markup over costs all of the incumbent firms are making economic profits on every unit they sell. As a result, the owners of each firm would prefer to have more buyers, if they could figure out some method, such as advertising or the other modes of competition, to sell more output *without* lowering the price. We return to these additional dimensions of how firms compete in section 9.16.

M-NOTE 9.10 Summary: Nash equilibrium results for the Cournot model as the number of competing firms vary

You can generate the results for monopoly ($n = 1$), duopoly ($n = 2$), oligopoly ($n > 2$, but not large), and unlimited competition ($n \to \infty$) using the equations below. In the equations below we use the inverse demand function $p = \overline{p} - \beta X$ and the measure of barriers to entry, b. We use the superscript N to denote a Nash equilibrium value, so that, for example $x^N(n)$ ("x super N of n") means the Nash equilibrium level of output of a firm when there are n firms in the market.

continued

Firm output (Equation 9.28): $x^N(n) = \frac{1}{n+1}\frac{\overline{p}-c}{\beta}$

Industry output: $X^N(n) = nx = \frac{n}{n+1}\frac{\overline{p}-c}{\beta}$

Market price: $p^N(n) = \overline{p} - \beta X = \overline{p} - \beta\left(\frac{n}{n+1}\frac{\overline{p}-c}{\beta}\right) = c + \frac{1}{n+1}(\overline{p}-c)$

Markup ratio: $\mu^N(n) \equiv \frac{p-c}{c} = \frac{1}{c}\left(c + \frac{1}{n+1}(\overline{p}-c) - c\right) = \frac{\overline{p}-c}{(n+1)c}$

Firm profits: $\pi^N(n) = (p-c)x = \left(\frac{1}{n+1}(\overline{p}-c)\right)\left(\frac{1}{n+1}\frac{\overline{p}-c}{\beta}\right)$

$$= \frac{1}{(n+1)^2}\frac{(\overline{p}-c)^2}{\beta}$$

Industry profits: $\Pi^N(n) = n\pi = \frac{n}{(n+1)^2}\frac{(\overline{p}-c)^2}{\beta}$

Equilibrium condition: Expected revenue per unit = cost per unit

$\hat{p} = (1-b)p = c$

Equilibrium number of firms: $n = \frac{p(1-b)-c}{bc}$

CHECKPOINT 9.6 Market dynamics

a Why does $\hat{p}(n)$ slope downward?

b What will happen to $\hat{p}(n)$, the number of firms, the degree of competition and the price markup when b decreases? Why?

c Why does an innovation that reduces the cost of production increase the equilibrium number of firms in the economy?

d If consumers' maximum willingness to pay increases, how does this impact the number of firms in the market and why?

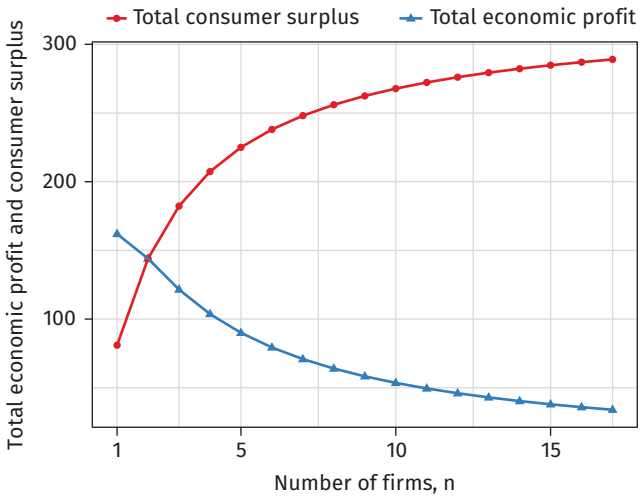
9.8 A CONFLICT OF INTEREST: PROFITS, CONSUMER SURPLUS, AND THE DEGREE OF COMPETITION

Another strategy that owners might pursue is to find ways of increasing b–barriers to entry–so as to reduce the number of firms in the industry. You can see from Figure 9.8 that the profits per unit produced–the markup–is greater the fewer firms there are competing.

While barriers to entry benefit the owners of the incumbent firms, they reduce the economic benefits of consumers, measured by consumer surplus. You can see this in Figure 9.12. For the monopoly case where $n = 1$ the red and blue dots show that profits (the blue dot = 162) is twice the red dot (consumer surplus = 81). You have already seen this result in the

! REMINDER Consumer surplus, a measure of consumer welfare made possible by an exchange, is the difference between each consumer's willingness to pay and what they actually pay for each unit of the good that they consume, generally summed over all purchasers of the good.

Figure 9.12 Conflicts of interest between firms' owners and consumers: consumer surplus and economic profit. Total consumer surplus increases and total economic profit decreases as the number of firms competing for the same market increases. The parameters used are the same as in the other figures: $\bar{p} = 20$, $\beta = 0.5$, $c = 2$.



top panel of Figure 9.4 where the profit rectangle is twice the size of the consumer surplus triangle.

Adding even just a single competing firm–the duopoly or $n = 2$ case–increases competition and lowers the price sufficiently to bring profits down somewhat and to substantially increase consumer surplus, so that profits and consumer surplus are equal at 144.

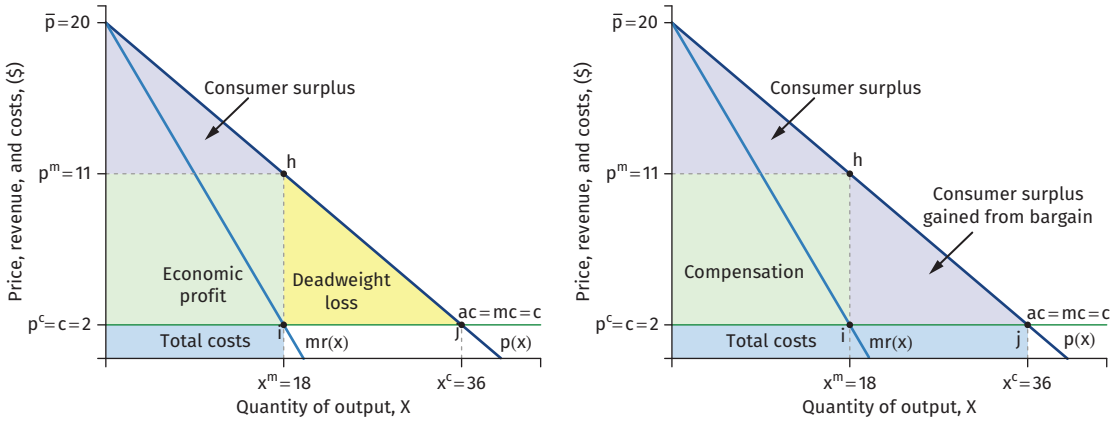
Notice something important: with $n = 2$ profits plus consumer surplus, namely 288 is greater than with $n = 1$ where profits plus consumer surplus is 243. This means that the conflict of interest between owners and consumers is not a zero-sum game: when the number of firms goes from 1 to 2 the sum of the gains to consumers and losses to owners does not sum to zero, it sums to 45. The same pattern persists as n goes to 3 and higher numbers: the consumers gain more in consumer surplus than the owners lose in reduced profits. Where did the extra benefits come from?

CHECKPOINT 9.7 Conflicts of interest Use the data in Figure 9.12 to show that the conflict of interest between consumers and firm owners is *not* a zero-sum game.

9.9 LIMITED COMPETITION AND INEFFICIENCY: DEADWEIGHT LOSS

The answer is that, in addition to redistributing income toward owners' profits and *away* from consumers, limited competition is also a source of

Figure 9.13 Monopoly and deadweight loss: why limited competition is Pareto inefficient. In panel (a) is the profit-maximizing output and price of a monopolist. Panel (b) illustrates the hypothetical case of the same firm producing twice as much and selling the resulting output (x^c) at the highest price feasible given the demand curve. If the firm produced x^c and the consumers as a group paid the owners of the firm an amount equal to the area of the green-shaded area in panel (b) labeled "compensation" the consumers would be better off by an amount equal to the area of the "deadweight loss" triangle, and the owners no worse off.



inefficiency. As the number of firms competing increases, the inefficiency is reduced so the total benefits—profits plus consumer surplus—increases. Erecting barriers to entry so as to limit competition is a strategy followed by firm owners that gives them a larger slice of a smaller pie.

Limited competition and deadweight loss

We measure the extent of the inefficiency by a quantity called the deadweight loss which you also encountered in Chapter 8 (e.g. in Figure 8.21). The deadweight loss represents the quantity of either economic profits or consumer surplus that could have been realized if "the firm" or any firm had produced more.

This is shown in Figure 9.13 (a) as the area of the yellow-shaded triangle for the case of a single firm—a monopoly like the one in Figure 9.4. Recall that the firm will produce the quantity of output at which its marginal revenue equals its marginal cost. It will then sell that output at the maximum price possible, indicated by point **h** on the demand curve in the figure. The total revenue of the firm is composed of two parts: costs (including the opportunity cost of the capital goods used) and economic profits, their quantities indicated by the area of the green and blue rectangles.

M-CHECK The areas and equations in Figure 9.13 (a) are:

Consumer surplus: $cs(x)$

$$= \frac{1}{2}(\bar{p} - p^m)x^m$$

Deadweight loss: $dwl(x)$

$$= \frac{1}{2}(p^m - c)(x^c - x^m)$$

Marginal revenue: $mr(x)$

$$= \bar{p} - 2\beta x$$

Inverse demand: $p(x)$

$$= \bar{p} - \beta x$$

! REMINDER Remember from Chapter 8 that *deadweight loss* is the feasible consumer surplus or economic profits that are not realized because price is above marginal cost so that too few units are produced and sold.

To see why the monopoly producing x^m units of output and selling them at the price p^m is inefficient, think of an alternative. Suppose hypothetically that the firm produced x^c and sold that amount at the maximum price it could, namely p^c which is also equal to the marginal and average cost. (Don't ask *why* the firm would do this, just imagine that it did.) Then there would be *no* profits and consumer surplus would be the entire area under the demand curve and above the cost price line, that is, the two purple triangles and the light-blue rectangle in panel (b) of Figure 9.13.

If this were to occur in reality, consumers would benefit and owners lose. So the hypothetical doubling of output is *not* a Pareto improvement.

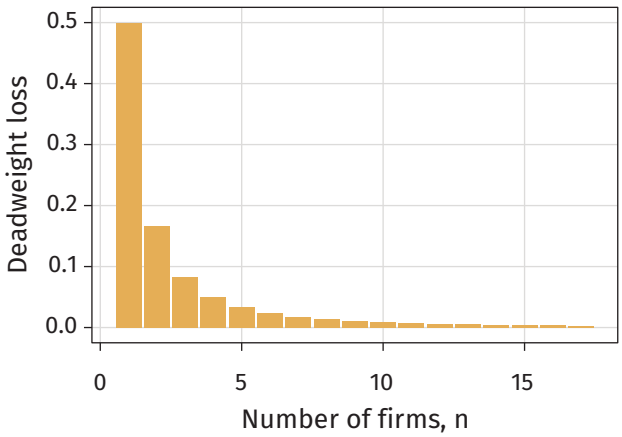
But it *could* result in a Pareto improvement *if* the consumers were to give up an amount of their increased consumer surplus sufficient to compensate the owners for their lost economic profits. (Again, do not ask why they would compensate the owners: this is another hypothetical thought experiment.) This is shown in Figure 9.13 (b). The consumers would have doubled their consumer surplus and the owners would have been *exactly* compensated for their lost economic profits, so they would be no worse off. The area of rectangle "compensation" is identical to the economic profit rectangle in panel (a) showing that the firm's owners are as well-off in panel (b) as in panel (a). The two purple triangles in panel (b) are the consumer surplus, double the amount in panel (a).

With this compensation for the lost economic profits of the owners, the hypothetical increase in production to x^c is therefore a Pareto improvement. This means that the monopoly output and price is *not* a Pareto-efficient allocation. The fact that we did not explain how the hypothetical increase in production to x^c could occur does not matter. We will see how that might be implemented shortly. All we need to do to show that some allocation is Pareto inefficient is that there exists some *other* allocation that is technically feasible, meaning that does not violate the basic facts given by the demand function and cost function, and that is Pareto superior to the allocation under consideration.

A Pareto-efficient allocation with price = marginal cost

Is the new hypothetical allocation itself Pareto efficient? To see that it is, think about why the monopoly allocation was inefficient: there was a deadweight loss arising from the fact that the price p^m exceeded marginal cost. If the firm were to produce x^c the price would equal the marginal cost, so there would be no deadweight loss. And so it would be impossible to find a technically feasible allocation that is a Pareto improvement over the allocation with x^c and p^c. Therefore x^c and p^c are a Pareto-efficient allocation.

Figure 9.14 Inverse relationship between number of firms and deadweight loss. As a market becomes more competitive and the number of firms increases, the size of the deadweight loss decreases. As the number of firms increases, the price moves closer to marginal cost, the markup decreases, and there is lower deadweight loss. Deadweight loss is indexed to being 0.5 under a monopoly when $n = 1$.



Are there any conditions in this model under which the Pareto-efficient allocation would actually occur, not hypothetically as a thought experiment but as a Nash equilibrium, given the relevant players' objectives and constraints?

There are, at least approximately. It will be a Nash equilibrium in which the price is approximately equal to marginal costs, or what is the same thing, where deadweight loss is approximately zero. We know from Equation 9.31 that the Nash equilibrium price varies with n, the number of firms according to:

$$p^{N}(n) = c + \frac{1}{n+1}(\bar{p} - c)$$

This means that the Nash equilibrium price p^{N} approximates marginal cost when the number of firms n is very large. This will occur when the barriers to entry b is close to zero because (as we know from Equation 9.36), the Nash equilibrium number of firms is given by the following:

$$n^{N} = \frac{\bar{p}(1-b) - c}{bc}$$

You can see from this equation that as b goes to zero (no barriers to entry) the numerator goes to $\hat{p} - c$ and the denominator goes to zero, so the equilibrium number of firms goes to infinity. This is the extreme case of what we call "unlimited competition."

Figure 9.14 illustrates how deadweight losses diminish and virtually disappear as the number of firms competing becomes very large. In Chapter 8 we showed that the supply and demand model with price-taking buyers and sellers–often referred to as *perfect competition*–also implements a Pareto-efficient Nash equilibrium. And it is for the same reason: that under those conditions the price would approximately equal the marginal costs of the highest cost producer.

The result is that perfect competition in the price-taking supply and demand model yields the same results as "unlimited competition"–that is large n–in the Cournot model of price-making competition.

Efficiency, competition, and inequality

Because owners of firms tend to be wealthier than are consumers on average, the extent of competition affects not only the *degree* of deadweight loss but also the extent of *inequality* in the economy. We have seen that by restricting its sales to x^m the owners of the monopoly gained a larger slice of a smaller pie, to the disadvantage of consumers.

A solution is to reduce the market power of the monopolist. Reducing barriers to entry and thereby increasing the number of firms competing would make the economy both more efficient and less unequal.

Notice from Figures 9.12 and 9.14 that the number of firms competing need not be very large to substantially *raise* the level of consumer surplus relative to profits, and *reduce* the extent of deadweight losses. We will see in section 9.11 that another way to implement an efficient outcome is to give the monopolist more power, not less, allowing it to make take-it-or-leave-it (TIOLI) offers to each consumer, in what is termed *perfect price discrimination*. You have already encountered similar cases in Chapters 5 and 10 where the actor with TIOLI power implemented a Pareto-efficient but highly unequal outcome.

CHECKPOINT 9.8 Deadweight loss and inefficiency

a In section 9.8, you saw that the conflict of interest between the consumers and firm owners in a market is not a zero-sum game. As n increases, consumers gain more consumer surplus than owners lose in reduced profits. Where do these extra benefits come from?

b In a competitive market, what is the price that will result in a Pareto-efficient allocation? Why is this price and subsequent level of output Pareto-efficient?

9.10 COORDINATION AMONG FIRMS: DUOPOLY AND CARTELS

When there is more than one firm, the conflict of interest is not just between the owners of the monopoly and the consumers. There are two conflicts:

- between the owners of the firms on the one hand and the consumers on the other; and
- between the owners of the firms who are competing to sell their goods.

In Figure 9.15 we represent both dimensions of conflict for the case where we have $n = 2$. At the Nash equilibrium, point **n** (identical to the case shown in Figure 9.6) the two firms each produce 12, so total output is 24, larger than the monopolist produced, namely 18.

Recall (from Figure 9.13) that consumers' surplus is maximized when a total output of 36 is produced. One possible allocation that would maximize consumer surplus is point **j** in the Figure 9.15 with each of the duopolists producing 18. The arrow from **n** to **j** shows that both firms producing *more* is better for consumers.

Along the line through points **j**, **k**, and **i**, the outputs of two firms sum to $\frac{(\bar{p}-c)}{\beta}$, which is equal to 36 in this case. Consumers would value points **k** and **l** as much as they do point **j**. Consumers do not care which of the firms produces more output as long as the total sum of output equals $\frac{(\bar{p}-c)}{\beta}$. At point **j** or any other point on this line the following are true:

- price is equal marginal cost, $p(X) = c$;
- consumers enjoy the maximum consumer surplus;
- deadweight loss is zero;
- economic profit is zero; and
- the allocation is Pareto efficient.

Somewhere along that line is where consumers would like the allocation to be.

But not the duopolists. The orange arrow shows that both firms producing *less* (than the Nash equilibrium) can raise profits to each. Just as the fishermen in Chapter 5 could do better if they restricted their overfishing the lake, duopolists could receive higher profits if they could *cooperate* so as to not "overharvest" their potential market by selling too many goods.

We have superimposed on Figure 9.15 the isoprofit curves of the duopolists. The yellow lens shows all of the combinations of the outputs of the two firms that are Pareto superior to the Nash equilibrium (**n**) *if we forget about the consumers and consider only the interests of the owners of the firms.* Restricting output of the two firms to just nine each (point **i**) would bring total output down to 18, the level the monopoly chose to maximize its profits. While consumers would like the outcome to move from **n** toward **j** the duopolists would like the outcome to move from **n** toward **i**.

EXAMPLE The most famous contemporary example of a cartel is OPEC (the Organization of the Petroleum Exporting Countries), which has regularly agreed on the number of barrels of oil each country would produce therefore affecting the world price for oil and gasoline or petroleum.

Figure 9.15 The duopolists' coordination problem and the maximum consumer surplus. If point **n**, the Nash equilibrium, is the status quo, owners of firms will seek to restrict output by implementing a point in the yellow Pareto-improving lens, ideally (for them) a Pareto-efficient point such as **i**. Consumers, in contrast, will be better off if output is increased, lowering prices, reducing deadweight loss, and increasing consumer surplus. For consumers the ideal level of production is given by the purple line; they are indifferent between points on the line, because at every point on the line price equals marginal cost and consumer surplus is maximized (eliminating both deadweight loss and economic profit for the firm's owners).

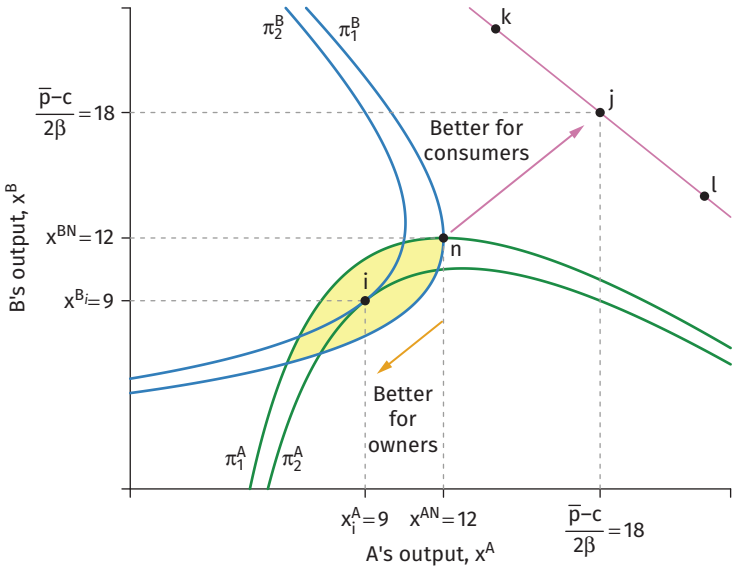
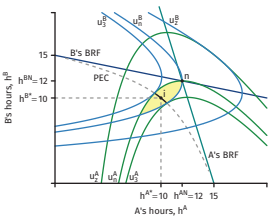


Figure 9.16 The overfishing coordination problem. Figure 5.12 from Chapter 5 shown here is about how many hours two fishermen will spend fishing on a lake, which is a common pool resource. They could both do better if they coordinated and fished less than they do at the Nash equilibrium (point **n**). As in Figure 9.15, the yellow lens shows all of the allocations that are Pareto superior to **n**.

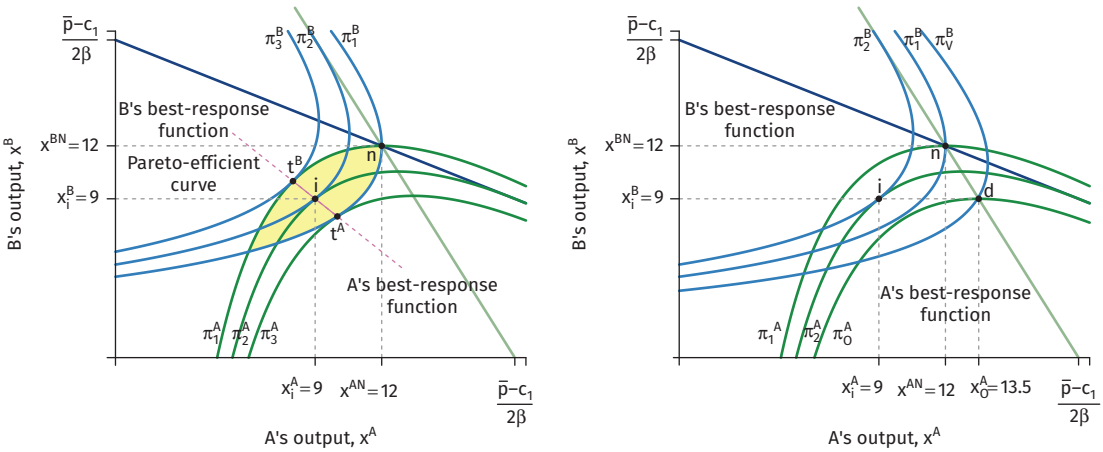


The duopolists' dilemma: A coordination failure among owners

From the owners' perspective, if they could agree to each produce an output of $x_i^B = 9 = x_i^A$, this would maximize their *joint* profits. Again, setting aside the interests of consumers, it would also be Pareto efficient as you can see from Figure 9.17 because the outcome lies on the Pareto-efficient curve.

But how could they enforce such an agreement (they could *not* get the government to enforce it because in most countries it would be illegal). You can see from Figure 9.17, if Firm A knows that Firm B will produce output x_i^B, Firm A's best response is to *increase* its output to the output x_O^A. Firm A behaves *opportunistically* by taking advantage of Firm B reducing its output (point **d**). When Firm A increases its profit to π_O^A which is substantially *higher* than the Nash equilibrium profit, A's opportunism results in Firm B's obtaining *lower* profit than, π_V^B where it is the *victim* of A's opportunism. The same reasoning applies to the owners of Firm B. Their best response is also to *violate* the agreement.

Figure 9.17 Duopolists' isoprofit curves and best-response functions. Firm A's output is on the horizontal axis and firm B's output is on the vertical axis. Looking at each firm's isoprofit curves, we can see that the equilibrium is Pareto inefficient with potential Pareto improvements. At the Nash equilibrium, each firm is on its equilibrium isoprofit line: π_1^A for Firm A and π_1^B for Firm B. If both firms could simultaneously *decrease* their output, then at least one firm would obtain higher profits, that is, at least one firm could move to a higher isoprofit curve (π_2 or π_3) while the other remained at their Nash equilibrium isoprofit line (π_1), or both firms could make higher profits, π_1 to π_2. Isoprofit curves π_O^A and π_V^B correspond to the extreme payoffs that the firms would receive when Firm A behaves opportunistically and Firm B is the victim of A's opportunism, as shown in Figure 9.18.



(a) Nash equilibrium and cartel choice for two firms

(b) Firm A behaves opportunistically

You can see from Figure 9.18 that the two duopolists face a Prisoners' Dilemma. A similar Prisoners' Dilemma type coordination problem occurs when there are many oligopolistic firms competing. On the basis of this reasoning the Cournot model reasons that firms will find it difficult to coordinate and form cartels.

! REMINDER An outcome is Pareto efficient when the firms' isoprofit curves are tangent to each other, meaning that the marginal rates of substitution are equal. When the isoprofits are tangent, no firm's owners can do better without making another firm's owners worse off and therefore there are no alternative outcomes that are Pareto superior to the tangencies of the isoprofit curve.

How owners address their coordination failure (consumers beware!)

A coordination mechanism that owners of firms use to solve their coordination problem is a **cartel** that negotiates output and price levels by dividing up markets among potentially competing firms. The firms would produce the total market output equivalent of the monopoly and each would obtain a share of the monopoly profit by selling at the monopoly price. They would

CARTEL A group of firms that collude to set output and/or prices in order to raise profits.

Figure 9.18 Cournot production: a arisoners' Dilemma game. A's payoffs are in the left-bottom corner of each cell; B's payoffs are in the top-right corner of each cell. Firm A's payoffs are ranked: $\pi^A_O > \pi^A_i > \pi^{AN} > \pi^A_V$. Similarly, for B: $\pi^B_O > \pi^B_i > \pi^{BN} > \pi^B_V$. The subscript "O" corresponds to "Opportunistic" and the subscript "V" corresponds to "victim of opportunism." For each player, their *opportunistic* profit is greater than their cooperative, reduced output payoff. As a result, for each firm, to choose the output on their best-response function (produce on the BRF) strictly dominates reducing output.

		Firm B	
		Reduce output	Produce on BRF
Firm A	Reduce output	$\pi^B_i=81$, $\pi^A_i=81$	$\pi^B_O=91$, $\pi^A_V=61$
	Produce on BRF	$\pi^B_V=61$, $\pi^A_O=91$	$\pi^{BN}=72$, $\pi^{AN}=72$

EXAMPLE For the duopolist, a cartel performs the function that the Impartial Spectator hypothetically, or government policies in reality might perform for the two fishermen seeking to overcome the overfishing problem: the cartel seeks to maximize the total profits of the two firms, just as the Spectator maximized the total utility of the two fishermen, except in this case the Impartial Spectator is not really impartial, as she is overlooking the reduced consumer surplus enjoyed by consumers.

still have a conflict, because having agreed on a total output (namely 18) each firm would prefer to produce more of it.

Cartel-type behavior was outlawed in the US by the Sherman Antitrust Act of 1890 and later antitrust acts, although US firms in some industries have strong incentives to limit competition if they can get away with it. In other societies cartels have sometimes become the normal, expected way of organizing production. But, because of opportunistic behavior, cartels that are not protected by law are often very fragile, as the Prisoners' Dilemma game above would suggest.

But, behaving as a cartel is not the only institution that firms adopt to try to obtain greater market power and increase their share of the rents on a market. Firms regularly *merge* with other firms (that is, to form one larger firm), in effect pooling their resources and becoming more like a monopoly. Firms also seek to *acquire* other firms either by buying them at a mutually agreed price or by setting up *hostile takeovers* through buying enough shares to name directors favorably disposed to the acquisition.

To sum up, a Nash equilibrium in the Cournot model with a limited number of firms is a Prisoners' Dilemma from the point of view of the firms typically resulting in their failure to maximize their joint profits by enforcing the price and quantity that a monopolist would choose.

Dynamic inefficiency

There is another possible source of inefficiency in monopoly pricing. Think about a firm that is considering developing a new product that will have few

close substitutes and hence will face cost and demand conditions similar to those shown in Figure 9.13.

From the firm's standpoint, the incentive to do this is the area labeled economic profit. But the profits the firm would gain do not measure the benefits that introducing the new product would yield. Consumers buying the product would also benefit in the form of consumer surplus. The profits accruing to the innovating firm fall short of the entire benefits of the introduction of the new product (consumer surplus plus economic profit).

The economic profit alone may not be sufficient to offset the development costs of the new product so the new product will not be produced. This is an example of *dynamic inefficiency* because it arises from the firm failing to make an innovation that would have increased the sum of consumer surplus and economic profit.

CHECKPOINT 9.9 Consumer surplus and the duopolists' coordination problem

a. In Figure 9.15, would consumers prefer to be at point **j**, **k**, or **l**? Explain your answer.

b. Why will duopolists find it difficult to coordinate and produce at point **i** in Figure 9.17?

9.11 PERFECT PRICE DISCRIMINATION: ELIMINATING DEADWEIGHT LOSS AT A COST TO CONSUMERS

Surprisingly, the deadweight losses associated with a monopoly's market power can be eliminated if we let the monopolist have a little *more* power, so that the monopolist can charge different prices to different individual buyers. This is called **price discrimination**. Price discrimination is an example of a situation in which the law of one price does not hold.

Some aspects of discrimination based on race, religion, sexual preference, and gender fall into the broad category of price discrimination. When an automobile dealer sells a car to a woman at a higher price than would have been charged to an otherwise identical man, we have an instance of price discrimination. But here, we consider cases where price discrimination is based on the buyer's willingness to pay, not her gender or some other characteristic.

PRICE DISCRIMINATION Selling the same product at different prices, for example, charging more to buyers with a greater willingness to pay.

✓ **FACT CHECK** A study of the markets for rice, kidney beans, sugar, tomatoes, and tortillas in Mexican villages found that those buying larger amounts paid substantially lower prices. For the tortilla market, for example, a 10 percent increase in the amount of tortillas purchased was associated with a 4.5 percent decrease in the price. This may have occurred because those buying small amounts were generally poor, and had fewer alternative places they could purchase the commodities. Richer buyers, purchasing larger quantities had better alternatives for shopping around.[5]

! REMINDER The participation constraint of a buyer represents the requirement that in order to voluntarily participate in a social interaction, such as choosing to purchase a good from a firm, an individual must receive a utility at least as great as the next-best alternative, which constitutes the buyer's fallback in the transaction. If a seller makes an offer that violates the participation constraint, the buyer will refuse it.

Perfect price discrimination and Pareto efficiency

Figure 9.19 shows how price discrimination will reduce the inefficiencies associated with monopoly or other forms of limited competition. To see how price discrimination might address the inefficiencies associated with monopoly or other forms of limited competition, take another look at Figure 9.13. Deadweight losses exist because the monopolist restricts how much he sells in order to maximize profits (return to Figure 9.13 to confirm this). Therefore, the price the monopolist charges exceeds marginal cost. Consequently, many consumers are willing to pay more for the good than the marginal cost that the firm would incur to produce it, which we can identify in Figure 9.19 because the demand curve represents the maximum price a buyer is willing to pay for each unit of the product sold.

But suppose the firm could make a private bargain with each of these consumers who are not buying the good at the price p^m. As first mover the monopolist would make the customer a take-it-or-leave-it offer charging a price equal to the consumer's maximum willingness to pay.

If this were possible, then the firm would produce more. The firm would produce and sell the good even to the consumer whose willingness to pay just barely exceeded the marginal cost. In this case, the price charged to the marginal consumer (the one with the least willingness to pay that exceeds the marginal cost) would barely exceed the marginal cost and there would be virtually no deadweight loss.

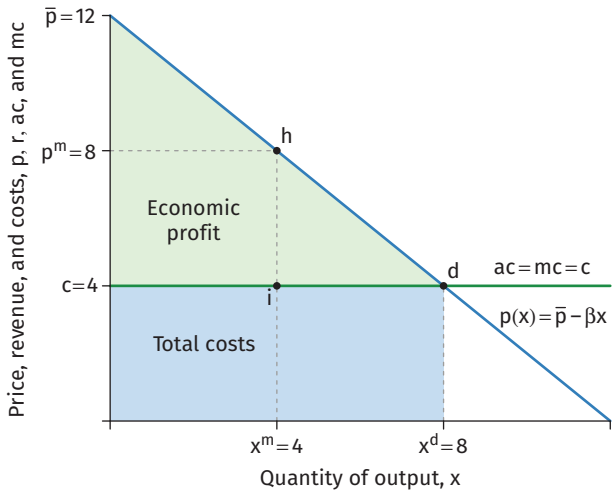
Of course if the firm could make a private deal with each of the consumers whose willingness to pay fell short of the price p^m in Figure 9.19, then it would also want to make a similar deal with all the rest of the consumers. But in these cases the firm would charge *more* than p^m because, as you can see from the demand curve, the remaining consumers (those to the left of x^m) have a willingness to pay that is greater than that price. If the firm could set a price for each unit sold so that the buyer receives no consumer surplus at all, this is called **perfect price discrimination**.

The perfectly discriminating monopolist is illustrated in Figure 9.19. The entire area under the demand curve represents the maximum sales revenue that the firm could get if it could somehow charge each consumer a personally chosen price for each unit (a different price for each consumer for every unit which that consumer buys) designed so that the consumer would be indifferent between buying or not buying. The result is that there are no benefits conferred on consumers in the form of consumer surplus. And there is no deadweight loss either.

Is the outcome when the monopoly practices perfect price discrimination a Nash equilibrium? You can check that given that the monopolist can

PERFECT PRICE DISCRIMINATION Occurs when a firm can make a separate take-it-or-leave-it offer to each individual buyer for each unit to be sold at a price equal to the consumer's maximum willingness to pay.

Figure 9.19 Cost curve, demand curve, and the perfectly discriminating firm. The perfectly discriminating firm charges each customer a price equal to the customer's reservation price. For example, the customer who buys the x'th unit of the good is charged p'. Because no buyer pays less than their maximum willingness to pay for the good, there is no consumer surplus; the perfectly discriminating firm appropriates the total economic rents from the transaction.



make a take-it-or-leave-it price offer and sets a price just slightly below the buyer's maximum willingness to pay, the buyer's best response to this price offer is to buy. Given the range of different people's willingness to pay, the monopolist cannot do better than a perfect price discrimination strategy.

Perfect price discrimination therefore results in Pareto-efficient level of output even by a monopoly because the price charged to the marginal consumer is equal to the marginal cost of producing that good. There is also another way to see this.

Remember that an outcome must be Pareto efficient if it is the result of one party maximizing his or her utility subject to a participation constraint (requiring that the other person's utility must not be less than some given amount). "Participate" here means "buy the good," and the consumer's fallback option is not to buy the good, and have more money for other purposes. Since the firm cannot secure buyers at prices above the demand curve, the demand curve is the participation constraint for the perfectly discriminating monopolist.

The result seems paradoxical: the inefficiencies resulting from the market power of a monopoly can be eliminated by giving the monopoly *even more power*, the ability to charge different prices to each consumer. The reason why this works is that when the firm charges a single price to all buyers, the owner faces a trade-off: he could sell more, but in order to

› EXAMPLE The perfectly discriminating firm is the ultimate price-making firm. So far the price-making firm can vary its output so as to affect the identical price at which it sells all of its goods. The perfectly discriminating firm can sell an additional unit of output at a price above marginal cost *without* reducing the higher prices paid by other buyers.

! REMINDER You have seen something like this in Chapters 4 and 5. When Ayanda had take-it-or-leave-it power, she was acting exactly as a perfect price discriminator appropriating for herself the rent from the interaction, constrained only by Biko's participation constraint. The outcome was Pareto efficient. The same was true when one of the fishermen had take-it-or-leave-it power as owner of the lake in Chapter 5.

do so he will have to lower the price; or he could charge more, but then fewer consumers will buy. Monopoly price-discriminating power liberates the owner from this trade-off posed by the demand curve.

Given the trade-off posed by the demand curve and the requirement to sell all goods at a single price, the monopolist restricts production and sales so as to allow a higher price. In other words, the reason why the monopolist finds a way to get a larger slice from a smaller pie is that (in the absence of perfect price discrimination) he cannot get the entire pie.

To sum up, perfect price discrimination by a monopolist:

- leads to a Pareto-efficient allocation since the monopolist will sell to every buyer whose willingness to pay is above the marginal cost of the goods produced;
- the result constitutes a Nash equilibrium; and
- distributes the entire rent to the monopolist (buyers receive no consumer surplus).

! REMINDER The word *distribution* refers to *how much* of the benefits or costs of any interaction each person gets. Perfect price discrimination results in greater inequality in the distribution of economic surplus from exchange because the owners of the firm gets the whole economic surplus, and buyers get none.

CHECKPOINT 9.10 The PC vs. the ICC

a Define the terms participation constraint and incentive compatibility constraint.

b Explain why, for the firm that can perfectly price discriminate, the demand curve is also the buyers' participation constraint, which is the relevant constraint on the firm's profit maximization?

c Why does perfect price discrimination eliminate all consumer surplus?

9.12 APPLICATION: PRICE DISCRIMINATION IN ACTION

If price discrimination can eliminate the gains from the exchange enjoyed by buyers by distributing their entire consumer surplus to owners' economic profits, then is it called "perfect?"

In science and other scholarly disciplines the word "perfect" does not mean "flawless." It often refers to some abstract idea, such as a "perfect gas" in physics which does not exist but which is a helpful simplification for understanding the important aspects of some processes. Perfect price discrimination is an abstract idea designed to clarify why monopolies are *in*efficient, not a value judgment about how good it is.

✓ FACT CHECK Price discrimination is often viewed as "price gouging" and it is generally condemned on ethical grounds. For example, 80 percent of Americans in a study by psychologist Daniel Kahneman disapproved of stores charging more for snow shovels after a snowstorm (when presumably people's willingness to pay is higher than usual).[6]

Although price discrimination benefits owners, firms do not typically act as perfect price discriminators: like perfect competition it is not a business practice that we expect to see often. There are three reasons why few, if any, businesses practice price discrimination in its pure form:

- *Information*: it requires information on potential buyers' maximum willingness to pay that is costly or even impossible to obtain.

- *No resale*: the monopoly firm would have to find a way to ensure that those who purchased the good at a low price could *not* resell it to those with a higher willingness to pay, thereby undermining the monopoly's high price sales.
- *Consumer objections*: buyers often react negatively to price discrimination, thinking that the firms should *not* be able to profit by charging different prices to different people, or to people living in different circumstances.

But even with these impediments, there is a price discrimination strategy that is feasible in many cases. When a firm has information that allows it to distinguish groups of potential buyers who have different willingness to pay, but not enough information to find out every individual buyer's reservation price, it can engage in group price discrimination.

The discriminating firm will charge a different price to members of each of the groups (or *submarkets*) depending on the firm's estimate of the willingness to pay of the members of each group.

Suppose the firm considers raising the price at which it sells to a particular group. If buyers in the group can meet their needs by purchasing a close substitute for the firm's good, then their willingness to pay for this particular product will be limited, and so a small increase in price will drive them away. The same will be true if the members of the group have little income. In this case their marginal utility of the limited amount of money left over for other purchases will be substantial, so that their willingness to pay will be limited.

If, on the other hand, buyers in the group have no good alternatives to buying from the firm, or if they are wealthy their willingness to pay will be greater and the firm can raise the price a lot and lose only a few customers. In this case, we would expect the price offered to the group to be high.

Competitive strategies like advertising the distinctive features of the product and consumer loyalty programs such as airline miles or discount cards are designed to deter customers from switching to their next-best alternative. Advertising by the firm is often designed to attract a group of buyers who are particularly attached to the advertised features of its product, or particularly taken by the glamour of using it, rather than some close substitute. Advertising also targets higher-income people, whose willingness to pay will be greater because their marginal utility of money left over is less.

Mac computers tend to be a lot more expensive than PCs, so the travel site Orbitz steered Mac users to more expensive hotel sites. Here the price-setter exploits the fact that rich people have higher willingness to pay. So lower-income buyers *might* see a lower price. But not necessarily. Discounts for the elderly (on average lower income) follow this logic.

For the identical stapler the Staples.com website charged a price of $15.79 to one buyer and $14.29 to another based on where they lived. They only lived a few miles apart. Office Depot told the *Wall Street Journal* (WSJ) that they use "customers' browsing history and geolocation" to personalize price-making. In the WSJ's study the areas that tended to get *lower* prices were *higher*-income areas where there was more competition from rival sellers.[7] In other words, the rich paid less and the poor paid more.

Airline companies charge fliers who stay in a city over Saturday night a lower fare for the same seat than fliers who return without staying over Saturday night. These consumers have a lower willingness to pay, and greater price elasticity of demand than other fliers. They have a higher elasticity of demand at any price because nonbusiness travelers typically have more flexibility in their travel schedules and routes, and more alternatives than business flyers. A business flyer who wants to stay over Saturday night will benefit from the lower fare, showing the airlines' ability to discriminate is limited by the information their customers reveal.

! REMINDER A Nash equilibrium is an allocation such that none of the actors could do better by altering their strategies, given the strategies adopted by others. This explains why, if actors are implementing a Nash equilibrium, they have no incentive to *leave the Nash equilibrium*. The Nash equilibrium concept says nothing about how they might *get there* in the first place.

CHECKPOINT 9.11 Perfect price discrimination and dynamic inefficiency
Explain why the dynamic inefficiency described in section 9.10—a socially valuable innovation not being introduced—would *not* occur if the monopolist considering the innovation were capable of perfect price discrimination.

9.13 RENT-SEEKING, PRICE-MAKING, AND MARKET EQUILIBRATION

The model of perfect competition introduced in section 8.12 describes a Nash equilibrium, but it does not explain *why* we might expect a market to be at or near the intersection of the supply and demand curves. Why would we give special attention to the market-clearing equilibrium point? The answer must be that other points are not Nash equilibria so that they will be disrupted by buyers or sellers (or both) changing the prices or quantities at which they are willing to transact.

Nonequilibrium prices and the short side of the market

To see this we study the cases illustrated in Figure 9.20. Suppose that for some reason the price is p^H, a price higher than the price given by the intersection of the supply and demand curves. To understand what this means remember that "supplied" does not mean "sold." It means produced and brought to market. Whether it is sold will depend on demand.

Figure 9.20 shows that at the price p^H the amount sold will be the amount demanded X^{DH}, because buyers *will* not be willing to buy more than this quantity. The quantity supplied is X^{SH} which exceeds the quantity demanded X^{DH}. Sellers therefore want to sell more than buyers want to buy

at p^H. This case is called **excess supply**, referring to the excess of goods produced and brought to market *but not sold*, remaining in warehouses or on the shelves rather than going into someone's shopping bag.

By contrast, when the price is $p^L < p^c$ buyers want more goods than the sellers want to sell, so there is **excess demand**, which might show up with people standing in lines to get limited numbers of goods or heading home with empty shopping bags. When either excess demand or excess supply exist, we say the market does not *clear*.

To see what happens in this case we need to introduce some new terms about the economics of *non-clearing markets*. Remember, to construct the supply and demand curves we asked a hypothetical question based on the idea that the buyer or seller could transact any quantity they wished at some given price. This will be true for both buyers and sellers at the price given by the intersection of the supply and demand curves. But when markets do *not* clear this "transact-any-amount-you-wish-at-the-going-price" assumption must be untrue of either the buyers or the sellers. When someone cannot buy or sell the quantity that they would like at the going price we say that they are **quantity constrained**.

We refer to two "sides" of the market with respect to demand and supply: the "demand side" are the buyers who demand the good; the "supply side" are the sellers who supply the good.

When the market does not clear, there is a **long side** of the market and a **short side** of the market.

- *Short side*: The short side of the market refers to the actors–buyers or sellers–as short siders, who are able to make all of the transactions they wish. They are on the side of the market where the number of *desired transactions* is least.

EXCESS SUPPLY Excess supply exists when, at the prevailing price, the amount supplied exceeds the amount demanded.

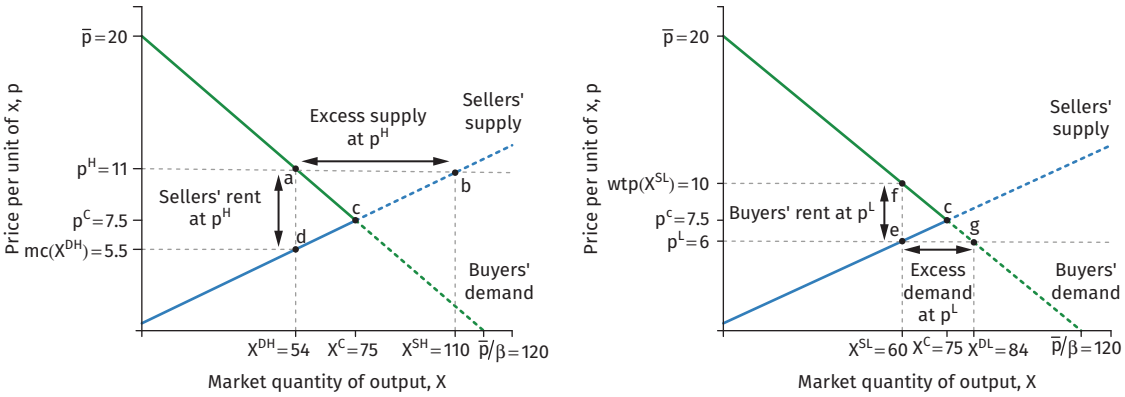
EXCESS DEMAND Excess demand exists when, at the prevailing price, the amount demanded exceeds the amount supplied.

QUANTITY-CONSTRAINED. An actor is quantity constrained if they are unable to transact the quantity they would like at the going price.

LONG SIDE OF A MARKET The long side of the market is the side—either supply or demand—on which the number of desired transactions is greater, given the price.

SHORT SIDE OF A MARKET The short side of the market is the side—either supply or demand—on which the number of desired transactions is least, given the price.

Figure 9.20 Excess supply and excess demand: rents in a non-clearing market. When $p^H > p^c$ (panel (a)) sellers would like to supply X^{SH}, but are able to transact with buyers only at the quantity *they* demand at p^H, which is X^{DH}. As a result, those sellers who sell to buyers gain an economic rent equal to the distance **ad**, which is the difference between their marginal costs at X^{DH}, $mc(X^{DH})$, and the price they obtain p^H. When $p^L < p^c$ (panel (b)) buyers would like to buy X^{DL}, but are able to transact with sellers only at the quantity sellers are willing to sell at p^L, which is X^{SL}. As a result, those buyers who buy from sellers at p^L gain an economic rent equal to the distance **ef**, which is the difference between their willingness to pay at X^{SL}, $wtp(X^{DL})$ and the price they pay p^L.



(a) Excess supply (b) Excess demand

EXAMPLE In colloquial English, to "get the short end of the stick" means to get a bad deal, to come up short in some bargain. But being on the short side of a market can be an advantage, as we will see when we study the market in labor in Chapter 11. In the equilibrium of the labor market there is an excess supply of workers so employers are on the short side of the market and workers both employed and unemployed are on the long side. This is also true in the credit market (Chapter 12), where there is excess demand for loans, so banks and other lenders are on the short side and borrowers are on the long side, some of whom would like to borrow at the going rate of interest but cannot.

- *Long side*: Some of the long-siders will be able to make the transactions they wish, but others will not. Some long-siders are quantity constrained. They are on the side of the market where the number of *desired transactions* is greater.

Either buyers or sellers can be on the short side of the market, depending on the price being above or below the market-clearing price. If the price is $p^L < p^c$ (Figure 9.20 (b)) so that there is excess demand, then sellers are on the short side of the market. When $p^H > p^c$ (Figure 9.20 (a)) by contrast, sellers want to sell more goods than buyers want to buy, so buyers, who can buy all they wish, are on the short side of the market. Table 9.2 summarizes these results about non-clearing markets.

Notice that in Figure 9.20 the supply curve above p^c is a dashed line, and the same thing is true of the demand curve below p^c. We do this to emphasize that these portions of the supply and demand curves are irrelevant to our analysis: we will never observe an outcome on the dashed segments. This is because the exchange is voluntary: there is no way that those who would like to buy or sell more at the going price can *force* others to exchange with them. As a result, it is the short side of the market that determines the quantity transacted: that is, buyers (the demand side) when $p^H > p^c$ and sellers (the supply side) when $p^L < p^c$.

Table 9.2 Excess demand and supply, rents, and quantity constraints. The two rows of the table refer to the case in which the price is higher (top row) or lower (bottom row) than the price that would clear the market.

Figure	Price	Excess supply or demand	Short side	Rents	Unable to transact
9.20 a	$p^H > p^c$	Excess supply	Demand (buyers)	Some sellers get economic profits	Other sellers get nothing
9.20 b	$p^L < p^c$	Excess demand	Supply (sellers)	Some buyers get consumer surplus	Other buyers get nothing

Rents in non-clearing markets

In Figure 9.20 the fact that the market does not clear is indicated by either *excess supply* or *excess demand*, the extent of which is measured in terms of quantity not bought or not sold. The quantity not sold is a *horizontal* distance in the figure.

The extent to which the market does not clear can also be measured *vertically* in the same figure, that is, by the difference between the price at which the good is transacted and the buyers' willingness to pay or the sellers' willingness to sell.

This means that in a non-clearing market there must also be *economic rents*, in the form of either consumer surplus or economic profits. These measure the extent of excess demand or supply in the vertical dimension, that is, in terms of the price.

At the price p^H the marginal cost of the last unit sold ($mc(X^{DH})$) is less than the price, so we have both:

- *excess supply* because $X^{SH} > X^{DH}$ (the horizontal dimension); and
- some (but not all) sellers who succeed in selling their output gain *economic profits* on their last unit sold equal to the vertical distance between p^H and mc at $X = X^{DH}$ in Figure 9.20 (a).

In the other case, at a price $p^L < p^c$, we have both:

- *excess demand* because $X^{DL} > X^{SL}$; and
- some (but not all) buyers who are able to purchase what they want gain *consumer surplus* (an economic rent) on the last unit purchased that is equal to the vertical distance between points **e** and **f** in Figure 9.20 (b), because they paid p^L for a good for which their willingness to pay (the height of the demand curve at X^{DL} or $wtp(X^{DL})$) is greater than the price.

Notice the pattern in the two examples above:

- *short side*: Those on the short side of the market are able to transact all that they wish at the given price; and
- *long side*: Those on the long side of the market fall into two different groups with decidedly contrasting outcomes: Those who succeed in making a transaction gain an economic rent while the rest of the so-called 'long-siders' get nothing, they are excluded from the market.

Rent-seeking and market equilibration

How does a competitive market adjust when the going price is such that there is excess demand or excess supply? For any kind of adjustment to occur the buyers and sellers *cannot* act as they might in the equilibrium of the perfect competition model, that is, as price-takers. As a thought experiment, if all of the buyers or sellers in the market described in Figure 9.20 were price-takers (and therefore do not change the price), then a price like p^H or p^L would persist forever.

But in a situation with excess demand or supply, being a price-taker is not the best you can do. In a competitive market for goods like that described in the figure, a non-clearing market is not a Nash equilibrium. Why?

- **Rent-seeking**: If economic rents exist when the market does not clear, then there will be opportunities to gain either consumer surplus or economic rent by changing the amount demanded or supplied, or offering a different price. These activities are called rent-seeking.
- **Equilibration**: Under most conditions, how the rents are obtained will result in prices and quantities *changing* so that the market eventually *equilibrates* and *clears*.

To see why, look at Figure 9.20 and imagine the price is p^H as in the left panel and you are selling the *last* unit you produce at *more* than its marginal cost.

What would you do? You would want to sell more goods. But you are *constrained* by the demand curve: you cannot sell more if you remain a price-taker and do not lower your price. So you become a price-maker

RENT-SEEKING Any activity undertaken to gain a rent for the actor is called rent seeking, including changing a price or quantity in a non-clearing market, creating barriers to entry to reduce competition, introducing an innovation to reduce costs, price discrimination, and lobbying or other political activities aimed at granting the actor some kind of legal or other advantage.

EQUILIBRATION Equilibration is the process of getting to an equilibrium from a nonequilibrium situation.

and acting as a rent-seeker, you would lower your price a little bit. The following will then happen:

- *Greater quantity sold*: You sell more as buyers switch to you because of the lower price you are offering (your goods are substitutes for the goods other firms produce).
- *Lower profit per unit*: You would make a slightly smaller rent on each unit (a lower profit per unit) because the price is lower (and therefore $p - c$ is lower).
- *Best responses by others*: Other firms losing business to you would do the same and lower their prices to best respond to your change in strategy.

The process would typically go on until the price fell to p^c. Something analogous happens if the price is lower than p^c. At price p^L it is the buyers who receive rents (their willingness to pay exceeds the price they pay). Put yourself in one of the buyers' shoes: you would like to buy more but cannot without paying more. So you would act as a rent-seeker and price-maker, and would offer to pay a higher price than p^L. Sellers would flock to you, and other prospective buyers would find that they too would have to increase their prices. Again the process will go on until the economic rents disappear, that is, until the market clears at p^c.

This is why the dashed portion of the supply and demand curves in the figure will not have any bearing on quantity transacted at the given price. All of the action must be on or between the solid lines, which encompass the *short side of the market*.

CHECKPOINT 9.12 **Short side vs. long side** Analyze the market outcome in Figure 9.20 (b) when $p = p_L$.

a. Explain what the initial economic rent would be when $p = p_L$.

b. Why would an outcome such as point **g** never be observed?

c. Would $p = p^L$ persist? Why or why not?

HISTORY **The long run** To stress the importance of knowing how an economy works when it is not at or near an equilibrium, in his 1923 *Tract on Monetary Reform*, John Maynard Keynes famously wrote that the "*long run* is a misleading guide to current affairs. *In the long run* we are all dead." He did not suggest that modeling processes that may take a long time before reaching equilibrium is pointless. In fact he uses the concept of the long run repeatedly in that book.[8]

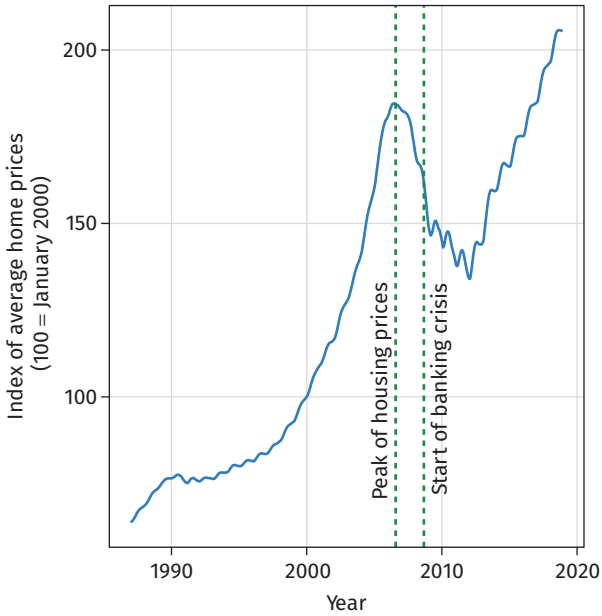
9.14 APPLICATION: WHEN RENT-SEEKING DOES NOT EQUILIBRATE A MARKET—A HOUSING BUBBLE

There are two reasons why we study how when the market is not in equilibrium the equilibration process works through the rent-seeking activities of buyers and sellers.

The first is that this process can take a very long time. To see why, think about a labor market in which wages for a given kind of labor are substantially higher in one region or city than in another. The market is not in equilibrium. The rent-seeking that might equilibrate this market requires

Figure 9.21 Housing prices in the US (1987–2018). The housing mortgage crisis started after the high price for houses was reached in August 2006. The subsequent banking crisis began in September of 2008 with the closure of Lehman Brothers, an investment banking firm. The data show the Case–Shiller index based on sales of single family homes.

Source: S & P Dow Jones Indices LLC (2020).



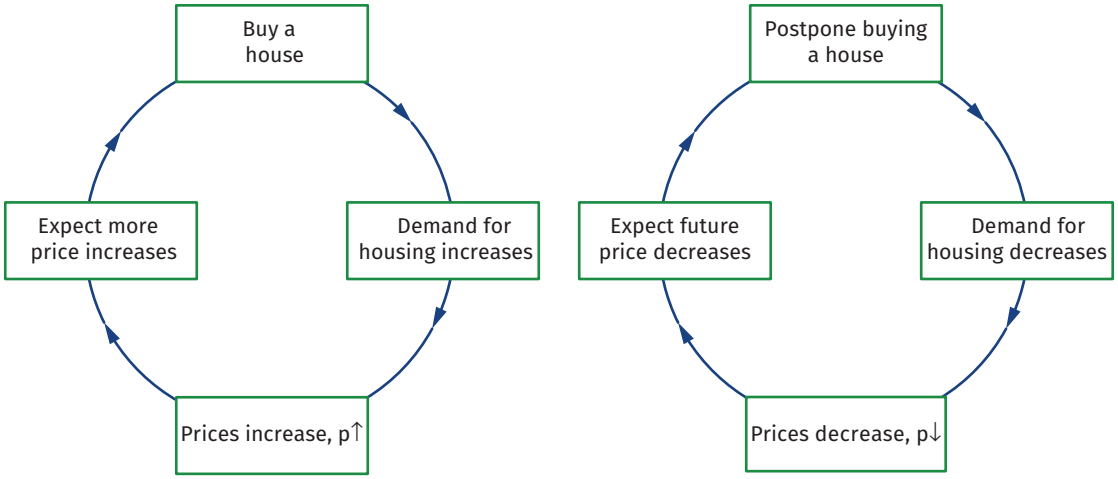
workers and their families in the low-wage regions to move to the higher-wage location. And moving requires them to leave their friends, neighbors, and relatives. Equilibration, therefore, often takes place over decades if not a generation.

As a result, the equilibrium of a market may be a poor guide to what we will find when we observe the economy empirically. But the idea of equilibrium remains informative. Even when the economy is out of equilibrium, the concept of equilibrium often provides predictions about what changes we should observe taking place: for example, workers moving to areas with higher wages or less unemployment.

The second reason to study the equilibration process is that sometimes it does not work. Rent-seeking may temporarily drive the market *away* from equilibrium over a long enough period to create a substantial amount of insecurity.

This is not just a theoretical possibility. It happens. Figure 9.21 shows the pattern of a bubble, then a crash for US houses: housing prices increased until August 2006, then started to drop a little bit, then crashed. They began to recover again in January of 2012 and have continued to increase since then.

Figure 9.22 The positive feedbacks leading to bubbles and crashes. A positive feedback occurs when some change causes further changes leading to the amplification of the initial change. In panel (a) an initial price increase (for whatever reason) leads to changes (following the arrows counterclockwise around the figure) that eventually feed back to cause a further price increase. A positive feedback process working in the opposite direction occurs in panel (b). Positive feedbacks are the basis of what people call virtuous cycles or vicious cycles.



(a) A housing bubble with rising prices

(b) A housing price crash

A housing price bubble

To see how equilibration fails, let's consider a particular market: housing. Let us assume that the housing market starts in equilibrium. For some reason the price of housing rises leading to more house construction and less house buying. There is now an excess supply of housing. If the rent-seeking process described above were to work, then owners of the unsold houses could do better by offering to sell at a lower price. Some of them would realize this and lower prices. Eventually the market would return to the earlier equilibrium.

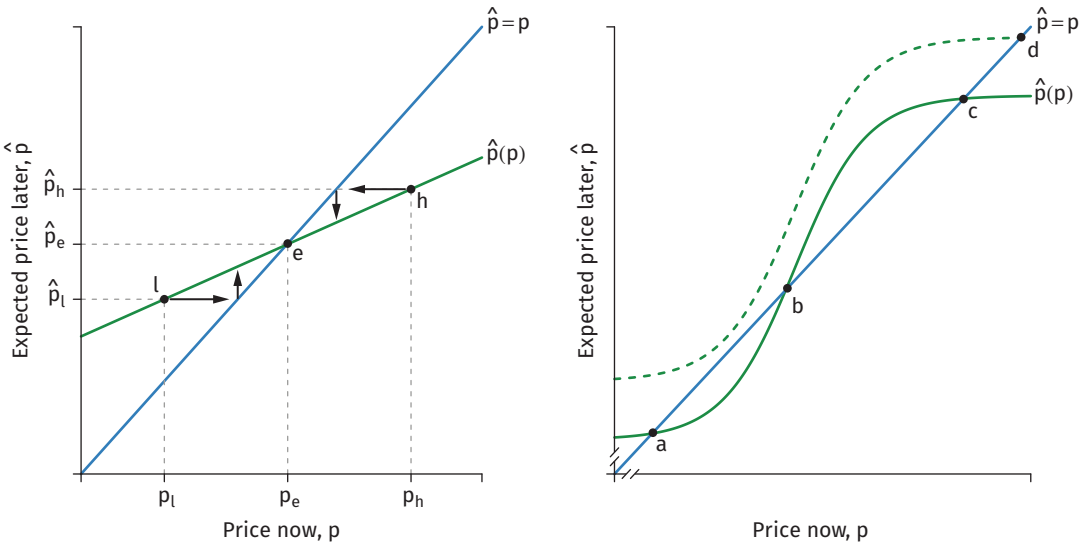
But rent-seeking participants in the housing market might think differently. Imagine you have just gotten your first well-paying job. You have two choices: buy a home now or postpone buying and continue renting. The increase in prices of houses leads you to postpone buying. This is how the equilibration process is supposed to work.

But if you think ahead, you might think as follows: "Housing prices just rose. I guess they may be even higher in the future; so both to save money, and to acquire an asset whose value is going to rise: I'll buy now." This logic is illustrated in panel (a) of Figure 9.22.

If this kind of economic reasoning were common—and there is no reason why it should not be—the increase in the price of housing would result in

! REMINDER You encountered tipping points already in Chapter 8 in the model of segregation of "Blues" and "Greens" in a housing market and Chapter 2 in the Palanpur game about plantings early or late.

Figure 9.23 The dynamics of price adjustment and tipping points. The horizontal axis in both cases is the price now (it could be of housing, stocks, or any good). The vertical axis is the price we expect in the next period. In both panels the green curve shows that the higher the price is now, then the higher we expect the price to be tomorrow. In panel (a), an initial low price leads people to expect a higher price tomorrow, which, when it occurs, leads them to expect an even higher price the period after that. Points **a** and **c** in panel (b) are similar to point **e** in panel (a). If prices are a little above or below the intersection, prices will adjust back to the intersection. These are stable equilibria. Point **b** is different. A price a bit higher than **b** leads people to expect an even higher price in the next period, and this leads prices to rise (moving away from point **b**), eventually stabilizing at point **c**. The breaks in the axes in panel (b) are because house prices do not start close to zero.



✓ FACT CHECK A housing bubble in the US, Spain, and other countries contributed to the 2007-2009 global financial crisis.[9]

an increase in the demand for houses. The shift in demand—if sufficiently great—would transform the market situation from the initial one of excess supply to excess demand. This would result in additional increases in housing prices, thereby fulfilling your expectation that your home value would go up. The result would be a sustained rise in house prices. This is called a *housing bubble*.

The term housing bubble is meant to suggest that prices do not increase forever. The bubble will *burst* if something happens that leads people to think that housing prices will fall. Then the reasoning runs in reverse.A fall in the price of housing indicates that buying now is a waste of money, prices will be lower later, and in any case owning housing is no longer a good investment. Demand for houses will collapse. Housing prices then crash.

Modeling a price bubble and a crash: Stable and unstable equilibria

The key to the bubble and the crash is that a change in prices can lead people to expect a future price change. To see how this works, in Figure 9.23 (a) we show the case where starting from an out-of-equilibrium price, the buyers and sellers converge on an equilibrium price. The green curve labeled $\hat{p}(p)$ tells us the price we will expect later if the price now is as given on the horizontal axis. This is the expected price function. So if the price today is the low price p_l, then the expected price later will be somewhat higher indicated by $\hat{p}_l$.

If our expectations were correct (which we assume), then that price becomes the new actual price now, indicated by the horizontal arrow to the $\hat{p} = p$. But if the price now is at that level, then the expected price (repeating the above process) will be still higher. This process will continue until the price is p_e in which case the price now *is also* the expected price later. So unless something changes from outside the model, there are no further changes in the price.

You can see that had the price been higher than p_e initially, then the process could have brought the price *down* to p_e. This process describes how market equilibration works, when it works. The price p_e is a stable equilibrium price.

In Figure 9.23 (b) we show a different price expectation function. There are two stable equilibria in panel (b), like point **e** in panel (a), at points **a** and **c**. But there is also point **b**. If the price is a little *above* the price at point **b**, then you can see from the green expected price function that people would *expect* the price to be even higher next period. And if the price were higher, then prices would continue to rise up to point **c**.

If the initial price is a bit *below* the price at point **b** then a similar process will lead prices to *fall* to point **a**. Point **b** is called a tipping point. At prices *above* the tipping point the economy 'tips' *up* to point **c**. At prices *below* the tipping point, it tips *down* to point **a**.

To see how a bubble can happen imagine that the economy is at point **a**. Now think about some change that would result in most people expecting prices to rise next period: for example the government just announced a program to subsidize housing costs for families with children, or banks just announced lower interest rates on mortgages, making it easier for people to borrow to buy houses.

This would shift up the price expectation curve to the dashed line in the figure. For every price today, the expected price later would be higher. The result would be the stable price equilibrium at point **a** would no longer exist. And the same is true of the 'tipping point' **b**. What would then happen?

The logic of Figure 9.23 (a) tells us that prices would continue rising until the economy reached point **c**. This is a price bubble.

! REMINDER Getting to an equilibrium In Chapter 5 we asked how the two fishermen might get to a Nash equilibrium if they were at some point not at the equilibrium. We showed how this could happen and also asked you to think of a case in which it would not happen. The "not happen" case is what we are studying here with the price bubble.

! REMINDER A value in some process—for example, a particular price of housing—such that the process is qualitatively different for values higher or lower than this value—prices rising above it or falling below it, for example—is called a tipping point.

› EXAMPLE Tipping point The temperature 0 degrees Celsius (32 degrees Fahrenheit) is a tipping point for water, above that point water is a fluid, below it it becomes a solid (that is, ice). Here, the value of 0 degrees Celsius is the "critical value" and the "qualitative difference" is how water transitions from being a liquid (its initial quality) to a solid (a different quality), therefore demonstrating a qualitative change.

If we are now at point **c** with high but constant prices, a downward shift in the price expectation function (eliminating the intersection at point **c**) would bring about a downward spiral of prices back to a low-price equilibrium like point **a**. This is the bursting of the bubble, a housing price crash.

Price bubbles and irrationality

Bubbles, crashes, and the instability they cause are often attributed to people's "irrationality." But notice there was nothing about your calculation above that was irrational in either the narrow sense of the term in economics or in the more common-sense meaning. You were using prices as information and "doing the best you can" in exactly the way that characterizes the constrained optimization process that is sometimes taken to be the definition of *economic rationality*.

What made the bubble possible was not stupidity or a gambling temperament, but instead two things: the facts that

- people (reasonably) took an increase in house prices as a signal of things to come, and that
- housing is a **durable asset**, whose future value is a critical concern when buying or selling.

Bubbles do not occur in markets for goods that are nondurable or do not last. You do not see people flocking to buy up as much fish as possible when the price rises, hoping to sell it for a profit later! (With improvements in refrigeration and storage, however, even this cannot be ruled out.)

CHECKPOINT 9.13 Tipping points

a In Figure 9.23 which price is a tipping point in the right panel?

b Is there a tipping point in the left panel?

c Is there some tipping point for water other than 0 degrees Celsius?

9.15 HOW COMPETITION WORKS: THE FORCES OF SUPPLY AND DEMAND

We have also seen how competition among rent-seeking buyers and sellers in a non-clearing market can equilibrate supply and demand. The fact that most goods are not durable (or are costly to store) means that bubbles are the exception, and movement toward a market equilibrium is more common. This shows that the supply and demand framework provides a

DURABLE ASSET A durable asset is one that remains valuable over a long period of time.

useful conceptual tool, even if the perfectly competitive model of supply and demand does not explain how price-taking buyers and sellers would ever *get to* a Nash equilibrium. The rent-seeking approach teaches an important lesson. Locating the intersection of a supply and a demand curve is not sufficient to understand how competition works and how buyers and sellers best respond to changes in supply, demand, and the degree of competition.

Changes in supply and demand affect the prices at which goods are exchanged by altering the fallback options of buyers and sellers. A seller's fallback (next-best alternative) in an interaction with a potential buyer is to sell to an alternative buyer or not to sell at all. A buyer's fallback is to buy from an alternative seller or not to buy at all.

! REMINDER A fallback is each buyer's or seller's next-best alternative to the current exchange or social interaction. The fallback may involve not exchanging or not participating in the interaction, or it may involve an alternative, such as buying from a different store, or selling to a different customer.

The price at which a seller can sell her good will depend on her own fallback and the fallback options of those to whom she could potentially sell the good. Similarly, the price at which a buyer buys a good depends on his fallback option and the fallback options of those from whom he could potentially buy the good.

Consider the three markets in Figure 9.24 where there is a single seller, Zenji, who can produce one unit of a good at a cost c. This cost is his minimum *willingness to sell*: he will not sell the good for less than it would cost him to produce it. (Because there is only one unit of the good produced we omit the "marginal" and "average" in describing its cost.) The good is not divisible so Zenji must sell one unit to just one buyer.

We will consider three potential buyers: Avanti (A), Bella (B), and Carlos (C). Figure 9.24 a is a **bilateral monopoly**, which we have encountered in the bargaining between Ayanda and Biko in Chapter 4: there is one buyer, Avanti (who could be called a monopsonist that is, a single buyer) and one seller, Zenji. In panel (b), Zenji has two prospective buyers, Avanti and Bella (A and B) with different maximum willingness to pay for the good ($\bar{p}^A > \bar{p}^B$). In panel (c) Carlos has joined the market, so we have three prospective buyers: A and C have the same willingness to pay for the good, and B's willingness to pay is lower than theirs ($\bar{p}^A = \bar{p}^C > \bar{p}^B$).

We start with market (a) the bilateral monopoly. If the buyer and seller exchange the good, then the total economic rent that results will be the sum of Avanti's consumer surplus and Zenji's economic profit. This is the difference between Avanti's willingness to pay and Zenji's willingness to sell: $\bar{p}^A - c$. Any price in the range $[c, \bar{p}^A]$ would satisfy both Zenji's and Avanti's participation constraints because both would do better than their fallback

BILATERAL MONOPOLY In a bilateral monopoly transaction there is a single transactor on each side of the market—one potential buyer and one potential seller.

! REMINDER Each of these scenarios is like an Ultimatum Game or a Dictator Game from Chapter 2 with different institutions determining whose actions determine the distribution of the economic surplus.

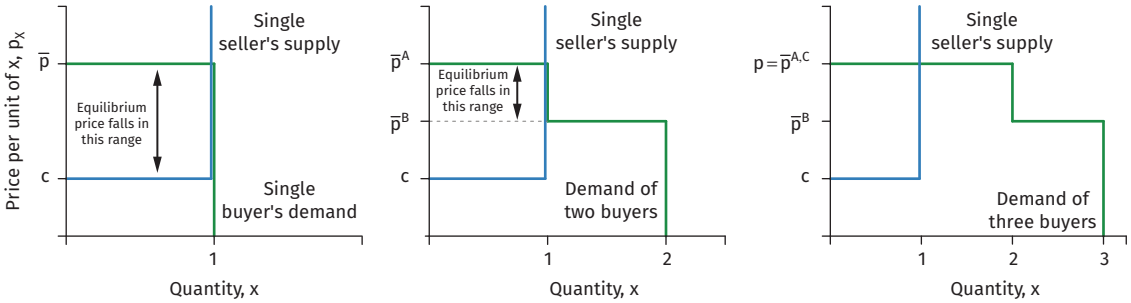
option. The lower the price, the more consumer surplus Avanti would get, and the higher the price, the more economic profit Zenji would get.

Which of the seller or the buyer obtains the larger share of the economic rent will depend on the rules of the game governing the interaction. For example if Zenji is first mover and can commit to a particular price, then he will charge $\bar{p}^A$, Avanti's maximum willingness to pay, or just a bit less. Avanti will accept and the total economic rent made possible by the transaction $\bar{p}^A - c$ (or virtually all of it) will go to Zenji in the form of economic profit. Avanti has barely improved on her fallback position, which in this case was to not buy the good, gaining a consumer surplus of zero.

Of course had Avanti been the first mover the results would have been exactly the opposite. She would have offered to buy the good at the price c, which is Zenji's fallback option, or a bit higher, capturing all of the economic rent in the form of consumer surplus.

How does an increase in demand affect the outcome of the interaction? This could take one of two forms. First, if Avanti for some reason came to value Zenji's good more, then her maximum willingness to pay would be greater than $\bar{p}^A$. Now supposing that the rules of the game had not changed

Figure 9.24 Supply and demand for one seller and the cases of one, two, and three buyers. Panel (a) is a bilateral monopoly with one buyer and one seller. We assume buyer A has the highest willingness to pay for the good. The transaction can occur at a price anywhere in the range between the seller's cost and the buyer's willingness to pay, $[c, \bar{p}^A]$. Panel (b) has one seller and two buyers, where buyer B's willingness to pay is lower than buyer A's willingness to pay, but higher than the seller's cost, $c < \bar{p}^B < \bar{p}^A$. The availability of B as an alternative buyer narrows the range of prices over which A and the seller will negotiate, since the transaction price p will now be in the range $[\bar{p}^B, \bar{p}^A]$. Panel (c) shows what occurs when a third buyer, C, with the same willingness to pay as A ($\bar{p}^C = \bar{p}^A$) enters the market.



(a) Bilateral monopoly (b) One seller, two buyers (c) One seller, three buyers

(Zenji is still first mover) then the result would be that Zenji would charge a higher price, and receive larger economic profit.

There is a second way the demand could increase: the entry of a second prospective buyer, Bella. Bella's willingness to pay changes Zenji's willingness to sell. Because Bella will purchase the good for any price less than or equal to $\bar{p}^B$, Zenji can credibly refuse to sell the good to Avanti for any price lower than $\bar{p}^B$. Zenji's minimal willingness to sell before was c the cost of the good but now it has increased to $\bar{p}^B$ because his fallback option in dealing with Avanti is no longer simply not selling the good, but instead selling it to Bella. Whatever the rules of the game are, as long as the exchange is voluntary, the result is to narrow the potential prices to the range $[\bar{p}^B, \bar{p}^A]$, and ensures that Zenji now takes at least $\bar{p}^B - c$ as economic rent from the exchange.

If Avanti is first mover, the best she can do is offer to buy the good from Zenji at the price $\bar{p}^B$, a higher price than when Bella was not in the market. The increase in demand resulting from Bella's entrance to the market changed the price because it improved Zenji's fallback option. Zenji could sell to Bella if he did not sell to Avanti.

Finally, panel (c) shows what happens when a third possible buyer, Carlos, who has the same willingness to pay as Avanti, $\bar{p}^C = \bar{p}^A$, joins the market. The seller, Zenji, now has the fallback option of selling the one unit either to Avanti or to Carlos at their common willingness to pay $p = \bar{p}^A = \bar{p}^C$.

Bella will not participate in the exchange once Carlos enters the market, because her willingness to pay is lower than either Avanti's or Carlos's. As a result Zenji would not agree to sell her the good at a price she would accept, even though his cost of producing the good is less than the minimum she would be willing to pay to acquire the good.

The arrival of Carlos means that Zenji's willingness to sell is no longer c when his fallback option was to not sell the good (and gain zero economic profit), and it is no longer Bella's willingness to pay $\bar{p}^B > c$ when his fallback option in dealing with Avanti was to sell to Bella. It is now $\bar{p}^A = \bar{p}^C$ because his fallback option in selling to Avanti is instead to sell to Carlos.

The presence of the fallback buyer for the seller—either Avanti or Carlos—ensures that Zenji will capture all of the economic rents in the form of economic profit.

The interactions illustrated in Figure 9.24 show that changes in demand (including an increase in the degree of competition among the buyers) alter the price by changing the fallback option of the seller.

This helps us to understand why in the perfectly competitive model consumers gain some consumer surplus. Perfect competition improves the fallback positions of buyers by giving them many sellers from whom they could purchase. Because a buyer from a particular seller in the perfectly

› EXAMPLE In Chapter 11 we will see that an increase in the demand for labor will affect the Nash equilibrium wage by altering the fallback option of employees. Employers will pay them more because workers' fallback options—finding another job—will be better, the greater is the demand for labor.

competitive model has the option of buying exactly the same good from another seller at the same price, their fallback position is exactly what they get in equilibrium. Rents are shared in the perfect competition model because buyers have access to something even better than a close substitute: an identical good at the same price from another supplier.

CHECKPOINT 9.14 **More sellers of the good** Imagine that instead of Avanti being joined by two other buyers of the good (Bella and Carlos), another seller, Yao, joined the market.

a Yao joins the market with costs, c^Y, which are greater than Zenji's costs, c^Z. Avanti is the only buyer. Draw the supply and demand curves. How has Avanti's fallback changed? Which seller will Avanti exchange with?

b How do these changes in supply and demand affect the prices at which goods are exchanged?

c What will be the division of economic rents between Avanti, Yao, and Zenji in part a. of this question if Avanti is the first mover?

9.16 THE "PERFECT COMPETITOR": RENT-SEEKING FIRMS COMPETING IN AND FOR MARKETS

HISTORY Friedrich Hayek was a leading critic of centralized economic planning and an advocate of limited government (see Chapter 14). But as you know from the head quote to this chapter, he was an equally vociferous critic of the model of perfect competition.

The Cournot model of the continuum of competition from monopoly to unlimited competition, together with the theory of price discrimination suggest a view of competition quite different from the model of perfect competition. In this alternative model of competition, buyers and sellers are not price-takers, rather they are price-makers and rent-seekers of the kind that Hayek envisioned in his view of competition in the head quote of this chapter.

HISTORY Our thinking is influenced by Makowski and Ostroy (2001), a modern-day restatement of ideas of Hayek, Ronald Coase, and Joseph Schumpeter. The term "perfect competitor" is from Makowski and Ostroy.

You have already seen an example of these active rather than passive traders in the case of the seller who can perfectly discriminate in the prices they charge, getting all of the consumer surplus, eliminating the deadweight loss associated with monopoly, and resulting in a Pareto-efficient outcome. The new insight here is that monopolies implement inefficient outcomes because of limited competition and because of the limits on price discrimination. The fact that monopolies limit output in such a way that some consumer surplus that is technically feasible is not realized is a consequence of the assumption that every unit must be sold at the same price, that is, a limitation on price discrimination.

If perfect price discrimination were possible, and in other cases like this, where active rent-seeking buyers and sellers lead to a Pareto-efficient outcome, we refer to the actor as the Perfect Competitor, substituting this term for "perfect competition." Of course, active rent-seeking buyers and

sellers are not *trying* to implement a Pareto-efficient outcome any more than the passive price-taking actors of the perfectly competitive model are.

To see how this active view of competition works, we need to broaden our view of how economic competition takes place. So far we have looked at two different ways to compete.

In the Cournot model, firms compete by setting output levels to maximize their profits taking account of the effect of their output choice on the prices at which they can sell the product, given the output level of other firms.

In the model of market equilibration by rent-seeking and price-setting we have just examined, firms compete by setting prices so as to get rents that are available when markets do not clear. But setting prices or quantities are just two of the competitive strategies firms adopt.

Erecting and heightening barriers to entry

Odd as the phrase sounds, an important way firms compete is by attempting to reduce the number of competitors. Recall that a firm's profit increases the fewer competitors there are in a market. We have already mentioned one method of reducing competition: forming a cartel (a group of firms that decide jointly on their level of output and pricing).

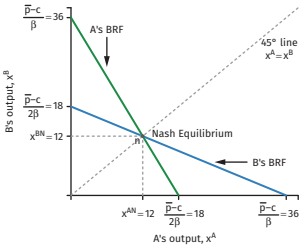
A second strategy to reduce the number of competitors is deterring entry by competitors either by buying the potential competitor or practicing what is called predatory pricing. Recall from Chapter 8 that predatory pricing is setting prices below cost for a period of time to discourage entry or drive out a competitor.

Entry deterrence by a monopolist in the Cournot model could be accomplished if the monopolist, say Firm A, when facing a possible entry by a competitor, Firm B, were to temporarily produce not at the monopolist's profit-maximizing level of output but at the level of output for which the competitor's best response is to produce nothing.

To see how this would work look at Figure 9.25. Because Firm B's best response to Firm A's selling more is to sell less, if Firm A is willing to temporarily forego making profits, it can induce Firm B to not produce at all, meaning, not to enter. This "entry deterring level of output" by Firm A is the x-axis intercept of B's best-response function. This is also the output level of the entire industry in the case of unlimited competition. The entry-deterring price would then be the competitive price, namely $p = mc$.

Intellectual property rights—patents, copyrights, and trademarks—provide a third strategy to deter entry. Patents and copyrights award the firm a monopoly over a device, process, work of art, or piece of knowledge or music. These have the objective of transforming public good—from which others cannot be excluded—into a club good, where the intellectual property right is designed to exclude others. Many firms aggressively convert their knowledge into intellectual property with the objective of excluding potential competitors.

Figure 9.25 Best-response functions for the two duopolists (repeated from earlier). The level of Firm A's output that would deter B from entering is $x^A = 36$. This is the value of A's output for which B's best response is $x^B = 0$.



! REMINDER A club good is non-rival (like a public good) but excludable (like a private good).

EXAMPLE In 2008, Amazon expressed interest in buying a small but rapidly growing e-commerce firm, Quidsi. Quidsi specialized in baby care, household goods, and beauty products. In 2009, Quidsi declined the offer by Amazon.com to buy them. Amazon immediately cut prices on diapers and baby products by 30 percent and shortly thereafter launched Amazon Mom offering free two day shipping for baby and household products. Quidsi's sales fell and their investors looked for ways to sell the business. In 2011, Quidsi was purchased by Amazon.[10]

REMINDER Product differentiation is a business practice aimed at making the firm's product appear more distinct from or less similar to substitute products.

Advertising and product differentiation

A monopolistically competitive firm—the sole seller of a particular good—can devote resources to advertising to shift the demand curve for its product up (increasing $\bar{p}$). Advertising can also increase brand loyalty so that consumers consider other similar products to be less good substitutes (less similar), thereby making the demand for the product less price-elastic. The effect will be to allow the firm to raise prices with a lesser opportunity cost of lost sales. Firms also invest in product design features that are primarily aimed at *product differentiation* rather than increasing the functionality of the product.

Consider, for example, the array of smartphones that exist on the market. Apple's iPhone uses iOS (a proprietary operating system) and the phone is sold as a high-end product, with competition from firms like Samsung with phones like the Samsung Galaxy S series using the Android operating system as a competitor. Apple adopts a variety of strategies (combined with advertising) to *differentiate* the iPhone from its competitors. (An indication of how important Apple's differentiation is can be seen in other firms' attempts to *copy* the iPhone.) As with other strategies, each firm will have a best-response product differentiation reaction to other firms. In some cases, firms will differentiate their products so dramatically that they form entirely new markets. Rent-seeking product differentiation leads to *market-making*, described below.

Lobbying

Firms engage in political activities to ensure an environment—taxes, property rights, interest rates, import tariffs, immigration rules, and other policies—that favors the owners' profits. Lobbying in some countries takes the form of contributions to political campaigns, hiring people to persuade elected officials of the firm's point of view, public advertising, and even vacation retreats for judges who might affect legal decisions affecting the firm. Incumbent firms also lobby for costly licensing requirements and other regulations that make entry of new firms more difficult.

Innovation

Improvement in technology, we showed in Chapter 6 (section 6.13), can create innovation rents for those who adopt the new methods first. These rents are temporary, ending only once competitors manage to adopt the same new technologies, or find equivalent alternatives.

Innovations need not be in technology; a new form of organization or institution may reduce costs and allow a firm to capture economic rents until others copy it. An example is outsourcing by firms, purchasing from other firms' inputs previously produced within the firm (such as the components of an iPhone, or a car's engine). This strategy allows firms located in high-wage countries to acquire inputs at low cost from countries

with lower wages. During the Industrial Revolution, the factory itself was an organizational innovation. Spinning yarn, weaving cloth, stitching cloths, and other kinds of work had previously been done mostly in people's homes under what was called the "putting out" system. The industrial factory brought hundreds of workers together under the direct supervision of a single management, creating new opportunities for cooperation and control in the labor process.

Market making: Competing for markets not in markets

Lastly, consider *market-making*, the creation of new markets to generate new rents. We look at two examples to illustrate the idea of market-making.

Let's go back to the mid-1950s and relate a story about how consumers got to have transistor radios. Pat Haggerty of Texas Instruments, having found out about the invention of transistors, managed to negotiate a license for the patent for transistors from Bell Labs. When he purchased the license for $25,000 dollars, transistors were being sold to the military for $16 each. Haggerty told his engineers at Texas Instruments that they would have to produce a transistor for $3 to make it possible to sell to consumers. Haggerty made this declaration in June 1954, telling his engineers they were going to market by November 1954—just in time for Christmas.[11] They made it. Texas Instruments started selling a transistor radio—the Regency TR-1—for $49.95. The Regency TR-1 came in four colors (black, ivory, Mandarin Red, and Cloud Gray) and it could fit in your pocket. At the time, the other radios on sale were all large, clunky pieces that took up space on the kitchen counter or living-room side table.

Few businessmen at the time thought people wanted a pocket-sized radio, but Haggerty knew better and engaged in remarkable market-making. By the end of 1955, over 100,000 units had been sold. Haggerty was also lucky that Elvis Presley's song "That's All Right" began blaring over the airwaves in 1954, and everyone wanted to hear it on their own transistor radio.

The Regency TR-1 was the iPod of the 1950s, taking advantage of the advent of rock and roll, and personalizing access to music. No one knew they wanted one, until they did.

Let's track back from the iPod of the 1950s to the iPad of 2010. Apple tried to sell a tablet once before in 1993—the Newton, but it tanked. Microsoft produced a tablet in 2000 called the Microsoft Tablet PC, but it didn't catch on either. Though many firms tried to produce a tablet that would satisfy customers, the market for tablets did not take off until the Apple iPad.

Apple's iPad was released in 2010 and many people wondered why consumers would want "a bigger iPhone that can't make phone calls." But these skeptics were wrong. Having released the iPad in April of 2010, by 2011 Apple had sold almost 15 million iPads. No one outside of Apple had predicted such success.

Figure 9.26 A Mandarin Red Regency TR-1 Texas Instruments transistor radio first marketed in 1954.

Photo Courtesy Joe Haupt.

HISTORY Hayek stressed that competition is a social relationship among people, objecting to "the . . . complete exclusion from the theory of perfect competition of all personal relationships existing between the parties . . . competition is in a large measure competition for reputation or good will."[12]

HISTORY According to Joseph Schumpeter (1883–1950) innovation occurs through a process that he termed *creative destruction*. Disruption, not equilibrium, is his central idea. The "creation" of the innovating firm "destroys" firms that do not keep up, giving the innovators at least temporary economic profits called *Schumpeterian rents*. These are competed away when other rent-seeking firms come up with new ideas or duplicate the ideas of innovating firms. Notice: the process of diffusion of innovations works because innovating firms do not have intellectual property rights sufficient to make them sole monopolists of the new idea.

Apple's *market-making* with the iPad paved the way for competition in a new market, with Microsoft reentering with the Surface Pro, Google selling the Google Nexus, Samsung the Galaxy tab, and Amazon the Amazon Fire Tablet (among others). Once Apple had made the market, other firms entered it in order to obtain a share of the economic rents on the tablet market.[13]

Since the advent of the Internet and the app-based economy, many new markets have also been created. From ride-hailing apps like Uber and Lyft creating new markets for rides that differed from standard taxi services, to Airbnb creating opportunities for homeowners and renters to match their preferences for renting out and staying in rooms or homes, the Internet and its ability to match buyers and sellers of goods has enabled many new markets to exist that previously did not.

CHECKPOINT 9.15 **Rent-seeking competition and perfect competition**
Contrast the strategies adopted by buyers and sellers in the perfectly competitive model with the rent-seeking activities of what we call "the perfect competitor."

9.17 APPLICATION: DECLINING COMPETITION AND INCREASING MARKUPS

The many forms of competitive rent-seeking just described may lead to a reduction in the degree of competition. In the four decades following 1980 this appears to have been the case in the US.

There are many ways that these trends can be measured. Focusing on the US, the revenue of the top 50 firms as a share of the industry total revenue has risen in many sectors including retail trade, finance and insurance, real estate and leasing, wholesale trade, transportation, and warehousing. Over this period, the number of new firms entering industries has fallen and as a result the age of firms has risen. Since the early 1980s, the fraction of total employment that was in young firms (less than five years old) has been cut almost in half.[14]

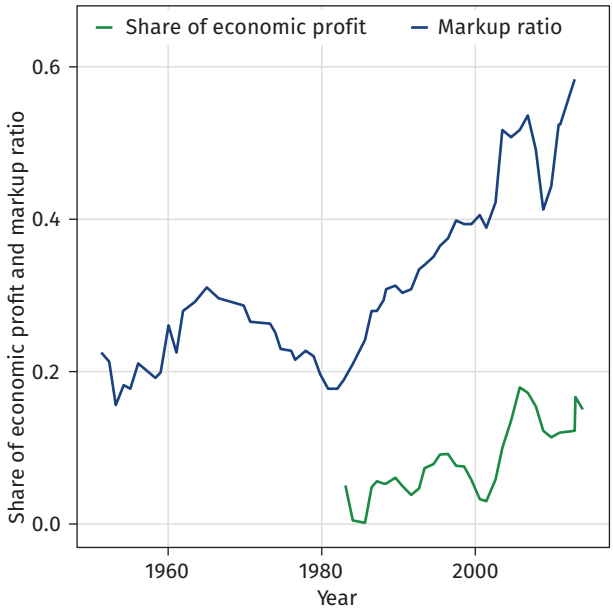
A better overall measure of the decline in the degree of competition in the US economy is the extent to which price exceeds marginal costs, which we have measured by the markup ratio, namely the excess of price over costs, divided by costs. The blue line in Figure 9.27 shows an estimate of the markup ratio across the US economy since the middle of the last century.

It rose during the period of strong demand from mid-century to the mid-1960s and then declined as the growth of the economy slowed during the period called stagflation. But from 1980 onward, it tripled, consistent with a dramatic reduction in competition.

As the markup measures the excess of a firm's sales revenues over its costs including the opportunity cost of capital, it is no surprise that

Figure 9.27 Rising markup ratio and economic profit as a share of total income. The markup ratio is $\mu = \frac{p-c}{c}$ as defined earlier. It equals 0 when $p = c$ and is greater than zero when $p > c$. The markup ratio is for the entire US private economy with estimates for each sector weighted by their share of total sales. The share of economic profits is defined as the share of accounting profits minus the share representing the opportunity cost of capital. The markup ratio and the profit rate move in the same way, suggesting a relationship between decreasing competition and increased profitability of firms.

Sources: Barkai (2020); De Loecker et al. (2020).



economic profit as a share of national income has also greatly increased. This is shown in the green line in the figure. Trends toward a less competitive economy in Europe are less pronounced, in large measure because the European Union has much stronger policies to ensure competition.

The Cournot model and other models of markets demonstrate the deadweight inefficiency and losses in consumer surplus that occur when a firm with a limited number of competitors faces a downward-sloping demand curve and therefore act as a price-maker.

To address these consequences, monopoly, duopoly, oligopoly, and other cases of limited competition governments have pursued **competition policies**, often breaking up large firms into smaller firms. By breaking up firms,

COMPETITION POLICIES Government policy and laws to limit monopoly power and prevent cartels or to otherwise regulate the process of competition. Also known as antitrust policy.

✓ **FACT CHECK** At the turn of the century economists regarded the European economy as less competitive than the US. But this relationship has dramatically reversed. In the telecommunications sector, for example, measures of competition decreased substantially in the US while rising in Europe. Consumers have felt the difference. The average monthly cost of broadband in the US in 2018 was double that in France and Italy.[15]

HISTORY Senator John Sherman called the act in his name "a bill of rights, a charter of liberty" and in 1890 on the floor of the US Senate said: "If we will not endure a king as a political power, we should not endure a king over the production, transportation, and sale of any of the necessities of life. If we would not submit to an emperor, we should not submit to an autocrat of trade, with power to prevent competition and to fix the price of any commodity."[16]

competition policies increase n, the number of firms. As a result, they lower the price markup over costs, reduce profits, and increase the consumer surplus.

Models and policies of this type have motivated antitrust legislation such as the Sherman Act (1890) in the US, and the breakup of the giant Standard Oil company by a decision of the US Supreme Court in 1911. Similar policies have been pursued by the European Commission, and Competition Commissions in South Africa, India, and many other countries. The interest of consumers is served by the increase in the number of firms as long as the smaller firm size is not associated with significantly higher cost.

9.18 APPLICATION: MODERN MONOPOLY, WINNERS TAKE ALL, AND PUBLIC POLICY

But there is an important difference between Standard Oil along with other giant firms such as US Steel and British Petroleum on the one hand, and a novel market structure exemplified by Amazon, Alibaba, Microsoft, Tencent, ByteDance (the owners of TikTok), and Facebook. The term "modern monopoly" has been used to describe these new giant firms but like the "conventional monopolies" with which they are contrasted they aggressively compete with other companies selling similar products.

Large is beautiful

What is "modern" about these firms is the extraordinary competitive advantages of large scale that they exploit. These advantages occur for two reasons:

! REMINDER Price markups on drugs Have another look at Figure 8.13: Prices of the pharmaceuticals shown there are between four and 30 times the cost of producing the treatments.

- *First-copy costs and near-zero marginal costs*. Many modern monopolies have substantial fixed costs and very low marginal costs. This is true of knowledge production and processing firms in which there may be substantial costs to produce or acquire rights to the information (this is called a "first-copy cost"), but close to zero costs to making it available to consumers after that. An example of this "zero marginal cost" characteristic of modern monopoly is Spotify and other streaming music services. The production of pharmaceuticals is another: the first copy costs of research are substantial, as are other fixed costs such as advertising and lobbying (to affect public policies in their favor). But production costs of the treatments are very modest.
- **Demand-side economies of scale**. As the examples of Amazon and Facebook illustrate, modern monopolies benefit from what are called

DEMAND-SIDE ECONOMIES OF SCALE occur when the value of a firm's product is greater to a buyer the more other buyers of the product there are.

"network-based economies of scale in demand." One of the benefits of buying on Amazon is that you can read reviews of the product you are interested in by hundreds of other Amazon users. People looking for an apartment for a weekend go to Airbnb because many apartment owners have posted their apartments there, which they would not have done unless lots of apartment seekers were logging on to Airbnb.

! REMINDER Economies of scale in demand are sometimes called "network economies of scale" because the value of being part of the network (e.g. of Facebook or Instagram users) increases the more other users there are.

A result of these novel forms of economies of scale is **winner-take-all competition**. This is a process of competition which results in a monopoly or near monopoly. As countless companies that challenged Amazon have discovered, the advantages of size are decisive.

To model modern monopoly, we have to modify our assumption (in the Cournot model) that average and marginal costs are independent of the output. Let's consider the extreme case of a firm or firms which have a substantial fixed cost c_0. Total costs (irrespective of output) are then c_0 and marginal costs are zero. Therefore, average costs are $\frac{c_0}{X}$. A single firm with this cost structure would behave just as the monopoly in the Cournot model, producing a quantity such that the *marginal revenue* equaled the marginal cost (namely zero). As with the monopoly in Figure 9.13, because price exceeds marginal cost the monopoly would produce less output than if *price* were equal to marginal cost (a price of zero).

What would be an appropriate policy to remedy the loss in consumer surplus resulting from the firm's monopoly status?

Competition policy for modern monopolies

The first thing that would come to mind is the traditional antitrust remedy: break up the monopoly into a sufficiently large number of firms so that prices and deadweight loss would fall, while consumer surplus would increase.

But this would make no sense. Why? Instead of a total cost of providing the goods equal to c_0 by a single firm, the costs would be the number of firms resulting from the breakup of the monopoly, n, multiplied by the fixed costs, or nc_0. Therefore, the average cost of the goods or services provided would be nc_0/X rather than c_0/X. Thus the average cost of a unit would double if there were two firms rather than one, quadruple if there were 4, and so on.

The challenge for public policy is that in the case of modern monopoly small firms are often high-cost firms. Since the 1980s the emergence of so-called superstar firms has meant a shift to firms with higher productivity of inputs and lower costs. Large size is cost-reducing because it spreads

WINNER-TAKE-ALL COMPETITION is a competitive process that results in a monopoly or near monopoly.

a single fixed cost over all the units produced. Table 9.3 summarizes differences between conventional and modern monopoly.

In the case of modern monopoly the following approaches have been proposed.

- *Public regulation or ownership*: Recognize that the cost advantages of large-scale production of this good or service along with the resulting winner-take-all competition make it unavoidably a monopoly. So let the government regulate the prices it can charge and other aspects of its business. This has been the conventional approach to what are called "natural monopolies" such as railroad and electricity networks.
- *Taxation of profits due to limited competition*: Taxation can mitigate the income inequalities associated with limited competition, but the market failure associated with the excess of price over marginal cost would remain.
- *Public funding of the first-copy costs*: Government support for research and cultural production, the results of which are freely available, can reduce the first-copy costs that contribute to the winner-take-all form of competition.
- *Competition for the market rather than in the market*: The objective of "competition *for* the market" policies is to seek to ensure that monopolies are owned and managed by people who will produce the best products at the lowest cost.

To see how competition for the market might work, think of an industry in which the sharply declining costs per unit of output sustains a winner-take-all kind of competition. Suppose a single firm is dominant in producing some good or service. The owners and managers of the firm are incompetent, resulting in inferior products or costs in excess of what is technically possible.

But they are able to deter entry by competitors through predatory pricing and other strategies. Then another group of potential owners, who would run the firm more profitably might consider buying a controlling share of the stock of the incumbent firm. Facilitating this kind of firm buyout is an example of competition for (rather than in) the market.

Government ownership of a monopoly rules out competition *for* the market among alternative possible owners. But the management of a government-owned firm may be replaced as a result of political competition. A publicly owned railroad offering poor service, for example, would be considered a political liability by the incumbent political party that would seek improvements. Neither competition for the market nor political competition can be expected to work as well as unlimited competition, where this is an option.

Table 9.3 **Conventional and modern monopoly.** The quotation marks in "natural monopoly" underline that while many new monopolies are in some respects like the older natural monopolies, like railroad networks, the limited competition they face is in part due to social not natural causes, for example, intellectual property rights.

Type	Examples	Characterization	Harm	Remedy
Conventional	GM, U.S. Steel Tata Industries BASF British Petroleum	Downward-sloping demand curve; stable competition among the few	Inefficiently limited buyer access to goods and services; foregone consumer surplus survival of high-cost firms	Objective: restore competition (in the market) by breakup of large units; price caps; public regulation/ownership of natural monopolies
Modern	Amazon Facebook Tencent/WeChat Alibaba Uber/Lyft Bristol Myers Squibb (pharmaceuticals)	A new "natural monopoly" high first-copy costs zero or low marginal costs; demand-side increasing returns; serial monopoly	Foregone consumer surplus when $p > mc$; centralized control over essential social infrastructure; limited competitor access to market; winner-take-all firm which may not be most efficient or highest value provider	Public regulation or ownership; support competition for the market and facilitate takeovers; auction of market dominance; corporate governance reforms; tax monopoly profits; publicly fund first-copy costs

CHECKPOINT 9.16 Demand-side economies of scale

a What are demand-side economies of scale? Give examples.

b Draw an inverse demand curve for a firm selling a product subject to demand-side economies of scale.

c On the inverse demand curve you just drew, pick a price $p = p^*$ and quantity $X = X^*$. How many people will buy the firm's product if $X < X^*$? How many people will buy the firm's product if $X > X^*$?

9.19 CONCLUSION

We have modeled and described how competition works as a result of strategic rent-seeking action by firm owners and managers. We have seen, paradoxically, that the rent-seeking actions that are essential to how competition works also include the anti-competitive construction of barriers to entry and other strategies to limit the extent of competition.

External effects have played an important role in our account of the competitive process, the sales of one firm putting downward pressure on the prices at which all firms can sell their products. We drew an analogy between the external effects of "overcrowding" the market with an earlier case: "overharvesting" fishing stocks. In both cases the external effects were among *symmetrical* actors: Fisherman A reducing the catch of identical Fisherman B, or Firm B's sales of its product reducing the price at which identical Firm A can sell its product.

In the next part of the book we will turn to a new kind of external effect. This occurs between asymmetrically located actors, called, respectively principals and agents. These include: bankers and borrowers, employers and workers, and owners of firms and their subcontractors. While the games we introduce to analyze these principal–agent relationships will be new, the results will be familiar: Pareto-inefficient Nash equilibria, mutual gains to be had through exchange, and conflicts over how these gains will be shared.

MAKING CONNECTIONS

Mutual gains and conflicts of interest over their distribution: Here, conflicts of interest occurred (a) among firms' owners over the distribution of economic profit and (b) between owners as a group and consumers over the distribution of the total rents (gains from exchange) in the form of economic profit as opposed to consumer surplus.

Institutions policies and rules of the game: How these conflicts are resolved depend on the extent of barriers to entry and other influences on the degree of competition that are affected by public policies such as intellectual property rights and competition policies.

External effects and coordination problems: The coordination problem facing the owners of competing firms is identical to the "Fishermen's Dilemma" encountered earlier. In both cases the production by each of the actors (fishermen, firm owners) reduces the benefits enjoyed by the others. A difference with the Fishermen's Dilemma setting is that in the Cournot model we also consider the interests of consumers who would be better off if the firms produced *more*.

Dynamics: The determination of the equilibrium number of firms in the industry and the resulting price markup over costs as a result of the extent of barriers to entry is similar to how we modeled the equilibrium number of people overexploiting the fishing stock. We will use the same model and the resulting equilibrium prices when we model the whole economy in Chapter 11.

A new benchmark model of the firm: We take the monopolistically competitive firm facing a downward-sloping demand curve and a constant marginal cost as our "default option" model. The perfectly competitive price-taking firm facing a horizontal demand curve, whose size is determined by a rising marginal cost function is a special case.

Rent-seeking, price-making, strategic behavior, and the perfect competitor: Profit-seeking is an active and strategic process of price-making (including price discrimination), product differentiation, innovation, securing a favorable legal and policy environment, and much more. This contrasts with the passive view of price-taking competitors, whose strategy sets are confined to the amount to be bought or sold. Price-making (and wage-making and interest-rate-setting) will play an important part in understanding the modern capitalist economy (Chapters 10 through 12 and 15).

Economics as an empirical science: Our new benchmark model is motivated by the evidence on cost functions in Chapter 8; it is put to work understanding recent increases in markups and the share of economic profits in total income in the US economy.

IMPORTANT IDEAS

competition	"the firm"	industry
unlimited competition	tipping point	markup and markup ratio
monopoly	duopoly	oligopoly
barriers to entry	marginal cost	equilibrium number of firms in an industry
profit maximization	marginal revenue	
best-response function	cartel	profit
demand-side economies of scale	Prisoners' Dilemma	merger
price-taking	market demand and supply curves	Pareto efficiency
perfect price discrimination	predatory pricing	price-taking equilibrium
equilibrium markup over costs	incumbent firm/entrant	expected price
rent-seeking and rent-seeker	fallback	product differentiation
market dynamics	long-run industry equilibrium	first-copy cost
consumer surplus	perfect competitor	opportunity cost of capital
product differentiation	"overharvesting" buyers	modern monopoly
network economies of scale	zero marginal cost	housing price bubble

MATHEMATICAL NOTATION

Notation	Definition
X	industry output
X^{-i}	total output of all firms except i
p	price
$\hat{p}$	expected price with barriers to entry
$-\beta$	slope of the industry demand curve
c	marginal cost
$r()$	revenue function of a firm
π	profits of a firm
Π	industry profits
n	number of competing firms in industry
b	probability of failure of a new firm
μ	markup ratio
ρ	opportunity cost of capital

Note on superscripts and subscripts: A, B: firms; N: Nash equilibria; m: monopoly; S: supplied; D: demanded; H: high; L: low.

"Not everything in the contract is contractual . . . the contract is not sufficient in itself but is possible only thanks to a regulation of the contact, which is social in origin."

Emile Durkheim,

De la Division du Travail Social (The Division of Labor) (1902)[1]

PART 3

MARKETS WITH INCOMPLETE CONTRACTING

Our translation to English of Durkheim's statement above can be further translated to simpler language: "contracts do not cover everything and they do not enforce themselves." In Part III of the book we study markets in which critical information about the goods or services being exchanged is either lacking or cannot be used in court, and as a result contracts cannot cover everything that matters in an exchange. Or contracts may simply be unenforceable.

The information problems and the incomplete contracts to which they give rise are illustrated by the difficulty of contracting for quality in Chapter 10. You purchase a made-to-order software package capable of executing a series of calculations and creating visual representations of the results. But how good, really, is the code? You and the software developer will often disagree. And there is no way that the exact capabilities of the code could have been specified in an enforceable contract (how good the visuals had to be, how quickly the algorithms involved would compute which kinds of results, and so on).

In situations like this the contract will necessarily be incomplete, and this fundamentally changes how markets work. This is why we devote separate parts of our book to markets with complete contracts (the more conventional subject matter of undergraduate economics texts) and those with incomplete contracts.

The employer and the employee in the labor market illustrate a similar problem in Chapter 11. The employer effectively rents your time, but he cannot purchase and contractually enforce exactly the tasks he needs you to do in order for him to make a profit. In Chapter 12, we show that banks and other lending institutions face similar problems of incomplete or unenforceable contracts in the credit market.

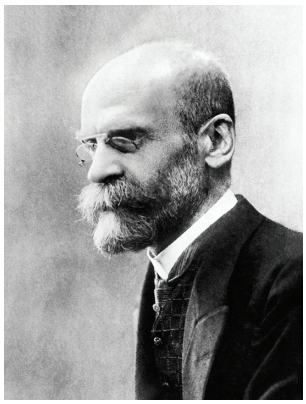
In these and other markets where enforceable contracts do not cover everything that matters in a transaction, mutually beneficial exchanges can nonetheless take place. Social norms may step in to facilitate mutually beneficial exchanges where contracts are incomplete.

If the employee's work ethic commits her to doing the best she can even when her employer is not looking, then the fact that hard work cannot be specified in a contract need not stand in the way of an exchange. Similarly, if the borrower's truthfulness in describing the project which the loan will finance commits him to not undertaking excessively risky options, the prudent use of loaned funds will occur, even if it cannot be enforced in a contract.

In these and other important markets, mutually beneficial exchanges are made possible by some combination of contract, social norms, along with a third element. This is the exercise of power by principals (lenders, employers) over agents (borrowers, workers).

Émile Durkheim (1858–1917)

Photo by IanDagnall Computing/Alamy Stock Photo.



HISTORY Emile Durkheim (1858–1917), considered (along with Max Weber, 1864–1920) a founder of sociology, stressed the importance of social norms (rather than the law) in making mutually beneficial exchanges possible.

How can a lender or employer—private individuals—exercise power in these interactions? They clearly cannot legally threaten or inflict physical harm as was the case, for example, when the "worker" was a slave. Moreover, because the exchange is voluntarily entered into, the borrower or worker has the option of simply not borrowing from this particular lender or not going to work for this employer. The agents, in other worlds are free to walk away from engaging in the transaction: they cannot be forced to participate.

The only real power the principal has is to terminate the relationship, that is for *him* to walk away. To make the threat of termination effective principals typically set prices (including wages and interest rates) so that agents receive a payment above their next-best alternative, that is, a rent.

Because the agent will lose the rent if the principal terminates the relationship, the worker, lender, or other agent has a good reason to work hard, use borrowed funds prudently, and otherwise to behave in the principal's interest, even if it is not required by contract. The threat of losing the rent is the basis of the principal's power.

The critical role of the agents' social norms (work ethic, truthfulness) and the exercise of power by principals in facilitating exchanges where contracts are incomplete are features that do not appear in markets with complete contracts. Another key difference is that where contracts are incomplete, markets typically do not clear in competitive equilibrium. A consequence is that unemployment—that is when labor supply exceeds labor demand—is to be expected even under ideal competitive conditions and need not be explained by "sticky wages" or other "market frictions."

Another consequence of the incomplete nature of contracts is that even with unlimited market competition the Nash equilibrium will be Pareto inefficient. This shows that coordination failures go beyond such problems as overexploiting nature, or markets with limited competition that you studied in the first two parts (respectively) of this book.

Taking account of the fact that contracts are often incomplete leads to new insights about how markets really work. And so, looking back to the head quote for Part II of this book, to Voltaire's amazement at how markets facilitate exchanges even among people who in other contexts might be at war, we add a caveat. This is that where contracts are incomplete, exchanges are facilitated if one of the parties (the principal) has power over the other party.

And we can underline Voltaire's observation: buyers and sellers care very much about the social norms of the people they are trading with. Trust matters. As Voltaire put it: "There the Presbyterian confides in the Anabaptist, and the Churchman depends on the Quaker's word." And so the exchange process is far from anonymous; it is often very personal. This is why when looking for a used car to buy, we typically go to somebody we know or a seller with a good reputation, not to a total stranger.

Because the exercise of power by principals over agents and social norms are essential to how markets work when contracts are incomplete, modern economists draw on the insights of political science, sociology, psychology, and the other social sciences. This is why we make occasional reference to these disciplines.

CHAPTER

10

INFORMATION

CONTRACTS, NORMS, AND POWER

In an economic theory which assumes that transaction costs are non-existent [that is, contracts are complete], markets have no function to perform and it seems perfectly reasonable to develop the theory of exchange by an elaborate analysis of individuals exchanging nuts for apples in the edge of the forest or some similar fanciful example . . .

Ronald Coase,
The Firm, the Market, and the Law (1988)[1]

DOING ECONOMICS

This chapter will enable you to:

- Understand the difference between complete and incomplete contracts, and why, due to limited information, many contracts are incomplete.
- Model an interaction of a principal and an agent in which their interests conflict concerning some aspect of the exchange that is not subject to a complete contract.
- Distinguish between "hidden actions" and "hidden attributes" as a source of the contractual incompleteness.
- Show in the hidden actions case that principals will set prices so that agents receive a rent, the market does not clear, and the outcome is Pareto inefficient due to the external effects of the agent's actions that are not subject to contract.
- Explain why, when contracts are incomplete, social norms and the exercise of power by principals facilitate mutually beneficial exchanges.
- Understand the conflict of interest between the principal and agent, and how the information available to the principal affects the distribution of the gains from exchange between them.
- See how the nature of the contract affects social aspects of exchange such as trust and repeated interactions.

10.1 INTRODUCTION: WHO INVENTED HARD RED WINTER WHEAT #2?

When we talk about the market for, say, wheat or toothpaste, what are we talking about? To a farmer, there is really no such thing as "wheat." There are literally hundreds of different species that we call "wheat," and until recently a farmer's crop might be a mixture of quite a few of them. This made buying and selling "wheat" difficult because while the farmer knows what he is selling, the buyer does not.

But we now have markets for a limited number of species and grades of wheat. These markets came about not because nature conveniently produced a standardized product called "hard red winter wheat #2" but because an economic organization—the Chicago Board of Trade—in the mid-nineteenth century adopted classifications and policies to create homogeneous categories of grain so as to facilitate transactions.[2]

Wheat once referred to a diverse collection of products, with size, genetic strain, and quality differing from one sack of wheat to another. For two farmers—Jones and Svenson—the supply, demand, and price for farmer Jones's wheat differed from the corresponding supply, demand, and price for farmer Svenson's wheat. The *markets* for the two farmers differed even though they both sold wheat.

But, thanks to the Chicago Board of Trade, different grades of white winter wheat, red winter wheat, spring wheat, and many other standardized categories came to be of such uniform quality that the ownership of grain no longer referred to any specific sack or particular lot of wheat, but to a contract entitling the owner to the delivery of a specified amount of some particular grade of wheat. As a result the biodiversity of wheat on the American Great Plains fell, but a limited number of new markets, resembling what one sees in an economics textbook, came into being.

Each type of grain has become a homogeneous good, for which traders can write *enforceable* and **complete contracts**—simply for an *amount* without worrying about the quality. The categories of grain are now like electricity: you purchase any amount of it by the kilowatt hour without caring or knowing which power plant generated it. You do not worry about the quality of the electricity you buy. What you buy is what you get. And if you paid for electricity that you did not receive, you can get your money back.

Grain is not unusual in this respect. Goods like Sugar Number 11, Corn Number 2 Yellow, or Light LA Sweet (that's crude oil) are not gifts of nature. Rather, they are created by a deliberate *process of standardization* to eliminate difficult-to-monitor differences in quality.

HISTORY Remarkably, this standardization of grain trading was accomplished by an entirely private body, the Chicago Board of Trade. Memberships in this body would themselves become marketable commodities before the nineteenth century ended.

COMPLETE CONTRACT A contract is complete if it (a) covers all of the aspects of the exchange in which anyone affected by the exchange has an interest, and (b) is enforceable (by the courts) at close to zero cost to the parties.

But the case of wheat is exceptional: it was possible by standardization to make reliable information about the quality of what was being purchased available to buyers, facilitating the exchange process.

But in most economic interactions there are some important bits of information that are known to some of the actors but not to the others. The language training instructor knows how successful he is in teaching fluency in some new language; but those signing up and paying for his classes have no idea of what quality of instruction to expect. The worker knows how hard she worked yesterday, her employer may not. The borrower knows how the loan will actually be used—prudently or recklessly, for example—the banker may not. These are examples (as you already know) of asymmetric information.

Other bits of information may be known to a buyer or seller, but not admissible as evidence in a court of law (that is, non-verifiable). In most legal systems, for example, the employer's account that he found the worker asleep at her desk, unless substantiated by witnesses, would not be considered to be verifiable and could not be used as evidence against the employee, for example to recover the wages paid to the employee for her nap time.

10.2 INCOMPLETE CONTRACTS: "NOT EVERYTHING IS IN THE CONTRACT"

Unlike electricity and red winter wheat #2 many of the commodities transacted in a modern economy are not homogeneous goods or services. The limited nature of information is the reason why, in many markets, contracts are *incomplete*: traders cannot easily verify quality, or they can observe the quality but such quality cannot be verified in a court of law or by some other third party to enforce the contract. If they are disappointed with the quality, there is often no way to get their money back. Think of hiring someone to care for an elderly relative or to take care of your children while you are at work. How can you know the quality of the care they are giving? And even if you later found out that they had been careless, could you go to court to get back the wages you paid them? Almost certainly not.

Contract enforcement by courts or other third parties is called **exogenous enforcement** because the enforcers are not parties to the exchange (they are *outside* the exchange). But for many exchanges this is not the case: the contract is incomplete.

EXOGENOUS ENFORCEMENT OF CONTRACT Exogenous enforcement of the terms of an exchange is done by courts or another third party—not the parties to an exchange themselves—and is a defining characteristic of a complete contract.

Complete and incomplete contracts share two features:

- *Mutual gain*: a transaction is based on the mutual expectation of gain by all the participants.
- *Conflict of interest*: there is a conflict of interest over how these gains will be divided among the parties to the exchange.

The key difference is that when contracts are incomplete the division of the gains is not enforced entirely by an external body—the courts—based on terms specified in the contract. The terms of the contract matter and courts may be involved, but the outcome of the exchange is also the result of strategic interactions among the participants involving rewards, punishments, and the exercise of power.

These include threats, promises, and the creation of incentives through offering repeated interactions. The preferences of the parties to an exchange also matter, for example a commitment to telling the truth about the condition of the used car you are selling, or a worker's intrinsic motivation to do high-quality work.

This is what Durkheim meant—quoted in the introduction to this part of the book—when he said that "not everything in the contract is contractual" and that market exchanges are socially regulated. We call this the **endogenous enforcement** of the terms of an exchange.

Students are experts on incomplete contracts. Suppose as a condition of your employment at a consulting firm following graduation you contracted to "learn microeconomics." How would it be determined that you had fulfilled your obligation? The contract might have been written "pass this particular course in microeconomics" but the employer would hardly be satisfied that this would guarantee that you were able to do the kinds of work they need. Or suppose you did poorly in the exam. Had you failed to learn microeconomics? Or did the exam not test your knowledge of microeconomics? Or did you have a particularly bad allergy attack the day of the exam? How could your employer ever enforce this contract?

Here are some other examples of exchanges under incomplete contracts.

- Owners of firms want managers to maximize the value of the owners' assets, but managers have their own objectives (first-class air travel, lavish offices, "on-the-job leisure") and managerial contracts fall far short of having an enforceable requirement to maximize the owner's wealth.

ENDOGENOUS ENFORCEMENT OF CONTRACT When the parties to an exchange—employers and workers, buyers and sellers, borrowers and lenders—themselves adopt strategies to ensure favorable terms of an exchange for aspects of it not covered by a contract, enforcement is endogenous.

- In many countries, families devote a sizable fraction of their budgets to purchasing educational services for their children, the quality of which is rarely specified in a contract (and would be unenforceable if it were).
- Parents, in turn, hope and expect that their children will care for them, if needed, in their old age; but there is no way to write this into an enforceable contract.

! REMINDER Information is asymmetric if something is known by one party but not by another. This affects the kinds of contracts that can be enforced because a party's information about their own attributes or actions may be *private information*.

Three of the most important examples of incomplete contracts are the subject of this chapter and the next two chapters. Here we consider the case where information on the quality of a good is known to the seller but not to the buyer—whether it is a used car or a piece of clothing provided to Benetton by a subcontractor. In Chapters 11 and 12 we study incomplete contracts in the labor market where information on the effort you put into your job is not readily available to your employer, and the credit market in which your promise to repay the money you borrowed may be unenforceable if you are broke. As these examples suggest, contractual incompleteness is the rule rather than the exception in economic transactions. Here are five reasons why.

› EXAMPLE The insurance company may not know that you are already ill when you seek to extend your health insurance coverage. This is a case of asymmetric information.

1. *Asymmetric or non-verifiable information*: Third-party enforcement of contracts requires information that is available to both parties and can be *verified* by third parties such as courts of law. Information is **verifiable** if it can be used in court to enforce a contract. Non-verifiable information such as hearsay, or even direct but uncorroborated eyewitness observation, generally cannot be used to enforce contracts. Information is asymmetric if something is known by one party but not by another.

› EXAMPLE The landlord may know that the poor condition of a rented apartment is due to the tenant's negligence; but it may be impossible to prove this in a court. This is an example of non-verifiable information.

2. *Time*: A contract is generally executed over a period of time as when a contract specifies that Party A does X now and Party B does Y later. But what if what B does later depends on other things that cannot now be determined? A complete contract must specify what the parties must do in every possible future situation or contingency or "state of the world." In general, people cannot completely specify these future states, and in any case, it is not ordinarily cost-effective to specify what to do in each contingency.
3. *Measurability*: Many of the services or goods involved in the exchange process are inherently *difficult to measure* or to describe precisely enough to be written into a contract. The restaurant owner would like his serving staff to interact in a pleasant manner with customers, but how can this be observed by the owner, and even if it were to be

VERIFIABLE INFORMATION Information is verifiable if it can be used in court to enforce a contract.

observed, how could it be measured or considered to be verified for use in a legal proceeding?

4. *Authority*: For some transactions there is no institution—no court or other relevant third party—capable of enforcing a contract. Many international transactions are of this type. For example, if a country defaults on its debt to international creditors, no third party enforces the claims on the debt. This has happened in a variety of countries internationally, such as Lebanon's default on $1.2 billion of Eurobonds in 2020, or Argentina's debt restructuring that took place repeatedly during the period 2005–2016.
5. *Motivation*: Even where the nature of the goods or services to be exchanged would permit a more complete contract, traders may favor a less complete contract for motivational reasons. Intrusive surveillance of workers by employers to establish verifiable information on work activities, for example, may backfire if the employer's distrust angers workers, leading to less satisfactory work performance.

! REMINDER In Chapter 2 the imposition of a fine on parents who were late in picking up their children from a day-care center is an example of "making the contract as complete as possible." Recall that the result was that more parents were late; completing the contract backfired.

As the final reason suggests, how incomplete a contract will be is in some measure a matter of choice. For example, where the parties to an exchange are trusting and trustworthy people committed to reciprocity in their dealings, perhaps having reciprocal or altruistic preferences like those we saw in Chapter 2, they may deliberately leave some important aspects of the exchange unspecified, even when the relevant enforceable contractual clauses could be written.

The first and fourth reason for incomplete contracts above—lack of verifiability and authority—make it clear that whether a particular good or service is subject to complete contracting will differ from one legal system to another. The completeness and enforceability of contracts depends on legal institutions in other ways as well. The ability of a lender to enforce a debt contract against a borrower may be greatly influenced by whether legal institutions include bankruptcy or other forms of limited liability that protect some of the borrower's assets from being taken by the lender, or, at the other end of the spectrum, imprisonment of delinquent debtors on behalf of creditors.

CHECKPOINT 10.1 The five reasons

a. Which of these five reasons why contracts are incomplete are involved in the examples given at the beginning of this section? Hint: more than one reason is typically involved.

b. Come up with your own example of an incomplete contract and explain why one or more of the five reasons apply to it.

10.3 PRINCIPALS AND AGENTS: HIDDEN ACTIONS AND HIDDEN ATTRIBUTES

A **principal–agent relationship** (also called an *agency problem*) arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent. The employer, for example would like the employee to work harder; the employee would like to go home a little less exhausted at the end of the day.
- *Incomplete contract*: the agent's actions or attributes are not known to the principal (or, if known, they are not verifiable) and so cannot be subject to enforceable contract. How hard the worker works cannot be specified in an enforceable contract.

EXAMPLE Many of the economic interactions between members of different economic classes—for example, employers and workers, landlords and tenant farmers—are principal–agent relationships.

Both conditions are necessary. If there were not a conflict of interest, then the agent would simply do what the principal desired (both would desire the same thing) without an enforceable contract. It would be as if the principal himself carried out the necessary action.

If a complete contract covered all of the agent's actions that mattered to the principal, then a conflict of interest would not make the relationship special: "purchasing" the agent's action would be no different from the principal buying some quantity of wheat or electricity. The models of exchange with complete contracts–used earlier in the book–would be perfectly adequate.

REMINDER In Chapters 8 and 9, we explained how buyers' willingness to pay and sellers' willingness to sell result in consumer surplus and economic profits—the mutual benefits from exchange. These models assumed that contracts were complete, such as for the purchase of wheat, or other commodities for which the quality was readily verifiable. We can use the same concepts for cases in which contracts are incomplete.

We can classify principal–agent problems into two categories:

- *Hidden actions*: These are things that an agent *does* that the principal has some interest in, but does not know (or lacks verifiable information about), such as the effort of an employee or the business practices of a borrower. An example is that a person whose home is fully insured against fire may take less care to avoid fires. Insurance companies call

PRINCIPAL–AGENT RELATIONSHIP A principal–agent relationship (also called an agency problem) arises when two conditions hold: (a) conflict of interest: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent; and (b) Incomplete contract: the agent's actions or attributes are not known to the principal (or, if known, are not verifiable) and so cannot be subject to enforceable contract.

this behavior **moral hazard**, and that term is sometimes used to apply to any hidden action problem.

- *Hidden attributes*: There are characteristics of an agent—what the agent *is*—that the principal has some interest in, but does not know (or lacks verifiable information about) such as which drivers are reckless, or which patients are seriously ill.

EXAMPLE Patronage or patron–client relationships are common in politics and have also been studied in principal–agent models. The patron—perhaps a political leader—is a principal who interacts with a subordinate, the client, who is the agent. The client provides difficult-to-monitor services (loyalty in political conflicts) in return for well-defined compensation (patronage jobs, money transfers, or access to public services).

To understand how mutually beneficial exchanges take place without complete contracts and how the benefits are divided among the parties, we need to use the concepts we have already developed—best response, first mover, fallback option, and non-clearing markets—to construct a new set of analytical tools called principal–agent models. Especially important among the methods you have already learned is the theory of repeated games introduced in Chapter 5. Most jobs are not one-shot interactions: they go on year after year. So the game between the employer and the worker is *repeated*, and what each party does in one period depends on what they expect to happen as a result later periods. Remember that in a repeated game what is called the stage game—like a simple Prisoners' Dilemma—is played more than once with the same players.

Models based on repeated games of this type are used to study transactions between employers and workers, lenders and borrowers, and a wide set of other exchanges as shown in Table 10.1. They range from the landlords and sharecroppers that we mentioned in Chapter 2 to a fundamental problem of democracy: citizens trying to control their governments.

In the right-hand column we list some of the strategies followed by principals to get agents to act in their interest. In the first row—about employers as principals and workers as agents—the term *contingent renewal* means that because employment is repeated game, the principal has the option to not *renew* the relationship (fire the worker) and whether this happens is *contingent* on (depends on) whatever information the employer gets on the worker's job performance.

CHECKPOINT 10.2 Asymmetries and hiddenness

a. Think back to previous chapters. What were the different kinds of asymmetries we saw between players in different games or social interactions? Identify each player's strategy sets, the extent of the information asymmetry between them, and their respective payoffs.

MORAL HAZARD If there is a conflict of interest between a principal and an agent over the agent taking some action that cannot be ensured by a complete contract, then the principal faces a moral hazard problem. Also referred to as the 'hidden actions' problem.

b. What is an example of a hidden **attribute** (characteristic of a person) that a consumer, worker, or borrower might want to keep hidden? Explain why.

c. What is an example of a hidden **action** (choice) that a consumer, worker, or borrower might want to keep hidden? Explain why.

Table 10.1 Principals and agents: hidden actions and attributes These are a few of the applications of principal–agent models. In this chapter we illustrate the model by the problem of difficult–to–measure quality of goods. In Chapters 11 and 12 we look at labor services and private debt. Political scientists have used similar models to study how citizens (the principals) can hold accountable public officials (their agents). This is shown in the public policy line, which is somewhat unusual because the citizen is the principal (there are many principals) and the government official is the agent. Contingent renewal here means that the official may not be renewed (not re-elected) if the citizens are not pleased with her performance.

Good or service	Principal/agent	Non-contractual action or attribute	Hidden action or attribute	Examples of strategies by principals to get agents to act in ways favorable to them
labor services	employer/worker	labor effort, care	action	contingent renewal
managerial services	owner/manager	effort, maximizing owners' profits	action	profit sharing, contingent renewal
private debt	lender/borrower	level of risk taken	action	collateral, shared control
sovereign debt (between nations)	lending government/borrower government	probability of default	action	trade sanctions, other intervention
goods	buyer/seller	product quality	action	contingent renewal
public policy	citizen/government official	policy choice and implementation	action	contingent renewal, referendum
residential tenancy	landlord/tenant	care of residence, local amenities	action	security deposit, contingent renewal
agricultural tenancy	landlord/tenant	labor effort and quality of land	action	shared residual claimancy
equipment rental	owner/renter	care of the equipment	action	deposit, ownership share in equipment
car insurance	insurer/insured	driving habits and competence	action and attribute	higher prices for younger drivers or those with previous accidents
second-hand cars	buyer/seller	quality of car	attribute	purchases only from sellers with good reputations
health insurance	insurer/insured	preexisting health of insured	attribute	required medical exam as a condition

Figure 10.1 George Akerlof (1940–) an American economist, has pioneered the introduction of insights from anthropology, psychology, and the other social sciences into economics as part of his studies of "gift exchange," "identity economics," and other topics. In 1970, he wrote a paper, "The Market for Lemons," in which he developed the model we present here. Now one of the most highly cited in economics and the main reason for Akerlof's Nobel Prize, before being published in the *Quarterly Journal of Economics*, the paper was rejected by both the *American Economic Review* and the *Review of Economic Studies* as "trivial" and by the *Journal of Political Economy* due to one referee's opinion that if Akerlof were correct, then much of economics would be wrong! Apparently the quality of an economics paper is at least as difficult to judge as is a used car.[3]

Photo by REUTERS/Alamy Stock Photo.

10.4 HIDDEN ATTRIBUTES AND ADVERSE SELECTION: THE LEMONS PROBLEM

A "lemon" (in American slang) is a used car you discover is defective after you buy it. A "peach" is a used car you discover works better and costs less to maintain than you expected after you buy it. The problem the existence of lemons poses to economic markets can be illustrated in a model of a used-car market where the principals are the prospective buyers, the agents are the sellers, and the hidden attribute (whether the car is a lemon) is known only by the seller.

Consider the following example (summarized in Table 10.2):

- Every day, ten owners of ten used cars consider selling.
- The cars differ in quality, which we measure by the monetary value of the car to its owner. Quality ranges from zero to 9,000 in equal steps: there is one worthless car, one worth 1,000, another worth 2,000, and so on. The average value of the cars is therefore 4,500.
- There are many prospective buyers, and each would happily buy a car for a price equal to its true value, but not more, which is their *willingness to pay* for a car.
- Sellers do not expect to receive the full value of their vehicle, but they are willing to sell if they can get even just a little more than half the true value, which is their *willingness to sell* for the car.
- The potential economic mutual gain—the sum of buyers' and sellers' surpluses—will be the difference between the willingness to pay and to sell, or half the price of the car.

Imagine that prospective buyers could ascertain the quality of each car, and approach each seller to bargain over the price. Then if sellers also knew each buyer's willingness to pay, by the end of the day all of the cars (except for the entirely worthless one) will have been sold at a price somewhere between their true value and half the true value. All mutually beneficial trades would have taken place.

But if potential buyers cannot ascertain the quality of any particular car that is for sale, the market will not work. Suppose that those who bought cars yesterday find out the true value of their purchase and post it on social media. Then today, the potential buyers will know the true value of the cars sold the previous day. They still do not know the true value of any of the cars for sale today. But they might reasonably adopt the rule that the most they are willing to pay for a car today will be the *average value* of the cars sold yesterday.

Table 10.2 The market for used cars and the choices of the buyers and sellers. Buyers are willing to pay up to the full value of the car. Sellers are willing to sell their cars if they can get more than half of the true value of the car. The surplus of a transaction—the mutual benefit from the exchange—is the willingness to pay minus the willingness to sell. The seller of the maximum value vehicle observes the estimated price and as this is below his willingness to sell, he will leave the market. The same is true for the second most valuable car owner, who will also leave, and so on as the days progress, sellers leave, and the market unravels.

Day	Number of vehicles	Maximum vehicle value ($)	Total value divided by number of vehicles	Estimated value	Choice of highest value vehicle owner	Choice of lowest value vehicle owner
1	10	9,000	$\frac{45,000}{10}$	4,500	leave	remain
2	9	8,000	$\frac{36,000}{9}$	4,000	leave	remain
3	8	7,000	$\frac{28,000}{8}$	3,500	leave	remain
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9	2	1,000	$\frac{1,000}{2}$	500	leave	remain
10	1	0	$\frac{0}{1}$	0	leave	leave

Now suppose that ten cars had been offered on the market on the first day. We use a proof by contradiction to show that, one by one, the highest-quality cars will drop out of the market, until there is no market in used cars. Consider the market on the second day:

- On the first day all the cars (as we assumed at the start) were put on the market and sold at their true value which is the highest price at which they possibly could have been sold.
- The average value of these cars was $4,500, so the most a buyer is willing to pay today for any car will be $4,500.
- At the beginning of the second day, each prospective seller expects a price of $4,500 at the most. Most of the sellers are happy: $4,500 is more than half the true value of their car.
- But one owner isn't pleased. The owner of the best-quality car ($9,000) would not sell unless the price exceeds half the value of his car: more than $4,500.
- On the second day the owner of the best car will not offer it for sale. No one with a car worth $9,000 will be willing to participate in the market on the second day.
- The rest of the cars will sell on the second day: their value averages $4,000.
- On the third day, buyers will know the average value of the cars sold on the second day, and will be willing to pay at most $4,000 for a car.

- The owner of the second day's highest-quality car (the one worth $8,000) will know this, and will not offer her car for sale on the third day.
- As a result, the average quality of cars sold on the third day will be $3,500. The owner of the third-best car will not put his car up for sale on the fourth day.
- And so it goes on, until, after ten days, only the owner of a lemon worth $1,000 and a totally worthless car will remain in the market.
- If cars of these two values sell on the tenth day, then, on the eleventh day, uninformed buyers will be willing to pay at most $500 for a car of any quality.
- Knowing this, the owner of the car worth $1,000 will decide she would rather keep her car than try to sell on the eleventh day.
- The only car on the market on the eleventh day will be worth nothing: the cars that remain on the market are lemons, because only the owner of a worthless car would be prepared to offer that car for sale.

EXAMPLE Here's an example of the problem of hidden attributes. If you really enjoy speed when you drive your car, you might consider buying more insurance. But if more risky drivers than safe drivers buy insurance, insurance providers have to cover the resulting higher costs to stay in business. Insurers will raise prices so that insurance will not be worth the cost for the safe drivers. This can result in a vicious circle of rising premiums and an increasingly risky pool of those willing to buy insurance. If the insurance company knew which drivers were risky and which ones were safe, they could design high-cost policies specifically for the risky drivers, while prudent drivers would be offered insurance at lower prices.

Economists call this process **adverse selection**, because the prevailing price selects which cars will be left in the market. The market of uninformed buyers *selects against quality*, and is *adverse* to the interests of potential buyers and sellers holding high-quality goods.

The lemons problem emphasizes the dependence of real-world markets on the nature of the information available to the exchanging parties and to the courts. Unless information is available to sustain the differentiation of products, contracts will converge to the lowest-common-denominator level for products. Lemons problems are particularly acute in markets for insurance and credit, where information is at a premium, but can occur in a wide variety of other situations.

Whether the information available is sufficient to allow markets to work will depend on the economic and legal institutions governing the exchange process. In the case of a commodity like #2 red winter wheat, the information required to sustain the market is provided as a kind of *public good* by the agency that sponsors the market and certifies the quality of the product.

There are many factors that could help to ensure that the lemons problem does not persist. Quality assurance bodies, like those for the wheat described above, are one such solution. Another is the introduction of legislation by third parties (states) to create compensation or liability rules such that those who purchase defective products in one-shot exchanges can receive full compensation or exchange their goods. This idea is demon-

ADVERSE SELECTION In a principal–agent relationship, adverse selection is a problem confronting a principal who lacks information on some relevant attribute of an agent such that the terms the principal offers may induce the agents with whom the principal could have most advantageously interacted to drop out. Also referred to as the 'hidden attributes' problem.

strated in the 'lemon laws' in most developed countries nationally, or across a variety of US states.

Another method by which parties can exchange to reduce the severity of the lemons problem is to repeat the interaction. When interactions are *repeated*, players can build up and sustain a reputation. Reputations are costly to build and maintain and if a player acts contrary to their reputation, then they may lose the investment they have built up in their reputation. Examples abound. In online markets, second-hand goods are sold by sellers with a star rating saying whether or not it is worthwhile purchasing from them and the rating information is symmetric. Car dealerships build up reputations for honest dealing and good-quality vehicles through word of mouth and customer satisfaction reports.

Finally, as we have seen in Chapter 2, people often wish to uphold social norms and to punish those who do not. Among these norms are honestly reporting the nature or quality of a good one is selling, providing no guarantees, but sometimes providing the social regulation of the contract that Durkheim mentioned as essential to buying and selling.

When we move from the problem of hidden attributes to hidden actions we encounter a new set of reasons why sustaining mutually beneficial trade may be difficult, even when mutual gains are technically possible.

CHECKPOINT 10.4 **Lemons problem and online transactions** The lemons problem explains why many potentially mutually beneficial used-good markets would not exist. But a great many of these markets do exist. Using what you have learned so far about incomplete contracts, explain what prevents used-sales markets (e.g. on eBay) from unraveling like the used-car market above.

10.5 APPLICATION: HEALTH INSURANCE

The lemons problem illustrates adverse selection due to hidden attributes. The problem is far more general than the used-car market. To see why, think about health insurance. Imagine hypothetically that you will be born into a population, but do not know whether you would be born with some serious health problem, or might contract such a problem later in life, or perhaps be entirely healthy until you die of old age.

Now imagine that before you were born, or even knew who your parents would be you were asked this question: Would you buy health insurance if the premium–that is, the cost of the insurance to you–which is the same for everyone, is just sufficient to pay for the medical services required across the entire population if everyone agreed to purchase it? This kind of hypothetical decision-making process is called making a decision behind a **veil of ignorance** because it is a thought experiment in

HISTORY The veil of ignorance is a thought experiment introduced in 1971 by the American philosopher John Rawls. It asks you to imagine making a choice—about an anti-discrimination law, immigration policy, insurance plan—while *ignorant* of your position—your gender, nationality, health status, for example—that will matter for how you fare under the decision you take. In Rawls's words, we should think about what a just society would be like as if "no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like."[4]

REMINDER The veil of ignorance is similar to the Impartial Spectator thought experiment that we used in Chapters 3, 4, and 5. Both ask you to think about what would be just by placing yourself in the position of others. Smith wrote: "the natural preference the every man has for his own happiness over that of other people, is what no impartial spectator can go along with."[5]

HISTORY One of the most controversial aspects of US President Barack Obama's healthcare legislation was the requirement that all citizens purchase insurance. Supreme Court Justice Antonin Scalia opposed this requirement as an unwarranted intrusion of the government into people's choices. Writing that if you can make people buy health insurance "you can make people buy broccoli."[6] The president's lawyer, Donald Verrilli replied that the market for broccoli was different from the market for health insurance.[7] Is Verrilli right? Are the two markets different in a way that might justify requiring everyone to buy health insurance while not requiring everyone to eat broccoli?

which you are invited to think about how you would act or what policies you would favor if you did not know the state of your actual health. Most people would be willing to purchase health insurance behind the veil of ignorance, because they would rather pay a premium representing the average costs of healthcare to the whole population rather than be individually responsible for paying for the treatment of a serious illness, which, even if it is very unlikely to strike them, would impose high costs that most families could not pay. The benefit of protecting oneself and one's family from a financial catastrophe (or the possibility that you can't afford healthcare when you need it) is worth the insurance premium on average.

But the thought experiment is unrealistic: we cannot sign people up for health insurance before they know how healthy they will be. The reason is that this would require signing them up for insurance before they are born. Though most people would buy fairly priced health insurance if they did not know about their future health status, the situation changes dramatically if they can choose whether to buy health insurance knowing something about their current and future health status.

Let's look at the situation from the standpoint of the insurance provider (the principal) paired with a prospective buyer of insurance (the agent):

- People are more likely to purchase insurance if they know that they are ill or likely to become ill. The average health of people buying insurance will be lower than the average health of the population.
- This information is *asymmetric*: The person buying the insurance knows more than the insurance company about how healthy they are.
- Insurance companies selling insurance to people who are sicker than average will be profitable only if they charge higher premiums than if everyone bought the same insurance.
- This will lead some—those who are reasonably certain that they are healthy and will remain so—to not purchase insurance.
- To remain in business, the insurance companies will have to charge even higher premiums. Eventually the vast majority of the people buying insurance will be those who know they already have or are likely to have a serious health problem.

This is a case of adverse selection again. The reasoning above shows why the hidden attributes of initial health status can result in an unraveling of the healthcare market. Why? Insurance companies make profits by insuring people who are healthy. Healthy people who want to buy insurance in case they fall ill in the future are priced out of the market, and will not buy insurance. In the extreme case, the health insurance premium will be so high that only people who know they are likely to become seriously ill will buy insurance (people with initial conditions of poor health).

In this case we have what is called a *missing market*. It is a market that could exist, but it would only exist if health information were symmetrical and verifiable (ignoring for the moment the problem of whether everyone would want to share their health data). Under those imaginary conditions, the market could provide benefits to both insurance company owners and people who wanted to insure themselves. Not having such a market is Pareto inefficient.

To address the problem of adverse selection due to asymmetric information and the resulting missing markets for health insurance, many countries have adopted policies of compulsory enrollment in private insurance programs or universal tax-financed coverage, such as the National Health Service in the UK, or similar services in Canada, France, and elsewhere.

When we move from the problem of hidden attributes to hidden actions we encounter a new set of reasons why sustaining mutually beneficial trade may be difficult even when mutual gains are technically feasible.

CHECKPOINT 10.5 **Applications other than health** Consider applications other than healthcare such as car insurance and home insurance.

a. What are the attributes people would want to keep hidden?

b. How does hiding these attributes affect the prices that insurers will charge?

c. Why might the market be missing or what might allow the market to be present?

10.6 HIDDEN ACTIONS AND MORAL HAZARDS: A CONTINGENT RENEWAL CONTRACT

Insurers, whether private or governmental, face problems other than hidden attributes. There is also the problem of hidden actions: buying the insurance policy may make the buyer more likely to take exactly the risks that have been insured against. A person who has purchased full coverage for his car against damage or theft may as a result take less care in driving or in securing his vehicle than someone who had not purchased the insurance.

Insurers typically place limits on the insurance. For example, insurance coverage may not apply or be more expensive if someone other than the insured person is driving, or if it is parked on a daily basis in a theft-prone location. These provisions can be written into an insurance contract and are enforceable. But, the insurer cannot enforce a contract about how fast you drive or whether you drive after having had a drink. These are the actions that are hidden from the insurer because of the asymmetric information: you know these facts, but the insurance company does not.

Here is a model of incomplete contracting in the case of hidden actions, inspired by Benetton, the global casual wear marketer. To understand the model, let's first explore some of the history of the Benetton company.

Benetton: Decentralized production

The Benetton family—three brothers and a sister born during the Great Depression and World War II—began with a small company selling sweaters to shops near the town of Treviso in northern Italy half a century ago. Benetton grew to become one of the world's largest designers, producers, and sellers of casual wear and other garments. A key to its early success was a highly decentralized system of production: the labor-intensive aspects of production—primarily sewing—were carried out by hundreds of small subcontractors, working to designs and schedules and with materials supplied by Benetton.[8] A few processes were done by Benetton itself—notably dying, performed at the last minute so as to keep the product in tune with fashion trends. In 2020 Benetton's subcontractors outnumbered their employees by a factor of 17 to 1.[9]

!REMINDER Be sure that you are now clear on the similarities and differences between *hidden actions* (sometimes called *moral hazard*) and *hidden attributes* (and the associated problem of *adverse selection*).

Importantly, *quality control* and marketing were centralized, performed by Benetton's own staff. Subcontractors who reliably produced goods of the specified quality benefited from permanent orders, quick payment by Benetton, and other benefits of their long-standing relationship with the company.

As a result of this decentralized subcontracting structure, Benetton saved on the costs of establishing its own production facility. The most valuable asset of the company was not garment-making factories but instead the trademark, the Benetton name itself (called an *intangible asset*), which, once the garments were acquired from the subcontractors, was of course attached to the product before sale. Many contemporary companies have similar structures, from manufacturing to design to developing code and online applications.

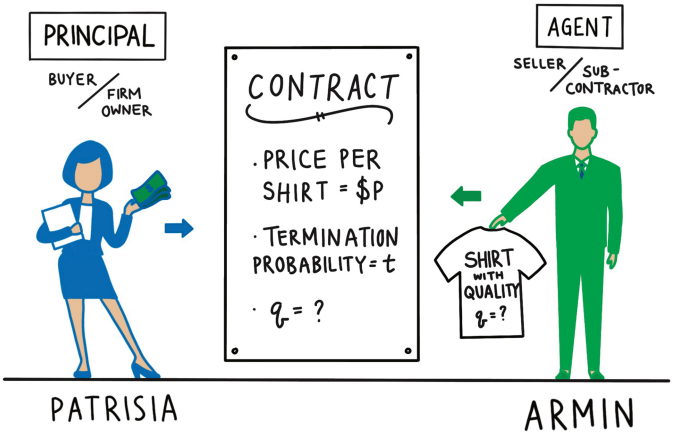
The Benetton model

To model these relationships we introduce a buyer (the *principal*, Patrisia) who purchases a good, say, a shirt, from a supplier (the *agent*, Armin, a subcontractor) for a price, p and then puts her trademark on the shirt and sells it to a consumer (who is not a player in this game). Their interaction is depicted in Figure 10.2.

The good—the shirt—has an aspect, its *quality*, q that is important to the principal, and costly or difficult for the agent to provide. Patrisia (who might be the Benetton company in this model) would like to pay a low price for a

Figure 10.2 Patrisia and Armin as principal and agent. Armin selling Patrisia a shirt with non-contractible quality (q) and Patrisia specifying a price (p) and termination schedule ($t(q)$).

Image credit: Anmei Zhi.



high-quality shirt, and Armin (one of the subcontracting shirt producers) would like to receive a high price for a shirt of lesser quality. So their interests *conflict*.

The principal–agent problem arises from this conflict of interest along with the fact that the quality of the shirt cannot be readily determined by Patrisia. She has the problem that shirts are sometimes rejected by customers, and Patrisia then has to refund the price when they return the good.

The challenge that the principal, Patrisia, faces is that she knows that Armin would prefer to produce lower-quality goods, but it is not cost-effective to determine the actual quality of each unit that she purchases. So she comes up with the following plan:

- *Repeated relationship*: Offer Armin an ongoing (or repeated) subcontracting relationship with her, so he can count on her buying his products year after year.
- *Termination probability*: But let him know that if he provides a shirt that is rejected by a consumer, she will terminate the relationship.
- *Enforcement rent*: pay him enough for the shirts he provides so that he does not want the relationship to be terminated, giving him a motive to provide quality shirts.

In other words she structures her interaction with Armin as a repeated game that she can terminate if she is not satisfied with the quality he provides.

Patrisia interacts with a large number of subcontractors like Armin. We simplify the problem by assuming that each period she purchases a given number of shirts (it could be one or one thousand) from Armin and each of the other subcontractors. So she is really choosing the price and the number of subcontractors from whom to buy.

! REMINDER In order to choose her strategy as first mover, the principal must consider what the agent will do in response. This is another example of backward induction in a sequential game, which you first encountered in Chapter 2.

Here is the structure of the game. The principal is first mover and seeks to maximize her profit by deciding on:

- *price*: the price to offer the agent (and the other subcontractors); and
- *quantity*: how many shirts to purchase.

The agent seeks to maximize the expected value of his utility over the duration of his doing business with the principal; he has just one decision:

- *Quality*: the quality of the good to supply.

The principal sets the price knowing the agent's best-response function, that is, the quality he will supply for every price she could offer. The agent chooses the quality to provide knowing the price the principal has offered and the probability that the transaction will be terminated for every level of quality he could provide.

To determine what the first mover (the principal) will do we need first to derive the best-response function of the second mover (the agent). The level of quality the agent supplies will depend on how valuable it is to the agent to continue the relationship with the principal. This is the rent that the agent will receive, and the reason why he will supply higher quality than he would provide otherwise.

CHECKPOINT 10.6 Hidden actions Think of markets (other than shirts) in which it is difficult for a buyer to determine the quality of a good or service prior to purchasing it, and in which information about the quality, even if known by the buyer, may not be verifiable.

10.7 THE EXPECTED VALUE OF THE TRANSACTION TO THE AGENT

Because this is a repeated game, how important it is to the agent that the relationship with the principal continue will depend on the following:

- the utility that the agent experiences in each period;

- the number of periods that the principal and agent will interact; and
- the utility the agent will experience if terminated by the principal, that is the agent's fallback option.

The agent's utility in a single period

The seller, Armin, prefers to receive a higher price p and to provide a lower level of quality q (which can take any value from 0 to 1, representing a range from a quality of 0 percent to 100 percent). Armin prefers to provide lower quality because it is costly—in terms of effort or care—for him to produce high-quality goods.

For concreteness we will express the general form of his utility function (below on the left) by a specific form (on the right below):

$$\text{Agent's utility} \qquad u(p,q) = p - \frac{\underline{u}}{(1-q)} \tag{10.1}$$

Equation 10.1 says the following:

- Because the marginal utility of the price, $u_p = 1 > 0$, Armin considers p to be a "good" something he would like more of because getting more money is valuable to Armin.
- Because the marginal utility of providing quality $u_q < 0$, Armin considers q to be a "bad," so he would rather not exert the effort required to provide $q > 0$.

The second term on the right hand side ($\frac{\underline{u}}{(1-q)}$), is Armin's *disutility of providing higher quality* shown in Figure 10.3. We can see three things:

- if he provides no quality (he just hands over a product with the lowest possible quality, i.e. $q = 0$) his disutility is $\underline{u}$;
- the higher the quality he provides, the more disutility he experiences: the curve is upward sloping; and
- the marginal disutility of quality is increasing: the curve is steeper at point **j** than at point **k**.

From Equation 10.1 you can also see that he will never choose to produce a perfect good (100% quality or $q = 1$) because at $q = 1$ his disutility is infinite or undefined.

There are two confusions to avoid here, one about slopes and the other about signs. First be clear about the difference between:

- the *height* of the curve which is the *disutility* of providing the level of quality indicated on the x-axis quality; and

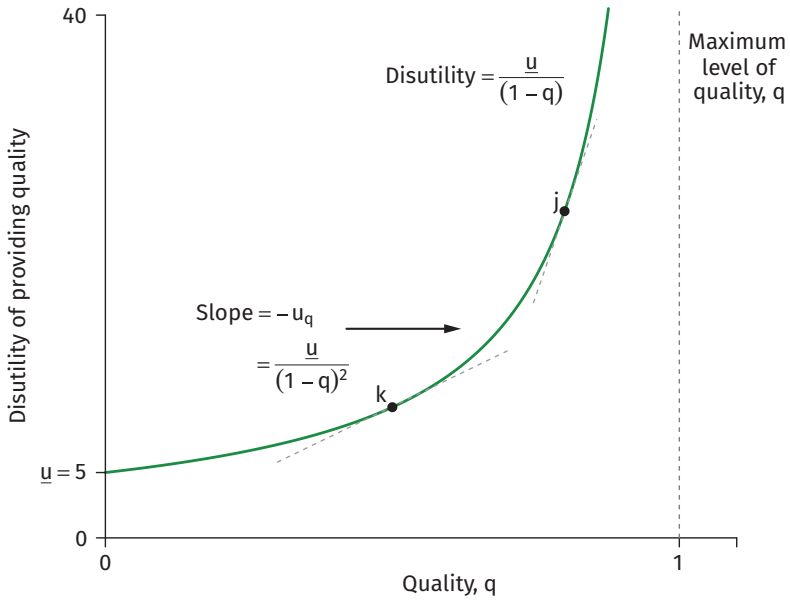
M-CHECK The marginal utility of providing quality is the derivative of u with respect to q, denoted u_q. It is negative because the second term in the utility function is preceded by a minus sign. When we refer to the marginal *dis*utility of quality—the slope of the curve in Figure 10.3—we mean the derivative of $\frac{\underline{u}}{1-q}$ (itself, without the minus sign) with respect to q which is $-u_q$, so it is positive. Given the agent's utility function and using the notation that $u_q = \frac{\partial u}{\partial q}$, we have

$$-u_q = \frac{\underline{u}}{(1-q)^2}$$

! REMINDER Equation 10.1 is another quasi-linear utility function of the type introduced also in Chapters 3, 4, and 7.

M-CHECK You can confirm that his disutility is infinite when $q = 1$ by substituting $q = 1$ into $\frac{\underline{u}}{1-q}$ and confirming that it would result in division by zero.

Figure 10.3 Armin's disutility of providing quality. The disutility of providing quality increases as the agent provides more quality. The slope of the tangency lines is the marginal disutility of effort at the corresponding level of quality provided. Comparing the slopes at points **j** and **k** shows that a small increase in quality will impose a greater disutility the higher is the quality he provides. In the figure Armin's baseline disutility of providing quality is set to $\underline{u} = 5$.



- the *slope* of the curve, which is the *marginal disutility* of quality, namely the effect on his disutility of providing a bit more quality when he is already providing the level indicated on the x-axis.

The second is the distinction between:

- the marginal utility of providing quality, u_q which is negative (providing quality reduces Armin's utility); and
- the marginal *di*sutility of quality which is just the same thing with a negative sign, $-u_q$, and so is positive.

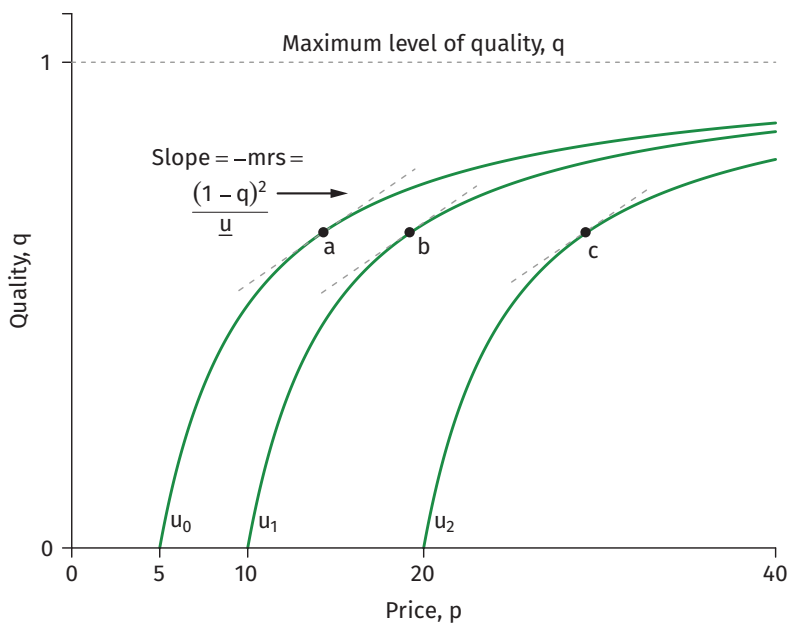
We can construct indifference curves for Armin based on his utility function: this is another example in which the choices include both a good (receiving the price) and a bad (providing quality). As a result, for any quality q on the vertical axis in Figure 10.4 Armin would prefer to obtain a higher price, p, which would place him on a higher indifference curve; and for any price, p, on the horizontal axis Armin would prefer to provide a lesser quality. The negative of the slope of the indifference curve is Armin's

marginal rate of substitution between more pay (the price he receives) and more quality. We derive this in M-Note 10.1.

We already know from Figure 10.3 and Equation 10.1 that increasing q reduces Armin's utility (increases his *dis*utility) and does so more and more as q approaches 1. This is the reason why in Figure 10.4 Armin's indifference curves become almost flat when Armin provides high levels of quality: Armin experiences an extremely high marginal disutility of providing any additional quality. At *already high* levels of quality providing additional quality (moving up in the figure) can be compensated only by a very large increase in price (moving to the right), as shown by the flattening of his indifference curves as q gets larger.

! REMINDER As in Chapter 3, when we are considering both a good and a bad, the indifference curves are upward-sloping. This is because having more of the bad can be compensated by having more of the good. In this chapter, though, the "bad", quality (q), is on the vertical axis and the "good" (price, p) is on the horizontal axis.

Figure 10.4 Armin's indifference map for "good" p and "bad" q. Armin's indifference map for "good" p and "bad" q for three values of u and $\underline{u} = 5$. With this quasi-linear utility function (Equation 10.1) his indifference curves are *horizontal* displacements of each other: their slopes depend only on the level of quality, not on the price. For instance, you can see that the slopes of the indifference curves at points **a**, **b**, and **c** in the graph are the same. These are points with the same level of quality but different prices. For the indifference curves shown, $u_2 > u_1 > u_0$.



M-NOTE 10.1 **The marginal rate of substitution for per-period utility**

Recall that the general form of Armin's utility function is:

$$u = u(p,q)$$

We want to find the changes in q and p that are consistent with no changes in Armin's utility—that is, staying on the same indifference curve. To do this, we totally differentiate Armin's utility function with respect to dp and dq and set the result equal to zero:

$$du = dq \cdot u_q + dp \cdot u_p = 0 \tag{10.2}$$

In Equation 10.2 we are using the notation that $u_q = \frac{\partial u}{\partial q}$ and $u_p = \frac{\partial u}{\partial p}$. Equation 10.2 requires that for any two points on an indifference curve, the utility difference associated with the difference in price $(dp \cdot u_q)$ is exactly compensated by the (opposite signed) utility difference associated with the difference in quality $(dq \cdot u_q)$, so that taking account of both effects the difference in utility between the two points is zero. We can rearrange Equation 10.2 to find $\frac{dq}{dp}$, which is the slope of the indifference curve:

$$\text{Slope of indifference curve:} \quad \frac{dq}{dp} = -\frac{u_p}{u_q}$$

Because the marginal utility of the price (p) is one, we have:

$$\begin{aligned} \text{Slope of Armin's indifference curve} &= -mrs(p,q) \\ &= -\frac{u_p}{u_q} \\ &= -\frac{1}{u_q} \end{aligned}$$

M-NOTE 10.2 **The marginal disutility of quality and the mrs with a specific utility function**

To find the marginal utility of providing quality, u_q, we differentiate Armin's utility function, Equation 10.1, with respect to q to find u_q:

$$\begin{aligned} u(p,q) &= p - \frac{\underline{u}}{1-q} \\ u_q &= (-1)(-1)\left(-\frac{\underline{u}}{(1-q)^2}\right) \\ &= -\frac{\underline{u}}{(1-q)^2} \end{aligned} \tag{10.3}$$

Using Equation 10.3, we can now find the marginal rate of substitution given for the general case by Equation 10.3:

$$\begin{aligned} mrs(p,q) &= \frac{u_p}{u_q} = \frac{1}{u_q} \\ &= -\frac{1}{\left(\frac{\underline{u}}{(1-q)^2}\right)} \\ &= -\frac{(1-q)^2}{\underline{u}} \end{aligned}$$

continued

The fact that p does *not* appear in the expression for the $mrs(p,q)$ means that the slope of the indifference curves depends *only* on the level of quality, not on the price as can be seen in Figure 10.4. This is because we have used a quasi-linear utility function that is linear in the price.

CHECKPOINT 10.7 Goods, bads, and axes

a. Consider Figure 10.3, what would happen to the curve if $\underline{u}$ increases or decreases (recall that this means that the disutility from working at all on the contract is changing).

b. Consider Figure 10.4 again, how would increasing (or decreasing) $\underline{u}$ change the indifference curves in $u(p,q)$? Why?

c. Substitute the values $q = 0.4$ and $q = 0.8$ into Armin's marginal rate of substitution, $mrs(p,q)$, given in M-Note 10.1. Assume $\underline{u} = 5$. What are the values for the mrs? How do you interpret them? (Be clear about what Armin is getting more of and what he is "willing to pay".)

The expected value of the transaction and the enforcement rent

To make her plan work, Patrisia has to do two things:

- get some information on the quality Armin provides as the basis for terminating the contract if necessary; and
- offer him a high enough price so that he receives a rent, and therefore he will prefer continuing doing business with Patrisia rather than being terminated.

She uses consumer complaints about quality as the basis for her termination decision and as a result she will terminate the relationship with Armin with probability $t = 1 - q$. This means that:

- If Armin provides no quality at all ($q = 0$), then there will surely be a complaint, so he will surely be terminated ($t = 1$).
- If Armin were to provide $q = 1$, then no consumer would ever complain, so he would be certain that Patrisia would continue buying from him ($t = 0$).
- If Armin provides $q = 0.5$, whether he gets terminated at the end of each period is a coin toss: 50 percent chance of keeping the contract and 50 percent chance of losing the contract.

If he provides quality of q, then at the end of the first period he will lose his contract with probability $1 - q$. If he is lucky and does not lose the contract, then the first period is just repeated—she offers the same price and so he offers the same quality. As a result, the next period will be his last with the same probability $(1 - q)$ and so on until his luck runs out and he is terminated.

How long will his contract with Patrisia last? Armin does not know for sure: he has to think about probabilities. To do this, imagine a game of

flipping a coin, in which you flip it once and if it came up heads you flip it again, and continue doing this until it comes up tails, at which point the game ends. If you did this many times, the game would sometimes end right after the first flip (because it came up tails); sometimes it would continue for many flips.

But on average how many flips would you expect to make, including the first one? The answer (which we show in the Mathematics Appendix) is two periods, which is just one divided by one-half, the probability that the game will end after each flip. So the expected number of periods that Armin's contract will last, $T(q)$ will be one divided by the probability of termination at the end of each period, or $t(q) = 1/t(q)$, where $t(q) = 1 - q$.

The value he gets from doing business with Patrisia is the utility he gets in a single period times the number of periods he expects to do business with her.

! REMINDER The term $v(p,q)$ is an expected value because it depends on a contingency that is not known, namely the duration of the contract in number of periods. For simplicity we refer to iso-value curves rather than iso-expected-value curves. When we use the term "value of the contract," we always mean expected value.

- The utility per period is $u(p,q)$.
- The expected lifetime (number of periods) of the contract is $T(q)$, which is equal to 1 divided by the probability that he gets terminated in a given period, or $T(q) = 1/t(q) = \frac{1}{1-q}$.
- So:

$$\text{Expected value} = \text{Utility per period} \times \text{Expected lifetime of contract}$$

$$\begin{aligned} v(p,q) &= u(p,q) \cdot T(q) \\ &= u(p,q) \cdot \frac{1}{1-q} = u(p,q) \cdot \frac{1}{1-q} \end{aligned} \tag{10.4}$$

This is how much he values being Patrisia's sub-contractor. To understand how motivated he will be by the threat of her ending the contract we need to know about his other options. How much he would like to keep his job is the difference between how much it is worth to him, v, and what he would be able to get if his relationship with Patrisia were to be terminated.

This next-best opportunity is his fallback option, which is finding another buyer and attempting to sell him the (presumably low-quality) good he has made. Later we will discuss his fallback option but for now we assume that his utility is zero if Patrisia terminates their relationship.

Because his fallback option is zero, it follows that v itself is Armin's **enforcement rent**, the loss of which he would like to prevent by supplying higher quality than he would otherwise do. If $v > 0$, then the following will hold:

ENFORCEMENT RENT In a principal–agent relationship an enforcement rent is the excess of the value of the transaction to the agent over the agent's fallback. The fear of losing the enforcement rent induces the agent to act in the principal's interest.

- It is a *rent* because it is how much his current situation is preferable to his next-best alternative.
- It is an *enforcement rent* because the rent motivates him to provide more quality than he otherwise would (he doesn't want to lose the contract).

In Figure 10.5 we show Armin's iso-value curves. Each curve is made up of all of the combinations of p and q that give the same values of v in Equation 10.3 with $v_4 > v_3 > v_2 > v_1 > v_0 = 0$. The iso-value curves differ from the single-period utility indifference curves in Figure 10.4: they are not uniformly upward-sloping. The upward-sloping portions of the iso-value curves are easy to understand. In Figure 10.5 Armin is indifferent between point **f**—providing high quality and getting paid a high price—and point **e**—providing lesser quality at a lesser price.

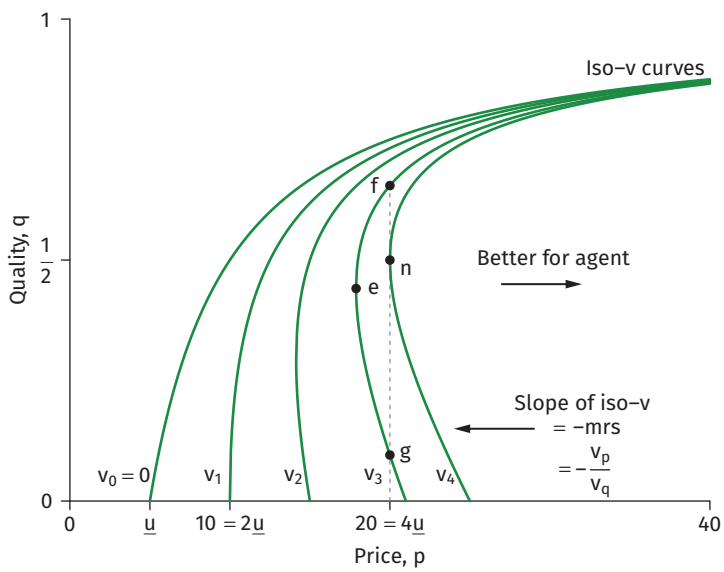
But why is Armin indifferent between point **e** and point **g**, where he is providing less quality and getting a higher price, which would seem to be a better deal than point **e**?

M-CHECK Using v_p and v_q respectively for the partial derivative of v with respect to p and q can derive the slope of the iso-value curve by the same method that we used in M_Note 10.1 to derive the slope of the per period utility indifference curve. It is $-mrs(p,q) = -\frac{v_p}{v_q}$.

Figure 10.5 The agent's iso-value curves. Armin's map of iso-value curves in the "good" p and "bad" q for five values of v using Equation 10.4 and $\underline{u} = 5$. Recalling that the marginal rate of substitution is the negative of the slope of the indifference curve,

$$\text{Slope}: \ \frac{\Delta q}{\Delta p} = -\frac{v_p}{v_q} = -mrs(p,q)$$

The iso-value curve labeled $v_0 = 0$ is the agent's participation constraint. The x-axis intercept of the participation constraint is at $p = 5$. The reason is that with $q = 0$ (meaning a point on the x-axis) and $p = 5$, the agent's per-period utility is zero (you can see this from Equation 10.4), so his value must also be zero, which is his fallback position.



The answer is that when Armin takes account of the likelihood of being terminated, providing more quality is not necessarily a "bad." The reason is that providing more quality will *increase* his chance of keeping his contract with Patrisia which (if $v > 0$) he values. At some low levels of quality, providing a little more quality will increase v because it will prolong the duration for which he receives the per-period utility. The iso-value function has a different shape from the single-period utility indifference curves in Figure 10.4 because the game is repeated, and Armin has an interest in continuing his relationship with Patrisia.

CHECKPOINT 10.8 Armin's iso-value curve

a. Use Armin's iso-value function to check that were he paid $2\underline{u}$ and provided $q = 0.5$, he would have $v = 0$.

b. What would happen to Armin's iso-value curves if his disutility parameter ($\underline{u}$) were to increase?

M-CHECK If we have a function like $u(p,q)$ then $u(p,0)$ means the utility if the price is p and the quality level is zero.

! REMINDER The incentive compatibility constraint is the agent's best-response function. It gives the combinations of the principal's actions—her choice of a price, a wage, or an interest rate, for example—and the agent's response. It shows the agent's actions that are compatible with the incentives and constraints offered by the principal.

10.8 THE AGENT'S BEST RESPONSE: AN INCENTIVE COMPATIBILITY CONSTRAINT

For any particular price that Patrisia offers, there are three things that Armin could do:

- refuse the contract and have utility $u = 0$;
- accept the contract but deliver $q = 0$ in which case his contract would be terminated with certainty, so he would receive $u(p,0)$ for a single period, and the value of his transacting with Patrisia would be $v(p,0) = u(p,0)$; or
- accept the contract, deliver some level of quality $q > 0$, and receive $u(p,q)$ for an expected number of periods, $\frac{1}{t(q)}$, and receive $v(p,q)$.

Armin will choose quality (q) to maximize his value (v), taking account of the fact the higher q will reduce the probability of termination t (remember at the end of any period, $t = 1 - q$). Suppose hypothetically in Figure 10.6 that the price offered by Patrisia is $p = 20$. Then we can think of Armin's optimizing problem in the following way, as we did in the other best-response functions we have derived in previous chapters. Starting at some low level of quality (point **g**) he would see that as he provides more quality, he reaches higher iso-value functions (not shown) until he reaches point **n**, the tangency of the iso-value curve labeled v_4 with the vertical line indicating the hypothetical price. If he proceeds upward–offering more

quality—then he will be crossing ever *lower* iso-value curves, such as the one at point **f**.

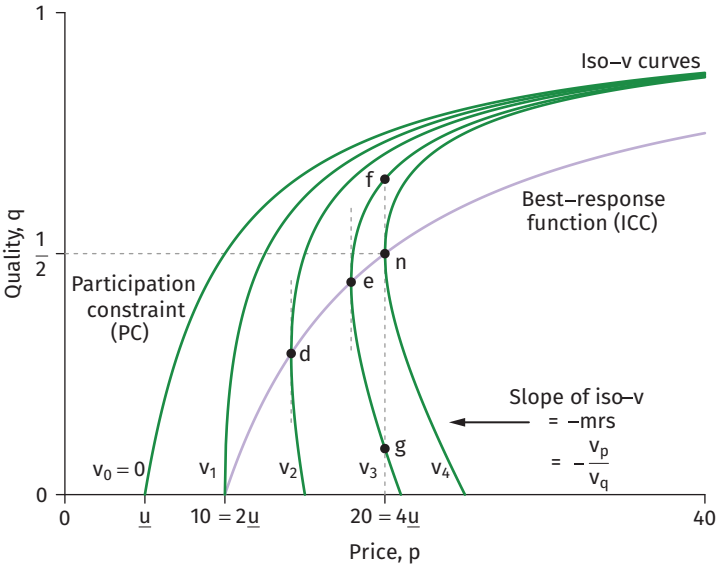
The quality the agent offers and price indicated by point **n** is one point on Armin's best-response function. Two others, constructed in the same way but at lower prices are points **e** and **d**.

Here is how we can derive an equation giving the agent's best response to the principal's price, that is, the best-response function shown as the purple line in the Figure 10.6. He will reject the zero quality option and want to supply more quality as long as the disutility of doing that—the marginal cost of quality—is less than the marginal benefit that he derives from providing additional quality. That is he will want to compare two *negative* quantities:

- *Marginal cost*: u_q, namely the *reduction* in his utility associated with providing more quality (the marginal utility of effort).

M-CHECK Equation 10.5 is the agent's best-response function for selecting q in response to the principal's choice of p. Below we use a specific utility function to derive a simple best-response function.

Figure 10.6 The agent's best-response function. The agent's best-response function is made up of all the points on an iso-value curve where the slope of the curve is vertical, as shown by three iso-value curves and corresponding points (**d**, **e**, and **n**) on the best-response function. Notice that where the best-response function intersects the horizontal axis the iso-value function is vertical. This is why at the price of $2\underline{u} = 10$ the best response of the agent is to provide $q = 0$. The best-response function is the incentive compatibility constraint or ICC. The segment of the horizontal axis between $\underline{u}$ and $2\underline{u}$ is the range of prices over which the price is high enough to get the agent to agree to a contract, but not high enough to get him to provide quality greater than zero.



M-CHECK This looks different from previous best-response functions, which showed the decision maker's action, amount of fishing time, for example, on the left-hand side of the equation, and on the right-hand side, the values determining the best response, including the others' fishing time. While it is more complicated, the meaning of Equation 10.5 is the same: the value of q satisfying this equation for each value of p is a best response.

! REMINDER Recall from Chapter 4, that when a first mover has price-setting power, but not take-it-or-leave-it power, the first mover maximizes their objective function (utility or profit) given the other person's best response or incentive compatibility constraint.

- *Marginal benefit*: $t_q \cdot v$, namely the *reduction* in likelihood of being terminated made possible by providing more quality times the effect of working harder on the probability of keeping the job.

So, as shown in M-Note 10.3 the level of quality that will maximize his value is that at which the marginal costs and marginal benefits are equal (remember both u_q and t_q are negative):

$$\begin{aligned} mc &= mb \\ u_q(q) &= t_q(q)v(p,q) \qquad (10.5) \\ \text{Marginal utility} &= \text{Reduction in termination probability} \\ &\quad \times \text{enforcement rent} \end{aligned}$$

The values of q satisfying this equation for each value of p is the best-response function.

To see what Equation 10.5 means, return to Armin's iso-value curves as shown in Figure 10.6. We show in M-Note 10.4 that if we use the specific utility function in Equation 10.1, then Armin's best-response function (for values of $p \geq 2\underline{u}$) is given by:

$$\text{Armin's best response:} \qquad q(p) = 1 - \frac{2\underline{u}}{p} \qquad (10.6)$$

This is the best-response function we derived in the figure, using $\underline{u} = 5$. You can see from Equation 10.6 why at prices lower than $p = 10 = 2\underline{u}$ Armin will not provide any quality at all. The agent would accept the contract and supply zero quality if the price were between $2\underline{u}$ and $\underline{u}$. This is because over that price range, the participation constraint is satisfied (so the agent "participates" in the contract) but prices in that range do not provide sufficient incentives for the agent to raise quality above zero. At a price below $\underline{u}$ the agent will not accept the contract.

CHECKPOINT 10.9 The best-response function

a. Explain why the best-response function is upward-sloping.

b. Why does it get flatter for higher levels of quality provided?

c. Imagine that there is some chance at the end of each period that Armin will have to move to another region (where he could no longer work as a subcontractor for Patrisia). How will this alter the best-response function? (Hint: How would this affect the expected duration of Armin's dealings with Patrisia. How would this affect $v(p,q)$?)

M-NOTE 10.3 The agent's best-response function for the general case

We know that Armin's (the agent's) value function is per-period utility times the number of periods that he expects the transaction to continue as given by the following equation:

$$\text{A's value} \qquad v(p,q) = u(p,q) \cdot \frac{1}{t(q)} \tag{10.7}$$

As a result, he will choose q—the only variable he controls—to find the quality level that maximizes his value given the price that is offered, considering the inverse relationship between the probability of termination and the quality of the good he provides. So, to find Armin's optimal choice, we differentiate his value function with respect to q and set the result equal to zero (imposing the first-order condition for a maximum). Notice, to find the first-order condition we need to use the quotient rule:

$$\begin{aligned} v_q &= \frac{u_q \cdot t - u \cdot t_q}{t^2} \\ \text{First-order condition:} \qquad 0 = v_q &= \frac{u_q \cdot t - u \cdot t_q}{t^2} \\ 0 &= u_q \cdot t - u \cdot t_q \\ u_q \cdot t &= u \cdot t_q \\ \text{Divide by } t\text{:} \qquad u_q &= t_q\left(\frac{u}{t}\right) \end{aligned} \tag{10.8}$$

But, recall that $v(p,q) = \frac{u(p,q)}{t(q)}$ (Equation 10.7), so we can rewrite the equation:

$$u_q = t_q \cdot v \tag{10.9}$$

This is the agent's best-response function, as can be seen if we write out Equation 10.9 more fully to make clear how u, t, and v depend on p and q.

$$u_q(p,q) = t_q(q) \cdot v(p,q) \tag{10.10}$$

Equation 10.9 requires that the marginal cost of providing more quality (remember, u_q is negative) is equal to the marginal benefit (the reduced chance of losing v). For any given value of p, the value of q that satisfies Equation 10.10 is the agent's best response $q(p)$.

Also recall from Figure 10.6 that the slope of Armin's iso-value curve is:

$$\text{Slope of iso-value}: \qquad \frac{dq}{dp} = -\frac{v_p}{v_q} = -mrs(p,q)$$

We know Armin reaches his maximum value when he provides the level of q such that $v_q = 0$. When $v_q = 0$, the denominator of the iso-value curve is zero and therefore the slope of the iso-value is undefined when Armin maximizes his utility, which means that the iso-value curve must be *vertical* at that point. This is why the point on each iso-value curve with a vertical slope makes up Armin's (the agent's) best-response function.

M-NOTE 10.4 The agent's best response with a specific utility function

Recall that the condition for the agent's choice of a level of quality to offer is:

$$\text{Marginal cost} = \text{Marginal benefit}$$
$$u_q = t_q v \tag{10.11}$$

Recall that $t = 1 - q$, so $t_q = -1$, and we can rewrite the condition given by Equation 10.11 using our specific utility function, termination probability, and value function:

$$-\frac{\underline{u}}{(1-q)^2} = -1\left(p - \frac{\underline{u}}{(1-q)}\right)\frac{1}{(1-q)} \tag{10.12}$$

To see how Equation 10.12 gives us Armin's best-response function, multiply both sides of Equation 10.12 by $-(1-q)$:

$$\frac{\underline{u}}{1-q} = p - \frac{\underline{u}}{1-q}$$
$$\frac{2\underline{u}}{1-q} = p$$
$$\text{Divide through by } 2\underline{u}: \quad \frac{1}{1-q} = \frac{p}{2\underline{u}}$$
$$\text{Raise both sides to the power } -1: \quad 1-q = \frac{2\underline{u}}{p}$$
$$\text{Isolate } q: \quad q = 1 - \frac{2\underline{u}}{p}$$

Thus we have Armin's best-response function:

$$q(p) = 1 - \frac{2\underline{u}}{p} \tag{10.13}$$

Since p is in the denominator of a negative term, you can see that an increase in the price offered by the buyer will make Armin provide a greater quality of the good.

! REMINDER In *minimizing* p/q the principal is seeking the *lowest cost* way of acquiring the goods she is to sell, just as did the cost-minimizing firms choosing technologies and input mixes to minimize their costs of production in Chapter 8. The only difference is that here the principal is *buying* the goods she will sell rather than *producing* them. In Chapter 11 we will study a similar problem in which the principal—an employer—is seeking to minimize the costs of production by finding the least cost way to motivate the workers she hires to provide the labor effort on which production depends.

10.9 THE PRINCIPAL'S COST MINIMIZATION AND THE NASH EQUILIBRIUM

The principal makes two decisions: how many units to purchase (which is equivalent to how many subcontractor–agents to engage) and how much to pay them for each unit. We have set up the problem so that we can focus on the second–the principal–agent problem. We have already analyzed models to address the first question of how many shirts she should sell to consumers, in Chapters 8 and 9.

To study the price-setting process in the principal–agent relationship, we proceed in three steps:

- Patrisia, the principal, knowing Armin's best response $q(p)$, determines the price p^N that will minimize the cost of acquiring quality (p/q) (we use the N superscript because this will be the Nash equilibrium price).

- Armin, the agent, best responds to the price offer by choosing the Nash equilibrium quality, q^N, the quality level that maximizes his value given the price (using his best-response function).
- When p^N is offered and q^N is the response, then the expected number of periods that Armin's relationship with Patrisia will last is $T^N = \frac{1}{t(q^N)} = \frac{1}{1-q^N}$.

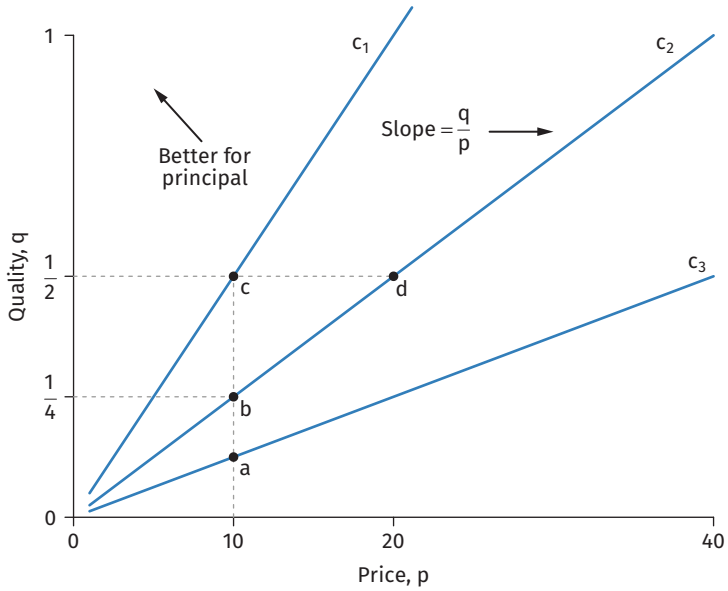
The principal minimizes costs

To show how the principal will set the price p in order to *minimize* the cost of acquiring quality, we show a variety of different potential ratios of price to quality to price (p/q) in Figure 10.7 as *isocost rays*. Along a given isocost ray, Patrisia has the same cost of quality, as can be seen by comparing points **b** and **d**. This is why they are called isocost ("equal cost") rays. Patrisia prefers isocost c_1 to c_2 to c_3 because $c_1 < c_2 < c_3$. Comparing points **a** on c_3, with **b**

! REMINDER A ray from the origin in (p, q) space is a line starting at $(0, 0)$.

Figure 10.7 Some of the principal's isocost curves. The principal's isocost curves for different ratios of quality for a given price, q/p. $c_3 > c_2 > c_1$. Comparing points **a**, **b**, and **c**, in which the same price is being offered, we can see that for the highest level of quality at **c** $(q = 0.5)$, the principal is on her preferred isocost line c_1. For lower quality, $q = 0.25$ at **b** and $q = 0.125$ at **a**, such as, the principal will be worse off on isocost lines c_2 and c_3 respectively. Along isocost line c_2, the principal is indifferent between points **b** and **d** because they result in identical cost of quality, but at different quality price bundles: at **b**, they are at the bundle $(q = 0.25, p = 10)$ with $q/p = \frac{0.25}{10} = 0.025$; at **d**, they are at the bundle $(q = 0.5, p = 20)$ with $q/p = \frac{0.5}{20} = 0.025$.

M-CHECK We say that the principal is indifferent between any of the points on the same isocost curve because the cost of quality is the same. We do not extend the isocost lines to zero, because $q/p = 0/0$ is undefined due to dividing any number by zero.



on c_2, and **c** on c_1, we can see that at the same price, the quality at point **c** is higher. So, steeper rays are better for the principal (corresponding to lower cost): the slope of any ray is q/p.

Think of these rays as similar to the indifference curves representing the objectives of a decision maker or the iso-profit curves representing the objectives of the owners of a firm. Here the isocost rays represent the cost-minimizing objectives of the principal in her relationship with the subcontractor. The slope of the isocost ray (that is q/p) is the negative of the marginal rate of substitution.

The Nash equilibrium price and quality

The isocost rays tell the principal which regions of the graph she prefers, roughly, which way is up. But she is constrained by having to provide the agent with incentives to implement the level of quality she desires. So the q she wants to motivate Armin to provide must lie on his best-response function. This equation—the incentive compatibility constraint—tells her what is feasible. In Figure 10.8 she would prefer to be, for example, at point **c** rather than at points **d** or **n**, because **c** is on a lower (steeper) isocost ray. But point **c** is infeasible because it is above Armin's best-response function.

The resulting cost-minimizing price will be somewhere on the incentive compatibility constraint, but where?

It is simpler to find the price that maximizes q/p than that which minimizes costs (p/q). The two are equivalent. So, putting her *objectives* together with what is *feasible*, Patrisia wants to find the p that will:

$$\text{maximize } \frac{q}{p}, \text{ where } q \text{ has to be such that } q = 1 - \frac{2\underline{u}}{p}$$

The price that accomplishes this is given by the following condition for cost minimization (as shown in M-Note 10.5):

$$\text{Condition for cost minimization:} \quad \frac{q}{p} = q_p \tag{10.14}$$

$$\text{slope of isocost} = \text{slope of agent's BRF}$$

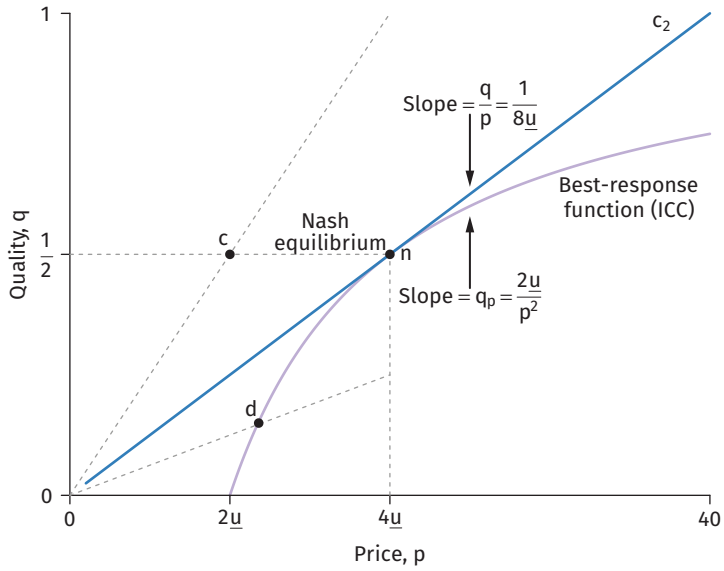
$$\text{Meaning} \quad mrs = mrt$$

This condition is shown in Figure 10.8: the price is chosen so that the slope of the agent's best-response function q_p is equal to the slope of the isocost ray with the highest ratio of quality to cost, q/p.

Equation 10.14, which requires equating the slopes of the best-response function and an isocost line, is another example of the $mrs = mrt$ rule.

- The slope of an isocost ray is the (negative of the) *marginal rate of substitution* between quality and price in the eyes of the principal, namely the increase in quality that would just compensate for a unit increase in the price.

Figure 10.8 The principal's cost-minimizing price and the agent's utility-maximizing quality provided. The Nash equilibrium, point **n** (p^N, q^N) is found at the tangency of the isocost line with slope q/p and the best-response function with slope $2\underline{u}/p^2$. Isocost line: $q = p/8\underline{u}$. Best-response function: $q = 1 - 2\underline{u}/p$.



- The slope of the best-response function is the (negative of the) *marginal rate of transformation* of paying a higher price into receiving higher quality.

The allocation (p^N, q^N) is a Nash equilibrium because it is a mutual best response:

- *Patrisia's choice*: given Armin's strategy choice—his best-response function—offering the price (p^N) is the best Patrisia can do (we know this because it is the result of the constrained maximum problem she just solved); and
- *Armin's choice*: given Patrisia's price offer, providing (q^N) is the best Armin can do (we know this because q^N is a point on his best-response function).

M-NOTE 10.5 The principal's cost-minimizing price

Here is the maximization problem: since p is the variable over which the principal has control, we *maximize* q/p by varying p. This is equivalent to *minimizing* costs (p/q).

$$\text{Maximize} \quad \frac{q(p)}{p}$$

continued

To find the price selected by this optimization problem, we differentiate this expression with respect to p:

$$\text{Quotient rule} \qquad \frac{d}{dp}\left(\frac{q(p)}{p}\right) = \frac{q_p \cdot p - q(1)}{p^2} \tag{10.15}$$

We then set Equation 12.16 equal to zero and solve for p:

$$\begin{aligned} \text{Condition for minimum cost:} \qquad \frac{q_p \cdot p - q}{p^2} &= 0 \\ \text{Multiply through by } p^2: \qquad q_p \cdot p - q &= 0 \\ q_p &= \tfrac{q}{p} \end{aligned} \tag{10.16}$$

Equation 10.16 is the first-order condition for the maximum quality per unit of price for the principal. The left-hand side is the slope of the agent's BRF and the right-hand side is the slope of the isocost ray.

With the specific utility function given in Equation 10.14, we have:

- *Slope of BRF*: The slope of the agent's BRF $q_p = \frac{2\underline{u}}{p^2}$.
- *Slope of an isocost ray, as above*: $\dfrac{q}{p}$

Therefore, we can rewrite Equation 10.16 for our illustrative utility function as follows:

$$mrt = q_p = \frac{2\underline{u}}{p^2} = \frac{q}{p} = mrs \tag{10.17}$$

Equations 10.16 and 10.17 can also be rewritten to show that the solution equates two effects of a small change in price (of one unit)

$$\text{\% increase in quality} = \frac{\Delta q}{q} = \frac{q_p}{q} = \frac{1}{p} = \frac{\Delta p}{p} = \text{\% increase in price} \tag{10.18}$$

Or what is the same thing: the marginal benefits of a price increase equal the marginal costs, both expressed in percentage terms.

M-NOTE 10.6 **Nash equilibrium: price, quality, and contract duration**

To find out what price Patrisia will pay to her subcontract or, Armin, we want to isolate p, so we multiply both sides of Equation 10.18 by p^2, to get:

$$\begin{aligned} \frac{q}{p} &= \frac{2\underline{u}}{p^2} \\ \text{Multiply by } p^2 \qquad pq &= 2\underline{u} \\ \text{Divide by } q \qquad \therefore p &= \frac{2\underline{u}}{q} \end{aligned} \tag{10.19}$$

We can then use Armin's best-response function to eliminate q, by substituting in $q = 1 - \frac{2\underline{u}}{p}$ into Equation 10.19:

$$\begin{aligned} p &= \frac{2\underline{u}}{q} \\ p &= \frac{2\underline{u}}{1 - \frac{2\underline{u}}{p}} \end{aligned}$$

continued

$$p\left(1-\frac{2\underline{u}}{p}\right) = 2\underline{u}$$

Dividing by p $$\left(1-\frac{2\underline{u}}{p}\right) = \frac{2\underline{u}}{p}$$

$$1 = \frac{4\underline{u}}{p}$$

Isolate p $$\therefore p^{\mathrm{N}} = 4\underline{u}$$

To find out how Armin responds, we substitute p^{N} into his best-response function:

$$q = 1-\frac{2\underline{u}}{p}$$

Substitute: $p^{\mathrm{N}} = 4\underline{u}$ $$q = 1-\frac{2\underline{u}}{4\underline{u}}$$

$$\therefore q^{\mathrm{N}} = 0.5$$

Finally we use q^{N} to determine the expected number of periods that Armin will sell to Patrisia and find that it is $T(q^{\mathrm{N}}) = \frac{1}{1-q^{\mathrm{N}}} = \frac{1}{1-0.5} = 2$

CHECKPOINT 10.10 **Armin's pride in the quality of his work** Suppose that Armin has come to have pride in the quality of the work he does, so the disutility of the effort it takes to produce high-quality goods that Armin experiences is cut in half, decreasing from $\underline{u} = 5$ to $\underline{u} = 2.5$. Redraw Figure 10.8, showing what happens to Armin's best-response function and the resulting Nash equilibrium ratio of quality to price, $q^{\mathrm{N}}/p^{\mathrm{N}}$.

Incomplete contracts, external effects, and an inefficient Nash equilibrium

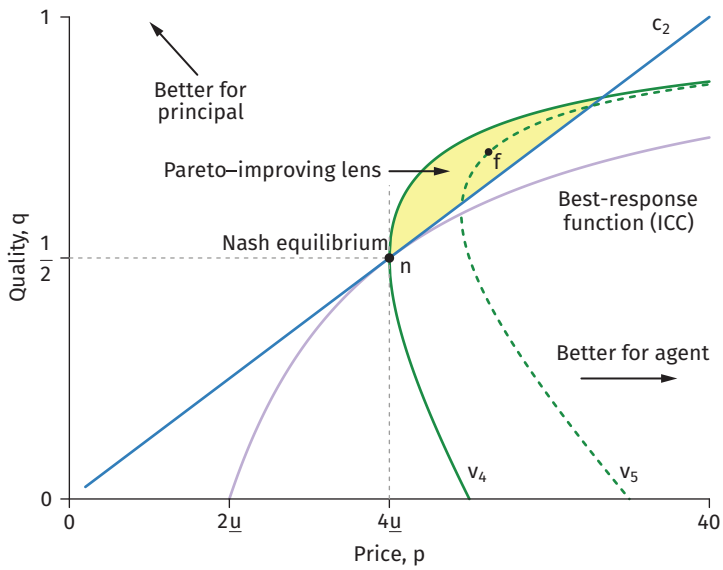
The Nash equilibrium allocation $(p^{\mathrm{N}}, q^{\mathrm{N}})$ is not Pareto efficient. We can see this in Figure 10.9. Remember that the principal prefers allocations that are higher and to the left (more quality for a lower price) and the agent prefers the allocations to the right. So any point that is both above the isocost line through the Nash equilibrium (better for the principal) and to the right of the iso-value curve v_4 (better for the agent), is a Pareto improvement over **n**.

In Figure 10.9 the shaded lens is the set of allocations that have these two properties and therefore are Pareto superior to the Nash equilibrium. You can see, for example, that point **f** is better for both Armin and Patrisia than is the Nash equilibrium, **n**. So **f** is Pareto superior to **n**. It follows that **n** is not Pareto efficient.

We know that points like **f** in the Pareto-improving lens must exist because the rule that the principal implemented to minimize her cost of quality–Equation 10.14–selected a price that equated the slope of her steepest possible isocost ray with the slope of the agent's best-response function. In other words she followed the $mrs = mrt$ rule.

But a Pareto-efficient outcome requires a different rule: equating the principal's marginal rate of substitution (mrs^{P}) with the agent's *marginal*

Figure 10.9 Pareto-improvements over Nash equilibrium with the contingent renewal incomplete contract. The shaded area shows feasible quality–price combinations such as point **f**, with lower costs (that lie on lower—meaning steeper—isocost rays) for the principal and also on higher iso-value curves for the agent, such as v_5. The Pareto-improving shaded lens extending upward and to the right of **n** *must* exist because the agent's iso-value curve at that point is vertical (it is on the agent's best-response function) and the principal's isocost ray at that point cannot be vertical. In the figure, the equation for the best-response function is $q = 1 - 2\underline{u}/p$ and the equation for the isocost line c_2 is $q = p/4\underline{u}$.



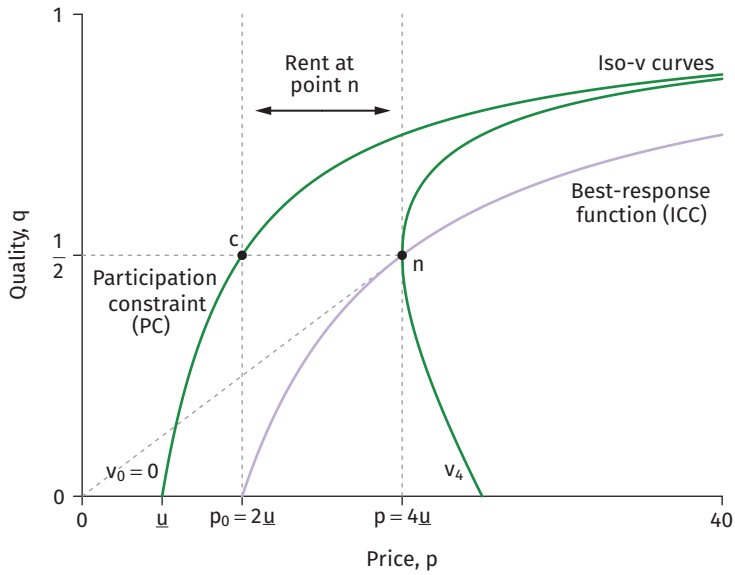
M-CHECK Remember, the slope of the iso-value function is $dq/dp = -\frac{v_p}{v_q}$. We know that $v_q = 0$ is the first-order condition for a maximum for Armin. So the point where the iso-value curve is vertical is a point on the agent's best-response function. Because **n** is on Armin's best response function we conclude that (as can be seen) the iso-value curve there is vertical.

rate of substitution (mrs^A). This would have required that the principal choose a price that would equate the slope of an isocost ray with the slope of the agent's iso-value curve.

You are already familiar with the reason why the Nash equilibrium is inefficient. Like the fishermen overexploiting their resource and diminishing each other's catch, the subcontracting agent, in deciding on his action—the level of quality to provide—is not taking account of the effect of his choice on another person, in this case the buyer-principal. The reason this occurs is that the contract is incomplete: it does not cover the quality he provides, which is then another example of an external effect.

We will see in section 10.11 that if the principal could just purchase quality (offering a particular p for some amount of q as if she were buying electricity or red winter wheat #2), then the result will be Pareto efficient. If the contract is complete, Patrisia minimizing the cost of quality by implementing the $mrs = mrt$ rule also unwittingly implements the $mrs^P = mrs^A$ rule.

Figure 10.10 The agent's rent at the Nash equilibrium. We can measure the rent that the agent receives at the Nash equilibrium by a horizontal comparison (that is, a price difference for a given quality) between the Nash equilibrium at point **n** and the complete contract outcome at point **c** on the agent's participation constraint.



The agent's enforcement rent

A key element in Patrisia's strategy, remember, and the reason why it works is that she pays him enough so that he wants to avoid being terminated. His enforcement rent is shown in Figure 10.10 as the distance between **n** and **c**. To see this, imagine Armin were at the Nash equilibrium, **n**, and we asked: hypothetically holding constant the level of quality he provides, how much less could he be paid and still be no worse off than at his fallback option, namely zero? To answer the question, we imagine moving to the left (lowering the price) from point **n**. We eventually hit the participation constraint $v_0 = 0$. At that point, **c** he receives no rent. So the difference in price between **n** and **c** namely, $2\underline{u} = v_4 - v_0$ is his per-period rent.

CHECKPOINT 10.11 Enforcement rent in another dimension In Figure 10.10 we measured the rent received by the agent in monetary units, that is the *price*. We could also measure his rent by answering the following question: At his Nash equilibrium price how much more quality could he provide without being worse off than in his next-best alternative? Identify this quantity in Figure 10.10 on your own sketched version of the figure.

! REMINDER You learned in Chapter 9 that the *short side* of a market is the side on which at a given price the number of desired transactions is the lesser of the two. The *long side* is the side on which the number of desired transactions is greater. The short side of a market may be buyers (as in the case of quality) or sellers (as in the case of financial institutions supplying credit, as we will see in Chapter 12).

! REMINDER A person or business is said to be credit constrained if: (a) they are excluded from borrowing entirely, or (b) they face limits on how much they can borrow, or (c) they pay extraordinarily high rates of interest when they do succeed in getting a loan.

10.10 SHORT-SIDE POWER IN PRINCIPAL–AGENT RELATIONSHIPS

The rent that the agent receives each period is the reason why terminating the exchange is an effective threat. If Patrisia cuts him off, Armin will not immediately find another principal to transact with on terms as good as his current transaction.

The key feature of these equilibrium contracts is that principals transact with agents who receive economic enforcement rents and *prefer* the current transaction to their next-best alternative. Because some agents receive enforcement rents in ongoing contracts, there must exist some other identical agents who are quantity-constrained, namely, the suppliers who fail to make a sale. If this were not the case, then immediately upon termination the agent could find another principal so the termination would not impose a cost on the agent. It is the fact that the market does not clear that makes the threat of termination effective.

In the Benetton model there are more agents looking for transactions than there are transactions being offered by principals. This is why, if Armin were terminated from his current transaction, it would take him time and a costly search process to find another. (For simplicity, we made the assumption that his next-best alternative is to get nothing, but as will see in section 10.13, this is not an essential feature of the model.)

Patrisia, the employer and other principals in similar principal–agent relationships, are on the short side of the market. Armin, the employee and other agents are on the long side; the long side includes agents who would like a transaction but cannot secure one. Armin fears being one of these: unable to sell his product.

We use the term short-side power to describe the power that Patrisia is able to exercise over Armin because she is on the short side of a market that *does not clear*. In Chapter 9 we studied out-of-equilibrium markets in which supply does *not* equal demand. In the Benetton model of principals and agents, when contracts are incomplete the market does not clear even when it *is* in equilibrium (this will be the case also in labor and credit markets).

Essential to Patrisia's ability to get Armin to do what she wanted is her ability to threaten to impose a major economic cost on him. This cost—called a sanction, or penalty—is part of our definition of **power**.

The definition can be applied to the Benetton principal–agent interaction. In equilibrium the following conditions hold:

POWER If, by imposing or threatening to impose sanctions on A, B can affect A's actions in ways that further B's interests, while A lacks this capacity with respect to B, then B has power over A.

- *Sanctions affect behavior*. Armin provides a shirt of higher quality than he would have in the absence of the Patrisia's threatened termination.
- *To the advantage of the person exercising power*. Patrisia benefits from this.
- *The relationship is asymmetrical*. Armin could not get Patrisia to act in a way beneficial to him by threatening her with termination.

To see why Armin did not have power similar to Patrisia, imagine that he threatened to end their relationship unless Patrisia raised the price above p^N. Patrisia would refuse to raise the price, knowing the following:

- because she is on the short side of the market, she could easily find another supplier (remember some of them cannot find a buyer); and
- it would not be in Armin's interest to carry out the threat because he would then have to find another buyer, and as a long-sider, this might not be possible.

! REMINDER A credible threat is one that will be in the interest of the actor to carry out, if the threat did not have its desired effect.

If Armin were to threaten to sanction Patrisia should she not raise the price (for example, to supply lower quality shirts), his threat would not be credible. Namely, it would not be in the interest of the actor to carry out, if the threat did not have its desired effect.

The definition of power emphasizes its use to advantage one person over the other. But the exercise of power is also essential to making mutually beneficial exchanges possible when contracts are incomplete.

Imagine, for example, that Patrisia were prevented from threatening to replace Armin with another supplier. Then he would provide only low-quality shirts, which she could not sell. As a result she would not purchase shirts from Armin. They would then *both* be worse off.

The case which we have analyzed—the exercise of power by the buyer (principal) over the producer (agent)—is just one example of power relationships that are sustainable as the Nash equilibrium of a system of voluntary competitive exchanges among private individuals. Other examples include the power that employers wield over employees, or lenders over borrowers, and the other hidden action examples in Table 10.1.

The principal–agent model of hidden actions shows the following:

- that the exercise of *power* is essential to the principal's strategy;
- that a *non-clearing market* is essential to the principal's ability to exercise power; and
- that the result of the principal's actions—along with the other actors in the market all acting independently—is a *Nash equilibrium* in which the market does not clear, thereby providing conditions for them to exercise power in their interactions with agents.

The last point is important and a bit counterintuitive. In a competitive environment, no principal, acting singly, can create markets that do not clear. But we have shown that, without intending to do so and without coordinating in any way, principals' choices–setting prices so that agents receive enforcement rents in order to minimize their costs–taken together create markets that do not clear.

In introductory economics, the exercise of power is commonly associated with limited competition in a market, for example what is sometimes called the "market power" of a price-setting monopolist. Markets that do not clear are typically attributed to government policies (such as rent control) or "market imperfections" such as "sticky prices" that do not adjust. The above three bullets show that where contracts are incomplete, the exercise of power and markets that do not clear occur in the absence of government policies and markets with unlimited competition–approximating the perfectly competitive model.

None of this would be possible if markets cleared. To see this, we imagine a case in which quality was subject to a complete contract in the Benetton model.

CHECKPOINT 10.12 Power and contract

a. Does Patrisia's contingent renewal contract with Armin satisfy the sufficient condition for Patrisia to be exercising power over Armin?

b. Is Patrisia's threat credible?

c. Think of a threat that a principal might make that would not be credible.

10.11 A COMPARISON WITH COMPLETE CONTRACTS

Suppose that Patrisia has discovered some magical device that can determine exactly the quality offered by Armin, and that this information is verifiable. So a complete contract is now possible. She can just name a price and the exact amount of quality that she would like in return. If he does not deliver the goods of the specified quality, he *does not get paid*.

Exchange with a complete contract

Armin will provide that amount, as long as she pays a price that makes Armin even just a bit better off than he would be without the contract. Of course she would not require him to deliver anything close to $q = 1$ because that would be so costly to him to achieve that she would have to pay a very high price in order for him to accept the contract.

The big difference that the complete contract makes is that Patrisia is no longer constrained by Armin's best-response function–the incentive compatibility constraint–but instead by Armin's participation constraint. This requires that Armin not be worse off agreeing to Patrisia's proposal than he would be were he to walk away, that is, to receive instead his fallback, which as before we assume to be zero.

She no longer has to provide him with a rent—a utility greater than zero, his fallback option—along with a threat to terminate the contract if the goods he has supplied are returned by disgruntled consumers. She can simply refuse to pay for the goods when he delivers them if they are of less quality than she has specified.

Her interaction with the agent is no longer repeated, it is a one-shot game. And it is no longer a principal–agent interaction: the conflict of interest over quality remains, but the second characteristic of a principal–agent relationship is now missing: quality is now something that can be enforced by contract.

! REMINDER In a principal–agent relationship, the participation constraint requires the principal to make an offer at least as valuable to the agent as what the agent would get—his fallback—were he to refuse the principal's offer. If the principal's offer is inferior to the agent's fallback, he will not participate in the transaction. Complete contracting in this context is the equivalent to Patrisia having take-it-or-leave-it power over Armin where she controls *both* p and q, whereas with incomplete contracting she only controlled p.

Nash equilibrium of the take-it-or-leave-it complete contracting game

Here is the new *complete contracting* game:

- Patrisia is first mover and she makes a take-it-or-leave-it offer to pay Armin p^C to purchase his product of quality q^C.
- Armin either *accepts* the offer and the exchange is carried out or *rejects*.
- In either case, this ends the game.

We use the C superscript to indicate the hypothetical complete contracting Nash equilibrium.

Because, due to the complete contract, she can just purchase quality, we can now interpret Armin's participation constraint as the minimum price at which he is *willing to sell* his products with the quality indicated by the *height* of the curve. So in Figure 10.11 he would be willing to sell his products with quality of $q = 0.5$ if the price were $2\underline{u}$. And the same goes for all of the price–quality combinations that make up the willingness to sell (participation constraint) curve.

Just as in the case where contracts are incomplete, her maximum will be where the ray from the origin with slope q/p is tangent to the constraint. But this is now the participation constraint (Armin's iso-value curve $v_0 = 0$) not the incentive compatibility constraint (Armin's best-response function). In Figure 10.11 (b) you can see that the outcome of the game is point **c**, with $q^C = 0.5$ and $p^C = 2\underline{u}$. This is the principal's offer that equates the slope of her isocost line with his iso-value curve, implementing the $mrs = mrt$ rule.

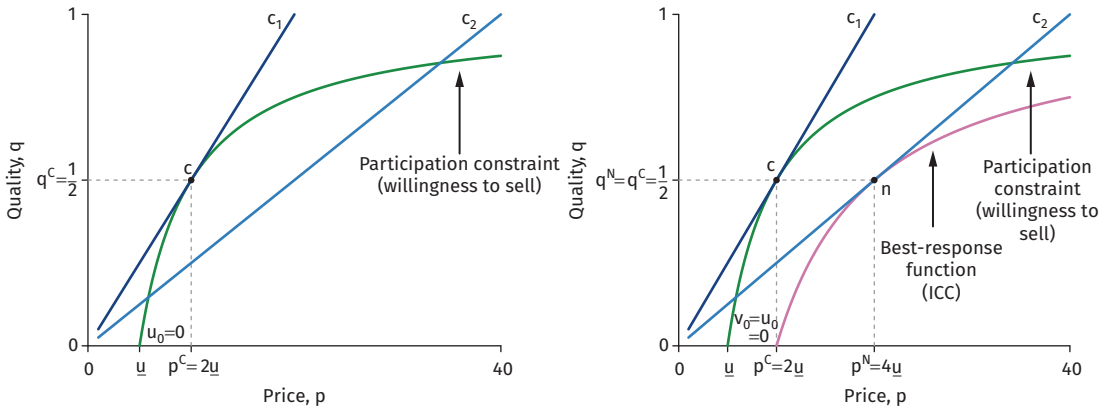
We can confirm that this is a Nash equilibrium:

- Given Armin's utility function and fallback option (described by his participation constraint) Patrisia is doing the best she can by *offering* the contract: $q^C = 0.5$ and $p^C = 2\underline{u}$.
- Given her offer and his fallback position, Armin is doing the best he can by *accepting* the contract.

Under the complete contract, Patrisia has paid half as much but received the same quality as when the contract was incomplete; so her miracle machine for detecting quality has cut her costs in half. Were Patrisia to offer the complete contracting price $p^C = 2\underline{u}$ under an incomplete contract, you

! REMINDER We showed in M-Note 10.1 that slope of the iso-utility indifference curve is u_p/u_q. When u_0 is the participation constraint on the principal's cost minimization this slope tells us how much greater quantity the principal can require if she grants a higher price. Like the slope of all constraints, its negative is a marginal rate of transformation. But this quantity is also the marginal rate of substitution giving the agent's relative valuation of the q and p.

Figure 10.11 A comparison of complete and incomplete contracts. In panel (a), the complete contracting game in which the principal either has take-it-or-leave-it power or simply price-setting power, the Nash equilibrium is the allocation where the lowest (steepest) feasible isocost ray is tangent to the agent's participation constraint, which occurs at point **c** with price and quality combination (p^C, q^C) (C for *complete*). The participation constraint in this case is $u = u_0 = 0$, the iso-utility indifference curve, rather than the iso-value curve because the game is one shot. In panel (b), the Nash equilibrium (p^N, q^N) is the point of tangency of the lowest feasible isocost ray with slope q^N/p^N, where q^N and p^N are the Nash equilibrium price and quantity, and the best-response function with slope $-2\underline{u}/p^2$. The bundle (p^N, q^N) is the *incomplete contract* solution. Summarizing, c_1 (complete contract): $q = p/4\underline{u}$; c_2 (incomplete contract): $q = p/8\underline{u}$. Participation constraint: $q = 1 - \underline{u}/p$. Best-response function: $q = 1 - 2\underline{u}/p$.



(a) The complete contracting game

(b) Both complete and incomplete contracting

can see from Figure 10.11 that Armin would provide zero quality (his best-response function $q(p) = q(2\underline{u}) = 0$).

Three characteristics of the Nash equilibrium with complete contract are:

- *Pareto efficiency*: Patrisia offered Armin a contract that implemented the $mrs = mrt$ rule, she also (without intending to do so) implemented the $mrs^P = mrs^A$ rule guaranteeing the Pareto efficiency of the Nash equilibrium. (The "Reminder" on the previous page explains why the two rules coincide.) It is not possible—by choosing some allocation other than $q^C = 0.5$ and $p^C = 2\underline{u}$—for Patrisia to be better off without Armin being worse off. We know from Chapters 4, 5, and 9 that when one actor optimizes subject to a participation constraint (rather than an incentive compatibility constraint) of another actor, the outcome is Pareto efficient. This follows from the definition of a Pareto-efficient outcome: if Patrisia has indeed minimized her costs subject to the requirement that

Armin have utility of at least zero, then it must be that any *other* allocation would make at least one of them worse off.

- *No rent for the agent*: The supplier's utility under the contract is identical to his next-best alternative. Remember we used the distance between point **n** and point **c** to calculate the rent that Armin received under the incomplete contract, so you can see that completing the contract transferred the rent from the agent to the principal.
- *The market clears*: The fact that the agent receives no rent means—by the definition of a rent—that he is no better off with the transaction than he would be at his fallback option, that is, if it were terminated. But this in turn means that his fallback option—what he gets if terminated—must be to immediately secure the same deal—providing q^C for ihe price p^C from some *other* principal. And for this to be true, it must be that there are no other agents just like Armin who are unable to sell their goods, because if there were, then he would be among them looking for a buyer, that is, worse off than he was with Patrisia. This is why the absence of rents indicates that markets clear.

M-NOTE 10.7 **The complete contracting outcome**

We use Armin's participation constraint to find the complete contracting outcome:

$$\text{Participation constraint} \qquad u(p,q) = p - \frac{\underline{u}}{1-q} = 0$$

$$p = \frac{\underline{u}}{1-q} \tag{10.20}$$

In the case of complete contracts, Armin's participation constraint is the same thing as his willingness to sell: for any given level of quality, Armin has a minimum price at which he is willing to sell some particular level of quality to Patrisia. This is given by Equation 10.20. To find the maximum value of q/p that Patrisia can get, we substitute the price p that is consistent with Armin's participation constraint (given by Equation 10.20):

$$\text{Maximize} \qquad \frac{q}{p} = \frac{q}{\frac{\underline{u}}{1-q}} = \frac{(1-q)(q)}{\underline{u}}$$

To find the price selected by this constrained optimization problem we first differentiate this expression with respect to q:

$$\text{Product rule} \qquad \frac{d\left(\frac{q}{p}\right)}{dq} = \frac{1}{\underline{u}}\left((1-q)(1)+(q)(-1)\right)$$

$$= \frac{1-2q}{\underline{u}}$$

We then set this equal to zero and solve for q^c. Then, by inserting this value into Equation 10.20 we find the value of p^c:

continued

$$\frac{1-2q}{\underline{u}} = 0$$
$$1 = 2q$$
$$q^c = \frac{1}{2}$$
$$p^c = 2\underline{u}$$

The quality of the product is the same as in the incomplete contracting case, but Patrisia is paying half the price she did in the incomplete contracting case. She is appropriating all the rents from the interaction with Armin, her subcontractor, and he is no better off than at his fallback option namely $u = 0$.

The price-setting game with complete contracts

In order to implement point **c** Patrisia did not need to specify the quality that she required Armin to deliver. She did not have to have take-it-or-leave-it power. To see this we modify the complete contracting game.

- Patrisia as before is first mover and can commit to some price at which she will pay for Armin's product depending on its quality: $p^C = \overline{p}q$. She is simply buying whatever quality he provides at the price given.
- Armin responds to her price either by rejecting the contract (if the price is too low) or by delivering a good of a quality level of his choosing.
- Patrisia measures the quality of the good and pays $p^C = \overline{p}q$.
- This ends the game.

Because the contract is complete, Patrisia knows that she will get exactly what she pays for, nothing more and (more important to her) nothing less. So she realizes that she did not need to have take-it-or-leave-it power—dictating the quality as well as the price—in order to transfer the entire rent from the transaction to herself. He could have offered a different contract simply specifying a price: she would purchase any q that Armin provided at a price $p = 4q\underline{u}$. This means that Armin could pick not only point **c**, but also any other point on the isocost line c_1 through point **c**. This line is now Armin's constraint, and he would like to be on the highest indifference curve as possible.

Which point would he pick? Recalling that points to the right and below are better for Armin (higher price and lower quality), point **c** is the point he would pick (all of the other points in c_1 are worse than **c**, in fact they give him negative utility).

So he would deliver his goods to Patrisia, she would measure their quality, finding q to be one-half, and pay him the price $2\underline{u}$ that is $p = 4q\underline{u}$, as promised.

CHECKPOINT 10.13 Why a complete contract matters

a. Explain why the availability of verifiable information about quality and hence the feasibility of a complete contract means that Patrisia is now constrained by Armin's participation constraint, not his incentive compatibility constraint.

b. How much per period would Patrisia be willing to pay to rent her "miracle machine" that verifiably measures quality, assuming that she purchases from 100 subcontractors in a period?

10.12 FEATURES OF EQUILIBRIA WITH INCOMPLETE CONTRACTS: SUMMING UP

Compared to a complete contract, the incomplete contracting contingent renewal Nash equilibrium has six important characteristics, shown in Table 10.3. These characteristics are general features of contingent renewal incomplete contracts, and do not depend on the utility function, the fallback option or any of the other special assumptions we have made in the Benetton model:

M-CHECK The fact that the quality provided is the same in the incomplete and complete contracting case, namely $q^N = 0.5 - q^C$ is what is called a "model specific result." It is not a general result but is a by-product of the particular utility function we have used. In general, the quality provided could be greater or less in the complete contracting case.

1. *Equilibrium rents*: Under the incomplete contract, the agent receives a rent above his next-best alternative. In Table 10.3 you see that he has a utility of $2\underline{u}$ for each period and that because he provides $q^N = 0.5$ the transaction is expected to last for the inverse of this, that is, two periods, giving him a rent of $4\underline{u}$.
2. *The principal is a price maker*: The reason why the principal does not treat the price as given is that she can benefit from changing the price, given the contractual incompleteness concerning the quality of the good. Price-making in the incomplete contracting context does not derive from any noncompetitive aspect of the assumed market structure, such as monopoly power. The Benetton model can include many principals and many agents and there are no barriers to entry, but contractual incompleteness means that principals can benefit by setting prices.

! REMINDER As we explained in Chapter 9 price-taking means taking some price (or wage) as given and not varying it as a way of doing better (increasing profits, reducing cost).

3. *Competitive equilibrium without market clearing*: Because the agent receives a rent, we know that unlike the complete contracting case, his next-best alternative cannot be to just walk across the street and contract with some other principal on identical terms. Instead, he will have to search for a new partner, during which time he will be without a transaction. But for this to be the case there must be other agents also searching for buyers: if he were the only buyer without a contract, then given the ordinary turnover from jobs (quits, deaths and other reasons for leaving a job) he would find a similar job without delay. So the fact that agents on the supply side of the market who are transacting with

Table 10.3 Complete and incomplete contracts: a summary of differences. The principal is the buyer of the good of variable quality; the agent is the producer and seller of the good. The numerical entries are from the model in which the agent's utility is given by Equation 10.1. ICC is the incentive compatibility constraint (the agent's best-response function). PC is the agent's participation constraint (the agent's utility in his next best option, which we set equal to zero). The incomplete contracting interaction is a repeated game; the complete contracting game is a one-shot game.

Contract over q, p	Incomplete	Complete
Constraint on principal's optimization	ICC	PC
Nash equilibrium price paid, p^N	$4\underline{u}$	$2\underline{u}$
Nash equilibrium quality provided, q^N	$\frac{1}{2}$	$\frac{1}{2}$
Expected duration of the interaction (periods)	2	1
Cost of quality to the principal, $c_q \equiv \frac{p}{q}$	$8\underline{u}$	$4\underline{u}$
Agent's utility per period	$2\underline{u}$	0
Enforcement rent?	$4\underline{u}$	0
Market clearing in equilibrium?	no	yes
Principal's short-side power?	yes	no
Durable interactions?	yes	no
Pareto-efficient Nash equilibrium?	no	yes

EXAMPLE A fever (elevated body temperature) is diagnostic of some kind of illness, most often an infection. The fever is evidence of the infection; not the cause. Similarly rents are diagnostic of excess demand or excess supply, not the cause.

REMINDER In Chapter 9 we showed that rents exist when there is excess demand or excess supply, so that there are quantity constrained buyers or sellers. Rents are a money measure of the more conventionally quantity-measured excess supply or excess demand. Correspondingly, when rents exist it must be true that the market does not clear.

principals are getting a rent is diagnostic about the state of the market: there must be excess supply. It is not the case that there is excess supply *because* agents are receiving rents. We will turn to the question—why does the market not clear when contracts are incomplete—in Chapter 11.

4. *Durable Transactions*: The principal (buyer) and agent (supplier) will interact over many periods, even though there are many identical buyers and suppliers. Competitive equilibrium with contingent renewal incomplete contracting will be characterized by a series of durable *bilateral* trading islands rather than a sea of anonymous traders engaged in one-shot interactions in spot markets.
5. *Endogenous claim enforcement through the exercise of short-side power*: The buyer (principal) in the contingent renewal contract minimizes costs by threatening to terminate the ongoing relationship with the seller (agent). Because of this threatened sanction, the agent acts in the principal's interests by providing a product of higher quality.
6. *Pareto-inefficient equilibrium*: Because the buyer maximized taking the supplier's best-response function as the (incentive compatibility) constraint, rather than the supplier's participation constraint, and because when contracts are incomplete the two constraints differ, the noncontractual equilibrium will not be Pareto efficient.

CHECKPOINT 10.14 Why can't they just agree?

a. If moving from the Nash equilibrium to point **f** in Figure 10.9 is a Pareto improvement (both Patrisia and Armin benefit), why don't they just agree to each change their strategies so that they can make the move?

b. Why do we know that a point like **f**, and the Pareto-improving lens, must exist?

10.13 INCOMPLETE CONTRACTS AND THE DISTRIBUTION OF GAINS FROM EXCHANGE

In the model discussed so far, the agent could choose any level of quality. By restricting the agent to two feasible quality levels and reducing the game to a one-shot game, we get a simpler model that focuses on the most important aspects of the problem and that we will find useful in other contexts, such as the employment contract in Chapter 11.

An incentive compatible price offer

In this stripped-down version of the scenario, Armin the agent and supplier of the variable-quality shirt may offer either low or high quality, at a disutility cost $\underline{u}$ and $\overline{u}$ respectively where $\underline{u} < \overline{u}$. In order to protect the reputation of the brand, if the quality is low, then Patrisia will not put her brand's label on the product (e.g. Benetton) and she will not be able to sell it. Patrisia (the buyer) never mistakenly thinks that a high-quality shirt is low-quality. But if Armin has provided her with a *low*-quality shirt she will detect this with probability t and Armin knows this.

Here is the game.

- *Principal offers a price*: Patrisia is first mover and offers a price p for Armin's product and announces that she will not pay if she detects that the good is of low quality.
- *Agent decides on the quality*: Given the price offered, Armin decides to produce high or low quality and delivers the good to Patrisia.
- *Agent may be terminated:* If he has delivered a low-quality good, then with probability t Patrisia detects this and refuses to pay, in which case Armin must sell the good to another buyer at a lower price.
- *Principal pays for the good:* If the good is of high quality, or if it is of low quality but not detected by Patrisia, she pays for the good.

To ensure that Armin will provide high quality, Patrisia must offer a price high enough so that his expected income from offering high quality, $p - \overline{u}$, is not less than his expected income when he offers low quality. Patrisia has to make it a best response for Armin to provide high quality.

The game with just two levels of quality is similar to the games in Chapter 1 where the players had just two choices, for example fishing ten

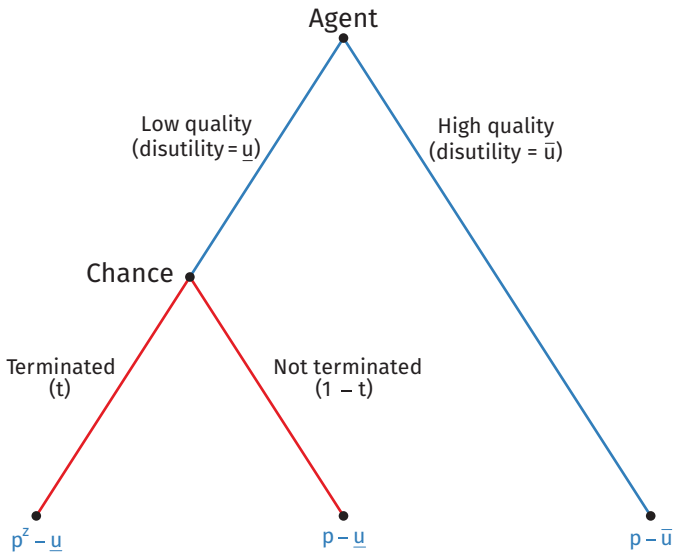
or 12 hours. So here the incentive compatibility constraint (ICC) that limits Patrisia's optimization problem is not a range of prices p that she must offer if she wants quality q. It is a single price that will make it in Armin's interest to choose high over low quality.

The interaction along with the agent's payoffs are shown in the game tree in Figure 10.12. The right-hand side branch gives the result for when the agent provides high quality: he bears the cost of producing high quality $\bar{u}$ and is paid the price p so his utility is $p - \bar{u}$.

The left-hand branch shows how the game proceeds if he produces low quality. There are two outcomes that might occur if he chooses not to produce high quality:

- *Agent's contract is terminated:* With probability t Patrisia detects the low quality, refuses to pay Armin, and he gets his fallback price p^z selling the good to some alternative buyer and so has utility $p^z - \underline{u}$.

Figure 10.12 The quality game, with the agent's payoffs. The game tree (like the one for the Ultimatum Game in Chapter 2) gives the order of play (from the top down), the actions taken at each node (branching point) in the tree, and the agent's payoffs that will result from each path through the tree. Here the agent moves first. But on the left branch of the tree it is not the principal who moves second: it is chance, namely the probability that low quality is detected. The principal has already committed to terminate him if his low quality is detected. When the branch that is followed is determined by chance it is conventional to say that "nature" moved (think back to our examples of expected payoffs when the contingency was "it rains").



- *Agent's contract is not terminated*: With probability $(1-t)$ Patrisia does not detect the low quality, so she pays him p and his utility is $p-\underline{u}$.

His expected income is the income he receives in these two events, multiplied by the probability of each of them occurring. This is the right-hand side of Equation 10.21.

$$\text{ICC:} \qquad p-\overline{u} \quad \geq \quad \overbrace{(1-t)(p-\underline{u})}^{\text{Not terminated}} + \overbrace{t(p^z-\underline{u})}^{\text{Terminated}} \tag{10.21}$$

To find the lowest price Patrisia can offer that will induce Armin to provide high quality, we rearrange Equation 10.21 to isolate p (as shown in M-Note 10.8) and find the Nash equilibrium price, p^N:

$$\text{Nash equilibrium price:} \qquad p^N \quad = \quad \frac{\overline{u}-\underline{u}}{t} + p^z \tag{10.22}$$

Patrisia will set the price p^N and Armin will provide high quality.

This is a Nash equilibrium because at the price p^N Armin would not do better by providing low quality, and given the incentive compatibility constraint based on what she knows about Armin's behavior, Patrisia cannot do better than to offer p^N. If she offered a higher price than p^N, she would be throwing away money. If she paid less than p^N, then he would produce low quality.

M-CHECK So that we could express Equation 10.23 as an equality rather than an inequality, we have assumed that as long as the expected income from providing low quality is not higher than the expected income from providing high quality, Armin will provide high quality.

From Equation 10.22 for the equilibrium price, the lowest price that Patrisia can offer compatible with Armin supplying high quality will be *higher*:

- the *greater* is the difference in cost to the agent $(\overline{u}-\underline{u})$ to provide high rather than low quality;
- the *greater* is the agent's fallback price, p^z; and
- *lower*: the more difficult it is for the principal to detect low quality, that is, the lower is t.

M-NOTE 10.8 **From the ICC to the Nash equilibrium price**

We find the Nash equilibrium price, by rearranging the incentive compatibility constraint, Equation 10.21.

$$\text{ICC:} \qquad p-\overline{u} \geq (1-t)(p-\underline{u}) + t(p^z-\underline{u})$$

We express this as an equality because the principal would never pay the agent more than necessary to secure high quality (as before we assume he will do what the principal intends as long as he cannot do strictly better doing otherwise):

$$p-\overline{u} = (1-t)(p-\underline{u}) + t(p^z-\underline{u}) \tag{10.23}$$

$$\text{Rearranging:} \qquad p-\overline{u} = t(p^z-p) + p - \underline{u}$$

continued

Add $\underline{u} - p$ to both sides and solve for p^{N}:

$$\underline{u} - \bar{u} = t(p^{z} - p)$$

$$\frac{\underline{u} - \bar{u}}{t} = p^{z} - p$$

$$p^{\mathrm{N}} = p^{z} + \frac{\bar{u} - \underline{u}}{t} \tag{10.24}$$

The agent's rent is the value of his transaction producing high quality, that is, $p^{\mathrm{N}} - \bar{u}$ minus what he would get if he produced low quality and was terminated, that is, $p^{z} - \underline{u}$. Or:

$$\text{Agent's enforcement rent} = (p^{\mathrm{N}} - \bar{u}) - (p^{z} - \underline{u})$$

$$\text{Replacing } p^{\mathrm{N}} \text{ by Equation 10.24} = \left(p^{z} + \frac{\bar{u} - \underline{u}}{t} - \bar{u}\right) - (p^{z} - \underline{u})$$

$$= \frac{\bar{u} - \underline{u}}{t} - \bar{u} + \underline{u} \tag{10.25}$$

$$= \frac{\bar{u} - \underline{u}}{t} - \frac{t(\bar{u} - \underline{u})}{t} \tag{10.26}$$

$$= \frac{1-t}{t}(\bar{u} - \underline{u}) \tag{10.27}$$

This is equal to the difference between the costs of providing high and low quality $\bar{u} - \underline{u}$ multiplied by $\frac{1-t}{t}$, which is the ratio of the probability of escaping termination to the probability of being terminated, if providing low quality.

Incomplete contracts, enforcement rents, and profits

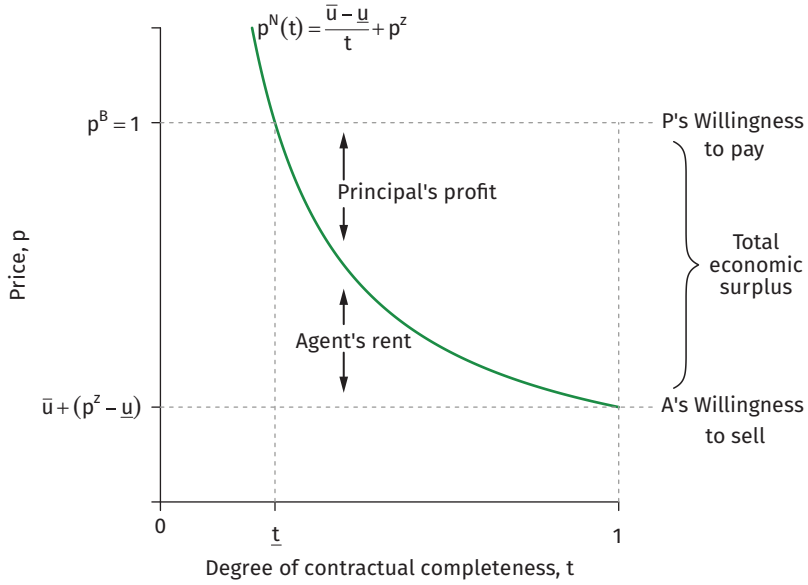
These results replicate some of the economics you have already learned about exchange with incomplete contracts. But in one respect this model goes further: it provides us with a measure of *how incomplete* a contract is, and to see why this matters. It gives us a measure of *how asymmetric* the relevant information is.

Armin knows if he has produced a low-quality good; but Patrisia will discover this only with probability t. So if we denote the quality of Armin's knowledge as 1 (100 percent), then the difference between this and the quality of Patrisia's knowledge (t) or $(1-t)$ is an indicator of the extent of information asymmetry. The accuracy of her information compared to his—t itself—is a measure of *how complete* the contract is.

We now can see the effect of contractual incompleteness on the distribution of income between the principal and the agent. To do this, let's assume that Patrisia can sell a high-quality shirt for some given price p^{B} (for Benetton). Then her maximum willingness to pay Armin for a good shirt is p^{B}. If she were to sell at that price, her profits would be zero. We do not need to consider the case in which she ends up with a low-quality good and has to (attempt to) sell that. If she pays Armin p^{N} he will not deliver low-quality goods.

Armin's minimum willingness to sell is the price at which he could have sold a low-quality good, plus compensation for the extra cost of producing high quality, or $\bar{u} - \underline{u} + p^{z}$. The difference between her willingness to pay p^{B} and his willingness to sell $\bar{u} - \underline{u} + p^{z}$ is the total rent to be gained from

Figure 10.13 Contractual incompleteness and the distribution of the economic surplus. The total rent made possible is the transaction difference between the principal's willingness to pay for a high-quality good and the agent's willingness to sell a high-quality good, or p^B minus $\bar{u} - \underline{u} + p^z$. We set p^B = 1 so that we can interpret the quantities shown as percentage shares of the price sold. Accounting profit and economic profit are identical because the buyer sells the good immediately following having purchased it from the supplier. So for the buyer (Benetton), there is virtually no opportunity cost of funds tied up in advance.



exchange of one unit. This is the quantity that will be divided up between them in the form of her profits and his rent. How it is divided up will depend on how complete the contract is.

Figure 10.13 plots Equation 10.22, showing that the Nash equilibrium price differs for different degrees of contractual completeness. It illustrates how the distribution of the gains from exchange depend on the extent of information asymmetry and contractual incompleteness. The vertical distance between the agent's willingness to sell and principal's willingness to pay is the "pie" that is to be divided up into the two "slices": the principal's profit and the agent's rent.

Equation 10.22 shows that if the contract were complete ($t = 1$ so that low quality could be detected with certainty), Patrisia need do nothing more than to meet Armin's participation constraint paying his minimum willingness to sell price:

$$\text{Complete contract price} \qquad p^C = \bar{u} - \underline{u} + p^z \tag{10.28}$$

Even at this low price he would provide high quality because in this case his best response to that price is to produce high quality. To see this, think about Armin's options. He could produce high quality and get $p^N - \bar{u}$ for sure, or produce low quality and get $p^z - \underline{u}$ for sure.

If the contract is incomplete, meaning $t < 1$, however, Figure 10.13 shows that Patricia will offer Armin a price greater than his willingness to sell. You can also see that the size of the resulting rent is greater, the more incomplete the contract is (going from right to left in the figure).

As expected, the more complete the contract, the larger is the share of the principal's profits. Do not conclude, however, that from the agent's standpoint the more asymmetric the information the better. There is some level of information asymmetry ($\underline{t}$ in the figure) beyond which the best the principal can do would be to pay the agent more than the price at which she can sell the trademarked good to a consumer. In this case there is no way for the transaction to satisfy the *principal*'s participation constraint. So she will stop purchasing shirts from suppliers, trademarking them, and selling them to consumers. With no exchange taking place there are *no* gains from trade to be shared.

We can use this simple model of the quality problem to better understand a fast-growing kind of work in many countries: the gig economy.

CHECKPOINT 10.15 Quality testing

a. Explain why having a perfect test of the quality of the goods that Armin produced accomplishes the same result as being able to enforce a complete contract.

b. What would happen to the curve $p^N(t)$ in Figure 10.13 if the price of Armin's fallback, p^Z increased?

c. How does an increase in p^Z change the distribution of rents between the principal and agent if t and p^B are held constant?

d. Why is an incomplete contract where $t < \underline{t}$ undesirable from the perspective of the agent?

10.14 APPLICATION: COMPLETE CONTRACTS IN THE GIG ECONOMY

A 'gig' for a musician or comedian is a single appearance for which they will be paid not by the hour, but an agreed sum for the performance. The *gig economy* is not about jokes and tunes, however, it refers to the combined activities of Uber or Lyft drivers, TaskRabbits, UpWorkers, Mechanical Turkers, and others who transport people and goods, home-assemble

online purchased furniture, and perform other well-defined tasks for which they are paid a fixed rate.

In many legal jurisdictions, gig workers are considered private contractors and not employees. They provide their own cars or tools and gain access to their gigs by means of a two-sided platform that connects those who will pay for the gig, and those who perform it. The company—Uber, for example—sets the prices and determines the number of people who are allowed to use their app.

The gig economy is a small portion (in the US not more than 2 percent of employment) even of those high-income economies where, for example, ride services like Uber and Lyft have made significant inroads against conventional taxi firms. The gig economy is growing because modern information technology makes it easier both:

- to match buyers and sellers—drivers and those needing a ride, or those needing a task done in their home—and others with the time and skill to do the job; and
- to define tasks sufficiently precisely—the exact time taken for a delivery for example—that gig workers can be paid by the task and not by the hour.

The second bullet means that app-based ride-hail, delivery, and other parts of the gig economy provide an illuminating contrast with the model of hidden actions with variable quality studied in this unit. The key difference is that in some cases the tasks performed are sufficiently well defined and easily measured so that a virtually complete contract is possible: if the person is not delivered from the hotel to the airport, the Lyft driver does not get paid; if the Ikea shelves purchased online are not assembled properly, the TaskRabbit does not make a penny.

This is equivalent to the simple model of the quality problem in which the principal is the person who has engaged the gig worker (the agent) to do a job. We can illustrate what this means with the simple 'two levels of quality' model just introduced. Think about some task, for example, assembling a bicycle that a person has just purchased in a kit. A Rabbit tasker might take on this job for an agreed-upon fee. If the task is not performed—the bicycle does not work properly when assembled, for example—this information will be available to the purchaser and the tasker will not be paid.

In terms of the model, the probability that low quality (a nonfunctioning bike) will be detected, t, is much closer to 1 than in conventional jobs. Figure 10.13 shows that if $t = 1$, then the agent—that is the gig worker—is paid her minimum willingness to sell and the principal—so that the owner of the platform—TaskRabbit or Uber—receives the *entire* economic rent.

A result is that gig performers in this economy face extraordinary economic insecurity: they are not guaranteed a fixed schedule of hours and pay, nor do they receive health insurance benefits, maternity leave, holiday pay, or pension contributions through their employer. Working full-time,

✓ **FACT CHECK** App-based urban transportation services like Uber use network-based software of the type used by eBay to profit from the massive expansion of smartphones with GPS capabilities. In New York City, three major app-based companies provide point-to-point service, in addition to Uber: Lyft, Juno, and Via. There are now twice as many Uber cars serving New York than there are licensed taxis. In 2017, these companies provided 160 million rides, and by February 2018, they were making twice as many trips as were conventional taxis. There are six times as many app-based vehicles providing service as there are taxis. Most app-based drivers work full-time. New York City is unusual in the number of "ride hail" companies operating there, but both Uber and Lyft operate in most US cities.[10]

the hourly earnings of the vast majority of app-based drivers for ride-hail services in New York City would place them below the official poverty level for New York.

The reason, also clear from the model and Figure 10.13, is that gig companies do not need to pay substantially more than the gig worker's next-best alternative. Beyond compensating the tasker for her time and trouble, they do not need to motivate the tasker to do the job as specified: if it is not done, the tasker will not be paid. A result is that the gig economy can often produce services at a lower cost and price than are available from conventional firms that, as we will see in the next chapter, pay their employees substantially more than their next-best alternative.

✓ **FACT CHECK** In reality, gig workers are often paid considerably *more* than their next-best alternative, as is evident from the rapid increase in the number of people ready to become Uber, Lyft, and other gig drivers.

The structure of the gig economy not only reduces the pay that workers receive, it also depresses their fallback option. An important feature of the gig economy is that the only way that drivers or taskers can get gigs is through the platforms owned by a few firms such as Uber, Lyft, TaskRabbit, Mechanical Turk, and others. This is equivalent to Patrisia being the only buyer or just one of few buyers to whom Armin could sell his goods. This means that those performing the gigs have no real bargaining power. If a Rabbit tasker objects to the terms, there will always be another tasker to take her place, but few if any other ways that the disgruntled tasker could find a gig.

HISTORY The gig economy resembles what was called "putting out" in eighteenth-century Great Britain. In this system, women spun on home-spinning wheels using cotton fibers provided by a "putting out" company. They were paid by the amount produced, just as gig workers are paid by the task completed. With many households ready and able to take up spinning, and in most cases just a single putting-out company, the spinners had little bargaining power, and were paid low prices for their products.

The computer platform that allows those who need a gig performed to connect to those performing the gigs makes possible substantial mutual benefits in putting together gig workers who have free time and the skills, a vehicle, or other equipment required with those willing to pay for a completed gig. But the distribution of these gains among those who hire the gig worker, the platform owners, and gig workers is highly unequal.

The gig work illustrates an important truth about the modern economy: the nature of our social interactions at work—whether they be personal or anonymous, long-lasting or ephemeral, for example—is strongly affected by whether the contract governing our transactions is complete or not. You have already seen a hint of this: the complete contracting game was a one shot, while incomplete contracts were modeled by a repeated game. Experiments confirm that the nature of the contracts influences the kinds of social interactions and social norms that are part of the exchange process.

CHECKPOINT 10.16 The gig economy

a. How do platforms such as Uber and Lyft make contracts between the principal and agent more complete?

b. Explain why platforms such as Uber, Lyft, and Mechanical Turk are on the short side of the market

10.15 APPLICATION: NORMS IN MARKETS WITH INCOMPLETE CONTRACTS

When contracts are complete, you get what you pay for. So, for a given price, there is little economic reason to be concerned about one's exchange partner's psychological makeup or moral commitments. If you do *not* get what you paid for, you get your money back at no cost to yourself. This is what a complete contract means: its terms are enforced–if necessary–by the courts, not by the parties to the exchange. And this is also why we care about who we interact with a lot more in cases where contracts are incomplete.

To see this, put all of the people with whom you have any economic interactions into two groups: those whose names you know and those whose names you do not know. You probably do know the names of your employer, your doctor, the person you consult with for legal advice, and perhaps your car mechanic. These are all exchanges in which the contract is substantially incomplete. Do you also know the name of the gas station attendant or the checker at the supermarket?

When contracts are incomplete, parties to an exchange–whether buyers or sellers–will favor social interactions where exchange is *personal* and *durable*. Exchange is *personal*, rather than *anonymous*, when the parties to exchange have personal knowledge of each other, such as personal histories and knowledge of whether the other party is trustworthy or untrustworthy. Exchange is *durable*, rather than one shot, when it results in long-term repeated interactions between the parties to exchange, such as when you regularly go to a hair stylist you like, or you trust a car mechanic or a babysitter you've known for a long time.

The relationship between contractual incompleteness and market structure can be seen in the contrasting structures of the rice and raw rubber trade in Thailand. Buying and selling in the wholesale rice market–where the quality of the product is easily determined by the buyer–buyers and sellers hardly know one another. This contrasts with the personalized exchange based on trust in the raw rubber market. In the raw rubber market quality is impossible to determine at the moment of purchase. As a result, buyers purchased rubber repeatedly from the same sellers rather than shopping around, a strategy that gave them the kind of short-side power that Patrisia has over Armin as a way of controlling the quality.

Similarly, in villages like Palanpur (in India), wheat and rice as well as seeds and fertilizer are standardized, easily measured commodities and are subject to relatively complete contracting. These inputs are bought and sold in region-wide markets in which transactions are governed by little more than the going price and the budget constraints of the participants. The markets are impersonal and anonymous.

› EXAMPLE According to Lisa Bernstein, in the diamond industry: "disputes are resolved not through the courts and not by the application of legal rules announced and enforced by the state … [but rather by] an elaborate, internal set of rules complete with distinctive institutions and sanctions."[11] Bernstein explains how diamond traders address the fact that quality is not easily determined by relying on an "internal set of rules," rules that have informally emerged from *within* the diamond industry.

HISTORY The economic historian Avner Greif (1994) analyzed the divergent cultural and institutional trajectories of the traders from Genova, Italy, and North African Maghrebi traders in the late-medieval Mediterranean from this perspective. The Maghrebi traders had what Greif terms a "collectivist" system of contractual enforcement whereby none of them would ever deal with anyone who had ever failed to fulfill their contractual obligations. The individualism of the Genovese traders, on the other hand, precluded the high levels of cooperation and loyalty to one another on which the Maghrebi system depended. But the limits of Genovese individualism also provided an impetus for their development and perfection of an ultimately more successful system of state and other third-party enforcement of contract terms.

By contrast, exchanges concerning labor, credit, the use of land, and the services of farm assets such as bullocks take place almost entirely within the village, and often within the same caste. Moneylenders in villages rarely extend loans to people they don't know or who don't live in the same village. The village markets in goods or services with *incomplete contracts* are *personalized*.

Experimental evidence

A number of experiments also show differences in behavior depending on the possibility of complete contracting.

Economists have investigated a variety of experimental markets to understand the decisions people make when buying and selling goods.[12] Experimental researchers can change the "institutions" under which participants interact. Remember, institutions are rules of the game, so in an experimental game, changing institutions just means changing the rules. For example, one experiment could have the institutional structure of *complete contracts*, and another *incomplete contracts*.

✓ **FACT CHECK** The sociologist Peter Kollock investigated "the structural origins of trust in a system of exchange, rather than treating trust as an individual personality variable" with similar results. Using an experimental design based on the exchange of goods of variable quality, Kollock found that *trust* in and *commitment* to trading partners as well as a concern for one's own and others' *reputations* emerges when product quality is variable and non-contractible but not when quality is subject to complete contracts.[14]

Economists Martin Brown, Armin Falk, and Ernst Fehr designed a market experiment to explore the effects of contractual incompleteness on patterns of trading. The good exchanged varied in quality, with higher quality more costly to provide. In the complete contracting condition, the experimenter enforced the level of quality promised by the supplier, while in the incomplete contracting condition the supplier could provide any level of quality (irrespective of any promise or agreement with the buyer).[13]

Buyers and sellers knew the identification numbers of those they were interacting with, so they could use information they had acquired in previous rounds as a guide to whom they would like to have as trading partners, and the prices and quality to offer. Buyers had the opportunity to make a private offer (rather than broadcasting a public offer) to the same seller in the next period, therefore attempting to initiate an ongoing relationship with the seller.

Very different patterns of trading emerged under the complete and incomplete contracting conditions. In the complete contract condition, 90 percent of the trading relationships lasted less than three periods (and most of them were one shot). By contrast, under the incomplete contracting condition only 40 percent of the relationships were fewer than three periods, and most traders formed *trusting* relationships with their partners.

Buyers in the incomplete contracting condition offered prices considerably higher than the cost of providing quality (just as in the principal-agent shirt quality model). When buyers were disappointed by the quality supplied, they terminated the relationship, withdrawing the implied enforcement rent from the supplier. Other differences are summarized in Table 10.4. The behavioral differences in complete and incomplete contracting treatments were particularly pronounced in later rounds of the game, suggesting that the subjects updated their behaviors according to experience.

Table 10.4 Contractual incompleteness and market social structure: experimental evidence. Experimental subjects interacted very differently when the quality of the "good" they were exchanging was certified (and enforced) by the experimenter (complete contracts) and when the quality was not known and was determined by the seller (incomplete contracts).
Source: Brown et al. (2004).

Structure of interactions	Complete contracts	Incomplete contracts
Duration	one shot	repeated
Offers	public	private
Price determination	haggling, offers rejected	price setting by short-sider
Traders	anonymous relationship	trust, retaliation for cheating
Market networks	many thin connections	bilateral trading islands

These experimental results suggest that there may be a two-way relationship between trust, reciprocity, and other social preferences, on the one hand, and the degree of contractual completeness on the other.

- *Long-standing relationships*: Where contracts are incomplete–as in the above experiment–economic interactions may endure over long periods during which people develop trusting and reciprocal relationships; while this is unlikely to be the case where contracts are complete.
- *Social preferences and contracts are substitutes*: Where people are trusting and reciprocal, making the contract "as complete as possible" may not be worth the legal costs and possible offense to one's trading partners. But if people are entirely self-interested, trying to complete the contract may be the only way to do business.

10.16 CONCLUSION

The most important organizations governing exchanges in modern economies are firms, whose managers in order to produce and market goods and services combine other people's labor and (what Adam Smith called) "other people's money," neither of which are subject to complete contracting. Labor and credit markets are typical of the many important exchanges in which what is transacted are not well-defined and easily measured objects, like the nuts and apples in Ronald Coase's example in the head quote for this chapter. In these markets the transaction involves something quite different and much more difficult to enforce (the promise to repay the loan, for example) or to measure (e.g. the promise to work hard on the job). Coase put it this way: "what are traded on the market are not, as is often supposed by economists, physical entities, but the rights

to perform certain actions . . . the objects of exchange are complex bundles of obligations and claims concerning who should do what under what conditions."[15]

In the next two chapters we use the principal–agent model you have just learned to study how the owners and managers of firms—as both employers of workers and borrowers from banks and other lenders—structure the rights to perform actions concerning other people's labor and other people's money, respectively.

MAKING CONNECTIONS

Limited information: Asymmetric and/or non-verifiable information about the quality of goods or other aspects of an exchange results in contracts that are incomplete (that do not cover all that matters to one of the parties to an exchange and/or they are unenforceable by the court), a common feature of modern economies that will be important in the remaining chapters.

Rules of the game: Incomplete contracts, external effects, and coordination failures: Because contracts are incomplete one or more actors will not take appropriate account of the effect of their actions on others; these external effects are similar to the environmental external effects (e.g. overfishing) in earlier chapters in that they result in coordination failures, that is, Pareto-inefficient outcomes.

Optimization limited by incentive compatibility and participation constraints: In principal–agent interactions the fact that contracts are incomplete means that the relevant constraint is not the agent's participation constraint but instead her incentive compatibility constraint. This means that the Nash equilibrium cannot be Pareto efficient because when the principal implements the $mrt = mrs$ rule as the solution of her constrained optimization problem the result is not a tangency between the indifference curves (or isovalue and isocost curves) of the two.

Mutual benefit and conflict over distribution: Like other economic actors, when principals and agents interact they each do so in order to gain something they value, so exchanges voluntarily entered into are mutually beneficial; but there is also a conflict about the distribution of these mutual benefits.

Rules of the game: Rents, contractual incompleteness, and inequality: The institutions governing exchange, including the extent to which contracts are incomplete, affect the distribution of the gains from exchange.

Price making: Principals are price-makers, not price-takers (and as we will see, other principals (employers and bankers) are also "wage-makers" and "interest-rate-makers"). They act as price-makers because contracts are incomplete, not because limited competition gives them market power, as was the case with monopolies, duopolies, and other settings with few competitors.

Non-clearing markets: In the Nash equilibrium of the market for the variable quality good there is excess supply (some agents are quantity constrained and unable to sell their products). Competitive markets that do not clear in equilibrium—including, as we will see, those with excess demand such as the credit market—are a feature of principal–agent models.

Power and social norms: When contracts are incomplete, the private exercise of power and social preferences such as trust and reciprocity (along with contracts) provide the basis for mutually beneficial exchange and also affect the distribution of these benefits between principals and agents.

Experimental evidence: Behavioral experiments clarify how the nature of contracts—complete or incomplete—affects the social structure of exchange.

IMPORTANT IDEAS

(in)complete contract	endogenous distribution of gains	verifiability
duration of contracts	measurability	termination
fallback	repeated games	strategic (a)symmetry
non-clearing market	power	social norms, trust, and fairness
information (a)asymmetry	agent	conflict of interest
principal	hidden actions	hidden attributes
adverse selection	moral hazard	lemons problem
insurance	contingent renewal contract	isocost ray
indifference curve	iso-value curve	endogenous claim enforcement
best-response function (BRF)	marginal rate of substitution	long side of the market
enforcement rent	price-making	short-side power
quantity constraints	durable transaction	$mrs = mrt$ rule
Pareto	Nash	$mrs^{\mathrm{P}} = mrs^{\mathrm{A}}$ rule
(in)efficiency	equilibrium	
disutility (of quality)	nforcement rent	

MATHEMATICAL NOTATION

Notation	Definition
q	agent's level of quality provided
p	price of the good offered by the principal
p^B	price at which the principal will sell the good
$u(\ )$	agent's utility function
$\underline{u}$	disutility of providing quality (parameter in A's utility function)
$\underline{u}$	(also) agent's disutility of providing low quality
$\bar{u}$	agent's disutility of providing high quality
$v(p,q)$	agent's expected value of the ongoing relationship with the principal
t	termination probability
$t(q)$	principal's termination schedule
T	expected duration of the agent's relationship with the principal $(= \frac{1}{t(q)})$

Note on superscripts and subscripts: N: Nash equilibrium; C: complete contract; z: fallback position.

CHAPTER

11

WORK, WAGES, AND UNEMPLOYMENT

If a workman moves from department Y to department X, he does not go because of a change in prices but because he is ordered to do so . . . for certain remuneration [the worker] agrees to obey the directions of the entrepreneur . . . the distinguishing mark of the firm is the suppression of the price mechanism.

Ronald Coase,
"The Nature of the Firm" (1937)[1]

DOING ECONOMICS

This chapter will enable you to:

- Explain why the employment contract is incomplete and how the labor market differs from markets in which contracts are complete.
- Understand how, when an employer hires a worker, both may be better off as a result, why there will be a conflict over the distribution of these mutual benefits, and how a more complete contract favors the employer in this conflict.
- Show how the employer chooses a wage to maximize profits and explain how, along with the threat of termination, the employment rents that workers receive motivate them to work hard and well.
- Show that in a competitive Nash equilibrium: (i) the employer exercises power over the worker; (ii) the wage and effort level are Pareto-inefficient; and (iii) involuntary unemployment will exist.
- Analyze wages, markups, unemployment, and profits using a model of firms' price- and wage-making in the whole economy.
- Understand the effect on a firm's hiring of the imposition of a minimum wage, including the conditions under which this will induce a firm to hire more workers rather than fewer.

Figure 11.1 Workers at the Ford plant in River Rouge, near Detroit, USA. The mural is one of many that Henry Ford's son and successor, Edsel, paid Diego Rivera, a member of the Central Committee of the Communist Party of Mexico, to create.

Detroit Institute of Arts, USA/Gift of Edsel B. Ford/Bridgeman Images.

Figure 11.2 Made by happy (or at least well-paid) workers: a Ford Model T. The guiding idea was simplicity. Ford once said that the buyer of the Model T can have "any color that he wants so long as it is black." And he wrote: "It will be constructed of the best materials, by the best men to be hired, after the simplest designs that modern engineering can devise. But it will be so low in price that no man making a good salary will be unable to own one."

Photo used by permission, Utah State Historical Society.

11.1 INTRODUCTION: HENRY FORD'S SHOCKER

On the morning of January 5, 1914, a little-known mechanic-turned-automobile producer named Henry Ford shocked his colleagues and competitors by announcing that he would pay his workers a minimum of $5 for an eight-hour day, at once shortening the work day and more than doubling the hourly rate of pay for the vast majority of his workers. Ford was not responding to insufficient labor supply: a reporter arriving that morning for the press conference at which the announcement would be made noticed a line of several hundred workers seeking employment.

In the weeks following the announcement, the queue outside the gates swelled to over 12,000, almost as many as were working inside. Remarkably, profits rose, supported by a more than doubled increase in output per hour of production labor. Ford would become a household name around the world and the combination of high wages and assembly-line work came to be called "Fordism."

For the lucky workers who had been in the right place at the right time, the basic facts of work life inside the plant changed beyond recognition. The previous year Ford's labor force had averaged 13,623. During the course of that year 50,448 had walked out the door; most had quit.[2]

There had been 8,490 fired. The year following the announcement, employment had grown by one-third, but the number quitting had fallen to one-tenth of its earlier level, and only 27 employees had been fired. Changes of this magnitude clearly cannot be explained by cyclical variations in supply and demand in the local labor market. It seems unlikely that Ford doubled the wage to attract better workers or to retain those workers in whom the company had invested expensive training. A Ford superintendent boasted that "two days is . . . ample time to make a first-class core molder of a man who has never seen a core-molding bench in his life."

Exactly why Ford raised wages and shortened the work day remains a mystery. More important, the success of his gamble is a puzzle, for it contradicts the view that profit maximization entails paying employees a wage as low as possible consistent with their showing up, that is to say, satisfying their participation constraint and nothing more. Ford's $5 day would not make sense in this model because $5 a day was much better than most workers' fallback option. The fact that people were lining up for jobs at Ford even before the wage increase tells us that even then they were paid much more than was required by their participation constraint.

11.2 EMPLOYMENT AS A PRINCIPAL-AGENT RELATIONSHIP

The most likely reason why Ford doubled the wage is that he understood that raising the wage can reduce the cost of labor. A principal–agent model explains how this seemingly paradoxical statement could be true. The key idea is that labor—the activity that produces cars—is not something you can

measure in hours on the job. Instead it consists of tasks done and, like the quality of goods in the Benetton model, these tasks cannot be written into a complete contract. Getting the job done may require paying workers a lot more than their fallback option.

Incomplete employment contracts

To see how a principal-agent model might explain why Ford's radical move worked, recall that a principal-agent relationship arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent.
- *Incomplete contract*: the agent's actions or attributes are not subject to an enforceable contract either because they are not known to the principal, or, if known, are not verifiable for some other reason.

The aspect of the exchange between Ford and his workers that fits these two conditions is the workers' effort on the job, completing the tasks required to produce Model Ts (the only car Ford produced at the time).

- The conflict arises because Ford profits if his employees work harder or faster, while the workers preferred to work at a slower pace both for their safety and in order to go home a bit less exhausted at the end of the day.
- But the workers' **effort**, manifested in literally hundreds of tasks performed per day, many as part of teams of workers, was not something that Ford could measure and write into an enforceable contract.

To see why the employment contract is necessarily incomplete, think back to three of the reasons why this is the norm rather than the exception in a modern economy given in Chapter 10:

- *Asymmetric or non-verifiable information*: Think about cyber-loafing: texting with friends and web surfing at work. The extent of this may be unknown to the employer, and even if it is known, the evidence of it may not be something that could be used to enforce a contract.
- *Time*: The worker takes a job today and the employer would like to renew her employment over a period of months or even years. The employer has no way to determine the tasks he would like her to do under all of the possible conditions that might arise over this period.
- *Measurement*: For most work tasks there are no measures of work done—e.g. quantity and quality of task completion—that are precise enough to be the basis of an enforceable contract. This is the case both because

LABOR (EFFORT) The amount of actual work devoted to production. Labor is measured in units of effort, not in hours.

! REMINDER So far, in Chapters 6 and 8, we have treated labor—measured in hours of employment—as an input into the production process. Employers, in our models of those chapters, could simply purchase this input much as they purchase kilowatt hours of electric power. But it is not hours on the job that produce output; it is tasks completed and the other dimensions of work actually done. This is not something that can be readily purchased because work is done by people who, having their own ideas about how they would like to spend their day, need to be motivated by some combination of carrots or sticks to do the work on which the employer's profits depend.

HISTORY Herbert Simon (1916–2001) who won a Nobel Prize in economics, provided the first model of the firm along these lines. In a 1951 article, he represented the employment contract as an exchange in which the workers transfer authority over their work tasks to the employer in return for a wage. Simon stressed the advantage to the employer of this arrangement given the unavoidable uncertainty about the tasks that would be required over the course of the contract, and therefore the high cost of agreeing to a complete contractual specification of the activities to be performed.[3]

EXAMPLE Tunisians farmers who worked for others some of the time for a fixed wage were half as productive as when the same person worked their own farm and owned the output of their own work.[5]

EXAMPLE What it means for a contract to be incomplete is illustrated by the difference between hiring someone to care for your child for an afternoon, and hiring an Uber driver to take you somewhere. If the Uber driver does not show up, or shows up half an hour late, he will not get paid. If you get home and your child is miserable, you may wonder if the babysitter cared well for her, but you will pay her anyway. Complete contract case: you were prepared to purchase a particular service from the Uber driver, and it was not delivered. Incomplete contract case: you hired the babysitter for a block of time, and hoped she would do a good job.

the tasks are difficult to measure and because their completion typically depends on the efforts of more than a single worker.

Evidence that the amount of work done in an hour can vary significantly, and that it depends on incentives come from some rare examples of what is called **piece-rate** compensation where the amount of work done *can* be written into a contract:

- When a pay system for auto glass installers in the US shifted from hourly wages to paying piece rates, output per worker rose by one-fifth.[4]
- When British Columbia tree planters were randomly assigned to piece-rate compensation—they were paid by the number of trees planted—they outperformed by 20 percent other planters randomly assigned to a fixed wage.[6]

But even with the substantial growth of the gig economy—where some of the contracts approximate piece rates—the number of piece-rate workers as a fraction of all workers in the US economy is not more than 5 percent. This is in part due to the fact that industries that once extensively used piece rates—clothing and shoe production, for example—now employ very few people. And sectors in which "work done" is almost impossible to measure—caring for others, personal security, knowledge production and distribution—have grown significantly.

We use the term "labor" (and the subscript l) to mean work done or effort; when we refer to the workers' time hired, we refer to "employment" and use the symbol h for hours.

A model of employment as a principal-agent relationship

Based on the key idea that work effort is not subject to contract, here is the principal–agent game between the employer and worker. The agent's action is now work effort per hour of employment e rather than product quality q (as in the previous chapter) and the principal will pay the agent a wage w rather than a price p. But the structure of the game is very similar to the Benetton model, as is clear from Table 11.1.

1. *The employer, the principal, is first mover*. He announces to the worker a wage w and the offer to renew the contract at the end of each period unless the worker is terminated for insufficient effort, which occurs with a probability that is inversely related to the effort provided by the worker, $t(e)$. We call the wage rate and termination schedule introduced in this step the employer's *labor discipline strategy*.

PIECE RATE Under a piece-rate contract, a worker is paid a fixed amount for each unit ("piece") of the product made.

Table 11.1 A comparison of two principal-agent models: quality and effort. What we have labeled the Ford model is also called the "labor discipline model" and sometimes the "efficiency wage model." We do not use the latter term because (as is the case with the Benetton model) the Nash equilibrium of the Ford model is not Pareto efficient. In both cases the interaction is a repeated game and the principal maximizes profits and the agent maximizes the value of the transaction.

	Benetton model	Ford model
Principal	Buyer (e.g. Benetton)	Employer (e.g. Ford Motor Company)
Agent	Producer and seller	Worker
Non-contracted action by agent	Quality of the good provided, q	Effort and care by the worker, e
Contingent renewal	Termination of subcontract, $t(q)$	Termination of employment, $t(e)$
Agent's fallback option	Another buyer after searching	Unemployment insurance then another job after job search
Principal sets	Price and termination, p, $t(q)$	Wage and termination, w, $t(e)$
Short side of the market	Demand (excess supply of goods)	Demand (excess supply of labor)

2. *The worker*, one of a group of identical workers who may be employed by the principal, responds to the employer's offer by selecting a level of effort to expend per hour of employment, e.
3. *The principal then chooses the total hours* of workers' time to employ, h, the agent is employed, and production takes place.
4. *Termination probability*: At the end of the period the worker is renewed with probability $1 - t(e)$ or terminated with probability $t(e)$. If terminated the worker receives her fallback option, which for now we assume to be zero. If the worker is terminated this ends the game.
5. *Non-termination*: Conditional on the worker's not being terminated, repeat steps 1-4 above with the values of w, e, and the termination schedule unchanged.
6. Repeat the previous step until the worker is terminated, which ends the game.

We call this the **labor discipline model** (or the Ford model in recognition of the car maker's $5 day). To understand the game we need to ask about each of the above steps: What is each actor attempting to accomplish and what do they know at that particular stage?

EXAMPLE Another piece of evidence that employment contracts do not cover everything that the employer expects to get by hiring a worker: "work to rule" is a trade union strategy in which workers perform those tasks that are specifically required in the contract, and nothing more. Very little production takes place under these conditions.

LABOR DISCIPLINE MODEL A model that explains how employers set wages so that workers receive an economic rent (called an employment rent), which provides workers an incentive to work hard and well in order to avoid job termination.

1. The employer selects the labor discipline strategy e, and $t(e)$ to *maximize* his profits (which requires minimizing the average cost of a unit of effort w/e), *knowing* the worker's best-response function $e(w)$.
2. When selecting a level of effort to perform the worker is *maximizing* the expected value of her employment $v(w, e)$, *knowing* the employer's labor discipline strategy as well as her fallback option.
3. When selecting the level of employment to hire (h) the employer is *maximizing* his profits and *knows* the cost of a unit of effort (which is the result of the labor discipline strategy he has implemented), how worker effort contributes to output, and the demand for the firm's product.

We take up these three steps in the next section. But because the employer has to know the worker's best-response function before deciding on a labor discipline strategy, we take up step 2 before turning to step 1 and finally step 3.

CHECKPOINT 11.1 **Time and toil** What is the conflict of interest between the employer (principal) and worker (agent)? List three primary reasons for why the contract between an employer and worker might be incomplete.

11.3 NASH EQUILIBRIUM WAGES, EFFORT, AND HIRING

Because the Ford model is very similar to the Benetton model and uses the same mathematical functions, you can quickly learn to use the model by studying the comparison in Table 11.1 and the representation of the model in Figure 11.3. (The overview of the Ford model in M-Note 11.1 will also help.)

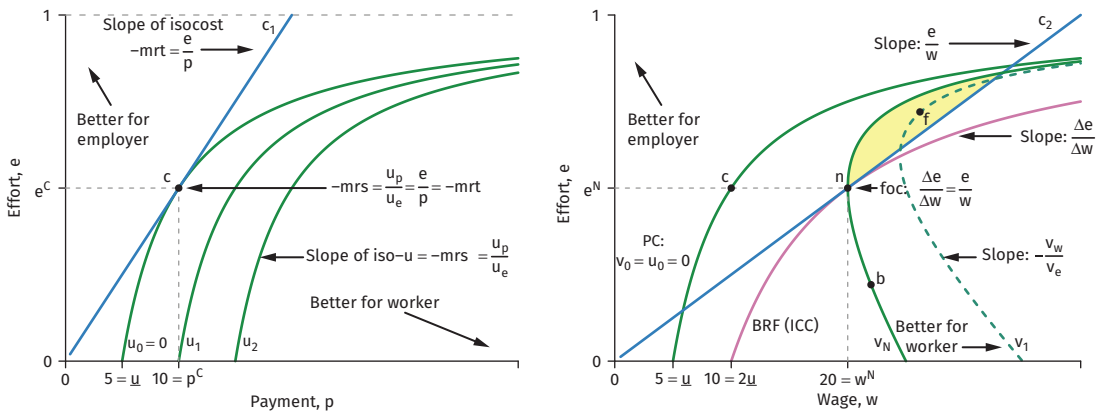
! REMINDER Similar to Figure 10.4 showing the subcontractor's per-period utility based indifference curves for price and quality, because to the worker the wage is a good and effort is a "bad," the indifference curves are upward-sloping. This means that starting at any point on one of the curves, to compensate for a small increase in effort (a move up in the figure) the level of pay must increase (a move to the right) in order for the two changes taken together to leave the level of utility unchanged.

The workers' iso-value curves and best-response function

The two panels in Figure 11.3 contrast a case where complete contracting is possible and where the contract is incomplete. Shown on the left, the contract could be complete if workers are paid by the amount they produce—piece rates—so that the employer is effectively purchasing their effort. On the right we show the general case in which the workers' care and effort on the job are not enforceable in a contract. We will contrast the results in these two cases in table 11.2.

The figure shows how the Benetton model introduced in Chapter 10 can be adapted to represent the relationship between an employer and a worker. The horizontal axis in both panels is the payment to the worker whether it be in the form of a wage or a price (for example if piece-rate payment is possible). (Ignore for the moment the rays from the origin labeled c_1 and c_2.) In Figure 11.3 (a), we show indifference curves representing the per-period utility function of the worker. Included is the curve labeled u_0 which gives the combinations of e and p such that the worker's utility is zero, that is equal to her next-best alternative or fallback option.

Figure 11.3 Comparison of complete and incomplete labor contracts. The per-period utility-based indifference curves in the panel (a) along with the employer's termination function are the basis for the multi-period iso-value curves in panel (b). The point of maximum value for a given price (like point **n**) gives us the worker's best-response function $e(w)$. The rays from the origin show the best the employer can do under the incomplete contract (point **n** in panel (b)) and the complete contract (point **c** in panel (a)). The yellow-shaded area is the Pareto-improving lens showing all of the pairs of wages (w) and effort (e) that are preferred by both employer and worker over the Nash equilibrium, (w^N, e^N). We explain this lens and the complete contracting case in Section 11.5.



(a) Utility-based indifference curves and the complete contract

(b) Iso-value curves and the incomplete contract

In the panel (b) we show two of the workers' iso-value curves. Points **b** and **n** of the curve v_N for example show combinations of the wage and effort by the worker that have the same expected value to the worker, taking account of both the per-period utility of those two combinations of effort and pay but also the effect of working harder (at point **b**), namely extending the expected duration of the job (by reducing the probability of termination).

The worker expects the same value from point **n** as point **b**, even though she is working harder at point **n** for a lower wage, because she expects that she will be terminated sooner if she is working less hard, i.e. at point **b**.

The point where the iso-value curve is vertical (for example point **n**) gives the effort level (the vertical axis coordinate of the point) that is the worker's best response to the wage (the horizontal axis coordinate of the point).

The best-response function is composed of points like **n** where the iso-value curves (not shown) are vertical. The worker's best-response function is the incentive compatibility constraint (ICC) on the employer's profit-maximizing strategy.

$$\text{Worker's best-response function (ICC)} \qquad e = e(w) \tag{11.1}$$

! REMINDER Similar to the Benetton model (see Figures 10.4 and 10.5) in contrast to the per-period utility-based indifference curves, the iso-value curves are not entirely upward-sloping. This is because the iso-value curves take account of the long-term consequences of supplying more of what the principal would like, and avoiding being terminated. In this longer-term perspective, effort can be both a bad or a good.

! REMINDER For the Benetton model in the previous chapter we showed that the points at which the iso-value curves are vertical are also points on the agent's best-response function (see M-Notes 10.3 and 10.4).

In Figure 11.3 (b) the iso-value curve labeled v_0 gives the combination of effort levels and wages such that the worker receives a value of zero, which is the same as his fallback option. This curve is identical to the curve labeled u_0 in panel (a) because in order for the job to be worth zero to the worker, it must be that the utility of having the job for a single period (shown by the iso-value curves in panel (b)) is also zero.

M-NOTE 11.1 **Review of model setup and solution with specific utility and profit functions**

The functions and solutions for this labor discipline model are set up and solved almost identically in Chapter 10. Here is a brief review of the setup and solutions. Make sure you are able to reproduce these results yourself. You can check the M-Notes in Chapter 10 especially 10.3 and 10.4 for guidance.

Complete contract model: We represent the worker's pay by p because with the complete contract the worker is paid by the number of pieces produced, not by the hour.

- Employer chooses price p and verifiable effort level e to maximize $\frac{e}{p}$ subject to the worker's participation constraint (PC)
- Worker's utility function: $u(p,e) = p - \frac{\underline{u}}{1-e}$
- Worker's participation constraint: $u(p,e) \geq 0$
- Nash equilibrium effort level $e^c = \frac{1}{2}$, price $p^c = 2\underline{u}$

Incomplete contract model: Here the worker's pay is an hourly wage, w.

- Employer announces the termination schedule $t(e) = 1 - e$ showing how the termination probability is less when greater effort is provided
- Employer chooses wage w and hours h to maximize profits subject to the worker's best-response function
- Given w, $t(e)$, and her fallback option ($= 0$), the worker chooses effort level e to maximize expected utility
  $v(w,e) = u(w,e)\frac{1}{t(e)} = \left(w - \frac{\underline{u}}{1-e}\right) \times \frac{1}{1-e}$
- Worker's first-order condition (varying e to maximize v): $u_e = t_e v$
- Worker's best-response function: $e(w) = 1 - \frac{2\underline{u}}{w}$ (see M-Note 10.4)
- Employer varies p to maximize $\frac{e}{w}$, subject to $e(w)$, the worker's best-response function (the employer's ICC)
- Effect of wage change on effort: $\frac{de}{dw} = \frac{2\underline{u}}{w^2}$
- Employer's first-order condition (Solow condition): $\frac{de}{dw} = \frac{e}{w}$
- Nash equilibrium wage $w^N = 4\underline{u}$, effort level $e^N = \frac{1}{2}$
- Letting $\underline{u} = 5$, we have $w^N = 20$ and $c_l^N \equiv \frac{w^N}{e^N} = 40$

We solve for equilibrium hours hired h^N later in the chapter.

An employer's labor discipline strategy: The Solow condition

We turn now from what the worker wants to what the employer cares about, the amount of effort he gets from a worker for every euro of pay that he gives, that is, e/w. This is the slope of a ray from the origin: along any such ray the effort per euro paid is constant. Two are shown: labeled c_1 in panel (a) and c_2 in panel (b). These are identical to the isocost rays from the Benetton model, the slope of which is q/p or the amount of quality that the supplier provides per euro of price paid to him.

In each figure, not shown, is an entire map of these isocost rays. Think of each ray as a representation of the objectives of the employer, like an indifference curve or an iso-profit curve. As was the case with these other maps of a decision maker's objectives, the negative of slope of an iso-cost line is the marginal rate of substitution between workers' effort and workers' pay.

The employer would like to be on a steeper isocost ray, meaning higher e/w or what is the same thing lower cost of a unit of effort w/e. But points above the worker's best-response function are *not* incentive compatible: there is no way that the wage and effort combination given by these points could come about because they require the worker to choose a level of effort that they would not choose at the wage in question.

Taking account of the incentive compatibility constraint, the employer would like to minimize the cost of a unit of effort (c_l) as follows:

$$\text{Minimize the cost of effort} \qquad c_l = \frac{w}{e} \tag{11.2}$$

$$\text{Subject to the ICC} \qquad e = e(w) \tag{11.3}$$

In Figure 11.3 (b) this means finding a point on the worker's best-response function (the employer's incentive compatibility constraint or ICC), that is, on the steepest possible isocost ray. This will be point **n**, where the isocost ray is tangent to the best-response function, that is, where the slopes of the two lines are equal:

$$\text{Slope of isocost ray} = \frac{e}{w} = \frac{\Delta e}{\Delta w} = \text{Slope of ICC} \tag{11.4}$$

$$\text{or, what is the same thing} \qquad mrs = mrt$$

where the mrt is the marginal rate of transformation of wages paid by the employer into effort performed by the worker. Equation 11.4 (which is derived in M-Note 11.5) gives us the solution to the constrained cost minimization optimization problem shown in Equations 11.2 and 11.3. It is the rule that tells the employer the wage that will minimize his cost of a unit of effort, called the Solow condition after the macroeconomist Robert Solow, who first demonstrated it.

The condition can be restated as: choose the wage such that the marginal effect of raising the wage is equal to the average level of quality per euro of wage spent. This is point **n** in Figure 11.3 (b). In M-Note 11.1 we use a

HISTORY The version of the labor discipline model presented here was developed by one of us (Bowles) to try to make sense of the movements of wages and labor productivity—called the great productivity slowdown—during the late 1960s and 1970s. Its initial purpose was not academic at all but instead was the basis of advice requested by a number of trade unions and public interest bodies seeking to understand the end of "the golden age of capitalism." Other variants of the model—that developed by Shapiro and Stiglitz for example—were motivated by the desire to provide Keynesian ideas about unemployment with a microeconomic foundation without making ad hoc assumptions such as "wage stickiness" (the tendency of wages to maintain their levels despite recessions).[7]

specific utility function with numerical parameters to show how the Solow condition gives us the profit-maximizing wage that the employer will offer.

You can confirm that the wage and effort level given by the Solow condition are the Nash equilibrium of the labor discipline and effort provision part of the game, so we now give them the N superscripts.

- *Wage*: Given that the worker has adopted the strategy described by her best-response function $e(w)$, the best the employer can do to minimize the cost of effort is to select the wage w^N.
- *Effort*: Given that the employer has offered w^N, the best the worker can do to maximize the value of her job is to provide e^N.

This completes the first step of the employer's profit-maximizing process: finding the labor discipline strategy that minimizes the cost of effort. Now, knowing the cost of a unit of effort w^N/e^N the employer will choose how much effort to allocate to producing the firm's output. To do this the employer must determine the number of hours of workers' time to hire h^N.

CHECKPOINT 11.2 **Nash equilibrium work and wages** Explain why point **n** in panel (b) of Figure 11.3 is a Nash equilibrium.

! REMINDER The hours worked in a day makes a big difference to individual workers and their families, and as you know from Chapter 7, the hours worked during the course of a year differs substantially among countries and changed markedly over the course of the twentieth century.

11.4 THE EMPLOYER'S PROFIT-MAXIMIZING LEVEL OF HIRING

An employer would normally face two questions concerning hours: how many hours a day will each worker work, and what is the total number of hours to be hired (this will determine the number of workers to hire, given the length of the working day). For simplicity we address only the second question, so h is just total hours hired by the employer.

Hiring hours, and employing effort

But hours of workers' time is not what produces the goods the employer wishes to sell; that is done by workers' effort. So we distinguish between:

- the number of hours of workers' time hired by the employer h called *hours hired*; and
- the total amount of actual work devoted to producing goods, which will be the effort provided by each worker in an hour times the hours hired, or $he^N = l$ called *labor employed*. Labor is measured in units of effort (sometimes called "efficiency units" which is why this model is sometimes called efficiency wage theory), not in hours.

To determine the hours hired the employer makes use of two analytical tools about which you already know:

- the *demand function* for the firm's product, showing the maximum number of units of the good x that can be sold at price p or $x(p)$; and
- the *production function*, showing the combinations of the amount of capital goods k and labor l that can produce each level of output x, or $x(k, l)$.

As is the case in selecting the level of output that you studied in Chapters 8 and 9, the owners of the firm maximize profits by producing a level of output such that marginal revenue equals marginal cost. To see what this implies for the employer's hiring, we introduce a new term, the **marginal revenue product of labor**, which is the change in total revenue associated with a small change in labor hours hired. As is shown in M-Note 11.2 this is the marginal product of labor (derived from the production function) multiplied by the marginal revenue (derived from the demand function).

When deciding on how much labor to employ, therefore, the employer weighs two things:

- **marginal cost of labor**: the effect on total costs of using more labor, that is, effort in production, which is the cost of effort per unit employed or $c_l = w^N/e^N$, against
- *marginal revenue product*: the benefits of hiring that worker.

Here the marginal cost of labor does *not* depend on the amount of labor hired, so the average and marginal cost are the same. In section 11.12 we introduce what is in some cases the more realistic case–termed monopsony–in which the cost of employing labor increases the more the employer hires. But for now c_l is fixed by the cost-minimizing labor discipline strategy of the employer.

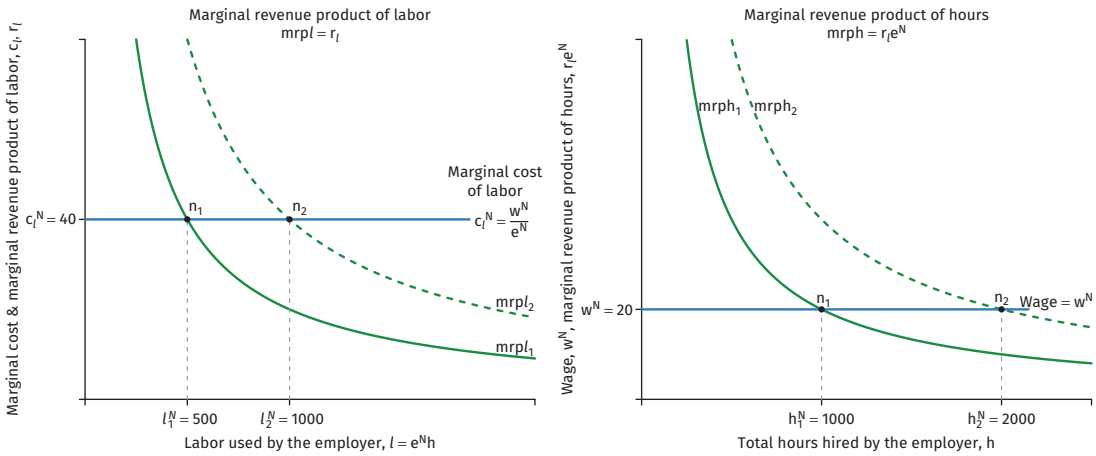
If the marginal revenue product of hiring more exceeds the marginal cost of hiring more, the employer will hire more. He will continue hiring more until it reaches the level of employment so that the marginal revenue product of labor equals the *marginal cost of labor* (*mcl*).

We illustrate this case in Figure 11.4 (a). The horizontal line labeled $c_l^N = w^N/e^N$ is the cost of effort determined by the employer's choice of a cost-minimizing labor discipline strategy. The downward-sloping solid curve shows the effect of employing more labor (by hiring more hours, given e^N) on the marginal revenue product of labor.

MARGINAL REVENUE PRODUCT OF LABOR The marginal revenue product of labor is the change in total revenue associated with a small change in labor employed.

MARGINAL COST OF LABOR The marginal cost of labor is the change in total wages paid associated with employing (a small amount) more labor, that is, effort.

Figure 11.4 Employer's hiring decision. We drew these figures using the parameter values and results from M-Note 11.1. In panel (a) we show the firm's decision about how much labor (that is, the total amount of effort which is equal to the amount of effort per hour multiplied by the number of hours hired) to employ as an input into the production process, given that the employer's decision on the wage rate w^N and the resulting amount of effort e^N per worker hour h^N has determined the cost of a unit of effort c^N. In panel (b) we show the same profit-maximizing problem in terms of the hours of workers' time hired. The marginal revenue product of hours is the marginal revenue product of labor times the amount of effort per hour e^N. The cost of an hour of a worker's time is just the wage w^N.



(a) Employer's profit-maximizing level of labor employed

(b) Employer's profit-maximizing level of hours hired

If the employer employs less labor than l_1^N, then the marginal revenue product of labor is greater than the marginal cost, so the employer would make greater profits by employing more labor. Similarly if the employer were employing more than l_1^N then he would see that the marginal cost of labor is greater than the marginal revenue product of labor, so it would profit by cutting back on employment. So the profit-maximizing level of employment is l_1^N.

In Figure 11.4 (b) we translate the decision on how much labor (total effort or eh) the employer will employ into how many hours of workers' time he will hire. Panels (a) and (b) provide different presentations of the same information: in panel (a) we see the employer's decision on how much of the input–labor–to use (l_1^N), while in panel (b) we see the way that he implements that decision (given the Nash equilibrium level of effort $e^N = \frac{1}{2}$) by hiring some number of hours of workers' time h_1^N. (M-Note 11.4 provides a numerical example.)

The level of hiring, h_1^N is a Nash equilibrium because w^N, e^N and hence c_1^N are all Nash equilibria. Hiring h_1^N is the best the employer can do, given

those values and the employer's production function and demand function (which determine the marginal revenue product of labor curve).

M-NOTE 11.2 The marginal revenue product of labor (effort)

With k and l being, respectively, the amount of capital goods and labor used to produce a level of output x, the information relevant to the problem is:

- the *production function* $x = x(k, l)$
- the *inverse demand curve* for the firm's product $p = p(x(k, l))$, and
- the firm's *total revenue* $r(k, l) = p(x(k, l))x(k, l)$

In doing the differentiation below, keep in mind the two places where the labor variable l appears as an argument in a function: the inverse demand function and the production function.

The marginal revenue product of labor is defined as the change in total revenue associated with a small change in labor devoted to production, that is, the derivative of the revenue function with respect to l:

$$\text{MRP of Labor} = \frac{\partial r}{\partial l} = \underbrace{\overbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial l}x}^{\text{Revenue lost due to lower price}} + \overbrace{p\frac{\partial x}{\partial l}}^{\text{Revenue gained due to increased sales}}}_{\text{Marginal revenue product of labor}}$$

The marginal revenue product can be decomposed into two terms, one negative and the other positive. The first term on the right-hand side is the negative effect of employing more labor (and producing more) on the price of the product at a given level of sales (because the demand curve is downward-sloping). The second term on the right-hand side is the positive effect of employing more labor and producing more revenues (at a given price). Together these two terms constitute the marginal revenue product of labor.

We can rearrange this equation so that it reads:

$$\text{MRP of Labor} = \frac{\partial r}{\partial l} = \underbrace{\overbrace{\left(\frac{\partial p}{\partial x}x + p\right)}^{\text{Marginal revenue (increase in } x)} \overbrace{\frac{\partial x}{\partial l}}^{\text{MPL}}}_{\text{Marginal revenue product of labor}}$$

from which we see that the marginal revenue product of labor is the product of two terms:

- the marginal product of labor
- the marginal revenue associated with an increase in x

The marginal product of labor time and the wage

From Figure 11.4 (b) we can see that the wage given by the Solow condition determines the level of hiring by the firm. As a result it also determines the marginal revenue product of labor hours. If the cost-minimizing

wage rate for the firm had been higher, the marginal revenue product would have been higher at the employer's profit-maximizing level of hiring.

The dashed lines in Figure 11.4 show the effect of an *increase* in the marginal product of labor (which also increases the marginal revenue product of labor in panel (a) and the marginal product of worker hours in panel (b)). The effect is to increase the level of hiring, so as to depress the (shifted up) marginal revenue product of labor now to the level of the wage.

The wage does not change. The wage was determined by the cost-minimizing labor discipline strategy (the Solow condition), and that is not affected by the productivity of labor.

M-NOTE 11.3 **Profit-maximizing employment of labor, given the cost of a unit of effort** $c_l \equiv \frac{w^N}{e^N}$

The relevant information here is the same as in the previous M-Note except that we now introduce the following:

- the Nash equilibrium (average and marginal) cost of a unit of labor (effort) is $c_l^N = \frac{w^N}{e^N}$ given by the Solow condition
- the cost of a capital good is c_k and
- the employer's total cost is $c_l l + c_k k$
- the employer's *profit* is its total revenue − total cost: $\pi = p(x(k,l))x(k,l) - c_k k - c_l l$

To find the level of labor to use in production that maximizes profits, we differentiate the profit function with respect to l, and set the result equal to zero. So, using the chain rule for differentiation we have:

$$\frac{\partial \pi}{\partial l} = \underbrace{\overbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial l}x}^{\text{Revenue lost due to lower price}} + \overbrace{p\frac{\partial x}{\partial l}}^{\text{Revenue gained due to increased sales}}}_{\text{Marginal revenue product of labor}} - \underbrace{c_l}_{\text{Marginal cost of labor}} = 0 \tag{11.5}$$

If this condition (Equation 11.5) is not satisfied it means that either:

- the marginal revenue product of labor exceeds the marginal cost of labor in which case the employer should employ more labor; or
- the opposite, in which case the employer should employ less labor.

We can also rearrange Equation 11.5 to find the following:

$$\frac{\partial \pi}{\partial l} = \underbrace{\overbrace{\left(\frac{\partial p}{\partial x}x + p\right)}^{\text{Marginal revenue (increase in } x)}\overbrace{\frac{\partial x}{\partial l}}^{\text{Marginal product of labor}}}_{\text{Marginal revenue product of labor}} - \underbrace{c_l}_{\text{Marginal cost of labor}} = 0$$

M-NOTE 11.4 Finding the profit-maximizing hours of labor (effort) used, with specific functions

Here we use a specific revenue function to determine the profit-maximizing number of hours to hire. To do this we first determine how much total effort the owner would like to devote to production, l^N. Assume that the employer is using some given level of capital stock, $\bar{k}$. Holding constant $\bar{k}$ and varying the amount of labor, the revenue function is:

$$r(l) = 20,000 \ln(l)$$

As was shown in Equation 11.5, the employer's profit-maximizing level of employment is such that the marginal revenue product equals the marginal cost of labor. Thus, the employer's first-order condition is:

$$\text{Marginal revenue product of labor} = \frac{\partial r}{\partial l} = \frac{w^N}{e^N} = \text{marginal cost of labor}$$

We—and the employer—already know (from M-Note 11.1) the Nash equilibrium values: $w^N = 20$ and $e^N = \frac{1}{2}$. Thus, recalling that $\frac{\partial ln(x)}{\partial x} = \frac{1}{x}$ we can solve for l^N, the Nash equilibrium level of labor (effort) to use:

$$\frac{\partial 20000 ln(l)}{\partial l} = \frac{w^N}{e^N}$$
$$\frac{20000}{l} = \frac{20}{\left(\frac{1}{2}\right)}$$
$$l^N = 500$$

Knowing that $l = eh$, we can use l^N and e^N to solve for equilibrium hours hired:

$$l^N = e^N h$$
$$500 = \frac{1}{2}h$$
$$h^N = 1,000$$

CHECKPOINT 11.3 Effects on the level of labor used and hours hired

Now think about what happens if we change some of the parameters in the model:

a. How does an increase in the cost of a unit of effort affect the amount of labor that the employer devotes to production?

b. Redraw Figure 11.4 to illustrate the change in labor used and the hours of workers' time that the employer hires.

M-NOTE 11.5 Employer's first-order conditions: the general case

We've presented the employer's choice of labor discipline strategy and work hours as separate problems. But, in reality, the profit-maximizing employer solves both simultaneously. To see this, we will work with a general model.

As before, the employer chooses the wage w and hours h, given the worker's best-response function $e(w)$, and gets profit π. We will assume that the

continued

employer is using some given level of capital stock, $\bar{k}$. Holding constant $\bar{k}$ and varying the amount of labor, the profit function of each employer is:

$$\pi = \text{total revenue - total cost}$$
$$\pi = r(e(w) \cdot h) - wh - c_k\bar{k}$$

As we have defined $l \equiv he$, we can rewrite the profit function as:

$$\pi = r(l) - wh - c_k\bar{k}$$

Because the cost of the capital goods used ($c_k\bar{k}$) is a constant it will not affect the profit-maximizing labor discipline strategy and level of hiring. Given the worker's best-response function $e(w)$, employers will now choose w and h to maximize profits. Differentiating profits with respect to wages and hours and setting the results equal to zero, and denoting $r_l = \frac{\partial r}{\partial l}$, we have two first-order conditions:

$$\pi_h = r_l e - w = 0 \tag{11.6}$$
$$\pi_w = r_l h e_w - h = 0 \tag{11.7}$$

In Equation 11.6 the term $r_l e$ is the marginal revenue product of hours of labor, so the equation requires that the level of hiring be such as to equate this to the wage, as is shown in Figure 11.4(b). We can use both equations to derive the Solow condition:

$$\pi_h = r_l e - w = 0 \implies r_l = \frac{w}{e}$$
$$\pi_w = r_l h e_w - h = 0 \implies r_l = \frac{h}{he_w} = \frac{1}{e_w} \tag{11.8}$$

Combining these two expressions

$$\text{for } r_l \text{ we get: } \pi_h = 0 \implies \frac{w}{e} = \frac{1}{e_w}$$

$$\text{Giving the Solow condition:} \quad \frac{e}{w} = e_w \tag{11.9}$$

Equation 11.6 ($\pi_h = r_l e^N - w^N = 0$) can be rearranged to determine the profit-maximizing level labor effort used, shown in Figure 11.4 (a):

$$r_l = \frac{w^N}{e^N} = c_l^N$$

$$\text{marginal revenue product of labor} = \text{cost of a unit of labor}$$

11.5 COMPARING THE INCOMPLETE AND COMPLETE CONTRACTS CASES

To understand better the Ford model of employment under incomplete contracts, we now provide a contrast with a hypothetical case of complete contracting in which the employer can effectively purchase the worker's work, not just her time.

Complete contracting

A complete contract would require that the worker deliver some specified level of effort. This would be approximated by cases in which piece-rate

compensation—mentioned earlier—is possible for a few very routine tasks such as data entry in which a person could be paid by the keystroke.

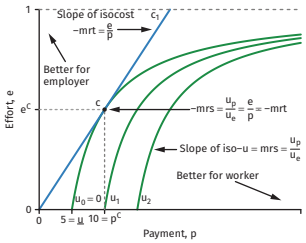
We show the complete contracting case in Figure 11.3 (a) (reproduced in the margin for reference). As in the case of the incomplete contract, the employer would seek to minimize the cost of a unit of effort. So, just as in Figure 11.3 (b), he would seek to be on the steepest isocost ray. But now he would just offer a price for some quantity of effort, rather than hiring the worker by the hour, and he will offer the lowest price consistent with her willingness to sell.

The employer's cost minimization would be constrained by the worker's minimum price at which she is willing to sell effort. To determine this remember that the worker's next-best alternative if not interacting with this employer is to get nothing. So for each level of e provided, the worker's minimum willingness to sell is given by the indifference curve u_0. This is the participation constraint limiting the employer's cost-minimizing attempts. The point on that participation constraint that is on the steepest possible isocost ray is indicated by point **c**, that is, where the isocost ray is tangent to the participation constraint.

Table 11.2 summarizes the contrast between the hypothetical case of complete contracting and the more realistic incomplete contracting case. Concerning the model setup in the first four rows, in both cases the principal maximizes profits, is free to set any price or wage that it wishes, and is competing with other firms.

The major difference in the model setup is that the constraint on profit maximization differs: as you have just seen, it is the participation constraint if the contract is complete and the incentive compatibility constraint (the worker's best-response function) when the contract is incomplete. A major difference between the two models follows from this.

Figure 11.5 The complete contracting case from Figure 11.3.



! REMINDER Point **c** in Figure 11.3 is derived in exactly the same way as that we derived the Nash equilibrium in the Benetton model, that is point **c** in Figure 10.11.

M-CHECK The fact that the level of effort provided in the Nash equilibrium of the complete contracting case, e^c, is the same as in the incomplete contracting case is a coincidence due to the functional forms we are using. In alternative formulations we model, we could have $e^c > e^N$ or $e^c < e^N$.

With incomplete contracts workers receive employment rents

The employer offers the worker a wage high enough that the worker prefers to keep the job rather than lose it, given her fallback option. The reason is that if the principal offered a transaction that was no better than the agent's next best alternative (i.e. a transaction on the participation constraint), then the agent would receive no rent and would not care if the transaction ended. The result would be that the effort supplied by the worker would be whatever she pleased (maybe zero) for there would be no fear of job loss.

The existence and the magnitude of the rent can be seen in Figure 11.3: it is the distance **cn**. To see this, we ask: How much is she better off at the Nash equilibrium than in her reservation position, namely $v = 0$? By comparing the iso-value curve at the Nash equilibrium with her iso-value curve at her

Table 11.2 Complete and incomplete contracts: a summary of model structure and characteristics of the Nash equilibrium. The first four lines refer to model structure while the remaining four refer to results. Remember ICC means incentive compatibility constraint and PC means participation constraint. Notice that the worker is better off (receiving a rent) in the Pareto-inefficient incomplete contract case.

Model characteristics and results	Incomplete	Complete
Employer maximizes	Profits	Profits
Subject to (employer's constraint)	ICC	PC
Competition among firms and workers?	yes	yes
Flexible wages and prices?	yes	yes
Employment rent for worker?	yes	no
Market clearing (no unemployment)?	no	yes
Pareto-efficient Nash equilibrium?	no	yes
Employer's short-side power?	yes	no

participation constraint, we can see that the answer is yes, she does get a rent.

To determine the size of the rent we ask: Supposing hypothetically that she is somehow constrained to work in the Nash equilibrium e^N, no matter how much or little she is paid, how much less could she be paid and still be no worse off than at her participation constraint? The answer is given by the distance **cn** in Figure 11.3.

CHECKPOINT 11.4 Employment rent in two dimensions The distance **cn** is a money measure of the employment rent, but it can also be measured in effort. If the worker is being paid w^N, how much more effort could she provide and still be no worse off than her fallback option? Use Figure 11.3 to show this quantity.

The labor market does not clear: There is excess labor supply

The existence of the employment rent means that the labor market does not *clear*: there are identical workers who are without a job and for whom $v = 0$ who would prefer to be employed receiving v^N, but are unable to make a transaction with an employer.

Those workers unable to make a transaction are *quantity-constrained*, that is, they are unable to purchase or sell as much as they want—as many hours of labor—at the going terms of exchange. The hours supplied by workers at the equilibrium wage exceeds the hours demanded by employers, which means there is excess supply of labor hours or unemployment.

We explained in Chapter 10 that the rent is evidence *that* the market does not clear, not the reason *why* it does not clear. We will explain why later in the chapter.

CHECKPOINT 11.5 Undercutting To check that point **n** in Figure 11.3 (b) is really a Nash equilibrium we need to consider the unemployed, and whether they—like the employer and the worker—are doing the best they can. Suppose an unemployed worker went to an employer and promised to work as hard as a member of his current workforce but at a slightly lower wage. How would the employer respond? Why would this unemployed worker *not* get the job?

Pareto inefficiency

The Nash equilibrium exchange (e^N, w^N) is *Pareto inefficient*. This is true for exactly the same reason that the outcome in the Benetton model of Chapter 10 is Pareto inefficient: the principal in both cases is maximizing profits subject to a best-response function (incentive compatibility constraint), *not* a participation constraint. Given that the contract is incomplete in both cases, the participation constraint is not the relevant limit on the principal's cost minimization.

This is another case of the principal implementing the $mrs = mrt$ rule so as to maximize profits. But you already know that a Pareto-efficient outcome requires the tangency of the isocost lines (representing the principal's objectives) and the iso-value curves (the worker's objectives), so that $mrs^P = mrs^A$ (using superscripts P for principal and A for agent rather than A and B for two actors).

We can see from Figure 11.3 (b) that the Nash equilibrium allocation (e^N, w^N) is not Pareto efficient. The yellow Pareto-improving lens shows possible allocations that like point **f** are both preferred by the employer, who would get more effort for a lower wage, and also preferred by the worker, who would be on a higher iso-value curve. This Pareto-improving lens must exist as long as the isocost ray and the iso-value curve are not tangent: and they cannot be tangent, as the worker's iso-value curve must be vertical if it is on her best-response function and for $w > 0$ the employer's isocost ray cannot be vertical. (We show the mathematics of this statement in M-Note 11.6.)

Exactly the same reasoning applies to aspects of the job other than the wage and the effort level, including workplace amenities such as flexible work hours, or a respectful and safe work environment. The employer will take *some* account of workers' preferences with respect to these amenities, because labor discipline depends on the worker having a lot to lose if she is dismissed. For example, if there is an inexpensive way to make the job more valuable to the worker–like installing air-conditioning, preventing managers from sexually harassing workers, or providing paid parental leave–then the employer will see this as a possible cost-cutting measure. The employer could introduce the inexpensive amenity and as a result be able to reduce the wage without reducing the worker's employment rent. But, the extent of "worker-friendliness" of the job dictated by profit maximization is not Pareto efficient.

! REMINDER The principal is using the *mrt* = *mrs* rule. The participation constraint that limits his optimization is an indifference curve of the worker. The *mrt*—the negative of its slope—is also the worker's relative valuation of effort and pay, that is, her *mrs*. So because the constraint is an indifference curve of the agent, not a best response function, by implementing the *mrs* = *mrt* rule the employer unintentionally also implements the $mrs^P = mrs^A$ rule

To see why, remember that the participation constraint is the worker's willingness-to-pay curve, and its slope—the marginal rate of substitution—is the worker's *own* evaluation of how much she values working less hard and being paid more. Because at point **c** —the complete contract case—the employer's isocost ray is tangent to the workers' participation constraint, it follows that the price of effort to the employer p/e is exactly the cost to the worker of providing that effort, the workers' marginal rate of substitution. This is why the Nash equilibrium is Pareto efficient if the contract is complete.

If the employment contract were complete, the employer would maximize profits by evaluating the importance of workplace amenities (relative to the wage or other things that the worker cares about) *exactly as the worker does*. The participating constraint thus requires the principal to take appropriate account of the effect of his decisions on the worker, thereby avoiding a coordination failure.

But this result does not hold when effort is not subject to contract. The reason is that in this case it is not the participation constraint that limits the employer's optimization process, but instead the worker's best-response function.

Workplace amenities are no different from wages in this respect. We have already seen that the profit-maximizing employer's offer (e^N, w^N) will be Pareto inferior to some other combination of e and w characterized by small increases in effort and wages (the points in the yellow-shaded Pareto-improving lens). The same reasoning applies to working conditions: a small improvement in workplace amenities accompanied by a small increase in effort would be Pareto improving.

M-NOTE 11.6 **The Nash equilibrium is not Pareto efficient**

We can show that the Nash equilibrium is Pareto inefficient using what we know from the first-order conditions of the employer and the worker, both of which must be satisfied for the outcome to be a Nash equilibrium. At the equilibrium, you can see in M-Notes 11.1 and 11.5 that the first-order conditions of the employer and the worker require:

$$\text{Employer sets the wage such that:} \quad \pi_w = 0 \tag{11.10}$$

$$\text{Worker sets the effort level such that:} \quad v_e = 0 \tag{11.11}$$

But it is also the case that:

$$v_w > 0 \tag{11.12}$$

$$\pi_e > 0 \tag{11.13}$$

These equations mean that at the Nash equilibrium values of w and e:

- the employer is indifferent to small variations in the wage (because he set the wage) but the worker strictly prefers a wage increase (Equations 11.10 and 11.12); and

continued

- the worker is indifferent to small variations in the effort level (because she set the effort level), but the employer's profits increase if the worker works harder (Equations 11.11 and 11.13).

Because both statements must be true at the Nash equilibrium, there exists some (sufficiently small) values denoted Δe and Δw (where Δ means a "small change") such that:

- $v(e^N + \Delta e, w^N + \Delta w) > v(e^N, w^N)$; and
- $\pi(e^N + \Delta e, w^N + \Delta w) > \pi(e^N, w^N)$,

where $\Delta e > 0$ and $\Delta w > 0$. So if the two could agree to a small increase in effort accompanied by a small increase in the wage, then both would be better off. But, they cannot realize this Pareto improvement because effort levels are not contractually enforceable.

! REMINDER If, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B, then B has short-side power over A. Short-side power is the power exercised by principals when they are on the short side of a non-clearing market.

The employer exercises short-side power over the worker

Recall two things from previous chapters.

- From Chapter 9 recall that when a market does not clear, the short side of the market—it can be either supply or demand—is the side on which the desired number of transactions is least. So in the labor market where there is excess supply (unemployment) it is the demanders—the employers—who are on the short side.
- From Chapter 10 recall that principals on the short side of a market can exercise power over those on the long side with whom they transact by threatening to terminate the transaction if the agent does not act in ways consistent with the principal's interest.

In the Ford model, the worker works harder than she would in the absence of the threat to take away her enforcement rent. By exerting more effort than she otherwise would, she contributes to the profits of the firm.

To see this go back to Figure 11.3 and suppose the employer set a wage of 5, which is equal to the disutility the worker experiences if she shows up at work and does nothing. The employer also threatens to terminate her employment if he is not satisfied with her level of effort. How hard would she work?

She would not work at all because she would be as well off without a job as with one if the wage is 5. When, instead, he pays her w^N she does work and contributes to the employer's profits. We can conclude then that:

- by threatening to impose a sanction on her (deprive her of her rent)
- the employer induced the worker to do something she would not otherwise have done (work) that

- furthered the employer's interests (raised his profits), and that
- the worker lacked the capacity to advance her interests by threatening the employer.

The last point holds because were she to threaten to terminate her relationship with the employer, he could find an identical worker among those looking for work, so this would inflict no cost on him. This confirms that the employer's relationship to the worker included the exercise of power, as defined.

CHECKPOINT 11.6 **A comparison** Use Figure 11.3 to show that the Nash equilibrium is Pareto inefficient in the incomplete contracting case and that the worker receives a rent, while neither of these is true in the complete contract case.

11.6 EMPLOYMENT RENTS AND THE WORKERS' FALLBACK OPTION

We have so far assumed that the worker's next-best alternative to having her current job is to have utility zero. We now turn to what the worker's fallback option really is.

Employment rents

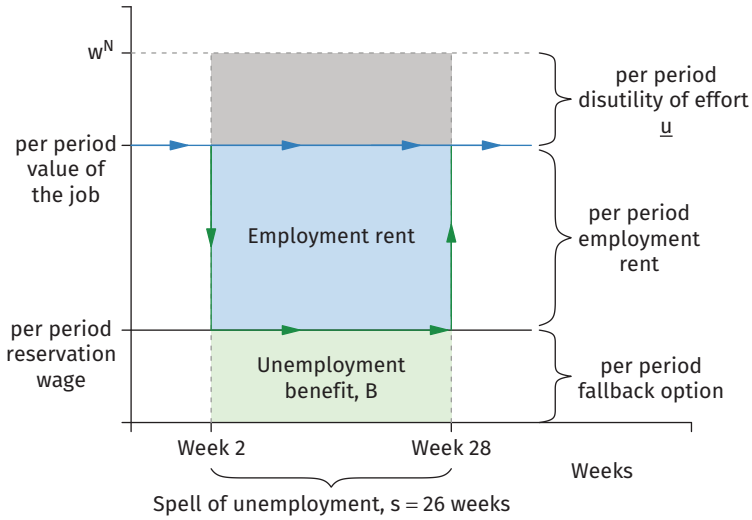
People without work can expect to receive assistance both from friends and family and from the government. We call this assistance the unemployment benefit. Those out of work also search for jobs, and most find employment after some period of time. We are interested in employment rents primarily because they explain the behavior of the employed (not the unemployed). But to do this, think what would happen if you were terminated for insufficient effort.

Imagine that you have a job paying w^N in which you are working at effort level e^N and you think hypothetically about what you would experience were you to lose your job. How would your life be different? A great many things would change, including that you might lose your friendships at work and perhaps be less respected by others; in the US you would most likely lose your health insurance. Here we focus on two important changes:

- *Lost income*: You would no longer have your wage w^N; instead you would have some kind of unemployment benefit B, but this would be less than your wage.
- *More free time and reduced disutility of work effort*: You would no longer be working harder than you would like.

The difference between how much you value what you have as an employed worker and what you *would have* if you lost your job is your *rent*. It is the maximum amount you would be willing to pay not to lose your job.

Figure 11.6 The opportunity cost of working and the employment rent. The blue arrows show the actual course of events of the person who has her job. The green arrows across are her hypothetical trajectory that she imagines would occur were she to lose her job. These two courses of events are identical before week 2 and after week 28, so the employment rent is the difference between the two over the period from week 2 to week 28.



The situation of the worker is depicted in Figure 11.6, which shows:

- the wage that the worker will receive if she remains employed;
- the disutility of effort that she will expend on the job at that wage; and
- the unemployment benefit per week that she would receive if her job were terminated.

We track over time both the employed worker and the hypothetical same worker were she to lose her job, considering the case in which her job is terminated at the end of week 2, and she experiences a spell of unemployment of 26 weeks, after which she finds another job on terms identical to the job she lost.

To calculate her rent per week on the job we compare two terms:

- the *value of her job* which is w^N if this is paid weekly minus $\underline{u}$ the disutility of effort when working at the level e^N for a week; and
- her fallback option, which is the weekly unemployment benefit B; also termed her *reservation wage*, as it is the least wage at which she would agree to take a job (meaning, to show up but not work). The difference between the two is her:

$$\text{Per week employment rent} = w^N - \underline{u} - B \tag{11.14}$$

✓ **FACT CHECK** The average length of a spell of employment in the US in September 2020 (half a year into the COVID-19 pandemic) was a little over five months. The expected length of a spell of unemployment varies over the business cycle. In the US just prior to the recession associated with the global financial crisis of 2008, for example, about one in six of the unemployed had been jobless for 27 weeks or more; two years later almost half of those without work had been unemployed that long.[8]

If she anticipates being out of work for a spell of s weeks, then her rent is:

$$\text{Worker's rent} = s(w^{N} - \underline{u} - B) \tag{11.15}$$

Losing your employment rent can provide strong motivation to act in accordance with your employer's wishes.

Employment rents are substantial

Estimating the size of the employment rent is a challenge. To do this we cannot simply compare the economic situation of workers currently employed with the unemployed because the unemployed are different people who, if employed, on average would earn less than those currently with jobs.

An entire firm closing or a mass layoff of workers provides a natural experiment that can help. When a factory closes because the owners or managers have decided to relocate production to some other part of the world, for example, virtually all workers lose their jobs not just those who might be particularly "unemployment prone."

Louis Jacobson and his coauthors exploited a natural experiment to estimate the magnitude of employment rents.[9] They studied experienced, full-time workers hit by mass job losses in the US state of Pennsylvania in 1982. In 2014 dollars, those displaced had been averaging about $55,000 in earnings in the year prior to losing their jobs. Workers who were fortunate enough to find another job less than three months after they lost their job took worse-paying jobs, averaging only $35,000: a loss of $20,000 in the first year after the firings. So our hypothetical story about the worker imagining a spell of unemployment followed by a return to an equally good job does not reflect what these workers experienced.

Four years later they were still making $12,000 less than other workers who had been making the same initial wage, but whose firms did not have mass firings. In the five years that followed their layoff they lost the equivalent of an entire year's earnings.[10]

! REMINDER in Chapter 3 we looked at studies measuring subjective well-being or "happiness." The studies here use the same measures to show loss of happiness, depression, and more from losing a job.

Another challenge in measuring employment rents is that taking account of how people feel about losing a job and being unemployed, rather than just the monetary losses, the rents may be considerably larger. The reason is the social stigma, indignity, and unhappiness resulting from being without work. One study using a data set allowing comparisons of the same individual when she is with and also without work, found that the amount of income that would compensate typical individuals for the social esteem and other costs of being out of work is actually greater than the income loss itself. So the true cost was more than twice the income loss.[11]

It is no surprise, then, that as the expected duration of a spell of unemployment increases reducing workers' fallback option and boosting job rents, workers work harder.

Evidence on job rents, unemployment, and work effort

Edward Lazear, chief economic adviser to former US President George W. Bush, investigated a single firm during the global financial crisis of 2008, to see how the workers reacted to the sharp increase in unemployment. The firm specializes in technology-based services, such as insurance-claims processing, computer-based test grading, and technical call centers, and operates in 12 US states.

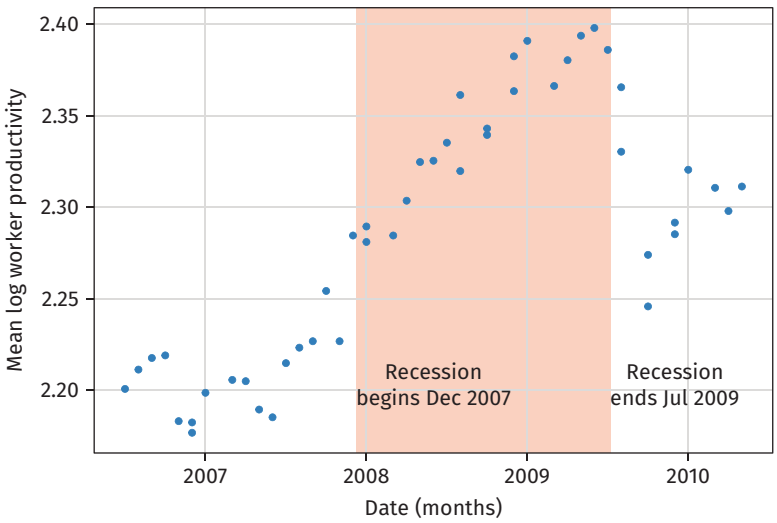
The nature of the work made it possible for the management of the firm to track the productivity of workers, which is a measure of worker effort. It also allowed Lazear and his colleagues to use the firm's data from 2006–2010 to analyze the effect on worker productivity of the sharp rise in unemployment during the global financial crisis starting in 2007.

Figure 11.7 shows the results. Productivity increased markedly as the duration of spells of unemployment rose during the financial crisis. This was particularly the case for branches of the firm located in places that experienced an especially large increase in unemployment.

This raises a question: As the expected spell of a bout of unemployment rose during 2008, employers could have cut wages and still maintained sufficient employment rents to motivate workers to work as hard as before. Productivity would not have risen but labor costs would have fallen due to the decrease in wages. Why did they not cut wages?

Figure 11.7 Labor productivity before, during, and after the global financial crisis. Monthly data are from a single firm in technology-based services during the period 2006–2010 where productivity can be measured by completion of computerized tasks. Productivity is measured as the log of mean worker output per hour. For workers with a job, productivity increases after the start of the recession and starts to decrease again after the recession ends.

Source: Fig 2. from: Lazear, E. P., Shaw, K. L, and Stanton, C., (2016) "Making Do with Less: Working Harder during Recessions," *Journal of Labor Economics* 34:S1, S333–S360. DOI: 10.1086/682406. Reproduced with permission.



Another economist, Truman Bewley wanted to know why employers typically do not cut wages even during recessions. He interviewed more than 300 business people, labour leaders, business consultants, and careers advisers in the northeast of the US. He found that employers chose not to cut wages because they thought it would induce an angry push back from workers, who would then work less hard, raising the cost of labor (meaning actual work done). The employers thought it would ultimately cost the firms more than the money they would save from lower wages.

CHECKPOINT 11.7 **The worker's fallback option** Why (in Equation 11.15) is the disutility of effort $\underline{u}$ subtracted from the difference between the weekly wage w and the weekly unemployment benefit B to get the employment rent?

11.7 CONNECTING MICRO TO MACROECONOMICS: A NO-SHIRKING CONDITION

The work speedup in response to the recession of 2007-2009 makes it clear that macroeconomic factors alter the environment in which the firm selects a labor discipline strategy and workers respond. This occurs because the macroeconomy affects the fallback option of workers.

We now introduce a model in which the worker's fallback option depends on the level of employment in the economy as a whole, and the unemployment benefit. The microeconomic problem of labor discipline—from the employer's perspective, getting workers to work hard and well—provides the basis for a model that connects wage-setting, work effort (and resulting productivity), and profits to the level of unemployment, unemployment insurance, and other aspects of public policy.

Incentive compatibility in the "no-shirking game"

The setup is as follows. The employer, who as before is the principal and first mover sets a "no-**shirking**" level of effort $\underline{e}$, and announces that he will terminate the worker without pay if she is detected providing less. The employer then figures out what wage is necessary to motivate the worker to work at the no-shirking level.

If there was a complete contract in effort, then the least "price" the employer could offer for $\underline{e}$ would be the unemployment benefit the worker would have received when not working (B) plus just enough more to compensate the worker for the disutility of providing the effort $\underline{u}(\underline{e})$ required by the employer. This is the worker's willingness to work.

$$\text{Willingness to work: } = B + \underline{u}(\underline{e}) \tag{11.16}$$

SHIRKING When a worker does not work as hard as the employer requires, economists call this "shirking."

This is the least the employer could pay such that if the worker worked at the specified level of effort, she would be no worse off than if she had not accepted the job (not "participated"). So we have:

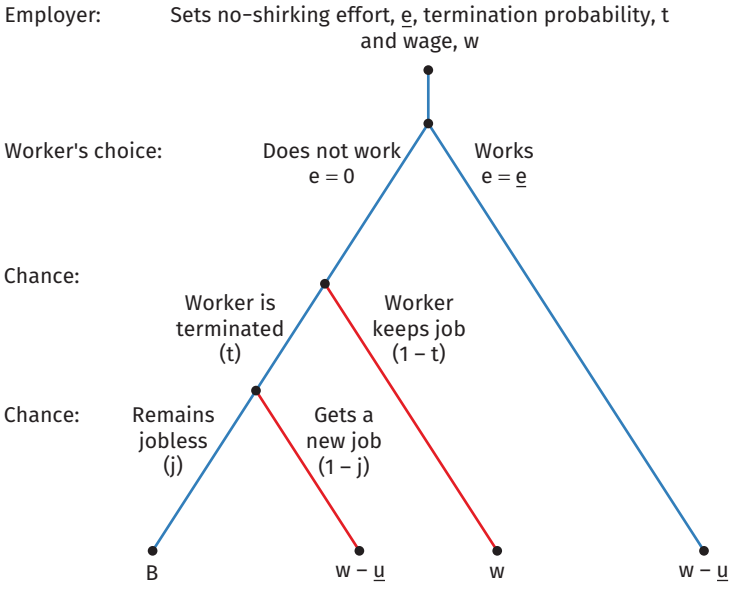
$$\text{Participation constraint: } w \geq B + \underline{u}(\underline{e}) \tag{11.17}$$

But because effort cannot be purchased under a complete contract and the employer cannot cost-effectively monitor all of the workers all of the time, the worker may be able to exert no effort at all and nonetheless get paid. So the employer has to provide incentives for the worker to work. Here is the game (as described by the game tree in Figure 11.8).

- The employer announces a wage, a "no-shirking" level of effort $\underline{e}$, and a system of monitoring her work that results in a probability t that the worker will be terminated if she does not provide $\underline{e}$.

REMINDER Earlier you studied incentive compatibility constraints expressed as a best-response *function* which gave for each level of effort the least wage that would motivate the worker sufficiently to provide it. Here there is just one level of effort—$\underline{e}$—so the ICC is an equation giving us just a single value, the least wage sufficient to motivate the worker to provide the no-shirking effort level when it cannot be enforced by a complete contract.

Figure 11.8 The sequence of moves in a one-shot game determining the payoffs to shirking. In response to the employer's statement of the no-shirking level of effort $\underline{e}$, the termination probability t and the wage w, she decides whether to exert $e = 0$ or $e = \underline{e}$. If she exerts effort $e = \underline{e}$, she obtains the payoff for working and incurring disutility of effort, $w - \underline{u}$. If she exerts no effort, $e = 0$, then there is a probability t she will be terminated and a probability $1 - t$ she will not. If she is not terminated, she gets the wage w (and incurs no disutility of effort). If terminated, she will remain jobless (be unemployed) with the probability j and get unemployment benefit, B, or she will be re-employed with the probability $1 - j$ and receive the same value of the job that she would have experienced had she not been terminated, $w - \underline{u}$.



- The worker decides to provide $\underline{e}$ or not (in which case she does not work at all, that is $e = 0$).
- If she provides $\underline{e}$ then with certainty she will not be terminated and will receive the wage w and experience the disutility $\underline{u}(\underline{e})$.
- If she provides less than $\underline{e}$, then one of two things may occur: with probability t, she is detected and terminated, or with probability $(1-t)$ she escapes detection and is paid w.
- If she is terminated, one of two things may occur: with probability j she does not find another job and receives the unemployment benefit B; or she finds another job with probability $1-j$.

This gives us the four possible paths through the game tree resulting in the four outcomes at the ends of the branches. In the last case (she is terminated but immediately is re-employed), the job that she gets is identical to the job from which she was terminated (workers and employers are identical). To see what this payoff is, we need to know the wage offered by the employer.

The no-shirking wage

This is called the **no-shirking wage**, that is, the smallest wage the employer can pay that will motivate the worker to supply effort of $e = \underline{e}$. To determine the no-shirking wage, the employer considers two numbers:

- *the worker's payoff if she does not shirk*: the wage she gets minus the disutility of effort $w - \underline{u}$; and
- *the worker's expected payoff if she does shirk*: this depends on the probability she gets fired (t), if fired, the likelihood she will remain jobless (j), and her unemployment benefit (B) if she remains without work.

To motivate the worker to work, the employer therefore needs to offer a wage high enough so that the payoff from working is not less than the expected payoff from not working, that is:

$$\text{Payoff to working at } e = \underline{e} \geq \text{Expected payoff to shirking at } e = 0$$

$$\text{ICC: } w - \underline{u} \geq (1-t)w + t(1-j)(w - \underline{u}) + tjB \tag{11.18}$$

Equation 11.18 is the *incentive compatibility constraint* ICC for the employer. As we have done in other models, we assume that the worker will provide the no-shirking level of effort if it is *as good as* shirking, so Equation 11.18 will be satisfied as an equality (not an inequality).

This means that $w - \underline{u}$ is the value of the job to the worker whether she decides to work at the level required by the employer (the left-hand side of

NO-SHIRKING WAGE The no-shirking wage is the wage that is just sufficient to motivate a worker to provide effort at the level specified by their employer.

the equation) or not (the right-hand side). So if a terminated worker finds another job, this is its value, as is shown in Figure 11.8 at the end of the did-not-work, terminated, found-another-job branch of the tree.

The right-hand side of Equation 11.18 is from the left hand ("shirking") branches of the game tree in Figure 11.8 and comprises the following probabilities, all given the fact that she has shirked:

- $(1-t)w$: *Not terminated*: She is paid the wage (w), weighted by the probability that she will not be terminated $(1-t)$ even though she did not work; plus
- $t(1-j)(w-\underline{u})$: *Terminated, immediately finds another job*: She is both fired and also gets a new job the value of which is $(w-\underline{u})$; this occurs with probability $t(1-j)$; plus
- tjB: *Terminated, remains jobless with unemployment benefit*: This is the unemployment benefit she will get (B) multiplied by the probability that she gets terminated *and* remains unemployed.

We can rearrange Equation 11.18 as a restatement of the incentive compatibility condition (ICC) called the *no-shirking condition*. To do this we isolate all the w terms (shown in M-Note 11.7) to find the no-shirking wage:

$$\text{Worker's ICC:} \qquad w^{\mathrm{N}} = \mathrm{B} + \underline{u} + \frac{1-t}{tj}\underline{u} \tag{11.19}$$

$$\text{No-shirking wage} = \text{Willingness to work} + \text{Employment rent}$$

Equation 11.19 tells us that the no-shirking wage will be higher:

- the greater is the unemployment benefit B;
- the greater is the disutility of effort $\underline{u}$;
- the smaller is the probability that shirking will be detected t; and
- the smaller is the probability that a terminated worker will remain jobless j.

We use the N superscript (for Nash equilibrium), because given:

- the worker's disutility of providing effort $\underline{u}$,
- her chance of being detected and fired if she does not work t,
- her chance of remaining without work if she is terminated, j
- and her unemployment benefit if she remains jobless, B

the employer setting the wage w^{N} and the worker providing effort $\underline{e}^{\mathrm{N}}$ is a Nash equilibrium because:

- w^{N} is the least wage the employer can offer consistent with the worker working (it minimizes the cost of effort)
- given the wage offer the worker cannot do better than to provide effort $\underline{e}^{\mathrm{N}}$.

M-CHECK If the fact that one of two possible events occurs does not influence whether the other occurs (the events are independent), then the probability that both will occur—called their joint probability— is $P_1 \times P_2$ if the two events will occur with probability P_1 and P_2. So the probability that both of these two events occur—she is terminated with probability t and she remains unemployed with probability j—is the product of the two probabilities, tj. This is called the joint probability of the two events.

Because both worker and employer are doing the best they can given the strategy adopted by the other, the no-shirking wage (w^N) and the employer's non-shirking standard ($\underline{e}^N$) are a Nash equilibrium.

M-NOTE 11.7 Rearranging the ICC to find the no-shirking condition

To find the no-shirking condition, we can rearrange the incentive compatibility constraint for the employer, Equation 11.18:

$$\text{ICC: } w - \underline{u} \geq (1-t)w + t(1-j)(w-\underline{u}) + tjB$$

The employer will pay the worker just enough for the worker to exert effort rather than shirk, and no more. So we will express the ICC as an equality:

$$w - \underline{u} = (1-t)w + t(1-j)(w-\underline{u}) + tjB$$

Collect all terms with w on one side:

$$\begin{aligned} w - (1-t)w - t(1-j)w &= t(1-j)(-\underline{u}) + tjB + \underline{u} \\ \text{Factor out } w\text{: } w\big(1-(1-t)-t(1-j)\big) &= t(1-j)(-\underline{u}) + tjB + \underline{u} \\ \text{Simplify: } w(1-1+t-t+tj) &= -t\underline{u}(1-j) + tjB + \underline{u} \\ wtj &= -t\underline{u} + t\underline{u}j + tjB + \underline{u} \\ wtj &= tj(B+\underline{u}) + \underline{u}(1-t) \end{aligned}$$

Solve for w:

$$w^N = \frac{tj(B+\underline{u}) + \underline{u}(1-t)}{tj}$$

$$w^N = B + \underline{u} + \frac{1-t}{tj}\underline{u}$$

This is the no-shirking condition shown in the text (Equation 11.19).

CHECKPOINT 11.8 No-shirking wage Consider Equation 11.19. What will happen to the *no-shirking wage* if the following changes occur?

a. The probability of getting caught and terminated (t) decreases?

b. The probability (j) that a worker will remain unemployed increases?

c. The unemployment benefit (B) increases?

d. The disutility of effort ($\underline{u}$) increases?

! REMINDER In Chapter 10 (section 10.13) we did a similar analysis on the distribution of gains between the agent (the seller) and the principal (the buyer, Benetton in the example). We extend that model here by taking account of how differing macroeconomic conditions affect the agent's fallback options.

11.8 INCOMPLETE CONTRACTS AND THE DISTRIBUTION OF GAINS FROM EXCHANGE

We can now see how the extent of contractual completeness affects the distribution of rents between the employer and the worker.

A measure of contractual completeness is the likelihood that a worker who does not work (sets $e = 0$) will be detected and terminated. To see how the completeness of the contract affects the extent of the worker's rent, return to Equation 11.19. Suppose hypothetically that the contract were complete, so $t = 1$: the worker would receive no employment rent. The employer would pay a wage just sufficient to satisfy the worker's participa-

tion constraint, given by Equation 11.17. This is because if the contract is complete the worker is not renting her time but instead effectively selling her effort, so the employer can pay the worker her minimum willingness to sell her effort. This is another illustration of the fact that when contracts are complete the participation constraint and the incentive compatibility constraint coincide.

To study the distribution of rents between the employer and the worker suppose that the per-period output produced by the worker providing the no-shirking level of effort is γ which can be sold for a price of 1 Euro. We assume there is no other input than labor effort, so wages paid is the employer's only cost. Then the wage divides the average revenue from one period of "no shirking effort" into the worker's per-period wage and the employer's per-period economic profit.

The total gains from the interaction between the employer and the worker are the worker's employment rent and the employer's profit. Both are rents, that is payments in excess of the actor's next-best alternative, which, in the case of the employer, we assume, is to receive zero profits. To compare the two rents we can rearrange the no-shirking condition to isolate the worker's employment rent:

$$w^{\mathrm{N}} - \underline{u} - \mathrm{B} = \frac{1-t}{tj}\underline{u} = \frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right) = \text{Employment rent} \tag{11.20}$$

Figure 11.9 shows how mutual gains generated by the worker's effort is divided. We can see that the more incomplete is the contract the larger will be the share of the revenue that goes to the wage (composed of the worker's willingness to work plus the employment rent).

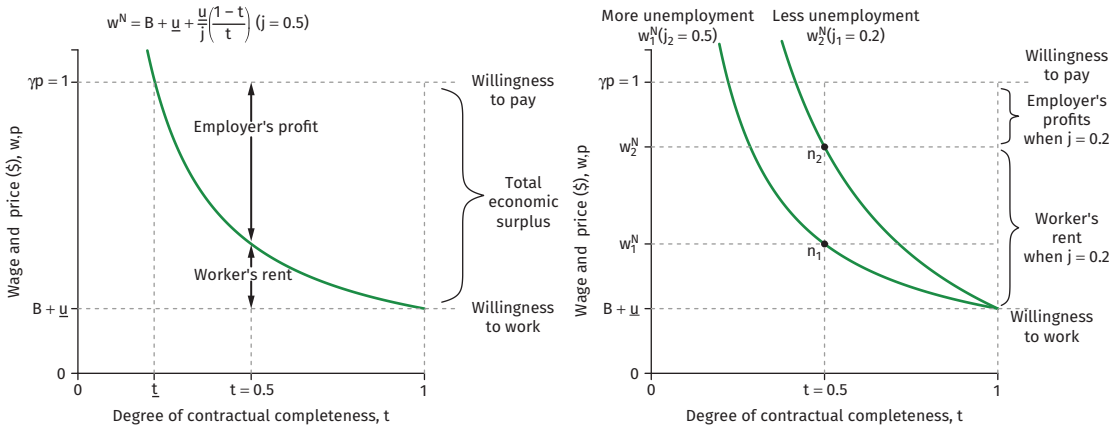
We can also see that there is some low level of contractual completeness $\underline{t}$ below which no mutually beneficial interaction between the worker and the employer is possible. The reason is that for $t < \underline{t}$ the no-shirking wage would exceed employer's revenue made possible by the workers' effort, that is $w^{\mathrm{N}} > \gamma p$. As a result, there would be no way that the employer could pay a wage sufficient to get the worker to work and then to sell the resulting product at a profit. As a result, no workers would be hired.

Comparing the two curves in panel (b), we can also see the effect of labor market conditions on the distribution of rents to the employer and the worker. If there is substantial unemployment so that the probability that the terminated worker remains without a job is 0.5 (as in panel (a), for the lower curve in the panel (b)) the employer gets most of the rent.

When that probability of remaining jobless falls to $j = 0.2$, the worker is less concerned about finding work if she is terminated, so the employer must pay her more to ensure that she works at the no-shirking level. In this case most of the rents go to the worker, not the employer.

CHECKPOINT 11.9 Incomplete contracts and the distribution of gains from employment Why is t a measure of how complete the contract is, and why does the worker's rent become larger as t falls? Why would the worker not prefer a situation in which $t < \underline{t}$?

Figure 11.9 Degree of contractual completeness and employment rents. The worker's minimum willingness to sell (provide the no-shirking effort level) is $B + \underline{u}$. The employer's maximum willingness to pay is $p\gamma$. Remember from Equation 11.20 that the worker's rent is $\frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right)$. In panel (a) the likelihood that the worker will remain jobless if they are terminated is $j = 0.5$. The figure shows for a given termination probability ($t = 0.5$) what the distribution of the surplus between the worker and employer will be given those values. Panel (b) shows what happens if the probability of remaining jobless is instead $j = 0.2$. Economic and accounting profits are the same in this case because labor is the only input into production and it is paid at the same time that the goods produced are sold, so the employer does not commit any funds in advance (as would be the case if the worker used some kind of capital goods).



(a) The distribution of total economic rent

(b) A decrease in the chances of not finding a job increases worker rents

11.9 APPLICATION: CONTRACT ENFORCEMENT TECHNOLOGIES

Charlie Chaplin's Great Depression film *Modern Times* ridiculed the efforts of "time and motion" men with stopwatches and clipboards trying to speed up the pace of factory work. Today, it is getting easier for employers to monitor how much effort a worker is putting in, at least in some jobs. The "items per minute" of checkout staff can be checked in real time by supervisors just by looking at their computer monitor (17 items per minute is the minimum to keep your job at one outlet).

Software called a 'keylogger' can record the keystrokes of data-entry personnel. In Chapter 12 you will read about the device installed in cars financed by an auto that allows the lender to remotely disable the starter of the car if loan repayments fall behind. These devices can make a contract more complete either by covering more of what the employer wants (keystrokes, items checked out) or by making the contract more enforce-

able (the remote ignition disabler). You know from Figures 10.13 and 11.9 that improving the contract in these ways will increase the share of the gains from exchange that will go to the principal—the buyer in the Benetton model or the employer in the Ford model. Correspondingly, as a result, a lesser share of the rents go to the supplier or the worker.[12]

A similar device has been in use now for decades—trip recorders on trucks. Trip recorders vastly improved the ability of employers to monitor the actions of the truck drivers.[13] The devices provided the company with verifiable information on the speed, idle time, and other details of the operation of the truck about which there was a conflict of interest between the driver and the company. For example, the cost of operating the trucks (paid by the company) is higher the faster the trucks are driven. Drivers preferred to drive faster than the cost-minimizing speed, and to take longer breaks. Drivers who owned their trucks—called owner-operators—were **residual claimants** on their revenues minus costs and so they internalized the costs of fuel and depreciation, driving at the cost-minimizing speeds and realizing significant savings as a result. Based on their lower costs, prior to the innovation of trip recorders, owner-operators successfully competed with company fleets, whose drivers were employees not owners.

When the trip recorders came in, companies were able to write contracts based on the speed at which the truck was driven, and to provide drivers other incentives to act in the companies' interests. By improving the companies' contractual opportunities the use of trip recorders induced drivers in trucks with recorders to drive slower and reduced the costs of operation of the company owned trucks. The result was a substantial decline in the market share of owner-operators.

The innovation of trip recorders therefore resulted in:

- **Less competition**: There was greater market concentration in the trucking industry, fewer firms and fewer owner-operators, and so less competition.
- **Inequality** Many owner-operators could previously make a good living, constituting a kind of 'middle class' of the trucking industry with incomes greater than the drivers employed by the large companies, but much less than the incomes of the owners and managers of those companies. By making possible a more complete contract between the companies and the drivers, the trip recorders contributed to inequality in the industry.

RESIDUAL CLAIMANT The residual claimant is whoever gets what is left over (the residual) from the revenue (or other benefit) of a project when all of the costs that have been contracted for are paid.

Unlike other on-board computers (the electronic vehicle managements systems, or EVMSs), the trip recorders provided no improvement in coordination between truckers and dispatchers, as the information was available to the company only on the completion of the trip. The sole purpose of the trip recorders was to make the contract more complete with respect to drivers' behaviors in which there was a conflict of interest between the drivers and the companies.

Monitoring knowledge work

Employers now use a vast array of software to monitor the work activities of employees: ActivTrak, InterGuard, Veriato 360, Teramind, WorkSmart, Prodoscore, and Work Examiner are all software packages used to track the activities of workers on-site in company buildings or at home when they log in to company servers.[14]

The software may monitor your use of email, messaging, browsing, and access to social media. Or, it may take a screenshot of your workstation and use your webcam to ensure you're at your workstation. The software will continue to take photographs every ten minutes to ensure you're on task, keep track of your keystrokes as you type, and monitor the movements of your mouse. When many white-collar workers were isolated at home during the COVID-19 pandemic Veriato advertised that its devices could provide their employer with information on "What hours employees are working, how much they're working, what they're spending their time on."

Though none of these technologies can provide verifiable information on all of the dimensions of effort, they provide data that employers can use to assess what kinds of worker activities correspond to worker productivity. In so doing, they make the employment contract more complete, re-allocating rents from the worker to the employer.

Monitoring and trust: When monitoring backfires

Another challenging facet of this problem is that workers respond differently when they are monitored relative to when they are allowed to remain autonomous and make their own decisions about their productive tasks. For example, a survey of German citizens during the COVID pandemic in 2020 found that more were willing to support government advice about social distancing, getting a vaccination, and installing a tracing phone app if compliance was voluntary than if it was enforced by law. The "control-averse" Germans tended to be those who did not trust the government or believe its scientific reports.

Economists have researched the choice of monitoring methods, something we have not explored here (as we just assume that termination is lower if you work hard or provide high-quality goods, without exploring how the employer acquires the relevant information). The research shows that when principals monitor agents closely and impose minimum quotas of

work the principals tend to activate workers' negatively reciprocal feelings and sow distrust among their workers. As a result, close or intrusive monitoring may backfire. Workers value their autonomy in nonmonetary ways, and so taking away that value reduces the value of the job to the worker and, as a result, diminishes the force of the employer's labor discipline strategy.

CHECKPOINT 11.10 Contract enforcement technologies Other than trip recorders, think of other technologies that allow employers greater information on the effort levels of their workers (whether or not the information is verifiable).

✓ **FACT CHECK** Evidence reviewed in Chapter 16 suggests, furthermore, that workers especially value knowing that their employer *could have* monitored them or penalized them in some way, but then refrained from doing so. In this way, workers reciprocate trust with effort, and distrust with shirking.[15]

11.10 EQUILIBRIUM UNEMPLOYMENT AND THE WAGE CURVE

We developed the no-shirking version of the Ford model by looking at the relationship between a single employer and a single worker. But that single worker is a member of a team of identical workers working for the employer. And we can extend the model further to represent the economy–wide labor market.

The no-shirking variant of the Ford model is connected to the macroeconomy because the likelihood of remaining unemployed if fired (j) depends on the level of unemployment in the economy–wide labor market. The more unemployed people there are looking for work, the less is the chance that any one worker finds a job.

The economy-wide wage curve

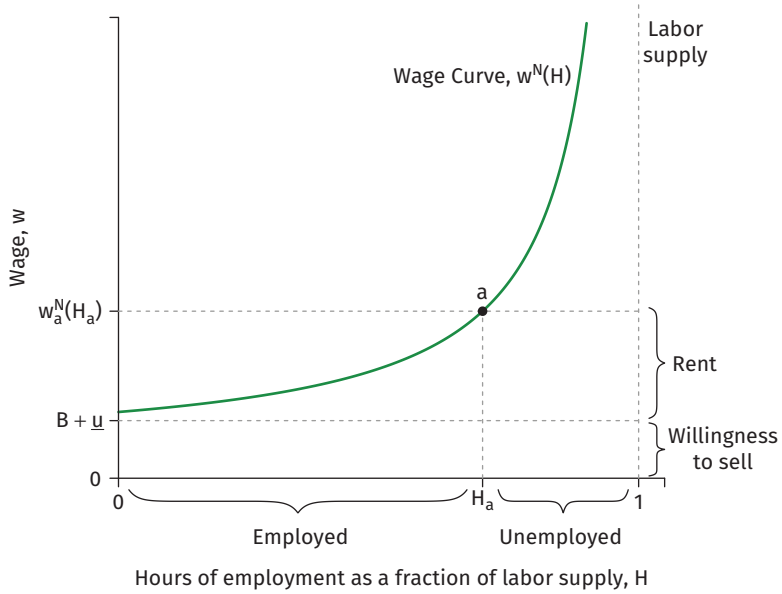
To represent this in our model we can write the probability of remaining jobless, j, as a decreasing function of the fraction of the labor force that is employed or H. If labor supply is 1 and $H = 1$ this means that everyone seeking work has a job, and $H = 0$ means that nobody is employed. Therefore, we write the probability of remaining jobless as $j = j(H)$ and we simplify by letting $j = (1 - H)$ so if everyone looking for a job is employed $H = 1$ then the terminated worker will immediately get a job with certainty ($j = 0$). (In M-Note 11.11 we show how a relationship between j and H (j decreasing as H increases) can be derived using information of how frequently people quit their jobs and how employers find new workers.)

To apply the no-shirking wage to an entire economy, we substitute the expression $j = (1 - H)$ into the no-shirking wage equation (Equation 11.19) to get:

$$w^{N} = B + \underline{u} + \frac{1-t}{t(1-H)}\underline{u} \tag{11.21}$$

We show the resulting economy-wide no-shirking wage in Figure 11.10. This is called the *wage curve*.

Figure 11.10 The economy-wide wage curve. The no-shirking wage $w^N(H)$ is the workers' willingness to work $B + \underline{u}$ plus the workers' employment rent. Notice that as unemployment falls (H rises) the rent required to motivate workers to provide effort increases. For a given fraction of employment, H_a, the corresponding wage will be $w_a^N(H_a)$ as shown by point **a**. Remember from Equation 11.19 that the worker's rent is: $\frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right)$ from which you can see that when $H = 0$ (so $j = 1$) we have the Nash equilibrium wage: $w^N(0) = B + \underline{u} + \frac{\underline{u}(1-t)}{t}$, which is greater than the workers' willingness to work, unless the probability that shirking will be detected, t, is equal to 1.



The wage curve exists

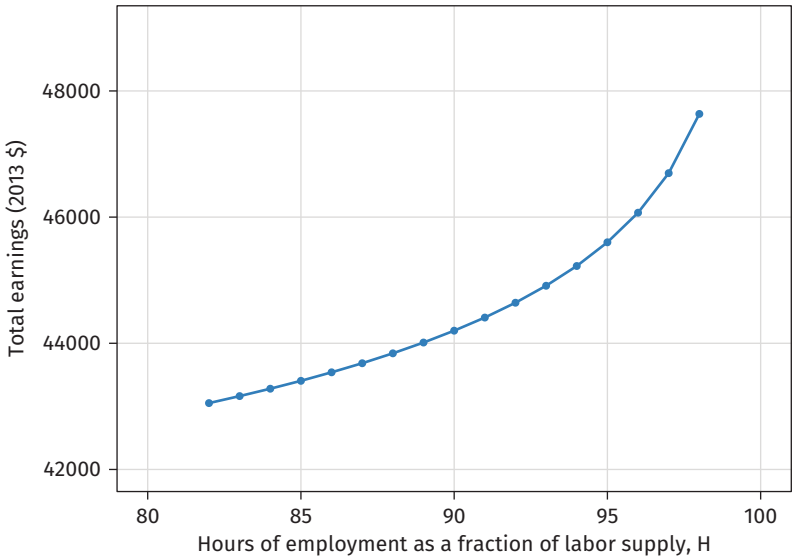
By using data on unemployment rates and wages in local areas and over different periods of time, economists can estimate and plot the wage curve for an economy. Real wages tend to vary with the level of employment as the labor discipline model predicts.[16] An example is shown in Figure 11.11, a wage curve estimated from data for the US.

As the labor discipline model predicts, Figure 11.11 shows that workers do better when more workers are employed–the wage curve is upward-sloping–and the effect of limited unemployment in pushing up wages is stronger, the less unemployment there is–the curve is steeper closer to $H = 1$.

Workers do better when unemployment is low not only because fewer workers experience unemployment, but also because those who are employed (the vast majority) are receiving higher wages. The level of

Figure 11.11 An economy-wide wage curve estimated for the US economy: 1979–2013. The sample of 2.34 million workers are males aged 26–64 over the years 1979–2013. Earnings are in 2013 dollars.

Source: CORE. The Economy.[17]



employment in the economy as a whole therefore affects the distribution of income between workers and employers.

Equilibrium unemployment: There is no market-clearing wage

Recall that in section 11.5 we said that the existence of rents in the labor market is *evidence that* the labor market does not clear. It does not explain why in the Nash equilibrium there will be excess supply—people seeking work at the going wage and not finding it. Here we explain why.

Equation 11.21 and Figure 11.10 give us the information we need to consider the question: In this model can there be an equilibrium in which the labor market clears?

The answer is no. The reason is *not* any of the following:

- *"Sticky" or inflexible wages*, so that workers' pay cannot adjust to excess supply of labor, that is, unemployment. Wages do adjust, which is what the wage curve shows.
- *Lack of competition* so that the unemployed cannot compete for jobs with the employed. They can, but as we saw, their offers to work as hard as the current workers for less pay will not be accepted by employers. (Review your response to Checkpoint 11.5.)

- *Trade unions* that sustain high wages making it unprofitable for firms to hire the unemployed. There are no trade unions in this model.

The reason why there will be unemployment in the equilibrium of this model is that a market clearing wage does not exist.

It appears from Figure 11.10 that as H approaches 1–market clearing, that is, no unemployment–the wage curve becomes nearly vertical. We can confirm that this is the case by evaluating the no-shirking wage when unemployment $(1-H)$ is zero, that is, $j = 0$:

$$w^{N} = B + \underline{u} + \frac{1-t}{t \cdot 0}\underline{u} \tag{11.22}$$

Equation 11.22 shows us that if H = 1 we would have to evaluate the wage, $w^{N}(1)$, with division by zero. So the wage would have to be infinite to motivate workers to supply effort if everyone who wants a job had one. This is impossible and so full employment is impossible in our model.

Equation 11.22 also allows us to understand why in the labor discipline model labor markets cannot clear in equilibrium, that is, why supply of labor cannot equal demand for labor in a market with incomplete contracts.

To see why this is the case let's imagine that in an economy composed of many identical workers and employers, labor supply equaled labor demand, so the markets cleared and there is no unemployment. The reasoning progresses as follows:

- *No unemployment*: *If* the labor market *clears*, supply of labor equals demand for labor and there would be no involuntary unemployment.
- *Immediate re-hiring*: Without unemployment, anyone who is terminated for insufficient work could immediately find a new job so $j = 0$.
- *Impossible infinite wage*: But, from Equation 11.19 or Equation 11.22 with no joblessness, $j = 0$, the lowest wage that would deter shirking would be $w^{N} = \infty$, that is, an infinite wage.
- *Firms shut down and create unemployment*: Firms cannot afford an infinite wage, so firms, shut down, put people out of work, so there would be unemployment.
- *Contradiction*: The existence of unemployment contradicts the initial premise that labor markets cleared.

The contradiction at the end of the chain of reasoning shows that labor markets cannot clear in the labor discipline model.

CHECKPOINT 11.11 The impossibility of full employment Use Equation 11.22 to explain why employment cannot be equal to labor supply in the Nash equilibrium of this model. In other words, why will there be an excess supply of workers?

11.11 THE WHOLE-ECONOMY MODEL: PROFITS, WAGES, AND EMPLOYMENT

If labor market clearing is not what determines the wage w^N and employment level H, what does? To answer the question we need to clarify some terms. In this chapter we use w or "the wage" to mean the real wage, or w^n/p, the nominal wage (w^n) divided by the price level (p). When we say that employers "set the real wage" we mean that they set the nominal wage for some given level of prices, giving the real wage w.

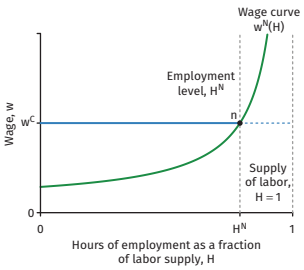
M-CHECK We want to determine two outcomes, the real wage w and the equilibrium level of employment H^N. But so far we have just a single equation, the wage curve. To determine the values of two variables we need two equations. The competition condition provides the second equation.

But to determine the level of the real wage and employment of the entire economy—not just a single firm—we use a model of product markets and goods prices as well as labor markets. The reason is that the real wage depends on the output and pricing decisions of the firm's owners, as well as their decisions setting the money wage. The Nash equilibrium real wage and employment level must be the result of best responses not only for workers and their owners as employers, but also for the owners as *sellers* of the products produced.

The wage curve gives us all of the possible Nash equilibria in the labor market. But the wage curve—by itself—cannot determine which of these combinations of w and H will occur.

To do this, we need one more piece of information. Think of this as another line in the wage curve figure, as shown in the margin as Figure 11.12 whose intersection with the wage curve will tell us which of the possible combinations of wages and hours of employment will be a Nash equilibrium.

Figure 11.12 Looking for a second equation in the wage curve figure.



To find this addition to the figure, we need an answer to the following question: Given the level of labor productivity and the extent of competition in product markets, what is the wage that will result in the number of firms neither increasing nor decreasing? This is the wage such that total employment in the economy is *constant*.

We will discover that there is only one such wage.

Barriers to entry and the competition condition

We use the model of competition in Chapter 9 to study a long-run equilibrium in which the number of firms is neither increasing nor decreasing. Because some firms relocate or cease to exist for reasons unrelated to the model, new firms must enter for the numbers of firms to remain constant. As you saw in Chapter 9, attempting to enter a market is a gamble: it requires investment that may not pay off if the entry attempt fails.

Whether the owners of a firm will attempt to enter depends on two things:

- the economic profit ($\hat{\pi}^E$) that the owners can expect to receive if they are successful, and
- the probability b that they will fail, not receiving that profit, but instead losing their investment.

REMINDER In Chapter 9 you saw that firms will attempt entry as long as the expected (accounting) profit rate of an entering firm is not less than the opportunity cost of capital or $\hat{\pi}^A \geq \rho$, or, what is the same thing, that the expected rate of economic profit $\hat{\pi}^E \equiv \hat{\pi}^A - \rho \geq 0$.

If there are few firms in the market there will be little competition among firms and prices will be high relative to costs. So expected profits will be sufficient to motivate owners to attempt entry. If there are more firms and competition is greater, prices will be lower so firms will not enter.

The **competition condition** tells us the unique relationship of prices to wages—that is the real wage rate w^c—that ensures the number of firms will neither increase nor increase. This wage provides us with the extra information that, along with the wage curve determines the total level of employment in the economy.

Recall from Chapter 9 that a firm considering entering the market will compare the costs of entering with the expected price at which they can sell their output if they avoid failure. We use the same notation: p is the price at which the good sells if the firm does not fail, w^n is the money (or nominal) wage, b (for *b*arriers to entry) is the probability that a firm attempting to enter the market will not be able to sell its product and fail, and γ is the amount of output produced by a worker in an hour when working at the no-shirking level.

The owner pays the workers at the beginning of the period and sells the resulting product at the end of the period. So the opportunity cost of paying the workers is the wage plus the opportunity cost of devoting the owner's resources to the wage rather than some alternative investment, or $w^n(1+\rho)$.

To simplify we assume that there are no inputs other than labor. So the wage bill is both the only cost and the level of investment the owner must devote to the firm. This is called "working capital" to distinguish it from the value of the capital goods used in production which here we abstract from. Because there is only one good in this economy—the firm's output—the real wage is the nominal wage divided by the price of the firm's output or $w = w^n/p$.

The cost of producing a single unit of output is the hours of labor required to produce one unit times the nominal hourly wage or $a_l \cdot w^n$, and the opportunity cost of paying this amount to hire the labor necessary to produce the good is $(1+\rho)a_l w^n$.

The firm will attempt entry if:

$$\text{Expected price} = p(1-b) \geq (1+\rho)w^n a_l = \text{opp. cost of attempted entry}$$

The labor time it takes to produce one unit is (by definition) the inverse of the productivity of labor or $a_l = 1/\gamma$.

COMPETITION CONDITION The competition condition is an equation giving the unique relationship of prices to wages—that is, the real wage rate w^c—such that the number of firms will neither increase nor decrease given the productivity of labor and the opportunity cost of capital.

Using this fact, expressing this as an equality and rearranging we have the result that real wage consistent with the competition condition, w^c is:

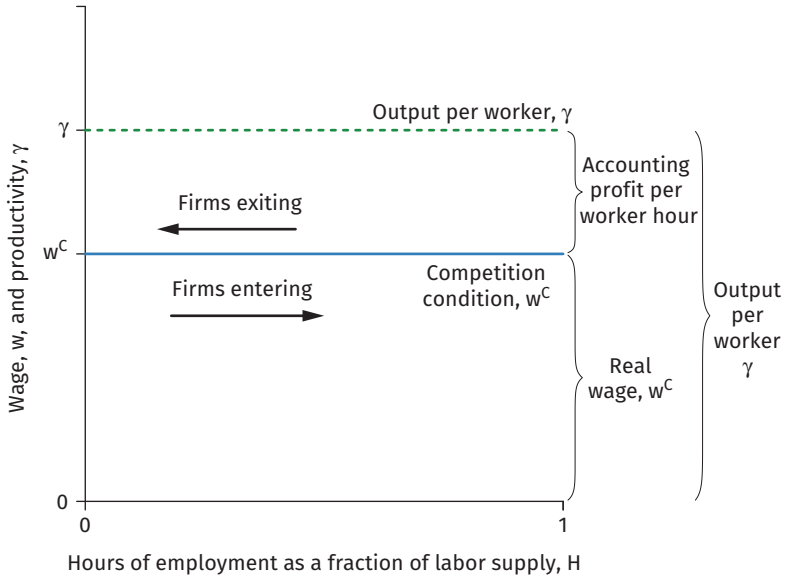
$$\frac{w^n}{p} = \frac{(1-b)\gamma}{(1+\rho)} = w^c \tag{11.23}$$

The real wage will be higher:

- the more competitive the economy is, that is, the lesser are the barriers to entry (b);
- the greater is the productivity of labor (γ), and
- the lower is the opportunity cost of capital (ρ).

Figure 11.13 shows the competition condition. If the wage is higher than w^c, then the expected (accounting) profit rate of a firm considering entering the industry will fall short of the opportunity cost of attempting to enter and so no firms will enter. Even if incumbent firms' profit rates allow them some economic profits on average, as in Chapter 8, some incumbent firms will fail for chance reasons unrelated to our model: a bad managerial

Figure 11.13 The competition condition and the real wage. Given the extent of barriers to entry, output per worker hour, and the opportunity cost of capital, the real wage indicated by the competition condition divides the output per worker between wages and profits in such a way that the number of firms does not change. It is called the competition condition because its level depends on the extent of competition, which is greater the lower are the barriers to entry so that the number of firms competing is greater. Greater competition (lower barriers to entry, more firms) allows a higher wage, shifting up the blue line.



decision, becoming a target of some other firm's predatory pricing, or the death or departure of some critical personnel. With some firms leaving and none entering, the number of firms will fall, and output and employment will decline. If the wage is lower than w^c the opposite process occurs. Firms will enter, producing more and hiring more labor.

M-NOTE 11.8 **Barriers to entry and the competition condition**

We can derive the competition condition expressing it as the requirement that the expected rate of economic profit of the firm attempting entry is equal to zero ($\hat{\pi}^E \equiv \hat{\pi}^A - \rho = 0$). Using the notation introduced in the text, we can write the expected accounting profit rate as the ratio of expected net revenues to the capital invested both expressed in per-unit terms. The amount of labor time (working at the no-shirking effort level) required to produce a unit of output is a_l. So:

$$\text{The accounting rate of profit} = \frac{p(1-b) - w^n a_l}{w^n a_l}$$

The numerator is accounting profits and the denominator is the investment by the entering firm, both expressed per unit of output. We equate this expression to the opportunity cost of capital ρ and rearrange the equation to get an expression for the real wage:

$$\begin{aligned} \frac{p(1-b) - w^n a_l}{w^n a_l} &= \rho \\ p(1-b) - w^n a_l &= \rho w^n a_l \\ p(1-b) &= w^n a_l (1+\rho) \\ \frac{w^n}{p} &= \frac{(1-b)}{(1+\rho)a_l} \end{aligned}$$

Taking account of the fact that the inverse of the labor input requirement for a unit of output, $a_l^{-1} = \gamma$, is the productivity of labor, we have the following:

$$\frac{w^n}{p} = \frac{(1-b)\gamma}{(1+\rho)} \equiv w^c = \text{the real wage that meets the competition condition}$$

which is the same equation as in the text, derived somewhat differently.

M-NOTE 11.9 **A special case: unlimited competition and zero profits**

If the economy were to be characterized by what Cournot called "unlimited competition," namely the complete absence of barriers to entry so that $b = 0$, then (continuing from M-Note 11.8) we have the following relationship between wages and the opportunity cost of capital, ρ:

$$\begin{aligned} w^c &= \frac{\gamma}{(1+\rho)} \\ \text{Raise to the power } (-1): \quad \frac{\gamma}{w^c} &= 1+\rho \\ \text{Isolate } \rho: \quad \frac{(\gamma - w^c)}{w^c} &= \rho \end{aligned} \tag{11.24}$$

The left-hand side Equation 11.24 is the accounting rate of profit, expressed in terms of one hour of labor, or hourly productivity minus the wage (that is

continued

profits per hour) divided by the capital invested per hour of labor (that is, the wage itself). So this is equivalent to:

$$\pi^{E} = \pi^{A} - \rho = 0$$

This expression is the competition condition for an economy with unlimited competition so that economic profits are eliminated. It is also called the **zero-profit condition**. We consider this to be a special case because as b approaches zero the number of firms in the industry (n) grows without limit, as can be seen from Equation 9.36 repeated here:

$$n^{N} = \frac{\bar{p}(1-b)-c}{bc}$$

EXAMPLE The economic effects of immigration are widely debated among the public. This interview from 2006 with Christian Dustmann (tinyurl.com/y3umsl4c), an economic historian who specializes in the effects of migration, captures this debate—in particular the impact of migrant workers in the British town of Swindon (from the CORE project. www.core-econ.org).

Prices, profits, wages, and employment

Putting the competition condition together with the wage curve in Figure 11.14 we have a model of the economy-wide labor and product market. But this is not a conventional model of demand and supply. It is true that:

- at wages above the competition condition employment contracts, below the competition condition employment expands; and
- on or above the wage curve workers supply the effort required by their employer, below the wage curve they do not.

M-CHECK The wage curve can also seen as an answer to the question: At any given real wage, what is the largest fraction of the labor supply employed consistent with workers being motivated to work at the effort levels dictated by their employers?

Be careful not to confuse the wage curve with a supply curve. The wage curve is not the "supply of labor." The wage curve is *not* the answer to the question: For each possible real wage, what is the amount of hours of labor supplied? Instead, the wage curve is the answer to the question: At each level of employment, what is the lowest wage consistent with workers providing effort on the job, that is, working? The wage curve divides the space in the figure into two regions.

- *No production*: At wages below the wage curve no production can take place because workers are providing no effort.
- *Feasible production*: At wages on or above the wage curve workers are working, so production is feasible (though firms may be leaving the industry if the wage is too high).

Similarly, the competition condition is *not* a demand curve. The competition condition does *not* provide the answer to the question: if the wage is w

ZERO-PROFIT CONDITION This condition requires that when barriers to entry are absent—the case of unlimited competition—expected economic profits are zero in a Nash equilibrium.

how many hours of labor will be hired? Instead, the competition condition divides the figure into two regions:

- At wages below the competition condition employment is *expanding* because firms are *entering*, as long as the wage is on or above the wage curve.
- At wages above the competition condition employment is contracting because firms are *exiting*.

We now bring together the competition condition and the wage curve. These two curves depict outcomes that satisfy two of the biggest challenges for a modern capitalist economy:

- *Investment*: In the real economy production requires more than the working capital to hire workers. The wage must be such that owners of firms have the incentive to invest—constructing machinery and buildings necessary to employ people (represented by the competition condition): this is ensured by being on or below the competition condition.
- *Work*: The wage and level of employment must be such that workers have the incentive to work hard and well. This is ensured by being on or above the wage curve.

These requirements cannot be ensured by government order: firms cannot be ordered to invest, and workers cannot be ordered to work. Because thousands—even millions—of people—each independently pursuing their own objectives—must act in ways consistent with these objectives, the incentives have to be right.

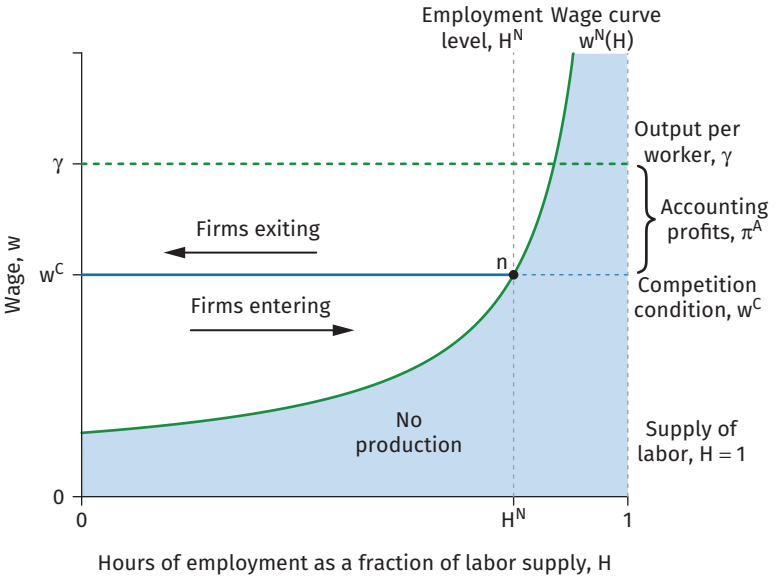
! REMINDER Equilibrium sounds like a good thing, better than disequilibrium. But, this model can explain conditions under which in equilibrium, wages would be extraordinarily low, unemployment substantial, and inequality unacceptably high. Here and in general, equilibrium means nothing more than that there are no forces that will change the situation other than exogenous changes, that is, from outside (external to) the model.

In Figure 11.14 the wage curve and the competition condition divide the space shown into four regions. In only one of them—on or below the competition condition and on or above the wage curve—are the conditions for both investment and work met. We call this the feasible production region. In the other three regions it is the case that either:

- firms are dis-investing, that is leaving so in the long run none will be hiring; or
- workers are not working, so firms will not be employing anyone; or
- both.

Narrowing down our attention to the feasible production region, are any allocations within it Nash equilibria? Recall that a Nash equilibrium is a situation in which, given the strategies adopted by others, one cannot do better by changing one's own strategy. In the economy as a whole, this means that at the equilibrium level of employment and real wage:

Figure 11.14 Equilibrium in the product and labor markets of the whole economy. Point **n** indicates the Nash equilibrium wage and employment levels, from which the level of profits and unemployment (that is $1-H^N$) can be calculated.



- *owners of firms could not make more profit by changing prices or nominal wages and thus altering the real wage*: they are paying the least wage consistent with workers working;
- *workers could not do better by working harder or less hard*: they are working at the level required by the employer, and this is the best they can do;
- *firms leaving the economy are exactly offset by firms entering; and no firm can benefit by expanding or shrinking its capital stock and hence their demand for labor.*

The first two points mean that the Nash equilibrium must be a point on the wage curve. The last point means that the Nash equilibrium must be on the competition condition. Therefore the Nash equilibrium is the intersection of the two curves.

Figure 11.14 shows that the labor market does not clear: the total level of employment, H, is less than the total hours of those seeking work $H=1$. When the proportion of the labor market H^N is employed there remains a proportion of workers represented by the complementary proportion $1-H^N$ who remain involuntarily unemployed. Those workers would like a job, but cannot get one.

We term this a long-run model because the process underlying the competition condition—the expansion or contraction in the size or number

REMINDER You might recall from an introductory macroeconomics course that aggregate demand is the economy-wide demand for goods and services rather than just demand in one industry or market.

of firms—requires building or scrapping equipment and buildings, a process that takes months or even years to complete.

Starting from the point **n** in the figure, shifts in aggregate demand in the economy associated with an export boom, or a collapse in investment, for example, will not shift the wage curve or the competition condition; but they can displace the economy away from the Nash equilibrium to higher or lower levels (respectively) of total employment than H^N.

The unemployment that exists at the Nash equilibrium is termed **structural unemployment**—meaning unemployment that results from the fundamental structure of the economy including its technology and institutions as represented by the two equations of our model. On any given day or year the level of unemployment may exceed or fall short of structural unemployment as a result of the business cycle or anything else that results in the actual unemployment rate being different from the structural rate. The difference between structural and realized unemployment is called cyclical unemployment.

CHECKPOINT 11.12 **Total employment** Explain the effect in the long run on total employment of the following:

a. A change in technology making it easier to detect a shirking worker.
b. An increase in the unemployment benefit.
c. An increase in labor productivity.
d. An improvement in the expected profits to be earned by investing in some other economy.
e. An increase in the supply of labor (for example by immigration).

11.12 MONOPSONY, THE COST OF INPUTS, AND THE LEVEL OF HIRING

In a **monopsonistic** labor market one or just a few firms hire a large fraction of the workers and the more workers they hire, the more they have to pay to retain and motivate workers to work. The owners of the firm are interested not simply in getting workers to show up (meeting their participation constraint): they want to provide incentives for the worker to work (meeting their incentive compatibility constraint). And because there are costs of

STRUCTURAL UNEMPLOYMENT Structural unemployment is the unemployment that results from the fundamental structure of the economy.

MONOPSONY A firm is a monopsony if it is the only buyer (or just one of a small number of buyers) in a particular market for some good or service.

finding and training new workers, the owners of the firm want to set wages so that few workers will quit their jobs.

You have already seen that employers are **wage makers**. Firms set wages so as to minimize the cost of acquiring the labor effort that is an input into their production. In the case presented in panel (b) of Figure 11.3 this labor-cost-minimizing wage is independent of how many hours of labor the firm hires, as can be seen in Figure 11.4. But for monopsonistic employers there is another dimension of wage-making: they know that their hiring decisions make the wage what it is.

A monopsonistic labor market: Wage-making

The distinguishing characteristic of the monopsonistic employer is that, taking the employer's eye view of the labor market, he can pay less if he hires fewer workers. This makes the monopsonistic firm similar to a monopoly, duopoly, or other firm selling on a market with a limited number of competitors: selling less is a profitable strategy, because it allows sales at higher prices.

The difference is that in the case of monopsony it is in the market for *inputs* that competition is limited and the objective of restricting *output* is to limit buying, and hence *reduce* the prices of inputs. Of course a firm can be both a monopoly and a monopsony, with limited competition in both the market for its output and the market for its inputs.

Why does restricting the amount of hiring lower the labor costs of a firm? Here are two major reasons.

- *Company town or neighborhood labor market*: A large fast-food chain may find that it employs a significant proportion of the available low-wage workers in the neighborhoods where it operates, particularly if the lack of low-cost public transport makes it difficult for workers who live in one neighborhood to work for firms located some distance away. In this case if a firm employs more in a neighborhood it may have to pay more.
- *Balancing quits and new hires*: Owners and managers face the following problem: they would like to maintain a level of employment that is just enough to produce the profit-maximizing level of output. But in any given period, say, a month, some fraction of workers will quit or retire. These workers will have to be replaced by new hires from among those applying for work. The firm will set the wage to balance the number of workers leaving and new hires to sustain the desired level of employment. But

WAGE MAKER Employers are called wage makers because they typically decide on ("make") the wage to offer to particular workers either unilaterally, or through bargaining with a trade union. The term is intended to contrast with what would be a wage-taking firm (like a price-taking firm) that cannot alter the wage or price to its advantage.

! REMINDER In the previous section and in Figure 11.14, we used H to denote the total hours of labor hired in the economy as a whole. Here we talk about a single firm and therefore use h.

because a larger firm will lose a greater number of workers in a month, in order to attract a sufficient number of replacements, a firm that seeks to maintain a higher level of employment will have to pay a higher wage.

For these two reasons, the lowest wage at which a monopsonistic employer can pay to motivate and retain workers is greater if he wishes to hire more, so the wage depends on the amount of hours hired or $w = w(h)$. The average cost of labor is the wage: in this model the workers are identical and they all are paid the same wage. This means that the firms do not practice price discrimination, that in this case would be termed wage discrimination (studied in Chapter 9). So the wage that all workers receive is the average wage.

M-CHECK The marginal cost of labor is: $\frac{\Delta w(h)h}{\Delta h}$

But for a monopsonist, the wage is higher if he hires more. This means that the effect on total costs of hiring additional hours of workers' time (marginal cost of labor) is greater than the average cost of labor (the wage). This is because hiring an additional hour adds to the total labor costs in two ways:

- the age paid to the particular worker for the additional hour; *plus*
- the effect of this additional hiring on the wages of all of the *other* workers currently employed by the firm.

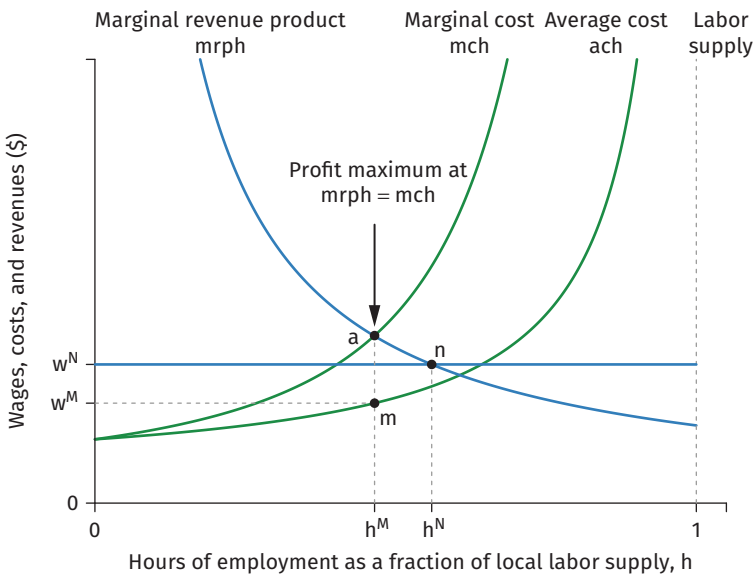
These two bullets appear in M-Note 11.10 as the two terms in Equation 11.25 labeled marginal cost of hours.

We illustrate this monopsony case in Figure 11.15. To facilitate the contrast with the no-monopsony case we also show the outcome of where monopsony is absent studied in Figure 11.4. In this case the employer had used the Solow condition to set a wage w^N, and was hiring under conditions in which the least-cost wage did not vary with its employment level. Equating w^N to the marginal revenue product of hours of labor at point **n** the employer hired h^N.

When monopsony is present, as in Figure 11.15, the employer will maximize profits by employing the number of hours such that the marginal revenue product of worker hours is equal to the marginal cost. Figure 11.15 differs from Figure 11.4 in that the marginal cost of labor is no longer the wage set by the Solow condition: it is now higher than the wage because the average cost of labor function $w(h)$ is rising with more hiring. Figure 11.15 shows that the monopsonistic firm hires up to the point that marginal revenue product equals the marginal cost of labor, that is, up to point **a**, where the firm hires h^M. From the figure you can see that the effects of monopsony are:

- *The firm hires fewer hours*: Relative to a labor market in which the firm is not a monopsonist, a monopsonistic employer will restrict employment, hiring $h^M < h^N$ as can be seen comparing points **a** and **n** in Figure 11.15.

Figure 11.15 The monopsonist's profit-maximizing level of hiring. Employment as a fraction of labor supply, h, is plotted on the horizontal axis, and marginal revenue product of hours hired, the wage per hour, and marginal cost of hours of labor on the vertical axis. For the firm without monopsony power the wage is determined by the Solow condition and is w^N and in that case the profit-maximizing level of employment (h^N) is given by the intersection of marginal revenue product (the downward-sloping blue line) and the horizontal blue average and marginal cost line w^N at point **n**. In the monopsony case, the marginal cost of hiring exceeds the average cost; therefore the condition for profit maximization ($mrph = mch$) is satisfied at point **a** with a profit-maximizing level of hiring of h^M. The employer pays a wage given by the point on the ach curve that corresponds to the level of hiring h^M, shown at the intersection at point **m** with the wage w^M. The monopsony therefore hires fewer hours of work ($h^M < h^N$) and pays a lower wage ($w^M < w^N$).



- *Lower wages paid to workers*: Because the monopsonistic employer hires fewer hours, workers who *are* hired are paid less than they would be if the cost of motivating and retaining labor were independent of the level of hours hired and if, as a result, the employer had hired more. Remember, the employer's hiring decision was given by point **a**—where the marginal revenue product of hours equals the marginal cost of hours. But the wage paid is determined by the average cost of hours, that is, point **m**. So $w^M(h^M)$, is less than w^N.

The model of the monopsonistic employer's hiring provides a lens with which to study a controversial topic: the effects of minimum wages.

M-NOTE 11.10 Labor hours hired by a monopsonistic employer

Here we study the number of hours of workers' time the employer will hire, assuming that the amount of effort per hour is given (equal to e^N as determined by the Solow condition). So the marginal product of an hour is just the marginal product of effort divided by a constant, namely e^N.

All of the information is the same as in M-Note 11.3, except that the total cost of hiring labor is no longer just wh because the wage rate, w, depends on how many hours of labor is hired.

So, the profit function is now:

$$\pi = p(x(h))x(h) - w(h)h$$

To find the profit-maximizing level of hiring, as in M-Note 11.3, we differentiate the profit function with respect to the amount of hours hired, h, and set the result equal to zero.

$$\frac{\partial \pi}{\partial h} = \underbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial h}x(h) + p(x(h))\frac{\partial x}{\partial h}}_{\text{Marginal revenue product of hours}} - \underbrace{\left(\frac{\partial w(h)}{\partial h}h + w(h)\right)}_{\text{Marginal cost of hours}} = 0 \tag{11.25}$$

The marginal revenue product of hours is unchanged, but the marginal cost of hours is now the average cost of hours—the wage itself—plus a term that captures the effect of hiring more on the wage. Because the marginal cost of hiring is higher than the average cost in the monopsonistic firm, but not in the case of the fixed wage, the profit-maximizing monopsonistic firm will hire fewer hours of labor.

CHECKPOINT 11.13 Monopsonistic hiring

a. What is monopsony? How is it different from or similar to monopoly?

b. Give examples of forms of monopsony other than in labor markets (think of Walmart or Alibaba).

c. Explain why the monopsonistic firm will hire fewer workers, or purchase fewer of other inputs, than an otherwise identical firm that is not a monopsonist.

11.13 MONOPSONY AND THE COST OF HIRING (NON-SHIRKING) LABOR

In January 1987, a headline in the *New York Times* read, "The right minimum wage: $0.00."[18] The minimum wage at the time was $3.35 ($ 7.79 in 2020 dollars). The *Times* editors wrote: "there's a virtual consensus among economists that the minimum wage is an idea whose time has passed. Raising the minimum wage by a substantial amount would price working poor people out of the job market."

They were not wrong about the consensus among economists at the time: economists generally opposed minimum wages. The demonstration

of how minimum wages were a misguided policy appeared in introductory textbooks using a standard intersecting supply and demand graph. That conventional figure showed that in the absence of any government intervention the market would *clear*, with the quantity of labor supplied equal to the quantity of labor demanded and anyone not working was doing so "voluntarily" as they would rather have leisure than work at the going market wage.

With a minimum wage imposed at a level above the "market-clearing wage," the thinking went, the supply of labor would exceed demand for labor and unemployment would result. We used a figure of exactly this type to teach what happens when prices are greater than market-clearing prices in Chapter 9. But, as we have shown in this chapter and in Chapter 10, labor markets and other markets with incomplete contracts generally do not clear in equilibrium.

Since the early 1990s economists specializing in the labor market have studied the effect of increases in the minimum wage on the level of employment, primarily in the US. One of the pioneering studies by David Card and Alan Krueger examined the 1992 increase in the minimum wage from $4.25 to $5.05 in the US state of New Jersey by comparing firms on either side of the border of New Jersey and Pennsylvania (which did not have an increase in the minimum wage). They found that the increase in the wage by almost 20 percent did not increase the rate of unemployment nor did it decrease the number of hours hired.[20]

Their method was to look at hiring before and after the wage increase in establishments in neighboring states, cities, or counties where a minimum wage increase was either present or absent. This approach became the standard way of evaluating these policies (making use of sophisticated statistical methods). To the surprise of many economists, the estimated job losses associated with a minimum wage increase have been either small or even nonexistent. In some cases, the data suggest even a small *positive* effect of the minimum wage on hiring.

As the evidence accumulated, economists began to reconsider their conclusions about the adverse employment effects of the minimum wage. In 2015 an elite panel of distinguished economists brought together by the Booth School of Business at the University of Chicago were asked whether they thought that raising the minimum wage to $15 (more than double the national minimum wage then in force) would result in a substantially lower employment rate for low-wage US workers relative to that under the status quo. Only 26 percent thought that it would.[21]

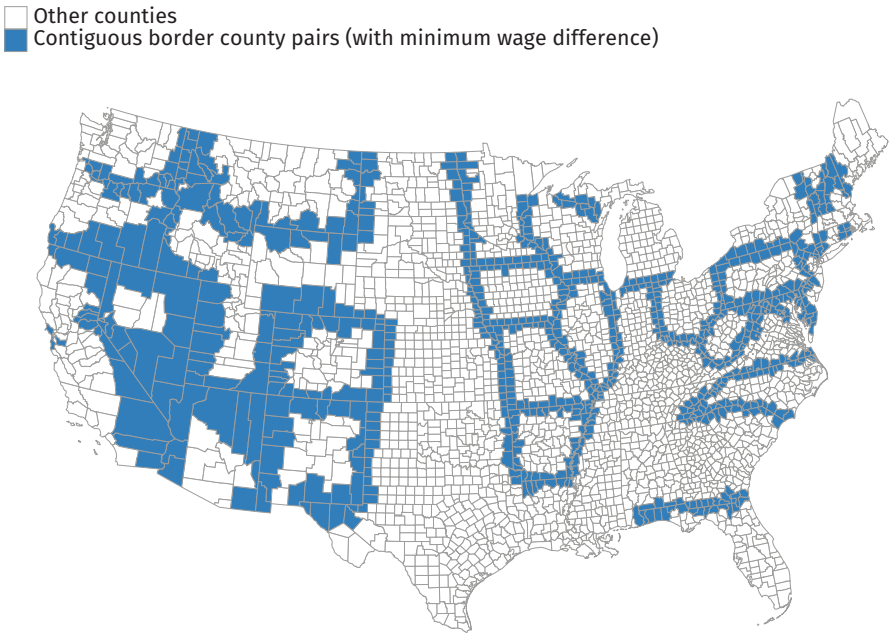
While the size (and even the existence) of the adverse employment effects of the minimum wage continue to be debated, many economists are now skeptical about the adequacy of the simple supply and demand model when

✓ **FACT CHECK** We have modeled monopsony—a market for inputs with a single dominant buyer—but the results are similar for an *oligopsony*, that is, a few wage making firms. Alan Krueger and Orley Ashenfelter found that many service-oriented US firms, such as H & R Block, McDonald's, Burger King, and Jiffy Lube, had "no-poaching" agreements that meant that the firms would not try to hire workers from other firms. By limiting workers' alternative employment opportunities, no-poaching agreements reduce workers' fallback option and allow the firms to pay lower wages without violating the no-shirking condition.[19]

HISTORY Paul Samuelson, the first American to win the Nobel Prize in economics, wrote a path-breaking introductory economics textbook in the aftermath of the Great Depression and World War II. In his chapter introducing the model of supply and demand among competitive price-taking buyers and sellers, he warns: "the demand for labor in the United States cannot be analysed by the methods of this chapter."[22]

Figure 11.16 A research design to determine the employment effects of the minimum wage. This map of the US shows (in blue) the paired counties that were compared in one study, in most cases those located on the boundary of two states with different levels and changes in the levels of the minimum wage.

Source: Dube et al. (2016).



applied to the labor market. We would like to see if the models you have learned so far provide any clues about why in response to the minimum wage the substantial hiring cutbacks expected by many economists have failed to materialize in the data. There are two.

- The *labor discipline model* (*Ford model*) shows that a wage increase will motivate workers to work harder, and this could partially offset the additional costs associated with a minimum wage increase.
- The *model of the monopsonistic firm's hiring* shows that the firm will restrict its hiring if employing more workers raises the wage. But, if the minimum wage is higher than the monopsonistic wage w^M over the range of hiring that it might want to do, then the minimum wage becomes the new marginal and average cost of hiring. This removes the monopsonistic motive to hire fewer workers, because hiring fewer workers would not allow the firm to pay a lower wage (because that would violate the minimum wage law).

Monopsony, labor discipline, and the no-shirking wage

The worker's best-response function (for example, in Figure 11.3 (b)) shows that paying higher wages will increase the effort workers provide; and this could partially offset the costs of a wage increase imposed by law. The same model also provides a reason why many employers are monopsonists, that is to say, why the cost of hiring labor would be less if the employer hires fewer hours.

The reason provided by the Ford model is similar to the "company town" logic above. The firm's own hiring:

- reduces the pool of unemployed workers in the local labor market and this
- decreases the probability j that a fired worker will remain unemployed, which
- raises the fallback option of employed workers (what they get if they are fired) which therefore
- raises the least wage consistent with the worker providing the no-shirking effort level.

This is why the average cost of labor, which in this model is just the no-shirking wage, increases with the level of hiring. If the average cost of labor is rising, then the marginal cost of hiring a non-shirking worker must exceed the average cost. M-Note 11.11 shows how the economy-wide wage curve can be repurposed to become the firm's average cost of labor curve in a local labor market. The result is that the individual monopsonistic employer faces an average cost of hiring labor curve that is a city-wide or other local version of the economy-wide wage curve (for example Figure 11.11, based on US data) introduced earlier.

M-NOTE 11.11 A monopsonist's cost of (non-shirking) labor

The cost of (non-shirking) labor in a local labor market.

To illustrate the idea of a monopsony with the labor discipline model mathematically, we consider the case in which there is just a single firm in a local labor market that hires a substantial fraction h of the relevant labor supply for this particular kind of employment in the location in question.

Similar to what we did when re-purposing the no-shirking condition for the economy-wide labor market we write the probability of the terminated worker remaining unemployed as a function of the level of employment, but here it is in the *local*, not the *national* labor market: $j(h)$. So, modifying the no-shirking condition (Equation 11.19) to take account of the fact that the probability of remaining unemployed depends on total employment, we have the following no-shirking wage:

$$w^{N} = B + \underline{u} + \frac{1-t}{tj(h)}\underline{u}$$

continued

✓ **FACT CHECK** Decio Coviello and coauthors studied the effect of minimum wages on the earnings and individual productivity of workers in a large US retailer that employs a substantial fraction of all department store workers. The (unnamed) firm operates in all 50 states and the researchers could study more than 10,000 workers in over 300 stores. The data allowed them to contrast wages and productivity in nearby stores which differed along two dimensions:

- one or more of the locations before and after a minimum wage increase; and
- locations in which the minimum wage increased or stayed the same.

The researchers found that an increase in the minimum wage of 65 cents caused individual productivity measured by sales to increase by 2 percent, consistent with the idea that higher wages provide incentives for workers to provide more effort.[23]

Worker quits and the probability of remaining unemployed

The no-shirking model did not take account of the fact that workers may quit, so we now need to introduce the idea of worker quits. Let the rate of current workers leaving the firm (quits) in any period be q, so the total number of quits will be qh. The firm needs to hire workers to fill the positions it has lost due to quits. The firm will therefore hire qh in every period to replace those who have left. Let us consider the case where a terminated worker may be rehired by his former firm perhaps in a different establishment: for example, getting fired at one restaurant branch may not stop you from getting a job at another branch of the same restaurant chain. We can then calculate the probability that a terminated worker finds a job. Take the firm's hiring qh each period and divide it by the number of unemployed $(1-h)$. This is the probability a worker finds a new job: $\frac{qh}{1-h}$. We also restrict $qh < 1-h$, which means that the firm's hiring (qh) must be less than the unemployment $(1-h)$.

The probability a terminated worker remains unemployed, is therefore:

$$j = 1 - \frac{qh}{1-h} \tag{11.26}$$

so as the firm's own hiring h increases, the worker's chance of *not* remaining unemployed also rises.

M-NOTE 11.12 **Why restricting hiring reduces labor costs**

To find the effect of the firm's hiring (a change in h) on the probability that an unemployed worker will remain unemployed (j) we differentiate Equation 11.26 with respect to the number of workers (h), giving:

$$\begin{aligned}\frac{\partial j}{\partial h} &= \frac{(1-h)(-q)-(-qh)(-1)}{(1-h)^2}\\ &= \frac{-q+qh-qh}{(1-h)^2}\\ &= -\frac{q}{(1-h)^2} < 0\end{aligned} \tag{11.27}$$

The fact that Equation 11.27 is negative means that if the firm hires more (increases h), j will be less, so, the likelihood of the worker remaining unemployed will decrease. And it also means that if the firm hires fewer hours the likelihood that the terminated workers will be unable to find a job increases. This illustrates the firm's monopsonistic status in the labor market: the fewer hours it hires, the worse will be the worker's fallback position, and hence the lower will be the no-shirking wage. (The case where the terminated worker will not be rehired in her former firm can be modeled in similar fashion. What matters is that Equation 11.27 is negative, which it will be in this alternative treatment.)

CHECKPOINT 11.14 The monopsonist's rising cost of (non-shirking) labor

Give two reasons why the cost of hiring an hour of labor may increase as the employer hires more.

11.14 THE EFFECTS OF A MINIMUM WAGE ON HIRING AND LABOR EARNINGS

It remains only to show that a monopsonistic firm–that is, one with a rising cost of hiring (non-shirking) labor–may find it more profitable to *increase* hiring when a minimum wage is imposed, rather than *decreasing* hiring.

There are two possible effects of interest:

- effects on the *total wages* paid to workers; and
- effects on the *level of hiring*.

Effects of the minimum wage on hours hired

With respect to the level of hiring, there are three cases to consider.

- *No effect*: the firm is already paying a wage higher than the minimum wage, in which case we say that the mandated wage is not binding.
- *Decreased hiring*: the minimum wage results in the firm hiring less labor.
- *Increased hiring*: the minimum wage results in the firm hiring more labor.

› EXAMPLE In this video (tinyurl.com/y4jgccg9) Arin Dube describes his study that found that, on average, raising the minimum wage increased the income of poor workers (from the CORE project. www.core-econ.org).

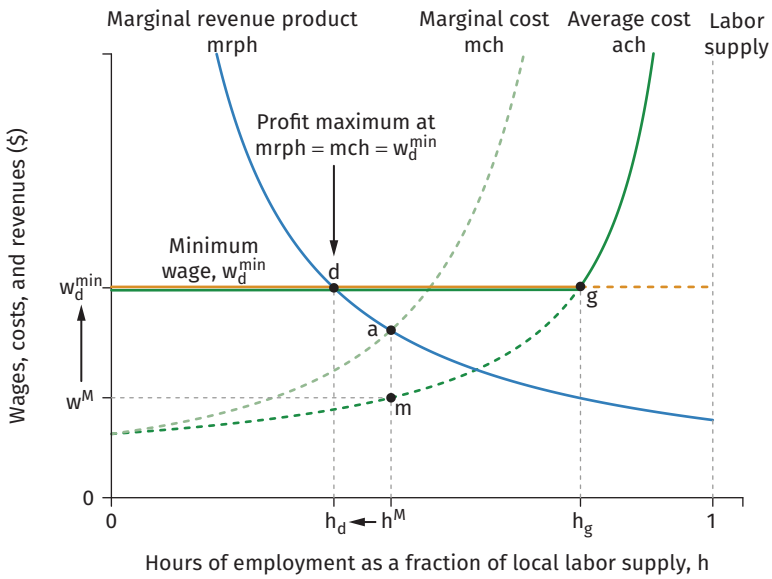
The first case will apply to many firms but it is not really of interest because the minimum wage law has no effect. Figure 11.17 shows the firm's hiring decisions where the minimum wage is binding and it results in a reduction in hiring. The dark-green curve–both dashed and solid portions–is the average cost of labor defined by the no-shirking condition. The marginal cost of an hour of labor is the light-green dashed curve. Why are the average cost of hours and marginal cost of hours curves at least partially dashed? The answer is the orange horizontal line–the minimum wage. As long as the average cost of hours–the no-shirking wage–is less than the minimum wage, then the average-cost-of-hours curve is no longer the least cost the employer can pay to employ (non-shirking labor). Paying $w^{M}(h)$–the average cost of hours–is sufficient to motivate the worker to work, but it is not legal. The employer is required by law to pay more than it needs to pay to avert shirking.

This means that the average cost of hours curve is no longer relevant. This is why it is dashed. The solid minimum wage is now the average cost of hours curve for levels of hiring less than h_g–the amount of hiring at which the average cost curve is equal to the minimum wage.

Because the average cost (meaning the minimum wage) is constant over this range $(0 - h_g)$ the marginal cost is equal to the average cost: hiring an additional hour increases total costs by the amount paid for the additional hour, w_d^{min}, and there is no effect on the wages paid to the other workers. This is the reason why the marginal cost curve is also dashed.

So to the right of point **g**, it is the minimum wage that is irrelevant, which is why that part of the line is dashed. Instead, we would have a solid

Figure 11.17 Monopsony hiring level with a minimum wage, Case I: a large wage increase reduces hiring. In this case, a minimum wage like w_d^{min} that is higher than the intersection at **a** results in a *decrease* in employment from h^M to h_d. The local labor supply is assumed constant and equal to 1. For any employment level less than h_g the minimum wage is the lowest wage the firm can offer by law and it exceeds the least wage that it would otherwise pay given by the average cost of hours of hiring. So the average and marginal cost curves for hiring levels less than h_g are irrelevant: the average and marginal cost of an hour of a worker's time is the minimum wage.



✓ **FACT CHECK** The model in which monopsony in the labor market explains why imposing (or increasing) a minimum wage may have a positive (rather than negative) effect on firm hiring predicts that where hiring is dominated by just a few firms we could see positive employment effects, and where monopsony is more limited, we would find a negative effect. This is exactly what a study of the retail sector of the US economy (think Walmart) found: where hiring was very concentrated, the minimum wage effect on employment was positive, and where concentration was limited the opposite was observed.[24]

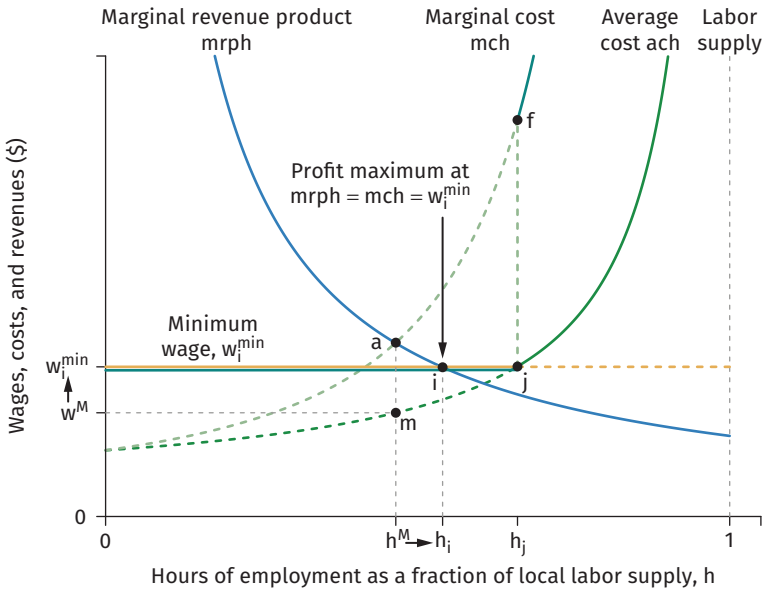
increasing marginal cost curve for employment levels above h_g (not shown). See Figure 11.18 for a case where the solid portion of the marginal cost curve is visible.

The effect of the minimum wage is, over some range of hiring, to break the link between firm hiring and the cost of sustaining labor discipline and hence the average cost of an hour. Over this range, the employer's incentive to restrict hiring so as to maintain lower hourly labor costs disappears.

In the first case (Figure 11.17) the minimum wage results in a decreased level of hiring, h_d relative to the monopsony's preferred hiring at h^M.

In the second case, in Figure 11.18 the minimum wage results in an *increase* in employment relative to the monopsony's hiring. The difference between the two cases is that the increase in pay over the wage that the monopsonist would have paid in the absence of the minimum wage in Figure 11.17 is much larger than the increase shown in Figure 11.18.

Figure 11.18 Monopsony hiring level with a minimum wage, Case II: a smaller wage increase increases hiring. Prior to the minimum wage, the monopsonist hires h^M as before equating $mch = mrph$. But the fixed wage given by the minimum wage w_i^{min} becomes the marginal (and average) cost of hiring. Marginal cost therefore intersects marginal revenue product at point **i** and determines the amount of labor hired at h_i. In this case, the minimum wage w_i^{min} is higher than the monopsony wage w^M and the amount of labor hired increases from h^M to h_i.



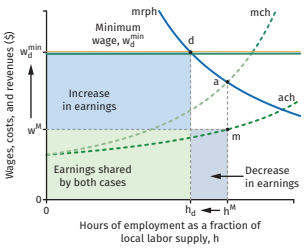
Effects of the minimum wage on labor earnings

We have shown that the imposition of a minimum wage (or an increase in the minimum wage) can result in either a decrease or an increase in the amount of hiring done by a monopsonistic firm. The other question about the effects of the minimum wage that we asked at the outset is whether taking account of both the change in hours and the change in the wage increases or decreases the total **earnings** of affected workers. There are two cases:

- If the effect is to increase the hours hired, then the total earnings of workers must rise (because in this case more hours are hired at a higher wage). This is shown in Figure 11.18.

EARNINGS This term—sometimes called "labor earnings"—refers to income from employment by a firm, government or some other employer, whether in the form of wages or salaries.

Figure 11.19 Change in worker earnings from a minimum wage. The minimum wage reduces hours hired, but those hours for which workers are hired earn more. The total earnings of low-wage workers therefore increases.



✓ **FACT CHECK** When minimum wage laws mandated increased pay for nursing assistants and other poorly paid nursing home staff in the US the health of those in their care dramatically improved. This is the finding of a study of 7,700 nursing homes in 1,136 counties, comparing the health of residents in nursing homes that had experienced wage increases that had not been experienced by nursing homes in neighboring counties. The author calculates that a 10 percent increase in the minimum wage nationally would lead to approximately 15,000 fewer nursing home deaths annually.[25]

- If the effect is to decrease the hours hired, then the effect on total earnings (hours times wages) could be either positive or negative. Figure 11.19 illustrates a case in which the gain in pay more than offsets the loss in hours, so total earnings increase.

The empirical evidence from the US and the UK finds that increasing the minimum wage raises the incomes of low-wage workers, and thereby reduces inequality. The reduction in inequality is one of the main reasons why many economists have come to have a more positive view of minimum wages as a labor market policy.

Effects of the minimum wage on labor effort and turnover

Firms' responses to an increase in the minimum wage go beyond a change (including a possible increase) in the level of hiring.

The imposition of (or increase in) a binding minimum wage will raise worker effort level for the following reason. Prior to the imposition of the minimum wage, the firm pays the no-shirking wage and workers are working at the firm's designated no-shirking level of effort. Now the firm is required to pay a higher wage, so it will pay more than is sufficient to motivate workers to work at the initial no-shirking level.

Faced with paying a wage that is higher than the no-shirking wage, the owners of the firm do not have the option to lower the wage (which would violate the law). So instead they can adjust upward the minimum ("no-shirking") level of effort the worker must do to avoid being terminated if observed by the employer. Given that they must pay the minimum wage in any case, this gains the employer greater effort per hour hired, at no cost.

The minimum wage thus inverts the logic of the no-shirking condition. Initially the no-shirking condition told the employer the least wage (w^N) it could pay and still motivate workers to provide effort at the level $\underline{e}$. With a minimum wage greater than w^N, the no-shirking condition, instead, tells the employer what is the greatest amount of effort and associated disutility $\underline{u}(\underline{e})$ that the firm could require as a condition for not being terminated consistent with the worker working rather than shirking.

At the higher imposed minimum wage, workers will have the motivation to conform to this new higher no-shirking level of effort. So, the cost of paying the minimum wage will be partly offset by the effort increase by workers.

A second effect—also favorable from the standpoint of the firm's owners—is that the higher wage paid by the firm will reduce the number of workers leaving the firm. This will save the firm on the costs of recruiting and training replacements.

But while the positive effect on effort and the reduction in turnover are both favorable to the firm's profit-making objective, they cannot fully offset the cost increase imposed by the minimum wage. If this were possible, it

would mean that the firm had been paying a less-than-profit-maximizing wage prior to the imposition of the minimum wage, and it would have already raised the wage just to gain higher profits.

A broader picture: The minimum wage in context

To understand the effect of the minimum wage in an entire local labor market or in the whole economy we would need to take account of additional likely effects.

- *Firm profits*: The unavoidable negative effects on the firm's profitability would in the long run be expected to lead to the exit of firms, which would both reduce hiring and could increase the degree of monopsony in local labor markets.
- *Changing demand*: The redistribution of income from relatively high-income profit recipients, who save a substantial portion of their income, to low-wage workers who spend most of what they earn, could increase aggregate demand, leading firms to hire more workers, at least in the short run.

While the positive effects on aggregate demand might be substantial for a minimum wage increase in an entire economy, they are unlikely to result in any significant increase in jobs resulting from a minimum wage increase enacted by a single city. The reason is that the increased earnings of those benefiting from the minimum wage would, for the most part, be spent on goods and services produced by workers in the rest of the world, not locally.

We return to a consideration of policies that can increase both wages and employment and reduce inequality in Chapter 15.

Eliminating policies that restrict labor market competition

Minimum wages are not the only policy that can effectively raise the earnings of low-wage workers. An alternative to imposing a minimum wage is to make labor markets more competitive so as to limit the monopsonists' ability to hold down wages by restricting hiring. Figure 11.20 shows the comparison of the competitive labor market in which employers cannot depress wages by hiring less, with the wage w^N and the case of monopsony with the wage $w^M < w^N$.

In Chapter 9 we described some of the rent-seeking strategies that owners of firms adopt to reduce product market competition and as a result increase profits. Similar competition-limiting strategies are adopted in labor markets. Chief among these are *non-compete clauses* that prevent either a current worker or a terminated worker from going to work for a employer competing with her current employer. Non-compete clauses were initially written into the contracts of a few scientists, top management, and other high-level employees who might leave the firm, bringing

✓ **FACT CHECK** Arin Dube and his coauthors found that for groups on which the minimum wage was expected to have a major impact (teens and restaurants) an increase in the minimum wage

- eliminated a substantial number of below-minimum-wage jobs;
- increased the number of somewhat above-minimum-wage jobs by about the same number; and
- very slightly increased jobs paid more than the minimum wage.

They also found that the minimum wage increase resulted in substantial reductions in quits as fewer people left their jobs, and a resulting reduction in new hires.

✓ **FACT CHECK** A non-compete clause in the contract of a $13 per hour packer at Amazon in 2015: "During employment and for 18 months after the Separation Date, Employee will not, directly or indirectly, . . . engage in, . . . manufacture, marketing, or sale of any product or service that competes or is intended to compete with any product or service sold, offered, or otherwise provided by Amazon . . . ." It's difficult to see how a worker fired by Amazon could get any kind of job at all without violating this contract, which may be the point.[26]

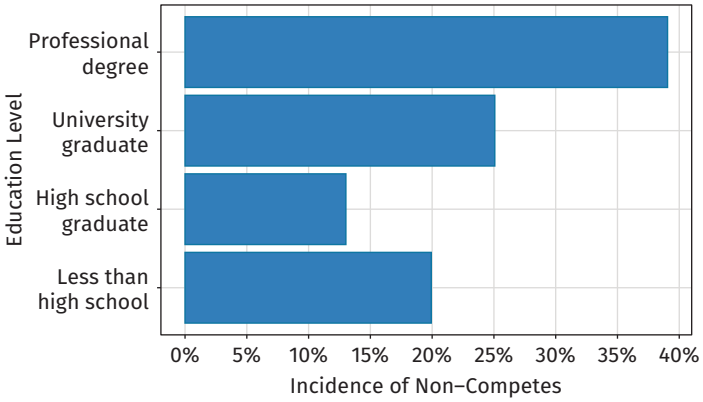
M-CHECK *Non-compete clauses and the probability of remaining unemployed.* A non-compete clause means that the terminated worker can apply to only a fraction $\mu < 1$ of the job vacancies. As a result, the probability of finding work following termination falls to:

$$\text{Probability of finding work} = \mu \frac{qh}{1-h}$$

which is less than without the clause, because $\mu < 1$. This is how non-compete clauses increase *j* the probability that a terminated worker will not find work and remain unemployed.

Figure 11.20 **Fraction of workers whose contracts have non-compete clauses.**

Source: Starr et al. (2019).



trade secrets to competitors. But in the US today roughly one in five workers has agreed to such a restriction.

While non-competes are more common among well-paid, highly skilled workers they are common also among workers like Amazon's packers, who are unlikely to have valuable trade secrets to pass on to a future employer seeking to compete with the marketing giant. About a fifth of personal care and services workers, for example are subject to such contractual provisions as are one in ten food preparation and serving workers.

EXAMPLE A "no-poach" clause in the contracts of McDonald's franchise holders commits them not to hire anyone who has worked for another McDonald's outlet within the past six months.[27]

These clauses prohibit, for example, a worker who was fired from or quit a job with McDonald's from taking any job in the fast-food business, substantially limiting her chances of finding work. We show the effect in M-Note 11.11. Prohibiting non-complete clauses in contracts would make labor markets more competitive and, by improving workers' fallback options, put upward pressure on wages.

CHECKPOINT 11.15 **Non-compete clauses and work effort** Suppose the owner of a firm is considering imposing a non-compete contract on his workers and he asks you to explain how it might affect the wage he will have to offer to motivate his employees to work. Use the equation for the no-shirking condition to provide a reply. Which variable (or variables) in the equation might the imposition of the non-compete clause affect?

11.15 CONCLUSION

Developments in both theory and empirical studies have given economists reason to think about not just the economics of labor markets, but also the sociology and politics.

The stimulus for much recent theoretical research on labor markets came from dissatisfaction with the microeconomic aspects of macroeconomic models of aggregate employment and unemployment. Macroeconomists exploring the microeconomics of Keynesian models of unemployment were prominent among the early innovators. Models based on incomplete contracting for effort or other aspects of the labor exchange explained how a competitive equilibrium could exhibit involuntary unemployment, thereby narrowing the gap between theory and empirical observation.

We do not know what Henry Ford had in mind when in 1914 he announced the $5 day wage. The fact that output per worker hour more than doubled following the increase suggests that workers' effort rose substantially. (Ford increased the level of supervision along with the wage, so the likelihood that slack work would be tolerated undoubtedly fell.) Whether the workers' increased effort was a response to the carrot of Ford's seeming generosity or to the stick of closer supervision and increased employment rents, we cannot say.

MAKING CONNECTIONS

Contracts: Between workers and employers, whether negotiated individually or collectively (by a labor union) cover important aspects of the interaction—hours, wages, and working conditions—but not others—how hard the person works, whether they quit, and many aspects of the quality of the work. Complete and incomplete contracts are contrasting institutions or rules of the game.

Price-making and wage-making: We have studied two categories of price-making: wage-setting by monopsonists facing limited competition in the labor market, and the choice of a labor-cost-minimizing wage as part of the employer's labor discipline strategy.

Institutions, fairness, and Pareto efficiency: The Nash equilibrium for interactions governed by incomplete contracts is Pareto inefficient, but the worker is better off than she would be with complete contracts.

Incentive compatibility, participation constraints, and constrained optimization: Because the labor contract is incomplete, the employer is constrained by the incentive compatibility constraint governing how hard the worker works, and in order to maximize profits pays the worker more than the minimum to just satisfy her participation constraint.

Conflicts over mutual gains made possible by exchange: The labor market is characterized by conflict over the mutual gains from employment. This conflict is not resolved in a contract, but is determined "on the ground" in the strategic interactions between the employer and the worker. A more complete labor contract favors the employer in this conflict.

Limited competition and inequality: As in Chapter 9, in this chapter, we found that limited competition among firms—here in the market for workers' time, not for the output of the firm—raises the profits of owners.

Non-clearing markets, rents, and power: Employers are on the short side of a non-clearing market; workers and the unemployed are on the long side. The employer's threat to terminate the rents that employed workers receive is the basis of the employer's power and labor discipline strategy.

The whole economy: Integrating the labor discipline model with the Cournot model of competition among firms in product markets provides the basis for analyzing wages, the markup, profits, and unemployment in the whole economy and to understand why the labor market does not clear and why limited competition on product and labor markets will result in lower wages and greater unemployment.

Evidence: Data on job rents, on how wages, and work effort vary with the level of employment, and on employers' monitoring of work effort are consistent with the predictions of the model.

IMPORTANT IDEAS

(in)complete contract	employer	labor discipline model
worker/employee	Solow condition	contingent renewal contract
self-employment	reservation wage	fallback
unemployment benefit	Nash equilibrium	short-side power
Pareto (in)efficiency	no-shirking condition	excess labor supply
long side of the market	workplace amenities	wage curve
marginal revenue product (hours)	whole economy labor market	opportunity cost of capital
zero-profit condition	labor supply	labor demand
minimum wage	competition condition	non-compete clause
no-poaching clause	monopsony	minimum wage
quit rate	Ford model	
disutility (of effort)	employment rent	

MATHEMATICAL NOTATION

Notation	Interpretation
u	utility function of the worker
$\underline{u}$	disutility of effort (parameter of worker's utility function)
y	income received by the worker
e	worker's effort
$\underline{e}$	no-shirking level of effort
p	price that the worker receives from a good she produces
w	wage that the worker receives
$\bar{w}$	worker's reservation wage
B	unemployment benefit (government payment to jobless worker)
c_p	employer's cost per unit of effort w/e
c^{N}	Nash equilibrium cost per unit of effort
t	termination probability
z	worker's fallback position
s	length of unemployment spell
$v(w,e)$	value of job
H	employment rate (whole economy)
h	hours of labor hired by a firm
l	total labor $(h \cdot e)$ hired by an employer
j	probability of remaining unemployed/jobless
u	unemployment rate
q	quit rate
$\pi(\ )$	employer's profit function
μ	markup ratio
γ	productivity per worker hour, which is a_l^{-1}
a_l	number of labor hours required for one of output, equal to γ^{-1}
ρ	opportunity cost of capital

Note on superscripts and subscripts: N: Nash equilibrium (incomplete contract); M: monopsony; other superscripts and subscripts refer to the hours or effort level of the worker or to particular points in figures.

CHAPTER

12

INTEREST, CREDIT, AND WEALTH CONSTRAINTS

[It] cannot be expected that . . . managers . . . of other people's money than their own . . . should watch over it . . . with vigilance.

Adam Smith,
Wealth of Nations (1776)

DOING ECONOMICS

This chapter will enable you to:

- Explain how legal institutions—the rules of the game affecting bankruptcy and limited liability—and the nature of the information available to lenders result in incomplete credit contracts.
- Understand why the Nash equilibrium of a lender-borrower interaction is Pareto inefficient, even with unlimited competition.
- Show that people without wealth may be excluded from the credit market, while wealthier borrowers will pay lower interest rates and can finance larger projects of lesser quality.
- Provide empirical examples of the credit market disadvantages of those with limited wealth.
- Explain how barriers to entry limit competition in the credit market, reduce the total rents available to lenders and borrowers, and redistribute the rents in favor of the lenders.
- Indicate the similarities and differences between the lender-borrower relationship and other principal-agent relationships you have studied, namely Benetton and the subcontractor, and the Ford Motor Company and the worker.
- Show how the credit market provides a way of understanding both the Keynesian multiplier and the way that that monetary policy can affect aggregate demand—particularly expenditure on housing and consumer durables.

12.1 INTRODUCTION: WHY MARY BOLENDER'S CAR WOULD NOT START

Her daughter's asthma was acting up and the 10-year-old had a fever of 103.5 degrees Fahrenheit (39.7 Celsius). Mary Bolender of Las Vegas (US) knew she had to get her to the hospital fast.

But that didn't happen.

Her car would not start. There was nothing wrong with the 2005 Chrysler van. Nothing except that the starter had been remotely disabled by the lender from whom she had borrowed the funds to buy the car. Her monthly check to the lender was three days late. The lender, C.A.G. Acceptance of Mesa, Arizona had agreed to lend her the funds only after installing a device that can be remotely activated to make it impossible to start the car.

In 2014, about two million vehicles in the US had these devices installed.[1] They do more than make the car unusable if you get behind in your loan repayment. Some devices send out loud beeps that become more aggressive as the due date for the check to the lender nears. The devices also can track the location of the vehicle: some include a "geo-fence" that informs the lender if the borrower is no longer driving to and from their job.

The tracking capabilities allow the lender to quickly repossess the car if the borrower does not pay up. For the lender this is a big improvement over the old days in which they hired 'repo men' (polite version: repossession agents) who would cruise neighborhoods looking for the vehicles of borrowers in default, and then tow them away.

To get her van back on the road, Bolender had to pay $389, money that she didn't have that day when her daughter needed treatment. (As far as we know, the child is fine.) "I felt absolutely helpless," she said. John Pena, the general manager of the company that installed the device in her van sees it differently. Without the new technology, he said "we would be unable to extend loans."

Bolender's experience indicates the possible misuse of the devices: on the day she needed to go to the hospital, the lender could not legally have repossessed her van, but they were able to disable it. But Pena, the lender, is right in stressing that without the devices many more people would be rejected when they need a loan, or would be charged extraordinary interest rates.

Borrowing to buy a car is a lot easier than borrowing to buy a meal or pay the rent. The reason? When you buy a car on borrowed money, the loan contract generally gives the lender the right to take the car if you do not repay on the contracted schedule. The car in this case is what is called **collateral**, that is, a transfer of ownership of something of value to

Figure 12.1 Interest rates on payday loans run in the many hundreds of percent annually.

Photo ©mikeledray/Shutterstock

› EXAMPLE Of course borrowers do not lack ingenuity, and ways of disabling the disabler have circulated on the Internet. Oscar Fabela left a film with a date only to find that his car would not start and the device was screeching like a burglar alarm. The two of them got the car going but the date was not impressed. "It didn't end well," he said. In response to borrower pushback, one device maker has produced a fake disabler (named the Decoy) that is installed visibly along with the hidden real device so that people will think they have turned the device off when it is actually still installed. See the following video for an explanation (tinyurl.com/y4gnseel).

COLLATERAL An asset that a borrower pledges to a lender as a security for a loan. If the borrower is not able to make the loan payments as promised, the lender becomes the owner of the asset.

the lender should the loan contract be violated by the borrower. Lending to someone to pay for food or to pay the rent is much riskier for the lender because there is typically no collateral. These are what are termed unsecured loans.

Where lenders find it difficult to ensure repayment or to seize the collateral, they typically do not lend money. But if they do, then the probability of not getting the money back is built into the interest charged. In New York City people in need of cash take out short-term loans to be repaid when their next pay cheque comes in. These payday loans bear interest rates between 350 per cent and 650 per cent per year. The legal maximum interest rate in New York is 25 percent.[2]

We know a lot about these practices in New York because in 2014 the "payday syndicate" offering these loans was charged with criminal **usury** in the first degree and the prosecution described their illegal dealings in some detail. In Illinois, the typical short-term borrower is a low-income woman in her mid thirties ($24,104 annual income), living in rental housing, borrowing between $100 and $200 and paying an average annual rate of interest of 486 percent.[3]

Seemingly extraordinary aspects of modern economies such as "payday loans" at usurious and illegal interest rates and "starter interrupt" devices installed in cars are explained by the incomplete or unenforceable nature of contracts. The promise to repay the loan is sometimes no more enforceable than an employee's promise to work hard and well. These two examples—payday loans and starter interrupt devices—are just a small window into the workings of a **credit market**, the subject of this chapter. In a credit market lenders (banks, payday lenders, and other financial institutions) provide loans to borrowers (people and companies) who commit at some future date to repay the amount borrowed plus an additional percentage of that amount, which is called the interest on the loan. The lenders are the supply side of the market (they are supplying the loans) while the borrowers are the demand side.

We will see that a sizable fraction of populations for which we have data are either unable to borrow at all or unable to borrow as much as they would like at the interest rates being charged by lenders. At the going price (the

USURY Unreasonably, unethically, or illegally high interest rates.

CREDIT MARKET In a credit market, lenders (banks, payday lenders, and other financial institutions) provide loans to borrowers (individuals and companies) who commit at some future date to repay the amount borrowed plus an additional percentage of that amount, termed the interest on the loan.

rate of interest) they are unable to borrow the quantity they would like (or even at all): they are called quantity-constrained borrowers.

Think how odd this is. You would definitely find it strange if you were told that half of all families were unable to purchase some commodity, say, milk at all, even if they had the necessary budget, or that they could purchase 2 liters but not 3. In the next section we will see that this is exactly what occurs in credit markets: large numbers of individuals are unable to "buy" any credit at all or the amount they would like at the interest rates being charged.

12.2 EVIDENCE ON CREDIT AND WEALTH CONSTRAINTS

In the credit market, many who would like to borrow:

- are excluded from borrowing entirely (unable to borrow), or
- face limits on how much they can borrow (can't borrow as much as they wish), or
- pay extraordinarily high rates of interest when they do succeed in getting a loan.

Those who cannot borrow at all are termed credit-market excluded. The credit-market excluded, along with those who face limits on how much they can borrow or face very high rates of interest are called **credit constrained**. Where borrowers face these limitations due to their lack of wealth or low income, they are termed **wealth constrained.**

How do we know if a family or business is credit constrained?

! REMINDER A person or business is said to be credit constrained if: (a) they are excluded from borrowing entirely, (b) they face limits on how much they can borrow, or (c) they pay extraordinarily high rates of interest when they do succeed in getting a loan.

Quasi-experimental measures of credit constraints

Economists look for situations that are like experiments in which we study the actions taken by people who are similar except that some unexpectedly get access to funds (for example an inheritance) and some do not. If a person is not credit constrained, then having extra funds should not lead to any change in business practices or behavior: if a change would have raised profits, then funds could have been borrowed for the purpose.

CREDIT CONSTRAINTS A person or business is said to be credit constrained if: (a) they are excluded from borrowing entirely, or (b) they face limits on how much they can borrow, or (c) they pay extraordinarily high rates of interest when they do succeed in getting a loan.

WEALTH CONSTRAINTS Any restriction on the kinds of contract one can engage in due to a lack of wealth is called a wealth constraint.

So if the surprise arrival of additional funds changes behavior, then it means that before their good luck they were credit constrained. Here is some evidence from the UK:

- In one study the inheritance of approximately $27,374 doubled a typical British youth's likelihood of setting up a business.[4]
- Another study found that people were much more likely to go into self-employment shortly after inheriting wealth, and that inheritance leads the already self-employed to increase the scale of their operations considerably.[5]
- Research on homeowners found that a 10 percent rise in value of housing assets that could be used as collateral in the UK increases the number of startup businesses by 5 percent.[6]

Here is another example. In 2011 the government of Nigeria invited aspiring entrepreneurs under the age of 40 to submit business plans for either startup firms or improvements on existing firms, offering the most promising proposals training in business and networking opportunities. Among those plans judged to be superior, a random selection would be offered an unconditional grant equivalent to almost $50,000. Twenty-four thousand plans were submitted and 720 plans were selected by a lottery from a group of these proposals that were judged to be well conceived.[7]

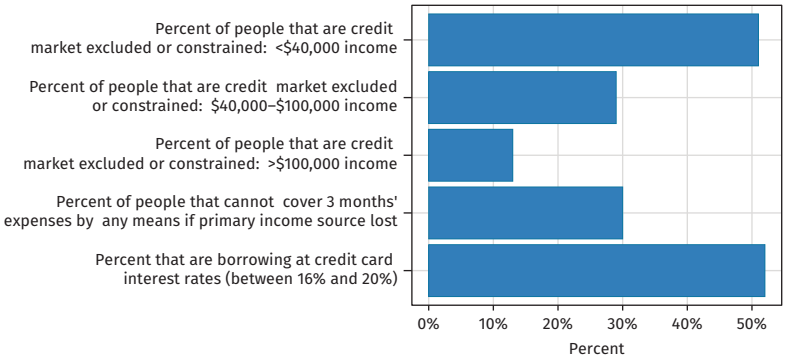
By comparing those randomly selected to receive the grant with those who were not selected, this program—called YouWiN!!—provides an ideal experiment for assessing the ways that credit-market exclusion and wealth constraints limit businesses. If the young entrepreneurs had been able to borrow all they wished at the going rate of interest, then the grant would have had no effect on how they ran their business. It just would have made them richer, allowing them to consume more, or to save and use their $50,000 to purchase stocks in other companies or lend it to others.

But, tracking what the companies did through 2015, that is not what happened. They had good projects that they previously could not finance or they had projects which they had succeeded in financing, but on a scale less they would have been profitable. The lucky winners used their grants to expand their businesses. They purchased more equipment and other capital goods and as a result hired more labor. New firms were created, and they were profitable and survived.

Even more dramatic evidence of credit-market exclusion comes from an experiment in Sri Lanka. A sample of 408 very small businesses were randomly divided into those that received a grant worth about US$100 and the control group that received nothing. The researchers then collected information on the subsequent investments, sales, and profits of the two groups. The firms that received the grants earned extraordinary profits equal to about 60 percent of their capital stock annually. If additional funds made profits of this magnitude available, these firms surely would have

Figure 12.2 Measures of credit market exclusion and credit constraints in the Us in 2019. These data are from US households surveyed in the 2019 and 2020 Survey of Household Economics and Decision making (SHED). This survey, conducted by the US Federal Reserve Board, measures the economic well-being of US households and identifies potential risks to their finances.

Source: https://www.federalreserve.gov/consumerscommunities/shed.htm.



borrowed if they could have (the rate of interest that banks charge in Sri Lanka between 12 and 20 percent). But they could not.[8]

Who is constrained and how many are they?

Another way to measure the extent of credit constraints is simply to ask people if they have been denied credit, or if they *believe* they would be were they to attempt to borrow. This survey-based evidence along with the quasi-experimental data provide an estimate of the extent of credit constraints.

The US Federal Reserve Board (the nation's central bank) surveyed US families in 2019 and 2020, with results for 2019 shown in Figure 12.2. (The results for 2020 may be atypical due to economic impact of the COVID-19 pandemic.) They found, for example, that among loan applicants with incomes less than $40,000 over half were credit constrained, that is unable to borrow at all, or not able to borrow the amount they wished. More than half of all families were paying rates between 16 and 20 per cent a year on their credit card debt (four or five times the rate of interest that homeowners were paying on their mortgages). Thirty percent reported they would be unable to cover their expenses from any source (including savings and borrowing) if they lost their primary source of income for three months (as many did during the pandemic).

Further evidence that credit constraints are common, even in high-income countries, comes from another study using a quasi-experimental strategy. It exploited the fact that credit card borrowing limits are often increased automatically (and from the card-holder's viewpoint, unexpect-

✓ **FACT CHECK** In studies of household surveys, one person in the household is typically designated as the main decision maker or *household head*. Differences in the characteristics of the household head have been shown to be relevant for a variety of economic outcomes, as the study in Italy shows.

edly).[9] If borrowing increases in response to these exogenous changes in the borrowing limit, we can conclude that the person was credit constrained.

Based on the borrowing behavior of US families, the authors found that "increases in credit limits generate an immediate and significant rise in debt."[10] Gross and Souleles estimate the extent of credit limits as follows:

> It is plausible that many of the one-third of households without bankcards are [credit-]constrained . . . Of the two-thirds with bankcards, the over 56 percent who are borrowing and are paying high interest rates (averaging around 16 percent) might also be considered liquidity-constrained, lacking access to cheaper credit. Combined with the households lacking bankcards, they bring the overall fraction of potentially constrained households to over 2/3.[11]

A study of Italian households found that those who did not borrow either because they were denied credit or believed they would be refused credit, were more likely to be larger poorer families, with an unemployed, less well-educated, female, and younger head of household.[12] Moreover, by comparison to families unlikely to face credit constraints, poorer, younger, families with more uncertain sources of income (self-employment rather than pensions, for example) tended to avoid holding risky assets, consistent with the view that credit-constrained people enjoy lower expected income from the investments they do make.

In sum, credit constraints are common in both high- and lower-income economies. A principal-agent model of lending and borrowing with an incomplete credit contract will explain why this is the case, and illuminate its consequences for how the economy works.

CHECKPOINT 12.1 **Credit constraints and COVID-19** What do the data in Figure 12.2 tell you about how U.S. families were likely affected by the 20.4 million jobs that were lost from March 2020 to June 2020 due to the COVID-19 pandemic.[13]

12.3 THE WEALTHY OWNER-OPERATOR CASE

To build up the model we start with two cases that do not involve borrowing with an incomplete credit contract:

- a person wealthy enough to own and operate a risky investment project without borrowing; and
- a person who must borrow the funds to invest in the project but (unrealistically) secures a loan with a complete contract.

The project we consider is risky in the sense that it may fail or succeed, and while the operator of the project (the borrower) can influence which of these is more likely to happen, there is no way to ensure that the project will with certainty yield a positive income.

We continue the analysis of risk in the next chapter, where we introduce what is termed "risk aversion" namely a preference for the "sure thing" over a risky bet with the same expected payoff. (A risk-averse person would prefer \$100 with certainty than a coin flip to see if she receives either nothing or \$200.) The lack of risk aversion is termed risk neutrality, which is what we assume in this chapter.

A risky project

Antonio invests in a project that requires an amount $\$\bar{k}$ to carry out. We will call this amount \$1 but it could represent \$1 thousand or \$1 million. Imagine that the "project" is a machine that has a dial on it by which Antonio can regulate its speed of operation, ranging from 0 to 1. The machine produces goods in proportion to the "speed" at which it is run. But the faster Antonio runs the machine, the greater is the likelihood that the machine will break and destroy itself and all of its output. Going forth, we will use the words "project" and "machine" interchangeably.

In the Benetton and Ford models we have represented the agent's objectives by a utility function. Here we treat his expected income as a measure of his utility, the quantity that he will maximize.

We assume that f, the probability that the machine will break (i.e. fail), is simply the speed at which it is run (so f represents both *f*ail and *f*ast). It surely will *not* break if it is not operated ($f = 0$) and it surely *will* break if it is run at top speed ($f = 1$). If the machine does not fail while in operation, it becomes worthless at the end of the period. The goods produced are available at the end of the period under the condition that the machine has not failed.

So, increases in the speed of the machine f represents *both*:

- *Higher potential income*: Greater income from the investment (higher output for higher f, therefore more goods sold which is the basis of Antonio's income).
- *Greater chance of failure*: Higher probability of failure when he runs the machine faster.

How Antonio evaluates this trade-off depends on how he feels about taking risks. Here we assume that Antonio wants to maximize his expected income (so he does not care about the risk, he is said to be "risk neutral").

Machines differ in how good they are. We let q, a positive constant, represent the *quality* of the project: higher-quality projects result in more

M-CHECK The probability of failure depends only on the speed at which the machine is run, *not* its quality (q). We could represent quality as a reduction in the probability of failure at any given speed, so that, for example, q varies from 0 to 1 and the probability of failure is $(1-q)f$. In this case the top-quality machine (that is, $q=1$) even when run at top speed would be indestructible! But, ruling out that case, this alternative model would produce similar results to what we have assumed here.

income for any given speed at which the machine is run. The project results in either:

- *Success*: income of qf if it succeeds; or
- *Failure*: 0 income if it fails.

The owner-operator's choice of a risk level

Antonio will sell the output of the machine, so we have Antonio's expected revenues from operating it (remember "Pr" should be read as "the probability of"):

$$\begin{aligned}\text{Expected revenues} &= (0\times\text{Pr. Failure})+(qf\times\text{Pr. Success})\\ \hat{r}(f) &= 0\cdot f+qf(1-f)\\ \hat{r}(f) &= qf(1-f) \qquad (12.1)\end{aligned}$$

His expected income $\hat{y}(f)$ from the project must include the opportunity cost of the \$1 that he could have invested in some other project, that is, $1+\rho$ where ρ is the opportunity cost of capital. If the owner-operator had not bought the machine and instead had invested the \$1 it cost at the risk-free interest rate ρ, he would have had (with certainty) $\$1+\rho$ at the end of the period.

His expected income from investing in the project is therefore

$$\begin{aligned}\text{Expected income} &= \text{Expected revenues} - \\ &\quad\ \text{Opportunity cost of the investment}\\ \hat{y}(f) &= qf(1-f)-(1+\rho) \qquad (12.2)\end{aligned}$$

✓ **FACT CHECK** Though a probability of failure of 50 percent may seem high for Antonio's project, in the US data from the Bureau of Labor Statistics say that roughly 50 percent of new businesses fail within their first five years of operation. About 20 percent fail within the first year. Businesses seeking to produce an innovation of some kind—a new app or a new pharmaceutical—are especially likely to fail. But our model does not attempt a realistic picture of risks, it is instead a way to think through the logic of incomplete contracting in risky situations.[14]

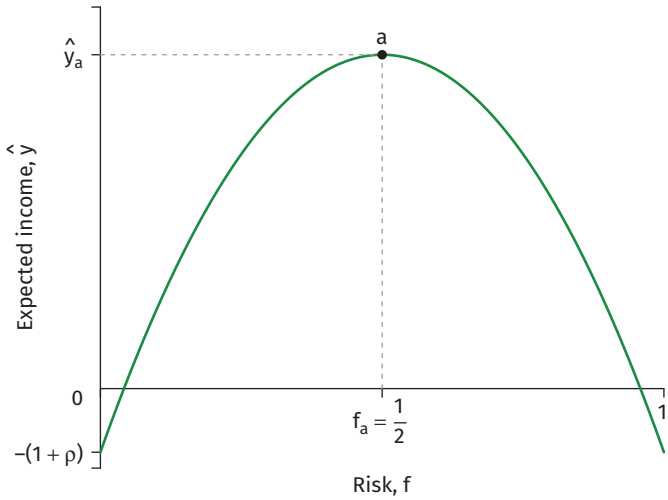
In Figure 12.3 we show Antonio's expected income from the project and how it depends on how fast he runs the machine.

Antonio is the owner-operator—he owns the machine and he owns any output that it produces—so he would vary the speed at which he runs the machine (f) to maximize his expected income on the project and set $f=f_a=1/2$ and (inserting this value in Equation 12.2) have expected income of $q/4-(1+\rho)$, as illustrated by point **a** in Figure 12.3, and shown in M-Note 12.1.

Having determined how fast he will run the machine if he invests in the project, Antonio now has to decide whether to do it. Because we have assumed that his next-best alternative is the risk-free return $1+\rho$ we can see from the figure that he makes an economic profit on the investment (income greater than opportunity cost of capital). So he should undertake the project.

You will see from Figure 12.3, and as is shown in M-Note 12.1 that the opportunity cost of capital ρ affects whether Antonio will undertake the project, but not the speed at which he runs the machine if he does.

Figure 12.3 Risk and expected income. If the machine is not operated at all ($f = 0$) then the expected income is negative, namely the opportunity cost of the \$1 project. Starting from a low level of risk, increasing the risk taken initially increases the expected income ($\hat{y}$). Eventually adopting an even more risky strategy leads to lower expected income. The maximum expected income, $\hat{y}_a(f) = q/4 - (1+\rho)$ is achieved when the operator chooses $f = f_a = \frac{1}{2}$. Read M-Note 12.1 to make sure you understand the calculations.



M-NOTE 12.1 Maximum expected income: the owner-operator case

Antonio will vary f to maximize $\hat{y}(f)$. To do this he uses Equation 12.2 and the first-order condition for a maximum, that is $\frac{d\hat{y}}{df} = 0$:

$$\text{Expected income} \quad \hat{y} = qf(1-f) - (1+\rho) = qf - qf^2 - (1+\rho)$$

$$\text{First-order condition} \quad \frac{d\hat{y}}{df} = q - 2qf = 0$$

$$\text{Isolating } f\text{:} \quad f = \frac{q}{2q} = \frac{1}{2} \tag{12.3}$$

So Antonio maximizes expected income by running the machine at $f = \frac{1}{2}$. Notice that the opportunity cost of capital ρ does not enter into his determining the speed that maximizes the income of the project.

We can now substitute $f = \frac{1}{2}$ into the expected income, $\hat{y}$:

$$\hat{y} = q\frac{1}{2}\left(1 - \frac{1}{2}\right) - (1+\rho)$$
$$= \frac{q}{4} - (1+\rho)$$

Therefore, Antonio's expected income is a positive function of the quality of the project.

CHECKPOINT 12.2 When would Antonio decide *not* to invest in the project?

a. Redraw Figure 12.3 showing a case in which Antonio would decide not to invest in the project.

b. If $\rho = 0.05$, what is the smallest value of q, the quality of the machine, such that Antonio would decide to invest in the project?

12.4 COMPLETE CREDIT CONTRACTS: A HYPOTHETICAL CASE

To introduce credit, we now consider the case in which Antonio has no wealth and therefore he can't self-finance his project (pay for the initial investment himself). Instead, he needs to borrow the funds from a lender, Parama. To allow us later to clarify the difference that the incompleteness of the credit contract makes, we begin with a hypothetical case in which the lender can include in the loan contract the degree of risk that the borrower takes. So the "speed dial" on the machine is not only visible to the lender, but also the information it shows is verifiable and can be used to enforce a contract.

Interest, repayment, bankruptcy, and limited liability

Parama, the lender, is in the business of making money, so she will want to make a profit from her loan. At the end of the period for which the loan is granted, in addition to requiring Antonio to repay the principal—the amount of the loan, \$1 in this case—she will require an additional amount, called the interest on the principal. The interest rate i is a percentage of the principal that is added to the amount the borrower is required to repay. At the end of the period, Antonio is required to repay Parama the amount principal multiplied by δ called the **interest factor** which is:

$$\begin{aligned}\text{Amount repaid} &= \text{Principal} + \text{interest}\\ &= \text{Principal} \times (1 + \text{interest rate}) = \text{Principal} \times (1+i)\\ &= \text{Principal} \cdot \delta \qquad (12.4)\end{aligned}$$

When the principal is \$1, as in the case we model, then the amount repaid equals δ itself, which is equal to $1+i$ (one plus the interest rate). But, just as operating the machine is risky for Antonio, lending money to him is also risky for Parama. The reason is that in most legal systems laws concerning bankruptcy and limited liability mean that if the project fails the lender may not be able to recover the loan. If the project fails, then the lender may not

INTEREST FACTOR The interest factor, δ, is one plus the rate of interest.

take the borrower's house or other assets except those specifically pledged as collateral for the loan. We simplify by assuming that if the project fails, Antonio pays back nothing: an aspect of the model that is crucial to what follows.

Antonio therefore has two possible incomes:

- *Project does not fail:* With probability $1-f$ the project succeeds, giving Antonio revenue of qf. He must also repay the principal his loan plus interest—that is, the interest factor (δ)—so his income is: $qf-\delta$.
- *Project fails*: With probability f the project fails, so he receives no revenue and, because of limited liability, he does not pay back the loan. His income is zero.

Antonio's expected income is the sum of these two incomes—that is, $qf-\delta$ and zero—weighted by the probability that each occurs:

$$\text{Expected income} \qquad \hat{y}(\delta,f) = \underbrace{(1-f)(qf-\delta)}_{\text{success}} + \underbrace{f\cdot 0}_{\text{failure}} \tag{12.5}$$

The lender's expected economic profits

Whether the project fails or succeeds, the opportunity cost of the funds Parama lends to Antonio is $(1+\rho)$. We assume that she, the lender, wishes to maximize expected economic profits, meaning the expected repayment of the loan minus the opportunity cost of the funds lent. Like Antonio there are two levels of income she may receive:

- *Antonio's project does not fail*. This occurs with probability $1-f$, and in this case Parama receives δ.
- *Antonio's project fails*: This occurs with probability f and Parama then receives zero.

So we have Parama's objective, to maximize:

$$\text{Expected profit} \qquad \hat{\pi}(\delta,f) = \overbrace{(1-f)\delta}^{\text{success}} + \overbrace{f\cdot 0}^{\text{failure}} - \overbrace{(1+\rho)}^{\text{opp cost}} \tag{12.6}$$

To summarize so far, here is the complete credit contract game (we use the C superscript to refer to the complete contracts case):

- Parama, the principal, announces both an interest factor δ^C and the speed at which the machine is to be run, f^C.
- Antonio, the agent, either accepts or rejects the contract.
- If he accepts, he operates the machine at the specified speed, and it fails with probability f^C.
- If the machine fails, both Parama and Antonio receive nothing.
- If it does not fail, they receive respectively δ^C and $qf^C-\delta^C$.

This ends the game.

! REMINDER Unfortunately, there are two meanings of the word *principal* that we have to use in this chapter and we don't want to confuse them. First, *principal* is used to describe the lender involved in a principal-agent relationship, such as the lender-borrower relationship we describe. Second, the word principal also has a particular meaning in *finance*: the principal is the initial sum or investment amount lent out by a lender to a borrower, such as the proportion of $1 that Parama lends to Antonio in our example.

! REMINDER Economic profit of a project is its income minus its costs including the opportunity cost of the funds devoted to the project. We use the term 'profits' here to mean economic profits, not accounting profits (which does not take account of the opportunity cost of the funds involved.)

Figure 12.4 The lender's iso-expected-profit curves. The points making up a given iso-expected profit curve indicate combinations of risk (f) and the interest factor (δ) that according to Equation 12.6 result in the same expected profit. The slope of the iso-expected-profit curve is $\frac{\Delta f}{\Delta \delta} = \frac{1-f}{\delta}$, as shown in M-Note 12.2. The iso-expected profit curves are upward sloping because, in order for expected profit to remain constant, an increase in f (the probability of failure) must be offset by an increase in the interest factor. This is also why a decrease in the probability of failure (e.g., a move from **c** to **k**) or an increase in the interest factor (e.g., a move from **c** to **j**) are better for the lender (resulting in their being on a higher iso-expected profit curve).

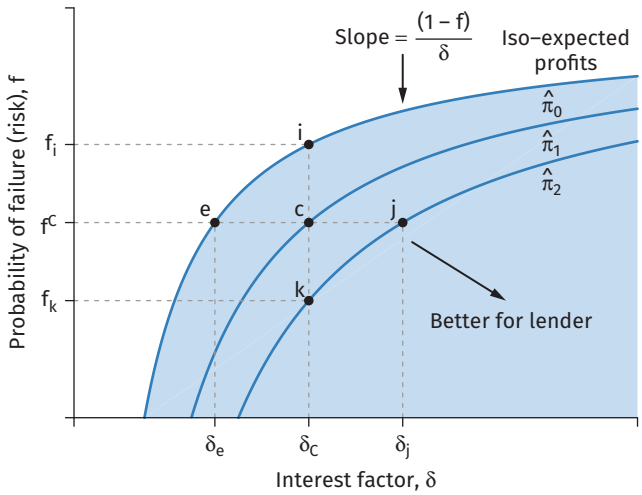


Figure 12.4 shows the lender's iso-expected-profit curves (indicated by $\hat{\pi}$) with higher expected profits represented by the curves that are to the right and lower. Why is this the case?

Remember, iso-expected-profit curves give Parama's evaluation of every point in the space, irrespective of whether that point could occur. So, to answer the question think hypothetically. Comparing points **e**, **c**, and **j** we can see that if the borrower were to run the machine at a speed f^C, the lender would have greater expected-profits if the interest factor, δ is higher, with $\hat{\pi}_0 < \hat{\pi}_1 < \hat{\pi}_2$ on the corresponding iso-expected-profit curves. Similarly, comparing points **i**, **c**, and **k** for a given interest factor δ^C, the lender would have a greater expected profit if the borrower were to take less risk.

The iso-expected-profit curve labeled $\hat{\pi}_0$ is special because it represents the combinations of f and δ such that the expected profit of the lender is zero (meaning accounting profits are just sufficient to offset the opportunity cost of the funds used for the project). The curve labeled $\hat{\pi}_0$ divides the space in the figure into two regions.

The blue-shaded area in the figure are the outcomes satisfying the lender's participation constraint (because for these combinations of f and δ her expected economic profit is positive or zero). She would not be interested in engaging in any lending on terms that lie outside the blue-shaded area. We will see later that this explains why some prospective borrowers are unable to secure a loan at any rate.

M-NOTE 12.2 The lender's *mrs*(δ,f) and the iso-expected-profit curve's slope

Using Equation 12.6, we can find the slope of the lender's iso-expected-profit curves:

$$\text{P's expected profit} \qquad \hat{\pi}(\delta,f) = (1-f)\delta - (1+\rho) \tag{12.7}$$

To find the slope of the iso-expected profit $\frac{df}{d\delta}$, we want to find the changes in f and δ that are consistent with no changes in the lender's expected profits, or staying on the same iso-expected-profit curve. To do this, we need to find the total derivative of the lender's expected profit function and set it equal to zero:

$$d\hat{\pi} = \frac{\partial\hat{\pi}}{\partial\delta}d\delta + \frac{\partial\hat{\pi}}{\partial f}df = 0$$

Or, using more compact notation of the partial derivative of $\hat{\pi}$ with respect to x, namely $\hat{\pi}_x \equiv \frac{\partial\hat{\pi}}{\partial x}$, we have:

$$d\hat{\pi} = \hat{\pi}_\delta d\delta + \hat{\pi}_f df = 0 \tag{12.8}$$

This means that, for two points on an iso-expected, profit curve, the difference in expected profits associated with a small difference in the interest factor ($\frac{\partial\hat{\pi}}{\partial\delta}d\delta$) is exactly compensated by the opposite signed difference in expected profits associated with the small difference in the risk level ($\frac{\partial\hat{\pi}}{\partial f}df$), so that the total difference in expected profits is zero.

We can use the derivatives ($\hat{\pi}_\delta$ and $\hat{\pi}_f$) of expected profit function (Equation 12.7) and rearrange Equation 12.8 to find $\frac{df}{d\delta}$, the slope of the iso-expected-profit curve:

$$\text{Slope of iso-expected-profit curve} \qquad \frac{df}{d\delta} = -\frac{\hat{\pi}_\delta}{\hat{\pi}_f} = \frac{(1-f)}{\delta} \tag{12.9}$$

Equation 12.9 says that for $f < 1$, the slope of the lender's iso-expected-profit curve is positive because $\delta > 0$. The lender's marginal rate of substitution is the negative of the slope of her iso-expected-profit curve, which we can therefore say is the following:

$$mrs(\delta,f) = -\frac{(1-f)}{\delta}$$

The borrower's participation constraint and the lender's profit maximization

The lender will propose a contract specifying both the speed at which the machine will be run f and the interest factor to be paid at the end of the loan

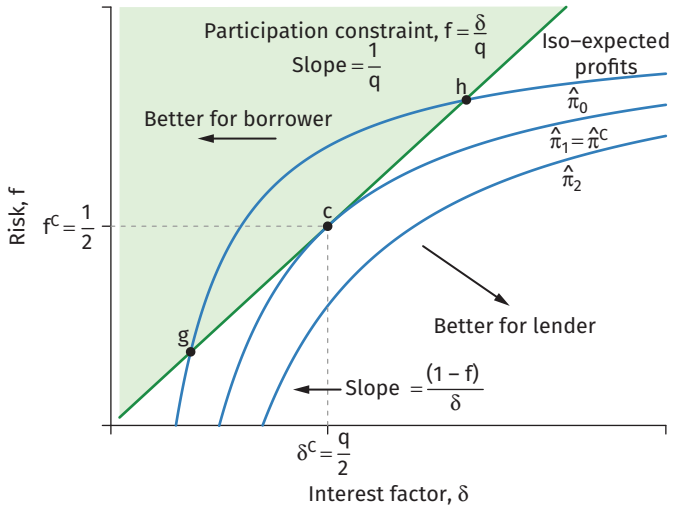
period, δ. The lender's offer must be sufficiently attractive to the borrower so that he will accept. This depends on what the borrower's other money-making opportunities are.

For simplicity we assume that Antonio has no other opportunities so his fallback option z is zero. So, using Equation 12.5, the participation constraint limiting the lender's contract is:

$$\text{Borrower's expected income} \geq \text{Borrower's fallback option}$$
$$(1-f)(qf-\delta) \geq z = 0 \tag{12.10}$$

The participation constraint divides the space of contractual terms in Figure 12.5 into hypothetical contracts that Antonio would accept (the green-shaded area) and that he would reject. Contracts whose terms (f and δ) satisfy the participation constraint as an equality are on the green

Figure 12.5 The lender's expected profit-maximizing choice of a contract (f^C, δ^C) when the contract is assumed complete. The lender receives higher expected profits at points that are lower and to the right. The figure shows that the best that the lender can do is to offer the borrower a contract indicated by point **c**, where her highest feasible iso-expected-profit curve is tangent to the borrower's participation constraint. At point **c**, the lender requires the borrower to run the machine at half speed ($f^C = \frac{1}{2}$), and the lender charges the borrower an interest factor of $\frac{q}{2}$. Points **g** and **h** both lie on the borrower's participation constraint, so the borrower is indifferent among them and point **c**. But, the single lender would not choose points **g** or **h** as they are not profit-maximizing (lying on $\hat{\pi}_0$) as the lender can make a take-it-or-leave-it offer to the borrower and increase their expected profits to $\hat{\pi}_1$.



ray from the origin. These contracts provide Antonio with the minimum expected income (that is, zero) sufficient for him to accept the loan.

As shown in M-Note 12.3 we can rearrange Equation 12.10 as:

$$\text{Participation constraint:} \qquad f \geq \frac{\delta}{q} \tag{12.11}$$

In what follows we express Equation 12.11 as an equality, or

$$f = \frac{\delta}{q} \tag{12.12}$$

We call this his willingness to sell because for any limit on Antonio's risk-taking (that is, f) that Parama would like to enforce in the contract, the participation constraint tells us the greatest interest factor that Parama can charge, given the quality level (q) of the project. Equation 12.12 tells Parama the following. For a given quality of the machine (q in the denominator), if she charges a higher interest factor δ, in order to continue satisfying Antonio's participation constraint, Parama must let Antonio run the machine faster (he will have a higher f because of her higher δ in the numerator). The negative of the slope of the participation constraint given by Equation 12.12 is the marginal rate of transformation of higher interest into faster speeds of the machine.

The lender can now implement the expected profit-maximizing contract by finding the point on the participation constraint (PC) that is also on the highest iso-expected-profit curve. This is point **c** where:

$$\text{Marginal rate of substitution} = \text{Marginal rate of transformation}$$

$$-\text{Slope of iso-expected-profit curve} = -\text{Slope of PC} = -\frac{\Delta f}{\Delta \delta}$$

$$-\frac{(1-f)}{\delta} = -\frac{1}{q} \tag{12.13}$$

As shown in the M-Check, we can use Equation 12.12 to replace the δ by fq in Equation 12.13, giving us the result that the lender will set $f^C = \frac{1}{2}$, as above using the C superscript to indicate the complete contract case.

Notice that the level of risk implemented when risk can be determined by the lender in an enforceable contract is identical to the risk chosen by the owner-operator. In M-Note 12.3 we show that this is the case because the problem solved by the owner-operator is mathematically identical to that solved by the lender who is able to enforce a level of risk by contract. The owner-operator's chosen level of risk is by definition Pareto efficient because he has maximized his expected income, so he cannot be made better off. And there is nobody else involved.

In the interaction between the lender and the borrower, the outcome is also Pareto efficient. This is because the lender maximized profit subject to a constraint given by the borrower's utility level (his participation constraint) and therefore the lender implemented a Pareto-efficient outcome. The lender implemented the $mrs = mrt$ rule equating the slope of

! REMINDER As before, for simplicity we assume that a contract is accepted if it is *not worse* than the alternatives.

M-CHECK From Equation 12.13

$$\frac{(1-f)}{\delta} = \frac{1}{q}$$

using Equation 12.12 to replace δ

$$\frac{(1-f)}{fq} = \frac{1}{q}$$

$$\frac{(1-f)}{f} = 1$$

$$(1-f) = f$$

$$\text{we get } f^C = \frac{1}{2} \tag{12.14}$$

REMINDER As we have seen in Chapters 4, 5, and 10, when one actor maximizes utility or profit subject to a constraint that the other's utility level not be less than a given minimum value, the result is by definition Pareto efficient. The participation constraint requires that the borrower's expected income (which is his utility) not be less than zero.

the isoprofit curve with the slope of the participation constraint, as shown in Equation 12.14.

But the marginal rate of transformation here is also the borrower's marginal rate of substitution, that is, the negative of the slope of the participation constraint. As a result, by implementing the $mrs = mrt$ rule the lender also unintentionally implemented the $mrs^P = mrs^A$ rule.

The risk levels that result in the two scenarios are the same. In both cases the decision maker–Antonio in the owner-operator case and Parama the lender in the lender-borrower case–was in a position to capture the entire gains–that is the total expected rents–from the transaction. In both cases the maximum possible total expected rent was $\frac{q}{4} - (1+\rho)$. They wanted the rent to be as large as possible.

The difference between the two cases is that in the owner-operator case Antonio had enough wealth to carry out the project himself so he got all of the rents. In the lender–borrower case he was without wealth and had to borrow funds from Parama; and she got all of the rents. Antonio's wealth (or lack of it) is what explains the difference in the distribution of the gains from the exchange.

Except for this important difference in who got the rent, complete contracting allows the lender to implement an outcome that mirrors what the borrower would do if he were both the owner and operator of the machine. In this sense, complete contracting erases the distinction between lender and borrower, and reinstates the world of the owner-operator. The results change when we turn to real world, incomplete credit contracts.

M-NOTE 12.3 **A complete credit contract mimics the owner operator**

To study the case of complete contracting shown in Figure 12.5, we start with the borrower's participation constraint Equation 12.10 satisfied as an equality, so:

$$\text{Borrower's expected income} \quad (qf-\delta)(1-f) = 0 \tag{12.15}$$

For Equation 12.15 to be true, either $(qf-\delta) = 0$ or $(1-f) = 0$.

- $(1-f) = 0$: If this is true, then it must be that $f = 1$, in which case the machine will fail with certainty. So this cannot be true.
- Therefore, the participation constraint can be rewritten $(qf-\delta) = 0$.
- Or, equivalently $\delta = qf$
- And rearranging this we have $f = \frac{\delta}{q}$.

With respect to the lender, we need to consider her expected profit function (see Equation 12.6):

$$\text{Lender's expected profit} \quad \hat{\pi}(\delta, f) = (1-f)\delta - (1+\rho) \tag{12.16}$$

continued

We can now substitute $\delta = qf$ into Equation 12.16 such that:

$$\hat{\pi}(f) = qf(1-f) - (1+\rho) \tag{12.17}$$

The function that we maximize for the lender here is the same as in the owner-operator case, namely Equation 12.2 and in M-Note 12.1. As in that case the lender will choose f to maximize profits given by Equation 12.17 and set $f = \frac{1}{2}$ (see Equation 12.14). Then inserting this value into Equation 12.17, we see that the lender has an expected profit of $\hat{\pi} = \frac{q}{4} - (1+\rho)$, the same as in the case of the owner-operator.

CHECKPOINT 12.3 **An offer you *could* refuse** Explain why Equation 12.11 is the participation constraint limiting the profits that the lender can make in the complete contracting case. If the lender makes an offer such that $\delta/q > f$, what will be the borrower's response?

12.5 THE GENERAL CASE: INCOMPLETE CREDIT CONTRACTS

A principal-agent model based on an incomplete credit contract can explain why lenders use devices like car disablers for borrowers in default or why so many families are unable to borrow the amounts they would like. We begin with borrower—Antonio—who has no wealth; later we will show why borrowers like Antonio are likely to be excluded from the credit market entirely.

The credit contract will be incomplete if either or both of two critical pieces of information are not known by (or unverifiable for) the lender.

- *Hidden attributes*: The quality (q) of the borrower's project—in our example how good the machine is.
- *Hidden actions*: The level of risk that the borrower takes (f)—in our example, how fast he runs the machine.

Either or both could be the case, but to focus on a concrete example (the speed of the machine) we assume that the quality of the project is known to the lender (no hidden attributes), but that the level of risk that the borrower takes is a hidden action.

Information about the speed at which the machine is run is either asymmetric (only Antonio knows it) or non-verifiable (Parama may know it as well, but cannot use it in court to enforce a contract). So the lender can no longer contractually set the degree of risk taken by the borrower, and can set only the interest factor, δ.

Here is the game:

- *The lender* is the principal, who knows the borrower's best-response choice of a risk level for each level of the interest factor ($f(\delta)$), and as

! REMINDER As in the Benneton model, the principal's name begins with a P (here, Parama, the lender) and the agent's name begins with an A (here, Antonio, the borrower). We will use "principal," "lender," and "Parama" interchangeably (and likewise "agent," "borrower," and "Antonio").

! REMINDER In Chapters 10 and 11 we presented similar games, first explaining the agent's best-response function (the quality of goods to provide, how hard to work) and then using the best-response function as the incentive compatibility constraint on the principal's choice of a price or a wage. We follow the same logic here.

! REMINDER Recall that a principal-agent relationship arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the principal's objective in such a way that there is a conflict of interest between the two.
- *Incomplete contract*: the agent's actions or attributes that are of interest to the principal are not subject to an enforceable contract.

first mover she determines and proposes the interest factor (δ) that will maximize her expected profits.

- *The borrower* (the agent) then selects the level of risk to take (the speed of the machine, $f(\delta)$).
- *Chance* then intervenes: the project either fails with a probability f or does not fail with probability $1-f$.
- *The borrower* then repays the interest factor to the lender if the project has not failed (the machine has not destroyed itself) and repays nothing if the project failed.

This ends the game.

To select her first move, the principal uses backward induction, anticipating what the agent, the second mover will do in response to each of the possible interest factors that the principal might offer. Because the lender (the principal) uses the borrower's best-response function as the incentive compatibility constraint for her expected profit-maximizing problem, we begin with that.

The borrower's best response

Given the interest factor set by the lender, δ, the borrower will choose f to maximize his expected income (repeating a slightly rearranged Equation (12.5)):

$$\underbrace{\hat{y}(\delta,f)}_{\text{Expected income}} = \underbrace{qf(1-f)}_{\text{Expected revenues}} - \underbrace{\delta(1-f)}_{\text{Expected repayment}} \tag{12.18}$$

The final term in Equation 12.18 makes it clear that by taking a loan from the principal, the agent is acquiring a kind of **insurance**: he is reducing the level of risk to which he is exposed. This is because:

- the level of risk taken is not enforceable by contract; and
- bankruptcy and limited liability laws make it unlikely that loans invested in failed projects will be repaid.

› EXAMPLE When you insure your car against theft or damage, you are paying the insurance company to transfer these risks from you to the company: the company, not you, will pay for a new car if yours is stolen or destroyed.

As a result, the lender bears some of the costs of the risky decisions made by the borrower: if the machine fails, the borrower does not repay the loan. So in choosing a risk level, the borrower does not 'own' all of the consequences of his choice. Recall that the risk level chosen when the decision maker bears all of the risk is one-half, as was the case with the owner-operator. But this is not the case for the borrower: because he does not bear all the consequences of the risk level that he chooses, he will run the machine at a value of f greater than one half. How much faster depends on the interest factor that the lender will choose.

INSURANCE Any costly action one can take that reduces the level of risk to which one is exposed.

The trade-off faced by the borrower is that running the machine faster so that it produces more revenue also increases the likelihood that it will fail and produce nothing. So running the machine faster:

- runs the risk of failure, and getting nothing instead of qf, the revenues the borrower would get if it did not fail; but it also
- increases the borrower's expected revenue by $q(1-f)$; and
- increases the probability that the amount owed to the lender, δ, will not have to be repaid (due to the failure).

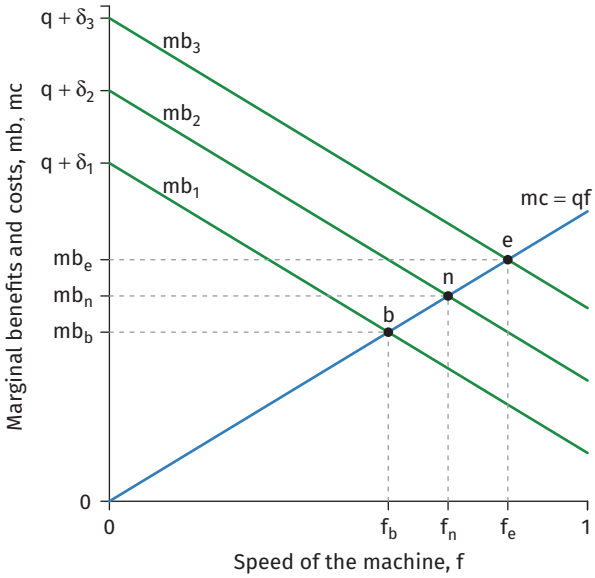
The expected income-maximizing level of risk is the value of f satisfying the following rule (derived in M-Note 12.4), expressed in terms of marginal benefits and marginal costs of increasing f:

$$\text{Marginal benefits} = q(1-f) + \delta = qf = \text{Marginal costs} \tag{12.19}$$

The rule is illustrated in Figure 12.6.

! REMINDER The fact that the lender bears some of the risk resulting from decisions made by the borrower is another example of a negative external effect, similar to the positive external effects that the worker's hard work or the quality of the subcontractor's good confers on the principals (the employer and the buyer) with whom they interact. Other negative external effects you have studied include the downward pressure one firm's sales had on the demand for another firm's products in the Cournot model, and the reduction in the fish caught by one fisherman resulting from the fishing of another.

Figure 12.6 Marginal benefits and marginal costs of choosing greater risk (f) given the interest factor (δ) determined by the lender. The marginal cost of running the machine faster is the revenue that will be lost if the machine fails. This is greater for higher levels of f because the machine produces more (if it does not fail) when it is run faster. The marginal benefit is the sum of the greater revenue made possible by running the machine faster (if it does not fail) plus the reduction in the probability that the loan will have to be repaid (if the machine fails). Marginal benefits are shown for three levels of the interest factor: $\delta_3 > \delta_2 > \delta_1$. The figure shows that, for a low value of δ, the borrower will run the machine slower than for a higher value of δ. Points **b**, **n**, and **e** here have counterparts in Figure 12.7, where we look at the same problem, but from a different vantage point.



To see how this equation determines the expected income-maximizing level of risk for the borrower to take, imagine that you are him. The lender has offered δ_1 and you are thinking about taking the lowest possible positive level of risk running the machine at a speed just barely above $f = 0$. The marginal cost of taking a little more risk is qf—the revenue of the project run at the speed f conditional on it not failing—that is what you would lose if the project failed. But this is almost nothing because you are running the machine so slowly that the revenues it generates if it does not fail are tiny.

On the other hand, the marginal benefits of greater risk would be substantial. At a very slow speed, running the machine a little faster would produce more revenue if it did not fail, and the probability of that occurring is very small. So the increase in your expected revenue from running the machine a little faster is $q(1-f)$. Second, by running the machine faster you reduce the likelihood that you will have to pay back the loan, saving you an expected value of δ. Comparing the marginal costs and benefits of taking a little more risk, you would definitely want to increase f.

Now consider the opposite case: you are running the machine just below the top speed $f = 1$. Running it even faster would increase the chance of it failing, and this would be a substantial cost because when run at near top speed, it produces a great amount of goods (if it does not fail). The marginal benefit of increasing f would still include the reduced likelihood of having to repay the loan. But the other marginal benefit, the increase in expected revenue from running it faster would be close to zero because it is already virtually certain to fail. Running the machine at near top speed cannot be the expected income-maximizing strategy for the borrower.

Similar reasoning eliminates all other levels of risk-taking except for the level of risk such that the marginal benefits for additional risk-taking equal the marginal costs. Equation 12.19—the intersection of the marginal benefits and costs lines in the figure—is the basis for the borrower's best-response function. We can rearrange it as follows:

$$\text{A's best response} \qquad f(\delta) = \frac{1}{2} + \frac{\delta}{2q} \tag{12.20}$$

Equation 12.20 says that if the lender offers the interest factor δ_1 then select the risk level f_b.

Equation 12.20 shows that:

- The borrower's best response to an interest factor of zero is to run the machine at a speed of one half, $f = \frac{1}{2}$, just as did the owner-operator and the borrower with a complete contract.
- The slope of the borrower's best-response function is $\frac{\Delta f}{\Delta \delta} = \frac{1}{2q} > 0$, so the borrower takes on more risk (higher f) if the lender chooses a higher interest factor.
- The level of risk, f, chosen by the borrower is less the higher is the quality q of the project (given the same δ).

- The best-response function is flatter ($\frac{1}{2q}$ is less) the larger the quality q. In other words, the higher the quality, the lesser is the effect of the interest factor in raising the risk level.

M-NOTE 12.4 The borrower's first-order condition

The borrower's best-response function, $f(\delta)$, is derived by finding for each interest rate that the lender may offer, the level of the borrower's chosen level of risk that will maximize his expected income $\hat{y}(\delta, f)$. To do this, we partially differentiate Equation 12.18 (repeated here) with respect to f and set the result equal to zero, or

$$\underbrace{\hat{y}(\delta,f)}_{\text{Expected income}} = \underbrace{qf(1-f)}_{\text{Expected revenues}} - \underbrace{\delta(1-f)}_{\text{Expected repayment}}$$

$$= qf - qf^2 - \delta + \delta f$$

$$\text{First-order condition:} \quad \frac{\partial \hat{y}}{\partial f} = q(1-2f) + \delta = 0 \qquad (12.21)$$

$$\text{Rearranging} \quad q - 2qf + \delta = 0 \qquad (12.22)$$

To derive the best-response function, Equation 12.22 can be rewritten to isolate the f term:

$$2qf = \delta + q$$

$$\text{Divide through by } 2q \quad f = \frac{\delta + q}{2q}$$

$$\text{Borrower's BRF} \quad f(\delta) = \frac{1}{2} + \frac{\delta}{2q} \qquad (12.23)$$

In the text we will use the fact that Equation 12.22 can be rewritten to represent the benefits and costs of running the machine faster:

$$q(1-2f) + \delta = 0$$

$$q - (qf + qf) + \delta = 0$$

$$q - qf + \delta = qf$$

$$\text{Marginal benefits} \quad = q(1-f) + \delta = qf = \quad \text{Marginal costs} \qquad (12.24)$$

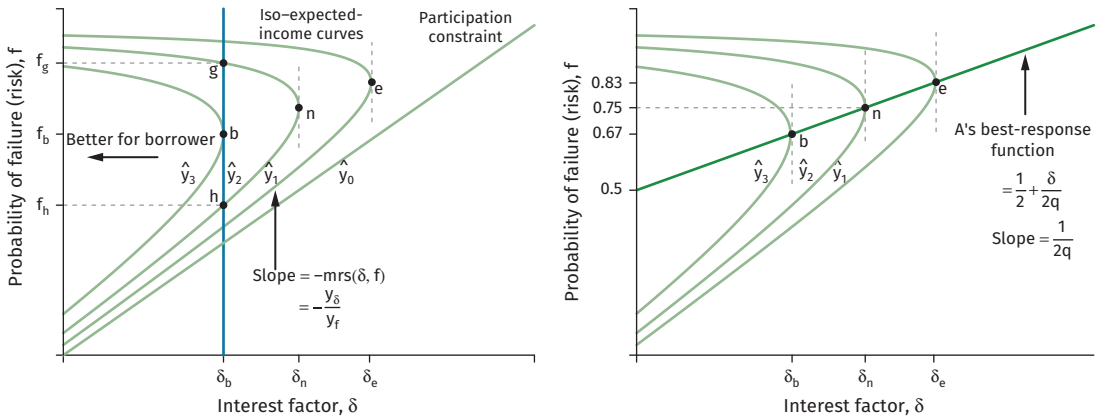
CHECKPOINT 12.4 Risk and quality Use Equation 12.20 to show that the borrower with a higher quality project chooses a lower level of risk.

! REMINDER Best-response functions In developing the Benetton model we first derived the subcontractor's (the agent's) best-response function by finding the quality level that equated the marginal benefits and costs of their action (quality of good supplied) in Equation 10.5. We then showed how the same best-response function could be derived from the agent's iso-expected-income curves in Figure 10.6. We follow the same procedure here. Having shown how equating the marginal costs and benefits of running the machine faster provides us with an expression for the borrower's best-response function (Equation 12.20), we now represent the same ideas using the borrower's iso-expected-income curves.

12.6 THE NASH EQUILIBRIUM LEVEL OF RISK AND INTEREST

We now illustrate the borrower's best response with the help of his iso-expected-income curves shown in Figure 12.7. Each point in the figure represents some hypothetical outcome of the interaction between the principal and the agent, that is, a combination of a degree of risk f and an interest factor δ (no matter how unlikely). An iso-expected-income curve (as in Chapter 10) gives us all such combinations that result in the agent having some given level of expected income.

Figure 12.7 The borrower's iso-expected-income curves and best response to the lender's choice of δ. Points **n**, **b**, and **e** in this figure are the same outcomes as **n**, **b**, and **e** in Figure 12.6. Successively higher rates of risk (speed of the machine) are adopted by the borrower as the interest factor goes from low, to moderate, to high. Remember the participation constraint (from Equation 12.12) is $f = \frac{\delta}{q}$.



(a) The borrower's iso–expected-income curve

(b) The borrower's best-response function

! REMINDER In Figure 12.7 the iso-expected-profit curves are rotated u-shaped, except for the linear participation constraint along which the expected income is zero (the borrower's fallback option). You already encountered a similar linear iso-curve in Chapter 8 Figure 8.7 in which the iso-profit curve for the firm for zero profits is linear, unlike the other iso-profit curves.

In panel (a), starting with the participation constraint where the expected income is zero, the expected income is greater for curves that are closer to the vertical axis, as indicated by the horizontal arrow, that is $0 = \hat{y}_0 < \hat{y}_1 < \hat{y}_2 < \hat{y}_3$. This is because for a given probability of failure, that is, comparing points on the curves horizontally, a lower interest factor raises the borrower's expected income. As we did in Chapters 1 and 4, we can think of the iso-expected-income curves as contours in a map of a hill that is sloping up to the left (the horizontal arrow is pointing up "hill").

For a given interest factor, that is, for points along a vertical line, e.g. **g**, **b**, and **h**, the comparison is more complicated. These points have counterparts in Figure 12.6 in which we compared the marginal benefits and costs of running the machine faster. At point **g** in both of these figures, for example, at the given interest factor δ_b, Antonio is running the machine too *fast*. In Figure 12.6 it was clear that the marginal benefits of running it faster fell short of the marginal costs. Here you can see that by slowing down the machine to f_b Antonio reaches a higher "contour" moving from iso-expected income of $\hat{y}_2$ at point **g** to $\hat{y}_3$ at **b**.

With the same interest factor δ_b, at point **h**, on the other hand, Antonio is running the machine too *slowly*. He can increase his expected-income (moving from $\hat{y}_2$ to iso-expected-income curve $\hat{y}_3$) by increasing the speed from f_h to f_b and moving to point **b**. Again, return to Figure 12.6 and notice that at **h** marginal costs are lower than marginal benefits and if he moves

to **b** they will be equal and, as shown in Figure 12.7 (a), he will increase his income from $\hat{y}_2$ to $\hat{y}_3$.

The borrower will not choose points such as **g** and **h** and will instead choose a point like **b**, **n**, or **e** at which his marginal benefits equal his marginal costs. For each of points **b**, **n**, and **e**, for a given level of interest factor δ_e, δ_n, or δ_b, the choice of corresponding risk level that maximizes that borrower's expected income is that point at which the iso-expected-income curve is tangent to the vertical line corresponding to that interest factor.

If you think of the line **gbh** as a "trail" across a shoulder of the hill shown by the contours, it rises from **h** to **b**, is flat at **b**, and then descends to **g**. So **b** is the highest point on the trail. That is why **b** is a best response. In the right panel of Figure 12.7, we see that the best-response function is made up of points like **b**, **n**, and **e** at which the iso-expected-income curve is vertical.

CHECKPOINT 12.5 The iso-expected profit

a. Explain in your own words why the iso-expected-income curves are upward-sloping for low values of f and downward-sloping for higher levels of f.

b. Use the borrower's best-response function (Equation 12.20) and what you know about the slope of the borrower's iso-expected-income curves to explain why the best-response function is made up of points where the iso-expected-income curves are *vertical*.

M-NOTE 12.5 The borrower's iso-expected-income curves

Here we:

- derive the borrower's marginal rate of substitution; and
- show why the indifference curve is vertical at any point on the borrower's best-response function.

The borrower's expected income (Equation 12.18, repeated here) is:

$$\underbrace{\hat{y}(\delta,f)}_{\text{Expected income}} = \underbrace{qf(1-f)}_{\text{Expected revenues}} - \underbrace{\delta(1-f)}_{\text{Expected repayment}} \tag{12.25}$$

The borrower's marginal rate of substitution.

We proceed as we have in finding the mrs in other models, e.g. in M-Note 12.2. Using the notation $\frac{\partial \hat{y}}{\partial x} \equiv \hat{y}_x$, we take the total derivative of the borrower's expected income function, and set it equal to zero:

$$d\hat{y} = \hat{y}_f df + \hat{y}_\delta d\delta = 0$$

$$\text{which, rearranged, is} \quad \frac{df}{d\delta} = \frac{-\hat{y}_\delta}{\hat{y}_f} = \text{slope of iso-exp-inc curve} \tag{12.26}$$

continued

REMINDER We have already (in M-Note 12.2) derived the $mrs(\delta, f)$, which is the negative of the slope of the lender's iso-expected-profit curves, and the slope of the borrower's best-response function which is the negative of the marginal rate of transformation or $-mrt = \frac{\Delta f}{\Delta \delta} = \frac{1}{2q}$.

The iso-expected-income curve is vertical on the BRF.

We know that at any point on the borrower's best-response function it must be that $\hat{y}_f = 0$ which is the first-order condition defining the BRF. But this means that the slope given by Equation 12.26 for any point on the best-response function is some quantity divided by zero, so the slope is undefined, which is to say the line is vertical.

The lender's expected profit maximization

The lender seeks to maximize her expected profits. But now that we have dropped the unrealistic assumption that she could enforce her chosen level of risk-taking on the borrower, she is bound by a tighter constraint than the participation constraint of the borrower (requiring merely that his expected income be at least zero). Because the contract is now incomplete, the lender will have to provide him with incentives to operate the machine more prudently than he otherwise would.

We show the interaction of lender and borrower in Figure 12.8. The lender is restricted to points on the borrower's best-response function. Profit increases as f is less and δ is greater (that is, down and to the right). She will select the interest factor for which the borrower's best-response function is tangent to her highest possible iso-profit curve, point **n** in the figure. This is the value of δ such that the marginal rate of substitution (from the principal's) iso-expected-profit curve is equal to the marginal rate of transformation (from the borrower's best-response function):

$$minus \text{ slope of iso-profit} = mrs(\delta, f) = mrt(\delta, f) = minus \text{ slope of BRF}$$

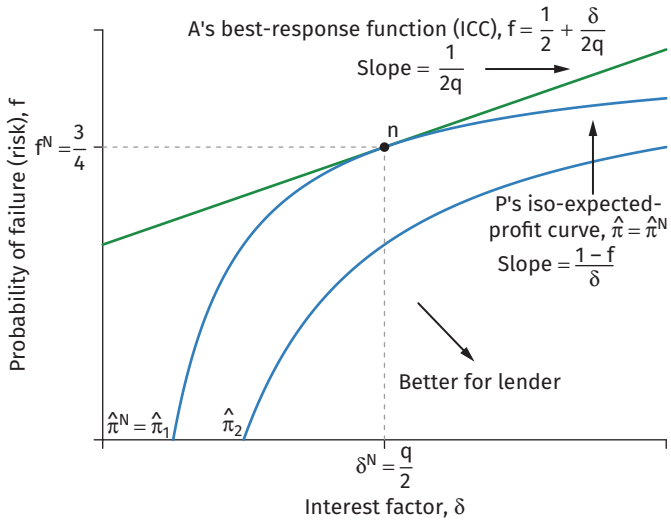
This rule for selecting the interest factor is equivalent to finding the δ such that the marginal benefit to the lender of raising the rate of interest (more repayment if the machine does not fail) is equal to the marginal cost to the lender (increased probability of no repayment due to the faster pace of the machine).

The marginal benefit of raising the interest factor for a given level of risk is $1-f$, that is, the probability that the interest factor will be paid. The marginal cost of raising the interest factor is $\delta \frac{\Delta f}{\Delta \delta}$ because raising δ increases the risk taken by the borrower by $\frac{\Delta f}{\Delta \delta}$, which results in a loss of δ to the lender if the machine fails. So we have:

$$\text{Marginal benefit } = (1-f) = \frac{\delta}{2q} = \text{Marginal cost} \tag{12.27}$$

We show in M-Note 12.8 that we can use the borrower's best-response function and this rule for lender's choice of the interest factor to determine the Nash equilibrium interest factor that the principal will choose, and the level of risk that the borrower will take in response.

Figure 12.8 The Nash equilibrium of the lender-borrower interaction. The agent (borrower) best responds to the principal (lender) by choosing a level of risk (f) on his best-response function based on the interest factor (δ) he is offered by the lender. The Nash equilibrium outcome is (δ^N,f^N) at the tangency of the borrower's best response and the lender's iso-expected-profit curve where the lender's $mrs(\delta,f) = mrt(\delta,f)$. This figure is similar to those in the previous two chapters showing the best-response function of the agent (the sub-contractor and the worker) and the iso-expected-profit curves of the principal (Benetton and Ford Motor Company). But there is a difference: here the line is the agent's best-response function (the incentive compatibility constraint), and the curves (the iso-expected profits) represent the principal's objectives.



Returning to Figure 12.8, we know that the outcome indicated by point **n**, namely (f^N, δ^N) is a Nash equilibrium (hence the N superscripts) because:

- Given the interest factor that the principal has offered, namely, (δ^N), then $f^N(\delta^N)$, is the best the borrower can do (it is a point on his best-response function); and
- Given the strategy adopted by the borrower as described by his best-response function $f(\delta)$, then δ^N is the best the lender can do (it is the solution of her constrained optimization process).

Notice in the second bullet, that the lender best responds to the *strategy* of the borrower—the best-response *function*—not to the *action* of the borrower, that is, the level of risk that he takes in the Nash equilibrium f^N. If it were the case that the borrower would choose f^N whatever level of the interest factor the lender chose, then, as we will see, she would charge a

M-CHECK We show the equivalence of $mrs = mrt$ and $mb = mc$ for the lender this way:

$$mrs(\delta,f) = mrt(\delta,f)$$
$$-\frac{(1-f)}{\delta} = -\frac{\Delta f}{\Delta\delta} \equiv f_\delta$$
$$(1-f) = \delta\frac{\Delta f}{\Delta\delta}$$

And because $\frac{\Delta f}{\Delta\delta} = \frac{1}{2q}$

we have $(1-f) = \frac{\delta}{2q}$

Marginal benefit = Marginal cost

Remember, marginal here refers to variations of the interest factor by the principal.

much higher interest factor. But this is not the case: the risk chosen by the borrower depends on the interest factor chosen by the lender.

M-NOTE 12.6 The participation constraint as a (linear) iso-expected-income 'curve'

In this note we explain why the iso-expected-income curve corresponding to the participation constraint (namely, for $\hat{y} = \hat{y}_0$) is linear in contrast to the other curves. We know that the participation constraint (PC) is:

$$\text{Participation constraint:} \quad f = \frac{\delta}{q} \tag{12.28}$$

$$\text{So slope of PC:} \quad \frac{df}{d\delta} = \frac{1}{q}$$

This means that the PC has a constant slope: it is a line. But the participation constraint is also an iso-expected-income curve, so we also know from Equation 12.26 that its slope is:

$$\frac{df}{d\delta} = \frac{-\hat{y}_\delta}{\hat{y}_f} \tag{12.29}$$

Can these two slopes be the same quantity (as they must be)?

$$\frac{df}{d\delta} = \frac{-\hat{y}_\delta}{\hat{y}_f} = \frac{1}{q} \tag{12.30}$$

Using the values of the two derivatives from Equation 12.25:

$$\hat{y}_f = q(1-2f)+\delta$$
$$\hat{y}_\delta = -(1-f)$$
$$\text{Plugging in:} \quad \frac{df}{d\delta} = \frac{-\hat{y}_\delta}{\hat{y}_f} = \frac{(1-f)}{q(1-2f)+\delta} = -\text{mrs} \tag{12.31}$$

Because our question is about points along the participation constraint we can use Equation 12.28 to replace f by δ/q in Equation 12.31:

$$\frac{df}{d\delta} = \frac{-\hat{y}_\delta}{\hat{y}_f} = = \frac{\left(1-\frac{\delta}{q}\right)}{q\left(1-2\frac{\delta}{q}\right)+\delta}$$
$$= \frac{1}{q+\delta-2\delta/q}\cdot\left(\frac{q}{q}-\frac{\delta}{q}\right)$$
$$= \frac{(q-\delta)}{q(q-\delta)}$$
$$\text{Slope of PC:} \quad \frac{-\hat{y}_\delta}{\hat{y}_f} = \frac{1}{q}$$

CHECKPOINT 12.6 The best response and *mrs*=*mrt* Confirm that the lender's choice of interest factor, $\delta^N = \frac{q}{2}$ by substituting the agent's best-response function: $f = \frac{1}{2} + \frac{\delta}{2q}$ into Equation 12.27.

M-NOTE 12.7 **Lender's profit-maximizing interest factor**

We used the condition for $mrs = mrt$ in Equations 12.13 and 12.27 to find the Nash equilibrium value of δ. Here we do the same derivation starting from the profit maximization of the lender by differentiating the profit function of the lender with respect to δ (the variable over which the lender has control) and setting it equal to zero. We maximize the lender's expected profit, $\hat{\pi}(\delta, f)$, subject to the borrower's best-response function. To do this, we write the lender's expected profits are as in (Equation 12.6), but f now depends on δ so we write f as $f(\delta)$, giving the new expected profit function:

$$\text{Lender's expected profits} \quad \hat{\pi}(\delta, f(\delta)) = \delta\big(1 - f(\delta)\big) - (1+\rho)$$
$$= \delta - \delta f(\delta) - (1+\rho)$$

Since f depends on δ, in the form of $f(\delta)$, we use the product rule in the first-order condition (recall that $f_\delta \equiv \frac{\partial f}{\partial \delta}$):

$$\text{First-order condition} \quad \frac{d\hat{\pi}}{d\delta} = 1 - \big(f(\delta) + \delta f_\delta\big) - 0 = 0$$
$$= 1 - f(\delta) - \delta f_\delta = 0$$

$$\text{Rearranging} \quad \frac{1 - f(\delta)}{\delta} = f_\delta \tag{12.32}$$

$$\text{Or,} \quad \delta = \frac{1 - f(\delta)}{f_\delta} \tag{12.33}$$

Equation 12.32 tells us that the lender's profit-maximizing interest factor is found by implementing the $mrs = mrt$ rule or, what is the same thing, the slope of the iso-expected-profit curve is equal to the slope of the best-response function.

M-NOTE 12.8 **Nash equilibrium risk and interest factor: numeric example**

We have two equations to determine two unknowns: (δ, f). They are:

- The borrower's best-response function, Equation 12.20: $f(\delta) = \frac{1}{2} + \frac{\delta}{2q}$
- The lender's first-order condition for choosing the interest factor, Equation 12.32: $\frac{1-f(\delta)}{\delta} = f_\delta$

We can now substitute in the borrower's best-response function $f(\delta)$ to find the interest factor that maximizes the lender's expected profit:

$$\frac{1 - (\frac{1}{2} + \frac{\delta}{2q})}{\delta} = f_\delta$$
$$\frac{\left(\frac{1}{2} - \frac{\delta}{2q}\right)}{\delta} = \frac{1}{2q}$$
$$\frac{1}{2} - \frac{\delta}{2q} = \frac{\delta}{2q}$$
$$\frac{1}{2} = \frac{\delta}{q}$$
$$\delta^{\mathrm{N}} = \frac{q}{2}$$

continued

This is the Nash equilibrium level of interest that the lender chooses. To complete the analysis we need to find the borrower's Nash equilibrium best-response level of risk ($f(\delta^N)$). We substitute δ^N into Equation 12.20:

$$\text{Borrower BRF} \qquad f(\delta) = \frac{1}{2} + \frac{\delta}{2q}$$

$$\text{Substitute in } \delta^N = \frac{q}{2} \qquad f = \frac{1}{2} + \frac{\left(\frac{q}{2}\right)}{2q}$$

$$f^N = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

The rules of the game matter: Total rents and their distribution

Table 12.1 summarizes the cases addressed so far, and demonstrates how the three different rules of the game support different results. Two points are important:

- *Distribution of the rents*: Both the owner who operates the machine without any other player involved, and the lender having a complete contract with a borrower get all of the rents (last column, first two rows). The borrower gets his fallback option (which is zero). But in the incomplete contract case (the third row), the borrower receives a share of the total rents.
- *Total gains from exchange*: The owner-operator case and the complete contract maximize the total rents available to the participants, meaning the output of the machine minus the opportunity cost of the funds used. This is because the person making the decision about the risk taken—the owner-operator or the lender—is the residual claimant on the income of the project. Because they get all of the rents (the previous bullet) they 'own' the consequences of their decisions. Comparing in the final column the third row with the first two, the total income and profit from the project are 25 percent less when the contact is incomplete. This is

Table 12.1 Nash equilibrium outcomes in the borrower–lender model depend on the rules of the game: ownership and the nature of the contract. Remember, for the lender's profit maximization process, the PC is the participation constraint for the borrower and ICC is the borrower's incentive compatibility constraint. The fact that the interest factor is the same under a complete and an incomplete contract is an artifact of the particular model we are using: the interest factor under the complete contract could be either higher or lower than under the incomplete contract.

Case: Differing rules of the game	Borrower's PC or ICC $f(\delta, q)$	Interest factor, δ	Risk, f	Expected rents per period: Borrower's expected income ($\dot{y}$)	Lender's expected profit ($\hat{\pi}$)	Total
Owner-operator (no loan)	—	—	$\frac{1}{2}$	—	—	$\frac{q}{4} - (1+\rho)$
Complete contract	PC: $f = \frac{\delta}{q}$	$\frac{q}{2}$	$\frac{1}{2}$	0	$\frac{q}{4} - (1+\rho)$	$\frac{q}{4} - (1+\rho)$
Incomplete contract	ICC: $f = \frac{1}{2} + \frac{\delta}{2q}$	$\frac{q}{2}$	$\frac{3}{4}$	$\frac{q}{16}$	$\frac{q}{8} - (1+\rho)$	$\frac{3q}{16} - (1+\rho)$

because the borrower does not own all of the consequences of his risk-taking, and as a result takes 50 percent more risk.

12.7 MANY LENDERS: COMPETITION AND BARRIERS TO ENTRY

While an urban neighborhood may have a single payday lender and many small towns just a single bank or money lender, most prospective borrowers–whether individuals or firms–can shop around. Many people, as we saw, use their credit cards as an alternative source of loans.

We have so far represented Parama as the sole lender from whom Antonio can borrow. We now introduce additional lenders competing with Parama in the credit market. To do this we embed the principal-agent model of the lender and borrower in a model of the entire market.

How many competitors there will be depends on the barriers to entry in the credit market and the opportunity cost of capital. As in Chapter 9, we take the perspective of a wealthy individual or firm considering entering the credit market. The entrant will consider the profits that they can expect to make if they successfully enter the market as well as the probability b that they will fail in their attempt, due to the barriers to entry. There are two categories of risk that the entrant faces: it may *fail* to transact any loans, or it may *succeed* in entering the credit market and then make a loan that is not repaid.

✓ **FACT CHECK** Barriers to entry in the credit market are anything that makes it likely that an attempted entry into the market—by a new credit card company, for example—will fail. Getting a license to operate as a bank, for example, requires meeting a set of demanding standards, often including a substantial level of assets. Entering the credit card lending market faces a challenging chicken and egg problem: merchants will not honor cards that few customers have, but customers will not have a card that few merchants recognize.

Suppose the entrant considers devoting some amount to their entry attempt, say, one million euros. Then their expected profits, in millions of euros, are:

- *if they fail to enter*, which occurs with probability b, they lose $1+\rho$, which is the opportunity cost of the assets they devoted to their project;
- *if they succeed in entering which occurs with probability* $(1-b)$ their expected revenues minus opportunity cost of capital is the same as the lenders already in the market, called the *incumbent firms*:

$$\text{Incumbents' expected profits} \quad \hat{\pi} = \delta(1-f)-(1+\rho) \tag{12.34}$$

where $(1-f)$ is the probability their loan is repaid.

They pay the opportunity costs of their entry attempt $1+\rho$ with certainty (whether they fail or succeed), while they have the opportunity to profit only if they do not fail. So the costs are *known* and the revenues are *uncertain*.

They will decide to attempt to enter as long as their expected revenues do not fall short of costs, so that their expected profits are not less than zero:

$$\text{Entrant's expected profits} \quad \hat{\pi}^b = \overbrace{(1-b)\left[\delta(1-f)-(1+\rho)\right]}^{\text{successful entry}} - \overbrace{b(1+\rho)}^{\text{failure}} \geq 0$$

$$\hat{\pi}^b = (1-b)\delta(1-f)-(1+\rho) \geq 0$$

The b superscript is a reminder that this is a possible entrant in the market facing barriers to entry. If this condition is not satisfied, firms will not attempt to enter the credit market. And because there is always some exit of firms for reasons outside our model, if:

$$\hat{\pi}^b = (1-b)\delta(1-f) - (1+\rho) \leq 0$$

then, no new firms will enter and the number of firms in the credit market will fall. The competition condition ensuring that the number of firms in the credit market will be constant is therefore:

$$\text{Competition condition: } \hat{\pi}^b = (1-b)\delta(1-f) - (1+\rho) = 0$$

$$\text{or, rearranged} \quad \delta(1-f) = \frac{1+\rho}{1-b} \tag{12.35}$$

Equation 12.35 is called the *credit market competition condition*. It is analogous to the competition condition for the whole economy based on the entry and exit of firms in the market for goods and services (in Chapter 11). In this chapter where there is no danger of confusing the two we will call it the competition condition.

We can express this condition (Equation 12.35) in terms of the expected profits that incumbent firms will make in equilibrium (subtracting $1+\rho$ from both sides):

$$\begin{aligned}\hat{\pi}^N(b) &= \delta(1-f) - (1+\rho) = \frac{(1+\rho)}{1-b} - (1+\rho) \\ &= (1+\rho)\left(\frac{b}{1-b}\right)\end{aligned} \tag{12.36}$$

Equation 12.36 is the expected equilibrium profit for incumbent lenders when the degree of competition among them is given by the quantity $(1-b)$. We use the superscript N here because the equation is based on the competition condition, which ensures that the number of lenders in the loan market is constant. Equation 12.36 shows that, as was the case with firms competing to sell the products they produced in Chapter 9, expected economic profits will be positive if there is limited competition in the credit market ($b > 0$).

If there are no barriers to entry so $b = 0$ (the case we call unlimited competition) the competition condition becomes $\delta(1-f) - (1+\rho) = 0$ which is called the *zero-profit condition* because in this case the expected profits of the entering and incumbent firms are equal and both equal to $\hat{\pi}^N(b) = \hat{\pi} = 0$. You have already seen this in Figure 12.4 where the iso-expected-profits curve labeled $\hat{\pi}_0$ is the competition condition for the case with unlimited competition (so that expected economic profits are zero).

In Figure 12.9 we show the competition condition. This is the iso-expected-profit curve for incumbent firms yielding sufficient profits so that the number of firms in the market *does not change*. The competition condition divides the space of credit market outcomes into:

M-CHECK We can derive Equation 12.36 as follows:

$$\begin{aligned}\hat{\pi}^N &= \delta(1-f) - (1+\rho) \\ &= \frac{(1+\rho)}{1-b} - (1+\rho) \\ &= (1+\rho)\left(\frac{1}{1-b} - 1\right) \\ &= (1+\rho)\left(\frac{1}{1-b} - \frac{1-b}{1-b}\right) \\ &= (1+\rho)\left(\frac{1-1+b}{1-b}\right) \\ \Longrightarrow \hat{\pi}^N &= (1+\rho)\left(\frac{b}{1-b}\right)\end{aligned}$$

! REMINDER The zero-profit condition requires that when barriers to entry are absent—the case of unlimited competition—expected economic profits are zero in a Nash equilibrium.

- outcomes with positive expected profits for an entering firm so that firms would enter and
- outcomes with negative expected profits for an entering firm so that firms would not enter and the total number of firms in the market would fall.

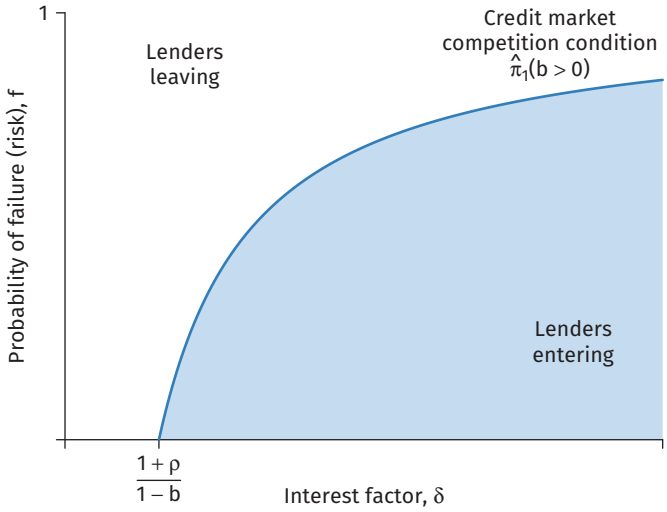
Credit market equilibrium with barriers to entry

An equilibrium in the credit market requires that the number of lenders (or the total amount of lending, if lenders are differing in size) be unchanging. So the outcome must be somewhere along the competition condition curve in Figure 12.9. But where? To answer this question we need another condition–some additional information about the relationship between δ and f– that we know must be true. This is provided by the borrower's best-response function.

The interest factor and risk level (δ^N,f^N) in the Nash equilibrium of the credit market as a whole–not just a single lender interacting with a single borrower, like Parama and Antonio–requires two things:

M-CHECK The competition condition provides a single equation giving the values of f and δ consistent with the amount of lending remaining constant. To determine the equilibrium value of these two variables, we need two equations, not just one. In Chapter 11 we saw that the wage curve alone did not determine the equilibrium level of wages and employment. The additional equation we added in that case was the competition condition based on barriers to entry in the goods market. Here we start with the credit market competition condition and will add a second equation based on the borrower's best response function.

Figure 12.9 The credit market competition condition for some given level of barriers to entry, *b*. The curve is composed of values of the interest factor δ and the level of risk f that satisfy Equation 12.35. The horizontal axis intercept shows that lesser barriers to entry and greater competition among lenders would be represented by a competition condition above and to the right of the one shown, as illustrated later in this chapter in Figure 12.17. In this figure, the level of barriers to entry is $b = 0.3$.



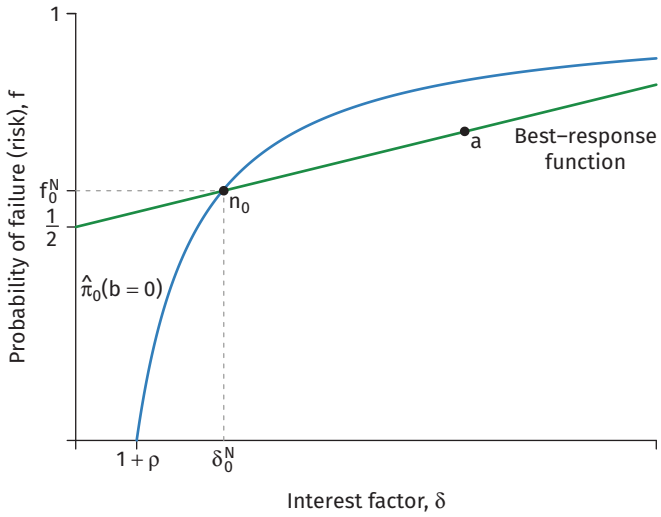
- *Incentive compatibility*: As before, the level of risk taken by borrowers must be incentive compatible given the interest factor charged by lenders, that is, it must be on their best-response function.
- *Competition condition*: And now, we further require that the number of firms in the credit market must be constant, so Equation 12.35 must be satisfied.

In Figure 12.10 we show these two equations and their intersection. You can see that this outcome satisfies both the competition condition (in blue) and the incentive compatibility constraint (in green). But are the lenders maximizing their expected profits?

To explore this, imagine that you are Parama or another lender, charging an interest factor of δ_0^N, with Antonio the borrower taking risk f_0^N. This outcome is at a point where your iso-expected-profit curve (the competition condition itself) is not tangent to the borrower's best-response function. So you might think that by charging a higher interest factor you could move to the right of $\mathbf{n_0}$ on the best-response function, to point **a** in the figure which is on a higher iso-expected-profit curve (not shown).

But competition in the credit market (even with barriers to entry) has changed the game. Antonio's fallback option is no longer to get zero as it was in his one-on-one interaction with Parama, but instead to borrow funds from another lender. And if Parama *did* raise the interest factor there definitely would be an alternative lender ready to lend to Antonio. So she could not charge more than δ_0^N and maintain her borrowers. This is another

Figure 12.10 Equilibrium in the credit market given by the borrower's best-response function and a competition condition. For this illustration we used $b = 0.3, q = 10$, and $\rho = 0.05$.



example of the analysis in Chapter 9: competition works by changing the players' fallback options.

CHECKPOINT 12.7 **Lenders entering and leaving the credit market** Explain why lenders will enter the credit market if the values of δ, f, and b are as indicated in the blue shaded portions of Figure 12.9.

12.8 WEALTH MATTERS: BORROWING WITH EQUITY

You know from the introduction to this chapter that a common practice of lenders is to require that the borrower provide some asset called collateral, the ownership of which will be transferred to the lender if the borrower does not repay. Collateral requirements are common when loans are provided for purchasing a home or a car (the home or car itself is the collateral asset). Those with limited wealth often cannot provide collateral for anything but a house or car loan, and for this reason have difficulty securing credit for other purposes such as starting up a business, or retraining to gain new skills.

EXAMPLE Wealth includes the market value of a home, car, any land, buildings, machinery, or other capital goods that a person owns, and any financial assets such as shares or bonds. Debts are subtracted—for example, the mortgage owed to the bank. Debts owed to the person are added. The flow of valued services associated with a home or a car include shelter and transport. Other forms of wealth yield incomes as with land (rent), capital goods (profits), and financial assets (interest).

Equity and collateral

An alternative arrangement, more common when the loan is to start or expand a business, is for the borrower to share in the risk of the project by investing some of his own **wealth** in the project. One's own wealth invested in a project is called **equity**. Lending to a borrower who has invested his own wealth in a project protects the lender to some degree from the borrower recklessly risking the funds.

Borrowers generally have some wealth, and if the expected **income** of the project is greater than the income from the individual's next-best alternative, it may be in the borrower's interest to invest equity in the

EXAMPLE A broader definition of wealth includes what is termed human capital, meaning the individual's skills, connections, and other capacities that contribute to a flow of labor income (called earnings). We will use the term *wealth* to refer to assets that may be used as collateral or equity, that is, excluding human capital.

WEALTH Stock of things owned (or the value of that stock) that yields a flow of income or other valued services to the owner.

EQUITY One's own wealth (rather than borrowed funds) invested in a project. There is a second entirely different use of the term, meaning the character of being fair, as in "an equitable division of the pie."

INCOME The largest amount that a household or person can consume over a given period of time without reducing the value of wealth (their stock of assets, minus any outstanding debt).

project or to provide collateral. This will alter the borrowing game between the lender and the borrower.

First, contrary to our assumption, the lender may not know q, the quality of the project. In this case, investment of the borrower's *own* wealth is a credible signal of the borrower's assessment of quality of the project, a *hidden attribute*. As we will see presently, in competitive equilibrium those with less wealth will need superior projects to obtain financing, so the borrower has an interest in overstating a project's quality in order to secure a loan. Knowing this, the lender would rather lend to a borrower who has risked his own assets by providing equity or collateral, which is a promising sign to the lender about the borrower's assessment of the quality of the project.

The second reason, and the one modeled here, is that the discrepancy between the objectives of the lender and borrower concerning the choice of the level of risk (this is the *hidden action*) would be reduced if the borrower invested in the project and thus shared some of the risk of failure with the lender. Investing equity or posting collateral does not make the loan contract complete, but we shall see that equity reduces the degree of conflict of interest between the lender and the borrower. It does this because the borrower now "owns" some of the consequences of his choice of risk.

Before studying the credit market as a whole, we begin with the two-person one borrower and one lender interaction in which the borrower has invested some equity in the project. As shown in M-Note 12.10 and 12.11, the analysis of a borrower posting collateral is qualitatively similar to the case when the borrower has equity (we model the case of collateral in section 12.14).

Best-response level of risk by a borrower with equity

The borrower, Antonio, now has wealth k. Other than this, the structure of the game between him and Parama is unchanged. His assets, worth k, could be invested in some alternative project (possibly a government bond) that after one period (year or other time unit) will with certainty yield him $(1+\rho)k$ in income. If Antonio invested these funds as equity in acquiring the "machine," he would then borrow the amount remaining to fund the \$1 project, that is, $1-k$. His expected returns (including the opportunity cost of the foregone returns on the alternative asset) would be:

$$\underbrace{\hat{y}(\delta,f)}_{\text{Expected income}} = \underbrace{qf(1-f)}_{\text{Expected revenues}} - \underbrace{\delta(1-f)(1-k)}_{\text{Expected repayment}} - \underbrace{(1+\rho)k}_{\text{Opportunity cost of capital}} \tag{12.37}$$

On the right-hand side, the expected revenues are the same as before the borrower invested in the project. But Equation 12.37 differs from Equation 12.18—the case of the borrower with no wealth invested in the project—in

two ways, both due to the fact that Antonio now has some of his own capital invested as equity:

- there is an opportunity cost of investing in this project (the last term on the right); and
- he has borrowed only $(1-k)$ so he will now repay Parama the lender $\delta(1-k)$ if the machine does not fail, rather than δ which was the case previously.

To simplify, we use the terms wealth and level of equity committed to the project interchangeably: agents devote all their wealth to the project, if they devote any.

The fact that Antonio has an alternative use of his funds does not affect the speed at which he will run the machine. But the fact that he now has less to pay back if the machine does not fail changes his incentives. Here is why. As before, Antonio, the borrower will select f to maximize $\hat{y}$ by equating the marginal benefits and marginal costs of operating the machine faster. But these are now (as shown in M-Note 12.9):

$$\text{Marginal benefits} = \text{Marginal costs}$$
$$q(1-f)+\delta(1-k) = qf \tag{12.38}$$

Comparing this rule for the borrower selecting f with the same rule where equity was absent (that is $q(1-f)+\delta = qf$) we can see that the right-hand side of Equation 12.38 is identical to when the borrower had no equity.

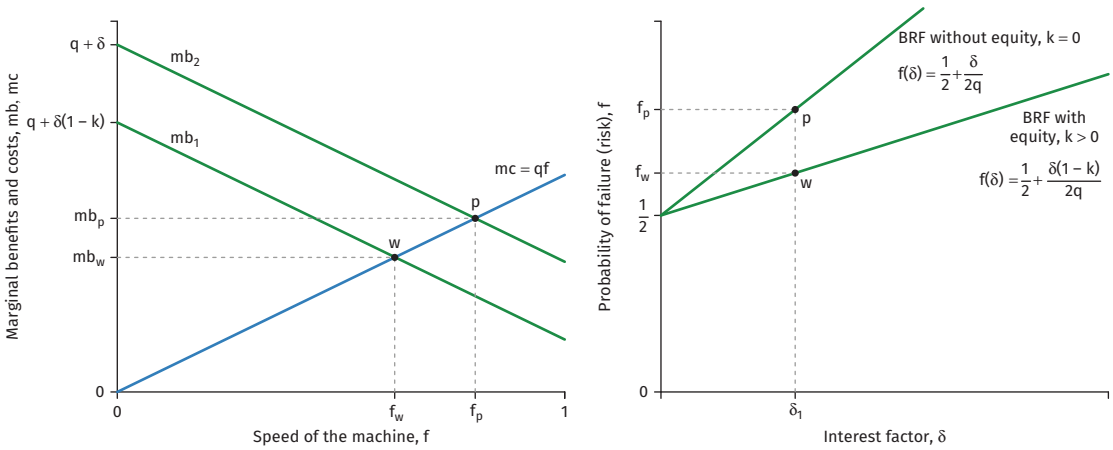
But the marginal benefit of increasing the speed of the machine is now less. Because the borrower has his own equity in the project the loan to be repaid is less, so the benefit of running the machine faster (increasing the risk of failure and not having to repay the loan) is also less. This is why having the borrower's equity in the project reduces the conflict of interest between the borrower and the lender.

The marginal benefit and marginal cost of taking on greater risk are depicted in Figure 12.11 in which we compare two cases. In one Antonio (the borrower) is poor and does not have any wealth to use as equity and the second in which he has wealth ($k<1$) which he invests in the project. The figure shows that, when he has wealth, Antonio's marginal benefits of running the machine faster are lower. Because of this, for a given interest factor δ_k, he will run the machine more slowly at f_w (he selects point **w**, for wealthy). When he is poor, if the lender selects the same interest factor, he will select point **p** (for poor) and a higher risk level at f_p.

Rearranging Equation 12.38–Antonio's rule for selecting f–we have his best-response function, shown in Figure 12.11:

$$\text{Best response with equity} \qquad f(\delta,k) = \frac{1}{2} + \frac{\delta(1-k)}{2q} \tag{12.39}$$

Figure 12.11 The choice of a risk level when the borrower has equity. In panel (a), the marginal benefit lines (labeled mb) are the left-hand side of Equation 12.38, with $k = 0$ (mb_2) and $k > 0$ (mb_1). Comparing points **p** and **w** for a given level of borrower's equity k and an interest factor (δ_1) determined by the lender, the borrower's expected-income-maximizing level of risk, f, will be lower when the borrower has equity. In panel (b), the borrower's best-response function with equity is similar to the best response without equity except that its slope is now $\frac{1-k}{2q}$ and so is lower (flatter) than when $k = 0$. Points **p** and **w** correspond to the choices of a poor and wealthy borrower respectively for a given interest factor (δ_1).



(a) Marginal benefits and costs of increased risk with and without equity

(b) The borrower's best-response functions with and without equity (k)

Equation 12.39 is identical to the best-response function where there is no equity except for the $(1-k)$ term (as you can see by setting $k = 0$.). Notice four things about this best-response function for the borrower with equity:

- A higher interest factor is still associated with greater risk-taking: as before, the best-response function is upward-sloping.
- As $k \to 1, f \to \frac{1}{2}$: for levels of equity close to complete equity financing ($k = 1$) he will approximate the prudent expected-income-maximizing risk choice of owner-operator ($f = 0.5$). If $k = 1$, the case is called "complete equity financing" and there is no principal. He borrows nothing.
- For higher levels of equity of the borrower the best-response risk level for any given interest factor is less (comparing points **p** and **w** in the two panels of Figure 12.11). This is because the marginal benefit of increasing risk is smaller.
- The borrower's best-response risk becomes less sensitive to the interest factor as his level of equity increases: its slope is now $\frac{1-k}{2q}$ rather than just $\frac{1}{2q}$. So, the best-response function is flatter.

M-NOTE 12.9 A borrower with equity selects the best-response level of risk

We start with Equation 12.37:

$$\begin{aligned}\hat{y}(f,\delta,k) &= qf(1-f)-\delta(1-k)(1-f)-(1+\rho)k\\ &= qf-qf^2-\delta(1-k)(1-f)-(1+\rho)k\end{aligned}$$

To find the best-response function we differentiate his expected income function with respect to f and set the result equal to zero:

$$\frac{\partial\hat{y}}{\partial f} = q-2qf+\delta(1-k)=0 \qquad (12.40)$$

Isolating f, gives the BRF $\quad f(\delta) = \dfrac{1}{2}+\dfrac{\delta(1-k)}{2q}$

Rearranging Equation 12.40 gives us the $mb = mc$ relation:

$$\begin{aligned}q-2qf+\delta(1-k) &= 0\\ q-(qf+qf)+\delta(1-k) &= 0\\ q-qf+\delta(1-k) &= qf\\ \text{mb} = q(1-f)+\delta(1-k) &= qf = \text{mc}\end{aligned}$$

CHECKPOINT 12.8 Equity Why would a bank prefer to lend funds to a borrower who has invested some of their own wealth in a project—equity, that is, $k > 0$—compared to one without equity ($k = 0$)?

12.9 EXCLUDED AND CREDIT-CONSTRAINED BORROWERS

Using this new best-response function so as to take account of the equity that the borrower has invested in the project, we now consider the Nash equilibrium of the entire credit market. Here prospective borrowers differ in the amount of wealth they are able to invest as equity in the project.

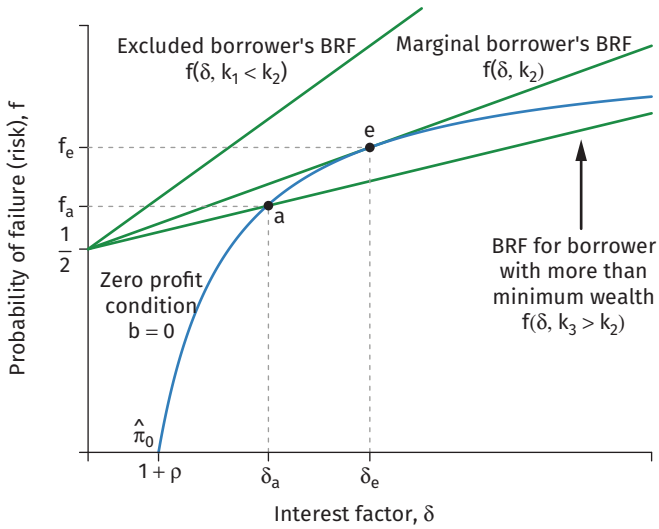
Remember for an outcome to be a Nash equilibrium, it must be a point on both:

- the borrower's *best-response function* which will vary with the amount of equity he has invested k; and
- the *competition condition*, which will vary with the degree of barriers to a prospective lender seeking to enter the credit market, b.

Credit-constrained and excluded would-be borrowers

We illustrate the market in Figure 12.12 for the case of unlimited competition. The competition condition with no barriers to entry ($b = 0$) is the iso-expected-profit curve labeled $\hat{\pi}_0$. This is the zero-profit condition. Also shown are three best-response functions for borrowers with differing amounts of wealth to invest in their project. In Nash equilibrium for the market we have the following:

Figure 12.12 Credit market exclusion and credit constraints. The competition condition (Equation 12.35) is: $\delta(1-f) = \frac{(1+\rho)}{(1-b)}$. Potential borrowers with $k_1 < k_2$ are unable to find a lender willing to lend to them. Those with more wealth ($k_3 > k_2$) pay a lower interest factor.



- *Limited wealth, excluded borrower.* The best-response function of those with limited equity (k_1) lie wholly above the competition condition. If lenders extended loans to these borrowers, the expected profits would be negative. As a result, borrowers with wealth k_1 are unable to borrow. They are the *credit market-excluded* or **excluded borrowers**.
- *Modest wealth, credit-constrained marginal borrower.* The borrower whose wealth, k_2, is just enough so that his best-response function is tangent to the competition condition at point **e** is called the **marginal borrower**. This is because he has exactly enough wealth to secure a loan. The interest factor and risk level for the marginal borrower will be $(\delta_2, f(\delta_2))$.
- *Wealthy, credit-constrained borrower paying a lower interest factor.* For the borrower with wealth $k_3 > k_2$ the Nash equilibrium is at point **a**, where you can see that he pays a lower rate of interest than the marginal borrower.

EXCLUDED BORROWER A borrower who is unable to obtain credit.

MARGINAL BORROWER A marginal borrower is a borrower with just enough wealth to secure a credit contract with a lender.

Wealth constraints on quality and size of projects

In addition to paying lower rates of interest, borrowers with more wealth will be able to finance larger projects and projects of lower quality. To see this take the lender's eye perspective on borrowers: Which would you rather interact with? The simple answer is you prefer borrowers who will not increase their risk-taking by very much if you charge a higher interest factor. You would like to find borrowers with flatter best-response functions. That is why lenders will not transact in any way with the excluded borrowers (their best-response functions are too steep).

Now recall that for the borrower who has invested k in the project whose quality is q we have

$$\text{Best response with equity} \qquad f(\delta, k) = \frac{1}{2} + \frac{\delta(1-k)}{2q}$$

$$\text{and, slope of the BRF} = \frac{(1-k)}{2q} = \frac{\text{amount of the loan}}{\text{twice the quality of the machine}}$$

Two conclusions follow:

- *Wealthier borrowers can finance inferior projects*: a wealthier person (higher k) can have a flatter best-response function and therefore receive a loan, even if the quality of the project q is less than the quality of the project of a borrower excluded because he has too little wealth.
- *Wealthier borrowers can finance larger projects*: the slope of the best-response function depends on the size of the loan, that is, the difference between the project size and the amount the borrower invests. So putting a given amount of equity into the project allows the wealthy borrower to increase the size of the project by the same amount, without affecting the slope of the best-response function.

CHECKPOINT 12.9 Wealth matters Explain why the borrower with k_1 will not get a loan, and the one with k_3 will pay a lower interest rate (in competitive equilibrium) than the one with $k_2 < k_3$.

12.10 COMPARISON OF COMPLETE AND INCOMPLETE CONTRACTS

In the credit market, just as in the labor market and the market for goods of variable quality (Chapters 10 and 11), the incompleteness of the contract between the principal and the agent leads to results quite different from what is the case in exchanges covered by complete contracts. We summarize the differences between the complete and incomplete contracting models in Table 12.2 and illustrate the two cases in Figure 12.13.

The borrower receives a rent

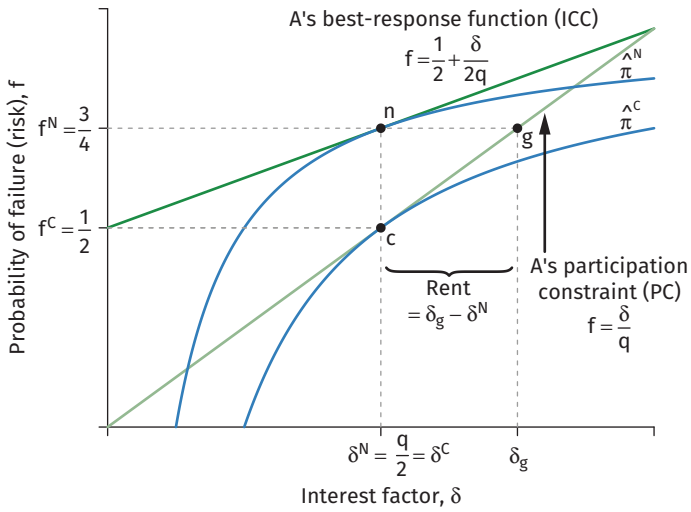
To see that the borrower receives a rent we consider his interaction with a single principal using Figure 12.13. You will observe that the Nash equilibrium (point **n**) is to the left of point **g** on his participation constraint. We know that expected income is higher to the left of the participation constraint (because the interest factor is lower) so he must expect an income greater than zero, his fallback option. So he receives a rent.

To determine how large his rent is we proceed as we did in Chapter 11, by asking: Hypothetically holding constant the level of risk he is taking in the Nash equilibrium (f^N) how much higher could the interest factor be without violating his participation constraint? (It does not matter that he would not hold constant his risk level if she raised the interest factor: this is another thought experiment.) The answer is given by the horizontal distance between points **g** and **n**, that is $\delta_g - \delta^N$.

The principal's profits are lower in the incomplete contracting case. From the figure you can see that $\delta(1-f)$ is $\delta^C(1-f^C) = \frac{q}{4}$ in the complete contracting case and only $\frac{q}{8}$ under the incomplete contract.

The total rent to be divided between lender and borrower is also smaller in the incomplete contract case than it is in the complete contracting case. The reason is that the borrower runs the machine at a speed faster than that which would maximize the expected total income from the project. But, even though the total rent is lower with incomplete contracting, unlike

Figure 12.13 A comparison of the complete and incomplete credit market contract outcomes. Point **n** corresponds to the incomplete credit contract Nash equilibrium with risk $f^N = \frac{3}{4}$ and interest factor δ^N. Point **c** corresponds to the complete contract outcome with risk $f^C = \frac{1}{2}$ and interest factor δ^C.



the complete contracting case, the borrower *shares* some of the rent when the contract is incomplete: the rent is not entirely taken by the lender as it was when contracts were complete.

The Nash equilibrium is Pareto inefficient

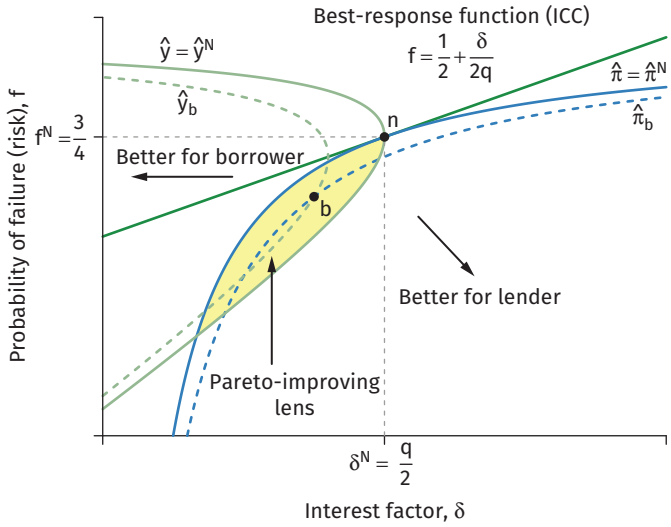
In the complete contracting case as we have already seen, the result must be Pareto efficient because the constraint on the lender's profit maximization is the borrower's participation constraint.

When the contract is incomplete, the borrower's best-response function not his participation constraint limits the lender's options. To see why the resulting Nash equilibrium (point **n** in Figure 12.14) must be Pareto inefficient notice that the lender is better off at points below and to the right of her iso-profit curve $\hat{\pi}^N$. The borrower is better off at points to the left of his iso-expected-income curve $\hat{y}^N$.

There is a set of points, such as point **b**, that are Pareto improvements over point **n**. At point **b** in yellow-shaded Pareto improving lens, the lender would earn higher profits ($\hat{\pi}_b > \hat{\pi}^N$) and the borrower would get higher expected income on $\hat{y}_b > \hat{y}^N$ showing that point **b** is Pareto superior to the Nash equilibrium at point **n**.

! REMINDER In M-Note 11.6 in Chapter 11 we used the first-order conditions for the principal's and agent's constrained optimization problems to demonstrate why the Nash equilibrium in a similar case (employers and workers) is not Pareto efficient.

Figure 12.14 Credit equilibrium with non-contractual risk level and Pareto-improving lens. At the Nash equilibrium, the borrower's iso-expected-income curve is $\hat{y}^N$ and the lender's iso-expected profit is $\hat{\pi}^N$. The yellow-shaded area is the Pareto-improving lens. At point **b** in the lens, the lender would earn higher profits on $\hat{\pi}_b$ and the borrower would get higher expected income on $\hat{y}_b$ showing that point **b** is Pareto-superior to the Nash equilibrium at **n**.



The Pareto-improving lens must exist for the same reason that contractual incompleteness led to inefficiency in the equilibrium of the employee-employer interaction in Chapter 11 or the interaction of buyer and seller of the variable quality good in Chapter 10. At the Nash equilibrium, point **n** in Figure 12.14 the lender's iso-expected-profit curve is tangent to the borrower's upward-sloping best-response function. This means that it cannot also be tangent to the borrower's iso-expected-income curve–as a $mrs^P = mrs^A$ condition for a Pareto-efficient outcome would require–because at point **n** the borrower's iso-expected-income curve is vertical (as we showed in M-Note 12.2 and in Figure 12.7).

CHECKPOINT 12.10 Pareto efficiency and Pareto improvements

a. Make sure that you can explain why it is the case that one person maximizing his profits or utility subject to the participation constraint of another must result in a Pareto-efficient outcome.

b. Can you explain why the set of Pareto-improving points in the Pareto-improving lens must exist when the constraint is the best-response function? *Hint*: think about the construction of the best-response function and what you know about the slope of the borrower's iso-expected-income curve at the Nash equilibrium, point **n**.

c. Explain each of the five differences between the complete and incomplete contracting case in Table 12.2 and why the difference in the contract produces the difference in outcomes.

Table 12.2 Complete and incomplete contracts: a summary of differences in the Nash equilibrium of the credit market. The ICC and the PC are the incentive compatibility constraint and the participation constraint, respectively. The case of the complete contract is shown in Figure 12.5; that of the incomplete contract is shown in Figure 12.8 and the figures in the remainder of the chapter. In Table 12.1 we show that the total rents are maximized in the complete contracting case but not in under incomplete contracts. If markets do not clear, then borrowers will fail to secure loans to finance projects that are identical to the projects of wealthier borrowers who have secured a loan. The two cases are contrasted in Figure 12.13.

Contract over interest and risk	Incomplete	Complete
Constraint for principal (lender)	ICC	PC
Rent for the agent (borrower)?	yes	no
Pareto-efficient Nash equilibrium?	no	yes
Total rents, maximized?	no	yes
Market clearing?	no	yes

12.11 WHY REDISTRIBUTING WEALTH CAN INCREASE THE SUM OF ECONOMIC RENTS

You have already seen (in Tables 12.1 and 12.2) that the Nash equilibrium in the credit market is inefficient in a second way, one that goes beyond the Pareto-efficiency criterion: the sum of the rents gained by the borrower and the lender is smaller than it could be. This is the case with just a single borrower and lender, but it becomes much more important when we consider an entire economy in which there are wealth differences among prospective borrowers and some borrowers lacking wealth are excluded entirely from the market.

Inferior projects of the wealthy are funded

These differences among potential borrowers cannot be efficient, as there will be some poor borrowers with good projects that will not be carried out because they are excluded from the credit market, while some wealthier borrowers will obtain financing to carry out inferior projects.

To see the consequences for the efficient use of resources, suppose that some given total amount of finance is available, normalized to one, to be divided among projects operated by a wealthy and a not wealthy borrower (these could represent many borrowers of each type). For simplicity we assume that the poor borrower has no wealth, and that the wealthy borrower will invest equity equal to one-half the size of the project in any project for which she is granted a loan. We order the projects of each from the best (highest q) to the worst (lowest q).

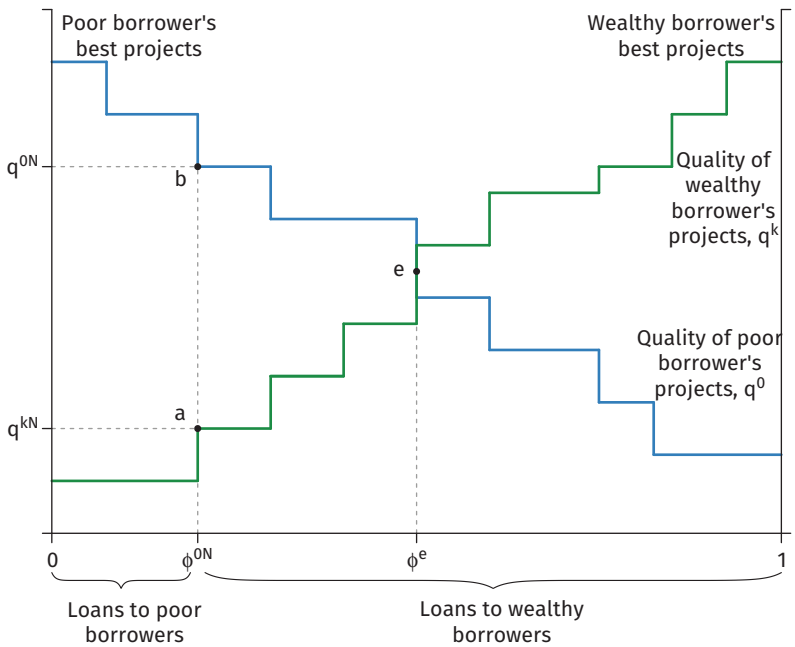
Figure 12.15 shows the array of projects of the two, the poor borrower's best project on the left, with projects of lesser quality arrayed in the step function descending to the right. The wealthy borrower's best project on the right, and his projects of successively lower quality are shown by the step function descending to the left.

The height of each step function is the quality of the particular project in question. The horizontal width of the step is the size of the project for the poor borrower and half the size of the project for the rich borrower (because he borrows only half of the project size, providing the rest with his own investment of equity).

In the figure ϕ^{0N} is the fraction of total loans received by the poor borrower (with $(1-\phi^{0N})$ the fraction going to the wealthy borrower). We use the N superscript because we know from the previous demonstration that, in the Nash equilibrium, the wealthy borrower will succeed in obtaining financing (a loan) for projects that are of lower quality than the best excluded project of the poor borrower.

To see why this reduces the total rents, think about the consequences of hypothetically shifting some of the loan funds from the rich to the poor, so that the rich person would not be able carry out his worst included project, and the poor person would be able to carry out his best excluded project. That would replace an inferior rich person's project with a superior poor

Figure 12.15 Efficiency losses due to borrower wealth differences. A total amount of credit (the length of the horizontal axis) is allocated between a borrower with no wealth (superscript 0) and a borrower with substantial wealth (superscript k) invested as equity in various projects wealth. In the Nash equilibrium a fraction of the total funds loaned ϕ^{0N} is allocated to finance the projects of the poor. The remainder $1-\phi^{0N}$ finances the projects of the wealthy borrower. The vertical axis measures the quality of the projects. The two step functions show that other than the size of the projects, both rich and poor have a similar distribution of quality of projects. This need not be the case, and it is unimportant in what follows. The average quality of the projects funded will be maximized if the best excluded project of the poor borrower is worse than the worst included project of the wealthy borrower, and the best excluded project of the wealthy borrower is worse than the worst included project of the poor borrower. This is not true at the distribution given by ϕ^{0N} because the quality of the worst included project of the wealthy borrower q^{kN} is less than the quality of the best excluded project of the poor borrower q^{0N}. The only distribution of funds for which the average quality of project maximizing condition is true is ϕ^{e}.



person's project, as you can see from the figure. This would increase the average quality of the projects funded because $q^{0N} > q^{kN}$.

You can also see that the same would be true of further redistribution of finance toward the poor person's project, until the poor person's projects received a fraction ϕ^{e} of the total.

Redistributing wealth to increase total rents

We called the redistribution of loans hypothetical, but it could be accomplished by a redistribution of the wealth of the two borrowers, so that they had equal assets. Their projects would then be treated equally in the credit market and the poor borrower would receive ϕ^e of the funds. The average quality of the projects funded would increase.

But there is a second reason why a redistribution of wealth could enhance economic rents, even in the absence of any change in the quality of the projects. If a poor former borrower were to become wealthy enough to finance the project herself, then she would "own" all of the consequences of her own risk-taking. As a result she would implement the expected income-maximizing risk level.

To see this, think of a particular borrower whose project has a quality q seeking a loan from a wealthy individual to finance the project. Suppose $q = 8(1+\rho)$, then in a credit contract like the ones we have studied, the two would transact as is shown in the top "before" line of Table 12.3. We assume that the lender's expected profit is just equal to the risk-free rate of return ρ and so, given the value of q we have assumed is equal to $((q/8)-1)$, while the borrower's expected income is $\frac{q}{16} = \frac{1+\rho}{2}$.

Now imagine that instead of the poorer person borrowing from the lender, the government confiscates the assets from its wealthy lender and gives this amount ($1) to the poor former borrower, who then operates the project as *residual claimant* at the owner-operator level of risk, that is, $f = \frac{1}{2}$. As a result the former borrower now has an income equal to $\frac{q}{4}$ or $2(1+\rho)$.

But at the same time, the government imposes a tax obligation on the beneficiary of this redistribution (the previously poor borrower), requiring him to pay $1+\rho$ at the end of the period. The tax must be paid irrespective of whether the machine fails. (The borrower might have to sell his car or mortgage his home in order to do this.)

! REMINDER The residual claimant is the person who owns what is left over after all of the contractually obligated costs are paid. In a firm, the residual claimants are the owners who own the revenues of the firm minus the costs of labor, materials, management and other inputs, as well as taxes have been paid. If a farmer pays a fixed rent and she farms and owns the crop that she harvests, then she is the *residual claimant* on the value of the harvest minus the rent and other costs incurred. She also owns the loss, if the harvest is worth less than the costs.

Table 12.3 Comparison of levels of income before and after redistribution for an owner of a machine and an operator of a machine. The "Before" line is the case illustrated in Table 12.1 for an incomplete contract and a borrower without wealth. The owner's income here is her accounting profit (we do not subtract the opportunity cost of capital, as we did in Table 12.1). The second line ("After redistribution, before tax") refers to the situation after the asset has been transferred from the initial owner to the operator but before he has paid the tax. The "After" row shows the result of redistributing wealth equal $1 to the erstwhile borrower, taxing the borrower an amount $1+\rho$ and transferring these tax revenues to the former owner of the asset, as compensation for the confiscation if her asset. We illustrate this case assuming that $q = 8(1+\rho)$.

	Total income	Initial owner's income	Operator's income
Before	$\frac{3q}{16} = \frac{3(1+\rho)}{2}$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$	$\frac{q}{16} = \frac{(1+\rho)}{2}$
After redistribution/before tax	$\frac{q}{4} = \frac{4q}{16} = 2(1+\rho)$	0	$2(1+\rho)$
After redistribution and tax	$\frac{q}{4} = \frac{4q}{16} = 2(1+\rho)$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$

The government then transfers these tax revenues to compensate the former owner for her loss, so she now has income $1+\rho$, the same as before. The beneficiary of the redistribution after paying the tax has an expected amount of $1+\rho$ for himself, and is therefore better off. (Recall he made only half this amount as a borrower, *without* the government intervention.) So the former lender is as well-off as before, and the former borrower is better off. The redistribution of wealth implemented a Pareto improvement.

There is nothing special about the numbers: all that is required to make a Pareto improvement possible is that total rents are larger when the former borrower becomes rich enough to run the project as the owner-operator. By extracting from the beneficiary the tax sufficient to pay compensation to the former owner irrespective of the fate of the project, the government was able to offer the equivalent of an enforceable loan contract to the beneficiary at the risk-free interest rate. What the asset transfer plus the tax accomplishes is to make the owner-operator of the project the residual claimant on all of the risk entailed by his choices (rather than being shielded from risk by the unenforceability of the promise to repay the loan).

The key to the success of the redistribution is that private transactions are governed by limited liability and bankruptcy laws that protect the borrower from risk by placing his other assets (car or home in the above example) beyond the reach of the lender seeking to enforce repayment. The obligation to pay the government's tax is not limited by these provisions. This accounts for the superiority of the owner-operator case, and allows for the Pareto-improving redistribution.

What this policy application shows is that redistributing wealth from richer to poorer people may allow a Pareto improvement, so that the poor benefit and the rich do not lose. But so far we have left out an important reason why redistribution of this type sometimes fail: the poor are probably more *risk averse* than the rich. They place a higher value than the rich on reducing the risk to which they are exposed. This being the case they may prefer lower expected returns on less risky projects. We return to the question of risk aversion and wealth or income redistribution to the less well-off in Chapters 13 and 15.

CHECKPOINT 12.11 Pareto-improving redistribution Imagine explaining to a friend studying economics who has not yet read this chapter how a redistribution of wealth can increase total rents and even be a Pareto improvement, in particular:

a. If there were an alternative outcome that would have benefited both the owner of the machine and the operator, why did they not simply make a deal to implement the mutually preferred outcome?

b. Why was the alternative outcome accomplished by a government when it did not happen by means of private exchange?

12.12 COMPETITION, BARRIERS TO ENTRY, AND THE DISTRIBUTION OF RENTS

We turn now from the effect of differing levels of wealth among borrowers and the effects of redistribution of wealth to consider another aspect of the credit market that is a subject of public policy. This is the degree of competition. We begin with a borrower without wealth, like the one depicted in Table 12.1, now shown in Table 12.4, contrasting the case of the one-on-one interaction between Parama and Antonio with a situation in which borrowers and lenders interact in a credit market with no barriers to entry or what we have called unlimited competition.

The difference between the final two rows of Table 12.4 underscores the effect that limits to competition has on the distribution of income. Looking just at the relationship between the lender and a single borrower, two differences are noticeable:

- *The total rents available to the borrower and lender*: Competition among lenders (the bottom row) results in a substantial increase in the size of the total "pie"–from 0.825 to 1.225. The reason is that competition forces lenders to charge lower interest factors–δ drops from 5 to 3–and in response borrowers adopt lower (more nearly total-rent-maximizing) levels of risk: f falls from 0.75 to 0.65.
- *The distribution of rents between borrower and lender*: Unlimited competition eliminates the rents of the lenders, they receive expected accounting profits equal to the opportunity cost of capital, all of the rents go to the borrower. As a result, their accounting profits, that is, their income derived from their loan of \$1 is $\hat{\pi}^A = \rho = 0.05$.

M-CHECK The second row in Table 12.4 is calculated by inserting the parameter values given in the table caption into the analytical expressions in the first row. The entries in the bottom row are calculated by inserting the expression for f in the top row into the zero-profit condition, that is, Equation 12.36 with barriers to entry, $b = 0$, and then solving for δ. We then used that value of δ in the borrower's best-response function to determine f.

Table 12.4 The difference that credit market competition makes: a numerical example of a lender and a single borrower This example is based on Table 12.1. The incentive compatibility constraint (ICC) in all cases is $f = 1/2 + \delta/2\mu$. For the numerical examples we used $q = 10$, $\rho = 0.05$, and for the last line barriers to competition, $b = 0$.

Case: Differing rules of the game	Interest factor, δ	Risk, f	Expected rents per period: Borrower's expected income, $\hat{y}$	Lender's expected profit, $\hat{\pi}$	Total
No competition (analytical expressions)	$\frac{q}{2}$	$f = \frac{1}{2} + \frac{\delta}{2q}$	$\frac{q}{16}$	$\frac{2q}{16} - (1+\rho)$	$\frac{3q}{16} - (1+\rho)$
No competition (numerical example)	5	0.75	0.625	0.2	0.825
Unlimited competition (numerical example)	3	0.65	1.225	0	1.225

The entries in Table 12.4 give the results for the lender and just one of her borrowers. To see how this affects the distribution of income in the economy as a whole, we need to take account of the fact that banks and other lending institutions interact with very large numbers of borrowers. So think about the lender, Parama, and the, let's say, 99 *other* borrowers to whom Parama has extended loans, like she did to Antonio.

The results in Figure 12.16 show that the elimination of competition increases the income of the lender by fivefold while cutting borrowers' incomes to approximately half of their level under competition. Total income is reduced by one-third.

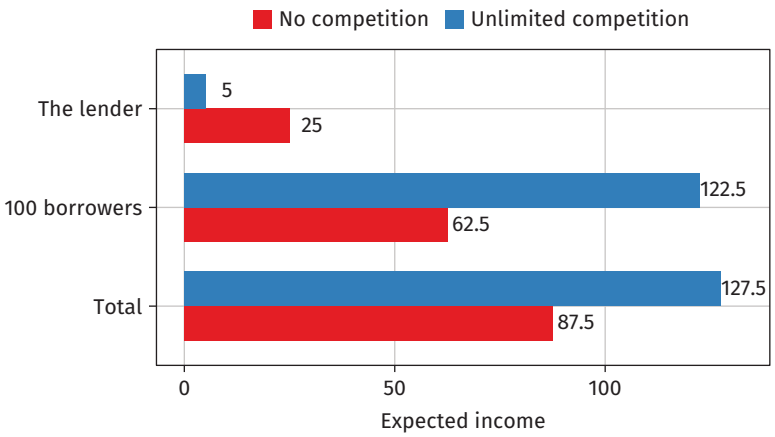
The contrast between these extreme cases–no competition versus unlimited competition–provide another illustration of the effect of limited competition: in the markets for goods and services

- increasing economic profits and reducing consumer surplus in Chapter 9 and
- reducing the real wage in the model of the whole economy in Chapter 11.

As well as:

- the effect of limited competition in the labor market–monopsony–in reducing employment and lowering wages in Chapter 11.

Figure 12.16 Effects of competition on total income of a single lender and her set of borrowers, and its distribution. The data for the figure come from the numerical examples in the last two rows of Table 12.4 with two additional pieces of information. We show the lender's income based on transactions with 100 borrowers identical to the single borrower shown in the table. And the lender's income is correctly measured by his accounting profits, namely $\pi^A = \delta(1-f)-1$, not by his economic profits ($\pi^E = \delta(1-f)-(1+\rho)$). So, for example, the top blue bar, the lender's income of 5 under unlimited competition (the case in which economic profits are zero) is 100 transactions times the size of the loan (1) times $\rho = 0.05$.



The comparison of the extreme cases–unlimited or no competition–suggests that public policies to reduce barriers to entry in credit markets would raise total income and reduce the inequality between lenders and borrowers. To see how this would work, in Figure 12.17 we study the credit market under two levels of competition:

- *barriers to entry*: the status quo with substantial barriers to entry ($b > 0$); and
- *unlimited competition*: a possible result of a government's competition policy in which barriers to entry are eliminated.

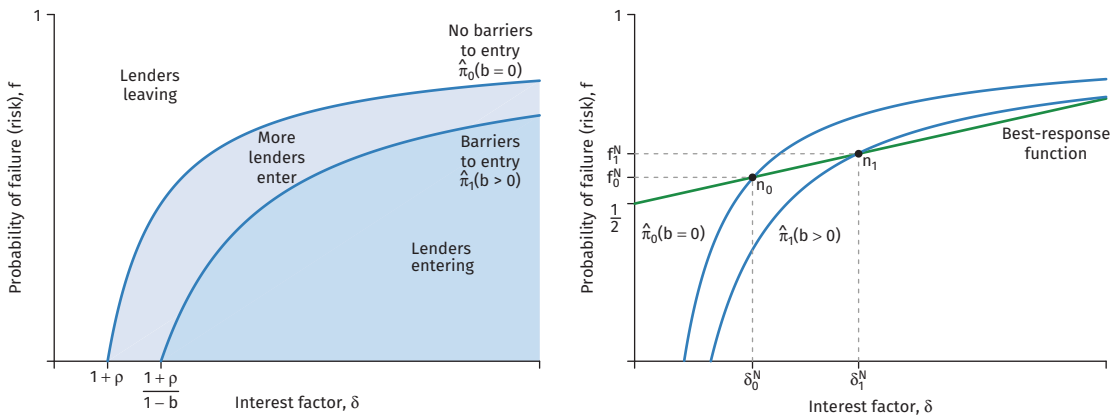
Panel (a) shows the competition conditions under these two assumptions. In panel (b), point n_1 is the status quo Nash equilibrium while n_0 is the outcome under unlimited competition, that is $b = 0$. If policies could be implemented to shift the competition condition to the left (as shown in panel (a)) the outcome would be:

- a reduction in the interest factor charged by the lender, from δ_1^N to δ_0^N,
- a reduction in the risk taken by the borrower, from f_1^N to f_0^N

resulting in

- an increase in total income (you know the total income is maximized at $f = 1/2$)

Figure 12.17 Nash equilibria with unlimited competition and barriers to entry. The green line is a borrower's best-response function. Two alternative competition conditions are shown, one for the case of unlimited competition ($b = 0$) and the other for positive barriers to entry. You can see that with greater barriers to entry the Nash equilibrium level of the interest factor and the level of risk taken are both higher.



(a) A decrease in barriers to entry shifts the competition condition

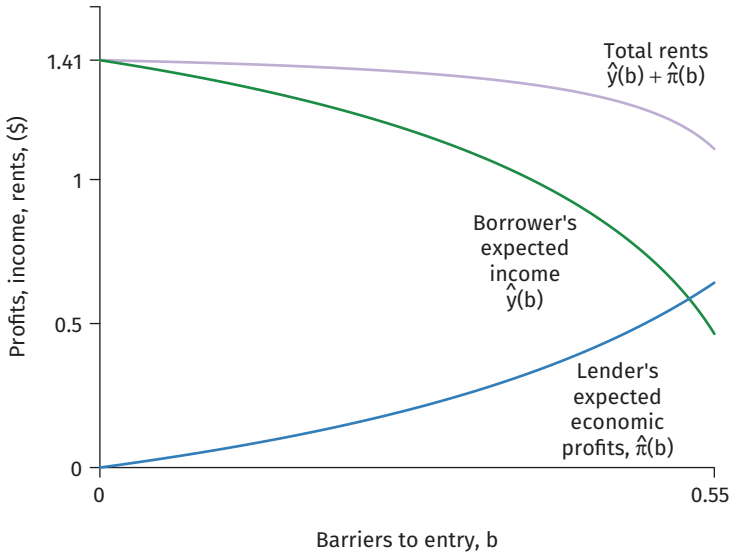
(b) Changes in barriers to entry change the Nash equilibrium

- an increase in the expected income of the borrower (you know from Figure 12.7 that expected income is higher at points to the left on the borrower's best response function), and
- a decrease in the expected income of the lender (because the competition condition has shifted to the left, n_0 is on a lower iso-expected-profits curve than is n_1).

In Figure 12.18 we generalize the insights of Figure 12.16 to show the effect of a range of values for the barriers to competition. We introduce a borrower's equity at the level of $k = 0.5$ and we show how variation in the level of barriers to entry affect the total rents and the expected income and profits of the borrower and lender respectively. As expected from the previous analysis, the greater are the barriers to entry (the less competition in the credit market):

- the less are the total rents (the income of the borrower and the economic profit of the lender);

Figure 12.18 Barriers to entry in the credit market and the distribution of rents between a lender and a single borrower. To show the effect of barriers to entry, we set specific values for the rest of the model and vary barriers to entry (b). In the figure the quality of the machine, $q = 10$, the level of equity is $k = 0.5$, and the opportunity cost of capital is $\rho = 0.05$. The total rents in the absence of barriers to entry shown in the figure (1.41) exceed the total rents in Table 12.4 for the case of unlimited competition because here the borrower has invested equity in the project and so he selects a lower level of risk, closer to the value ($f = \frac{1}{2}$) that would maximize the sum of rents.



M-CHECK In Figure 12.18, for $b = 0.58$, the profit rate of the incumbent firms would have to be so high that the borrower with $k = 0.5$ (the equity level we have assumed) would be the marginal borrower. For $b > 0.58$ there is no profit rate of incumbent firms that would be both sufficient to attract entering firms and consistent with incumbent firms transacting with borrowers.

- the greater is the economic profit of the lender; and
- the lesser is the income of the borrower.

CHECKPOINT 12.12 **Competition matters** Using Figure 12.17 explain how increased competition reduces the equilibrium interest factor.

12.13 APPLICATION: FROM MICRO TO MACRO—THE MULTIPLIER AND MONETARY POLICY

We saw in Chapter 10 that when account is taken of the incomplete nature of contracts, competitive markets need not clear in equilibrium; and in Chapter 11 this is true specifically of the labor market, which in long-term equilibrium is characterized by structural unemployment. Here we show how principal–agent models of the credit market provide additional foundations for a unification of microeconomics and macroeconomics.

The microeconomic foundations of the Keynesian multiplier

An essential concept for macroeconomic policy is the Keynesian multiplier. The Keynesian multiplier indicates the total effect on **aggregate demand** generated by a single unit of exogenous change in expenditure (for example in the form of a government transfer, investment, or net export demand).

Think about a fall in export demand, which results in fewer of the exported goods being produced and a reduction in employment, meaning some workers lose their jobs. The Keynesian model shows how this income shock to the employee then sets off ripple effects throughout the economy amplifying the initial effect.

This occurs because with reduced income, there is a second-round effect: the unemployed worker spends less money on goods and services, so the income of the local grocer now also falls, along with the incomes of other people from which she would have purchased goods and services had she not lost her job. A third round follows: the grocer purchases less from his suppliers, and so on. The multiplier is a measure of the extent to which an initial shock is amplified by successive rounds of reduced expenditure.

But if the unemployed worker could have readily borrowed sufficient funds to sustain her previous level of consumption until she found another job, the lost income of the unemployed worker would be where the process ends. There would have been no amplification of the shock. National income would fall by the amount of the reduction in export demand. There would be

AGGREGATE DEMAND The sum of expenditures on goods and services produced in a country, including demand from the rest of the world.

! REMINDER In Figure 12.2 you saw that about 30 percent of US households reported in 2019 that they could not cover three months' expenses by any means (including borrowing) if their primary income was lost.

no second- and third-round ripple effects. To explain the multiple rounds of additional expenditure that is essential to the Keynesian multiplier, it is common to assume that there are many "hand-to-mouth" households, whose consumption expenditure rises and falls with its income, which seems to be the case. By one estimate: "half of households follow the 'rule-of-thumb' of consuming their current income."[15]

The model of the credit market with incomplete contracts that you have learned provides one explanation of where this rule of thumb might come from and why these hand-to-mouth households are so common. A great many families are either excluded from the credit market entirely, or are unable to borrow except in small amounts or at prohibitively high rates of interest.

The microeconomics of monetary policy

The credit market model explains how **monetary policy** moderates the business cycle, for example by expanding aggregate demand during a recession. It does this by lowering opportunity cost of capital and thereby expanding lending to businesses and individuals.

In the US the Federal Reserve System sets a target federal funds rate of interest which effectively determines the interest rates charged by commercial banks throughout the economy. The same is done in other countries by the central bank, for example, in Germany the Deutsche Bundesbank, the Bank of England, or the Reserve Bank of India. The rate of interest at which banks lend is one of the main determinants of the *opportunity cost of capital*.

The amount of lending in the credit market is determined by the credit market competition condition (Equation 12.35).

$$\text{Credit market competition condition} \quad \delta(1-f) = \frac{1+\rho}{1-b}$$

which, recall, depends not only on the extent of barriers to entry, b, but also on the opportunity cost of capital ρ.

To see how this affects the macroeconomy, we use Figure 12.19 where the initial competition condition is $\hat{\pi}_{\rho_2}$ which corresponds to the initial interest rate, ρ_2. Now suppose the Federal Reserve System decides to lower the federal funds rate from ρ_2 to ρ_1 ($\rho_2 > \rho_1$). The decrease in the federal funds rate will shift the competition condition up and to the left as is shown by the new competition condition, $\hat{\pi}_{\rho_1}$.

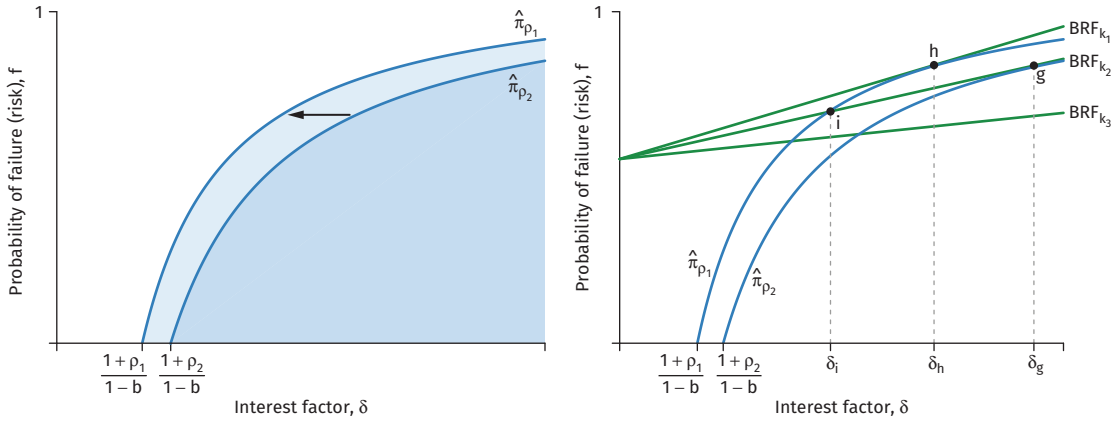
MONETARY POLICY Policies implemented by a central bank affecting the rate of interest at which businesses and others can borrow and the amount of borrowing, thereby regulating aggregate demand to moderate the business cycle and regulate inflation, are termed monetary policy.

There are two effects of the change in policy:

- *include previously excluded borrowers*: Some of the prospective borrowers with less wealth, k_1, who were previously excluded from the credit market, can now borrow. In Figure 12.19 (b) the reduction in the opportunity cost of capital was just sufficient so that her best-response function BRF_{k_1} makes her a new marginal borrower, at point **h** (paying an interest factor δ_h); and
- *lower interest for previously included borrowers*: The two borrowers shown who were able to borrow previously are now able to borrow at lower interest rates. The borrower with wealth k_2 for example was previously borrowing at the interest factor of δ_g and after the reduction in the opportunity cost of capital can now borrow at the lower rate of δ_i (comparing points **g** and **i**).

The effect of these two changes in the credit market will be to increase spending: the previously excluded borrower will now spend the funds borrowed, and the other two borrowers will now enjoy higher profits on the projects that they previously financed with their loans. They can also finance larger projects or projects of lesser quality.

Figure 12.19 The competition condition and the central bank interest rate. When the central bank's monetary policy decreases the opportunity cost of capital from ρ_2 to ρ_1, the competition condition shifts from $\hat{\pi}_{\rho_2}$ to $\hat{\pi}_{\rho_1}$ as shown in panel (a). In panel (b), we show the best-response functions of three borrowers with different levels of equity, $k_3 > k_2 > k_1$. At the initial opportunity cost of capital (ρ_2) the borrower with equity of k_2 is the marginal borrower and pays an interest factor δ_g corresponding to point **g**. With the opportunity cost of capital lowered to ρ_1, a borrower who was previously excluded prior to the policy change is now the marginal borrower in the credit market (at the tangency of $\hat{\pi}_{\rho_1}$ to the new marginal borrower's best-response function, BRF_{k_1} at point **h**). The interest factor that a borrower with wealth k_2 pays decreases from δ_g to δ_i which can be seen by comparing points **g** and **i**.



(a) Shift in competition condition when ρ decreases

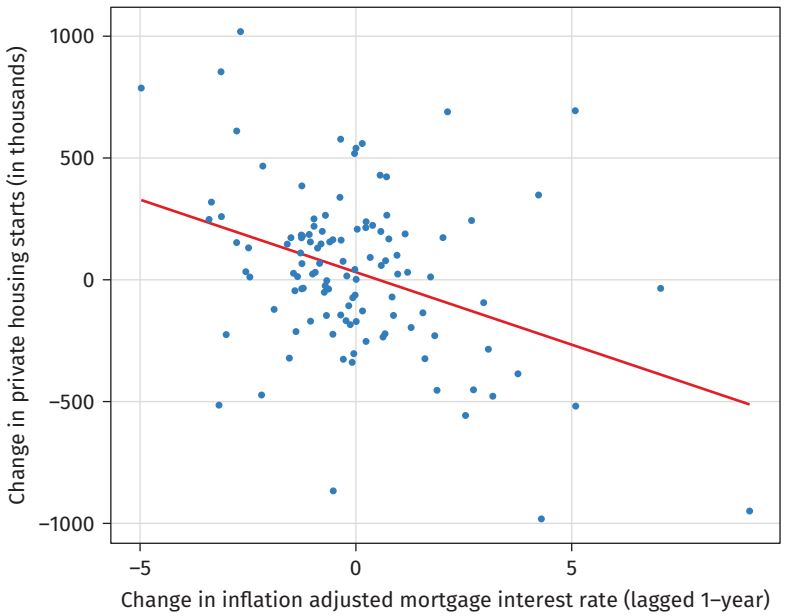
(b) Including previously excluded borrowers and reducing interest rates

We have modeled the effect of monetary policy on investment in projects similar to our "machine." The primary effect of lower interest rates on spending, however, is not to expand investments such as building new plants, office buildings, and equipment but instead on the construction of new homes and the purchase of cars and other consumer durables (typically bought on credit).

A slight modification of the credit market model is required to study housing and consumer durable credit, to take account of the fact that lenders typically require collateral rather than equity. (We do this in the next section.) But the underlying mechanism is the same: a reduction in the opportunity cost of capital (when the central bank lowers its lending rate) lowers the Nash equilibrium interest rate on mortgages and car loans, and allows some previously excluded borrowers to secure a loan.

Figure 12.20 illustrates this channel for monetary policy affecting aggregate demand. It shows that a decline in mortgage interest rates is associated with an increase in new home construction. The dots in the upper left, for

Figure 12.20 Mortgage interest rates and new private housing construction, US 1977-2006. The vertical axis measures the change in the number of new home units on which construction began between one quarter (e.g. Q1, meaning January to March) of the year and the same quarter a year earlier. The horizontal axis measures the change in the real mortgage interest rate over the previous year. The real mortgage interest rate takes account of the effect of inflation on the cost of repaying the loan. We also show best fit estimated (red) line based on these data. The slope of the red line (−59.53) means that a one percentage point increase in the mortgage interest rate is associated with a drop of close to 60,000 new housing starts for the following year.



example, show that in quarters that saw a big drop in the real mortgage interest rate over the previous year there was big increase in housing starts. As we shall see at the start of Chapter 13, the global financial crisis resulted in many people defaulting on home loans and going bankrupt, the central banks of the world stepped in by engaging in monetary policy to stimulate aggregate demand by lowering the central bank lending rate.

CHECKPOINT 12.13 The microeconomics of macroeconomic policy In Figure 12.19 show the effect of the reduction of the opportunity cost of capital on the interest rate paid by the borrower with wealth k_3.

12.14 APPLICATION: THE CASE OF COLLATERAL RATHER THAN EQUITY

We began with an example of a loan secured by collateral: Mary Bolender's car that was disabled by the lender. Because of the importance of collateral in supporting lending for car and home purchases (and therefore its key role in the effectiveness of monetary policy), we now return to this case, contrasting collateral with equity, the case in which the borrower has invested some amount k in the project.

The borrower has collateral of value k, and the structure of the game between him and the lender otherwise is unchanged. As was the case of the borrower without wealth,the borrower will borrow \$1 for the project, but now additionally will post k as collateral. The collateral may be the borrower's home if the loan is a mortgage for the purchase of the home. It could also be a car being purchased on credit. Therefore the borrower can still use the collateral until it is claimed, so there is no opportunity cost of devoting k to collateral unless the project fails. Thus, his expected income will consist of the following:

- As before, in the case the project succeeds (with probability $(1-f)$), the borrower receives expected returns minus expected repayment of the loan.
- Now, additionally, if the project fails (with probability f), the borrower will have to give the lender the collateral, k.

Putting this together, we have:

$$\hat{y}(f,\delta,k) = qf(1-f) - \delta(1-f) - fk \tag{12.41}$$

In M-Note 12.10 we show that the borrower's best-response function, that is derived by varying f to maximize the borrower's expected income, is

$$f = \frac{1}{2}\left(1-\frac{k}{q}\right) + \frac{\delta}{2q} \tag{12.42}$$

This differs from the best response function of the zero-wealth borrower in the absence of collateral: the entire function is shifted downward (the

constant now is less than $\frac{1}{2}$). The reason is that the marginal cost of running the machine faster now includes not only the lost revenue of the machine if it fails, but also forfeiting the collateral.

As a result, for every interest factor (δ) the risk taken by the borrower (f) is less, making it more profitable for a lender to extend a loan. This means that by posting collateral a borrower may avoid being excluded from the credit market, as the lender who disabled Bolender's car said.

The lender's expected profits are also different from the case of equity (Equation 12.6) or a loan that is not secured by collateral.

- the amount of the loan is now 1 (rather than $(1-k)$); and
- instead of getting nothing in the case that the project fails, the lender will receive the value of the collateral, k.

So expected profits are now

$$\hat{\pi}(\delta, f, k) = \delta(1-f) + fk - (1+\rho) \tag{12.43}$$

In M-Note 12.10 we show that the lender will maximize expected profits by choosing the interest factor below (where $f_\delta \equiv \frac{\partial f}{\partial \delta}$).

$$\delta = k + \frac{1-f(\delta)}{f_\delta} \tag{12.44}$$

The interest factor charged by the lender will now be higher than in the case without collateral (Equation 12.33). The reason is that the marginal cost of raising the interest factor is now less: while a higher interest factor will induce the borrower to run the machine faster as before (Equation 12.42), now in the case that the project fails the lender receives the collateral k instead of zero, as was the case in a loan without collateral.

Where the lender can require that the borrower post collateral, then, there are two offsetting effects:

- the risk taken by the borrower will be less (for any given interest factor charged by the lender); and
- the interest factor charged by the lender will be higher.

The resulting level of risk taken in the Nash equilibrium could be either higher or lower than for the case without collateral.

There are also offsetting effects of collateral on the distribution of rents between borrowers and lenders.

- by shifting downward the borrower's best-response function, posting collateral will allow some previously excluded borrowers to secure a loan, and to receive some positive level of expected income (that is, a rent); but
- for those able to secure a loan without collateral, the requirement to post collateral will increase the rents obtained by the lender and reduce the rents obtained by the borrower.

M-NOTE 12.10 The case of collateral: the borrower's risk decision

Here we derive the best-response function of the borrower who posts collateral rather than investing equity in the project.

The expected income of the borrower is:

$$\hat{\pi}(\delta, f, k) = \delta(1-f) + fk - (1+\rho) \tag{12.45}$$

Following the procedure in M-Note 12.4 we differentiate the borrower's expected income, Equation 12.41, with respect to f and set it equal to zero to maximize expected income, which gives us his first-order condition (FOC) and best-response function:

$$\frac{\partial \hat{y}}{\partial f} = -qf + q(1-f) + \delta - k = 0 \tag{12.46}$$

$$\text{First-order condition} \quad \underbrace{q(1-f)+\delta}_{\text{MB}} = \underbrace{qf+k}_{\text{MC}}$$

The collateral requirement increases the marginal cost of taking greater risk. Rearranging the first-order condition, Equation 12.46, to solve for $f(\delta, k)$ we have:

$$\begin{aligned} \text{BRF:} \quad f(\delta, k) &= \frac{1}{2} + \frac{\delta - k}{2q} \\ &= \frac{1}{2} - \frac{k}{2q} + \frac{\delta}{2q} \\ &= \frac{1}{2}\left(1 - \frac{k}{q}\right) + \frac{\delta}{2q} \end{aligned} \tag{12.47}$$

We can see that, instead of flattening the slope of the best-response function of the borrower as occurred in the case of equity, collateral shifts down the best response of the borrower without changing its slope.

M-NOTE 12.11 The case of collateral: the lender's choice of the interest factor

Writing the borrower's best-response function simply as, $f = f(\delta)$, because we are considering variations in k, the expected profits of the lender are now:

$$\begin{aligned} \hat{\pi}(\delta, f(\delta), k) &= \delta(1 - f(\delta)) + f(\delta)k - (1+\rho) \\ &= \delta - \delta f(\delta) + f(\delta)k - (1+\rho) \end{aligned} \tag{12.48}$$

Differentiating this function with respect to δ, setting the result equal to zero, and then rearranging the expression we have:

$$\begin{aligned} \frac{\partial \hat{\pi}}{\partial \delta} &= 1 - \left(f(\delta) + f_\delta \delta\right) + f_\delta k = 0 \\ \Longrightarrow \quad \underbrace{1 - f(\delta)}_{\substack{\text{Marginal}\\ \text{benefit}}} &= \underbrace{f_\delta(\delta - k)}_{\substack{\text{Marginal}\\ \text{cost}}} \end{aligned} \tag{12.49}$$

The left-hand side is the marginal benefit of increasing the interest factor, while the right-hand side is the marginal cost. More collateral reduces the

continued

marginal cost. Equation 12.49 can also be rearranged to show the lender's expected profit maximizing choice:

$$\delta = \frac{q+5k}{8}$$

Equation 12.49 can also be rearranged to require that the slope of the iso-expected-profit curve (the left-hand side) is equal to the slope of the best-response function (the right-hand side):

$$\frac{\partial \hat{\pi}}{\partial \delta} = \hat{\pi}_\delta = 1 - f(\delta) - f_\delta \delta + f_\delta k = 0$$

$$\frac{\partial \hat{\pi}}{\partial f} = \hat{\pi}_f = -\delta + k$$

$$\frac{\hat{\pi}_\delta}{-\hat{\pi}_f} = \frac{1 - f(\delta) - f_\delta \delta + f_\delta k}{\delta - k} = 0$$

$$= \frac{1 - f(\delta)}{\delta - k} - \frac{f_\delta(\delta - k)}{\delta - k}$$

$$-mrs = \frac{\hat{\pi}_\delta}{-\hat{\pi}_f} = \frac{1 - f(\delta)}{\delta - k} = f_\delta = -mrt$$

CHECKPOINT 12.14 Collateral and equity How does the case of lending with collateral differ from lending with equity?

12.15 APPLICATION: COTTON AS COLLATERAL IN THE US FOLLOWING THE END OF SLAVERY

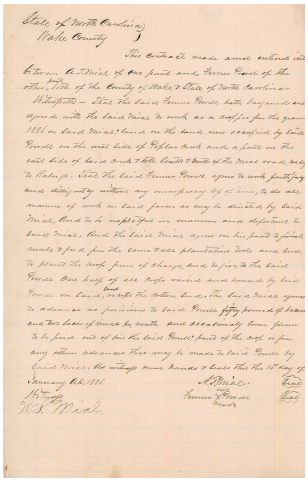


Figure 12.21 A share cropper's contract from North Carolina in 1886. It is signed with an X by the formerly enslaved African-American Fenner Powel (at the bottom, right) in which he agrees to give the landowner M. S. Mial half of the crop he grows and "to be respectful in manner and deportment to said Mial." There is nothing in the contract about Mial being respectful to Powel.

Courtesy of the State Archives of North Carolina.

In the US South, prior to the Thirteenth Amendment to the US Constitution that abolished slavery (1865), it was said that "cotton is king." But it was not until after the end of slavery that cotton truly ascended to the throne among crops. In the quarter of a century *following* the demise of slavery, the production of cotton relative to corn (the main food crop) increased by 50 percent.[16]

The intensification of the cotton monoculture puzzled observers at the time and since, as it coincided with a slight downward trend in the price of cotton relative to corn. Moreover, there were no changes in the technical conditions of production that would have offset the adverse price movement. In fact, the growth of corn yields appears to have outpaced cotton yields during this period.

Nor can the shift from corn to cotton be explained by changes in factor supplies: the cotton-growing regions of the South experienced a serious labor shortage following the war in part because former slaves reduced their hours of work, which should have led some farmers to abandon cotton in favor of corn, as the latter was a much less labor-intensive crop.

What then explains the growing dominance of cotton?

To answer this we need to investigate the structure of local *credit markets*. To finance the crop cycle, most farmers–mostly poor sharecroppers and rental tenants, many of whom were former slaves–purchased food

(including corn) and other necessities on *credit* during the growing season. Because there typically was a single merchant in each area, the prices at which the farmers accumulated their debt were inflated by the *monopoly power* of the merchant-lender.

The loans were repaid when the crop was sold at the end of the season. Most farmers lacked any substantial wealth that they could post as collateral, so the merchant-lenders secured their loans by means of a claim (called a **lien**) on the farmers' future crop in case of default. The crop itself would be the collateral, and like Mary Bolender's car, would be seized by the lender if the borrower was unable to repay the loan.

The system was therefore called a *crop lien system*. According to its most prominent researchers, Roger Ransom and Richard Sutch, the crop lien system favored cotton:

> In the view of the merchant, cotton afforded greater security for such loans than food crops. Cotton was a cash crop that could readily be sold in a well-organized market; it was not perishable; it was easily stored … For these reasons the merchant frequently stipulated that a certain quantity of cotton be planted … It was the universal complaint of the farmers that the rural merchants predicated his willingness to negotiate credit on the condition that sufficient cotton to serve as collateral had been planted.[17]

The crop lien system that came to prominence in the US South after the Civil War was an ingenious solution to the problem of providing credit to asset-poor borrowers. It substituted the farmers' unenforceable promise to repay the loan in the future by an action observable by the lender prior to the granting of credit, namely having already planted cotton on which the merchant had first claim.

Taking account of the relative resource costs and prices of the two crops, Ransom and Sutch estimate that the cotton farmer purchasing corn on credit could have increased his income by 29 percent by shifting resources from cotton to corn. But this was precluded by the fact that because the farmer had little wealth, he needed credit, and for the same reason, credit was conditioned on planting cotton. The result, according to Ransom and Sutch was that:

> The southern tenant was neither owner of his land nor manager of his business … his independent decision making was limited to the mundane and menial aspects of farming. The larger decisions

LIEN A lien is a property right in some good held by a lender to secure the repayment of a debt. Collateral is a form of lien.

> concerning land use, investments in the farm's productivity, the choice of technology, and the scale of production were all made for him.[18]

The story of how cotton became king in the aftermath of the end of slavery could not be told if the contract to repay a loan was complete and enforceable. Cotton was the lender's solution to a problem of an incomplete or unenforceable contract.

CHECKPOINT 12.15 **King cotton** In the US after the Civil War, why did lenders insist that as a condition of getting a loan former slaves were required to plant cotton rather than corn?

12.16 WHY AND HOW WEALTH MATTERS

A widely circulated legend has it that the F. Scott Fitzgerald once said to Ernest Hemingway "The rich are different from you and me." To which Hemingway is said to have replied "They have more money."[19] The first thing that is wrong with this charming conversation between the two great American authors is that it did not happen. But that is just the beginning.

Having greater wealth conveys quantitative advantages—it determines the location of one's budget constraint and gives you a larger feasible set of goods you can buy. But if contracts are complete that's all it does. The rich "have more money" than others.

In an ideal world of complete contracts, all participants in the economy would face the same contractual opportunities (and hence the same prices) irrespective of their holdings. The poor would be constrained to buy less than the rich, but they would transact on the same terms.

By contrast, where contracts are incomplete, wealth confers qualitative advantages including greater personal autonomy and less being subjected to the will of others—being an employer rather than an employee, for example. Substantial wealth also makes it more likely that an individual will be the merchant lender rather than the indebted sharecropper, or one of the Benetton siblings rather than one of their subcontractors. Substantial wealth, in other words, opens up opportunities to be principal rather than an agent.

We have seen that that in the Nash equilibrium of the credit market wealthier people:

- pay lower interest rates;
- can borrow more and so finance larger projects; and
- can finance projects of lesser quality.

People who lack wealth either:

- are not able to engage in contracts or projects that are available to the wealthy; or
- enter contracts on less favorable terms (higher interest factors) than wealthier borrowers.

Stepping back from our simple illustration of the "machine" as a project, the terms on which one can borrow spell the difference between being a principal rather than an agent in other economic interactions. And this often means experiencing one's economic life as one who makes decisions and gives directives as opposed to one who carries them out and who may feel fortunate to have a job at all. Therefore because credit contracts are incomplete, wealth differences have qualitative effects, excluding some and empowering others.

The most obvious reason why wealth influences the contracts one can engage in is that only those with sufficient wealth can undertake projects on their own as owner-operators, as illustrated by the first case we studied. Those with enough wealth to start their own businesses have the rights to their profits (they are residual claimants) and can control what they do. They own the results of their decisions and hence there are no external effects of their actions so their projects yield the maximum possible rents.

A second reason why wealth influences the type of contracts that one is offered is that wealth ownership reduces the conflicts of interest and misaligned incentives arising from contractual incompleteness in principal-agent relationships. Wealthier people have access to superior contracts because their wealth allows contracts that more closely align the objectives of principal and agent. An example is when the borrower has sufficient wealth to provide collateral or put her own equity in a project. The borrower who provides collateral or equity to his project experiences enhanced incentives for the following:

- to supply effort for the project to succeed;
- to adopt risk levels preferred by the principal; and
- to reveal information to the principal and to act in other ways that advance the principal's interests but that cannot be secured in a contract.

People lacking wealth may acquire education and other forms of human capital on less favorable terms than the rich. In the US and many other countries, university students from families with limited wealth accumulate massive debts, for example, the repayment of which often constrains them to major in fields of study with high expected incomes, even if their interests lie elsewhere. We will consider the problem of financing university studies—and the proposed policy of free tuition—in the next chapter.

Similarly, in residential housing markets, those with sufficient wealth are more often owners rather than renters and therefore benefit directly in increased values of their own property from the actions they take to

improve their property and their neighborhood. The asset-poor are more likely to be renters; unable to borrow the funds necessary to buy their homes.

CHECKPOINT 12.16 Why wealth matters Why do those with more wealth have better borrowing opportunities than those with less wealth?

12.17 CONCLUSION

While our model was about an imaginary "machine" that was an income-making opportunity that the borrower could choose to take (or not), the reality that the model was built to illuminate includes less benign circumstances. Examples are desperately poor people paying usurious interest rates on payday loans so as to be able to purchase necessary medicines, or former slaves taking on a new form of bondage in their subjugation to merchant-lenders, or people's educational choices and later life chances being limited by the prospect of paying off student loans.

In the next chapter, we continue the analysis of risk and differences in wealth, introducing the important fact that people prefer certainty over risk if all else is equal, that is, people are risk averse. But, as we have seen, taking risks is essential to making the best of one's economic opportunities. We also see that limited wealth is associated with risk aversion, resulting in a vicious circle of enduring poverty. Not only do wealthy people *have* more money, but they *treat* their money differently when it comes time to make investments with it and bear risk when doing so.

MAKING CONNECTIONS

Incomplete contracts: The rules of the game and limited information: Bankruptcy, limited liability, and the limited information available to the lender make the level of risk taken by the borrower and the promise to repay a loan difficult to enforce.

Mutual gains (rents) from exchange and conflicts over their distribution: Borrowing and lending—like buying and selling, and hiring and working—allow both parties to the exchange to do better than were the exchange not to occur. The outcome of the conflict over distribution of these rents depends on the rules of the game.

Competition, barriers to entry, and the distribution of income: As in the market for goods, barriers to entry limit the number of competing lenders, reducing the borrower's expected income, raising the lenders' expected profits, and reducing the sum of the rents available to the two parties.

Participation and incentive compatibility constraints and optimization: As in other principal-agent models, these two constraints limit both the borrower's and the lender's optimization process; which one will be relevant depends on the rules of game.

Principal-agent models and Pareto inefficiency: Like the labor market and the market for goods of variable (and difficult to monitor) quality, the Nash equilibrium of the lender-borrower interaction is not Pareto efficient. This is because, in selecting the interest factor and the level of risk to undertake, the lender and the borrower (respectively) implement the $mrs = mrt$ rule. The result is a violation of the $mrs^P = mrs^A$ rule for a Pareto-efficient allocation.

Quantity constraints: Those excluded from borrowing are, like the unemployed, quantity constrained.

Nash equilibria in two-person and economy-wide interactions: The equilibrium of the credit market requires lenders' interest factors to be a best response to borrowers' (risk-taking) best-response function and the number of lenders in the market must be such that firms' decisions to enter and exit the market result in a constant number of incumbent firms. You studied similar conditions for the Nash equilibrium in cases where many people are interacting, rather than just two, including the common property resource (the number of people fishing in Chapter 5) and the extent of competition in the goods market (the number of firms competing to sell their outputs in Chapters 9 and 11).

Evidence: empirical relevance and history: The market-excluded and credit-constrained borrowers predicted by the model are evident in empirical studies. The model helps understand the rules of the game imposed economic hardships on former slaves in the US after the Civil War.

Public policy: redistributing wealth, enhancing competition: The model allows an assessment of the effects of policies such as wealth redistribution and reducing barriers to entry both on the size of rents to be shared between borrowers and lenders and on their distribution.

Micro-macro: Incomplete credit contracts provide the microeconomic basis for understanding both fiscal policy (the Keynesian multiplier) and monetary policy (the effect of varying the central bank's target interest rate on the opportunity cost of capital and hence on the demand for housing and cars).

IMPORTANT IDEAS

(in)complete contract	credit	rent
risk	interest factor	interest rate
borrower/agent	lender/principal	fallback
owner-operator	unlimited competition	wealth
Pareto (in)efficiency	Nash equilibrium	credit-market excluded
entrant (lender)	opportunity cost of capital	equity
incumbent lender	monetary policy	competition policy
project quality	inequality	redistribution
residual claimant	marginal borrower	credit-market constrained
Keynesian multiplier	quantity constraints	competition condition
central bank	collateral	crop-lien system

MATHEMATICAL NOTATION

Notation	Interpretation
$\hat{y}$	expected income of a borrower
$\hat{r}$	expected return of the owner-operator
f	risk of the project (failure probability and speed of the machine)
q	project quality
δ	interest factor (one plus interest rate)
z	the borrower's fallback position
$\hat{\pi}$	lender's expected profit (incumbent firms)
$\hat{\pi}^b$	lender's expected profit (prospective entrant firm)
b	probability that a lender attempting entry will not succeed
k	borrower's equity in the project or collateral posted
ρ	opportunity cost of capital
ϕ	fraction of total funds loaned to poorer borrowers

Note on superscripts: C: Complete contract; N: Nash equilibrium (incomplete contract); other superscripts and subscripts refer to the wealth level of the borrower or to particular points in figures.

"Which of these systems [central planning or market competition] is likely to be more efficient depends on the question under which of them can we expect that fuller use will be made of the existing knowledge. And this, in turn, depends on whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ought to be used but which is initially dispersed among many different individuals, or in conveying to the individuals such additional information as they need in order to enable them to fit their plans in with those of others."

Friedrich Hayek,

"The Use of Knowledge in Society," *American Economic Review* (1945)

PART 4

ECONOMIC SYSTEMS AND POLICY

In this final part of the book we return to our starting point, applying some of the analytical tools and models you have learned to understand how the institutions making up an economic system work as a whole, and how they might be made better.

There remains one set of tools that you need to do this: the analysis of risk, and the related study of inequality. Risk, which we study in Chapter 13, will be important because understanding the success of capitalism in raising material living standards requires an analysis of a risky process at which capitalism as a system has excelled: innovation.

Inequality is related to risk because the analytical tools to evaluate risky choices that a single individual may make can be repurposed to study how we evaluate an unequal situation. For example a model of how a person evaluates a gamble where both a 'good outcome' and a 'bad outcome' are possible can be also be used to evaluate a situation where 'being wealthy' and 'being poor' are the lifelong "good" and "bad" outcomes for different people. Because, along with innovation, capitalism generates significant levels of inequality, having the tools to analyze both is essential to understanding today's world.

In Chapter 14 we consider ideal models of a decentralized market economy in which contracts are complete. Two quite different variants of this approach are called "perfect competition" and "efficient bargaining." They are similar in that under some conditions they implement Pareto-efficient outcomes. We call these models utopian because they rely on assumptions—unlimited competition, complete contracts, and efficient bargaining—that are remote from actual economies as we know them.

We also revisit one of the greatest debates in the history of economics between advocates of this ideal market economy and proponents of an equally utopian conception of central economic planning of the type practiced by the Soviet Union from the time of the first five-year plan in 1928 until the end of Communist Party rule in 1991. The epigraph above is from Friedrich Hayek's contribution to that debate in which he introduced an important bit of realism, namely, the idea that information is scarce and incomplete. On this basis he advocated markets over planning because of what he held to be the superior use that markets make of the necessarily limited information available to economic actors.

Then in Chapter 15 we provide a more empirically based model of how the capitalist economy works, one that takes account of elements missing from the utopian versions of perfect competition and efficient bargaining. We model a "second-best" world in which the idealized assumptions of the utopian models of central planning and perfect competition are replaced by more realistic starting points.

Capitalism, like central planning and other economic systems, is a way of coordinating how we produce and distribute the goods and services on which we live, and how we interact with each other and with nature in the process of doing this. What distinguishes capitalism from other economic systems is that it is based on privately owned profit-making firms hiring employees to produce goods and services for sale on a market in competition with other firms. Managers of firms direct the uses of what Adam Smith called "other people's money" (the investor's stake in the firm) and other people's labor (that is the employees' work), neither of which is typically subject to a complete contract.

We ask what it is about capitalism that accounts for "the great hockey stick of history," namely the sharp upturn of output per capita and rising living standards that occurred in many countries with the emergence of the capitalist economy. We show that in many economies and time periods capitalism has been an "innovation machine" in part because it is the wealthy owners of firms who are in a position to make risky decisions concerning new technologies and products. The benefits of innovation and sometimes unfair inequalities are thus inextricably linked in the capitalist economy.

Putting together our models of markets for goods and services, credit, and labor, we have a picture of how the capitalist economic system as a whole works. The results—the division of output between profits and wages and the level of employment, unemployment, and output—determine the level of income inequality.

Public policies based on the microeconomics of capitalism can contribute to human well-being by addressing market failure and unfair inequality. In the final chapter we return to our starting point in Chapter 1—the classical institutional challenge. We show how the tools you have learned can contribute to the design of public policies in pursuit of these objectives, and why these policies sometimes fail.

CHAPTER

13

A RISKY AND UNEQUAL WORLD

It is not certain that nothing is certain.

Blaise Pascal
Pensées (Thoughts) (1670)

DOING ECONOMICS

This chapter will enable you to:

- See how experimental methods allow us to study attitudes toward risk empirically, including gender differences in risk-taking.
- Use utility functions and indifference curves to understand why people buy insurance and how this can facilitate people taking risks in ways that raise their expected incomes.
- Explain why people with less income or wealth (and as a result with limited opportunities to borrow) may be especially reluctant to take risks, and how this will reduce their income on average, perpetuating inequalities.
- See that both insurers and the insured benefit when insurance is purchased and that conflicts will exist over the distribution of these mutual benefits.
- Apply the model of risky decision-making and insurance to questions of public policy, such as tuition for higher education and tax and transfer policies to reduce both risk exposure and inequalities.
- Identify limitations of the model of risky decisions including cases where probabilities of the occurrence of uncertain events are unknown and understand the alternative insights based on loss aversion and prudence.
- Pose questions about fairness and economic injustice with the help of feasible sets, indifference curves, and Adam Smith's Impartial Spectator.

13.1 INTRODUCTION: CRASHED

"My daughter purchased a home for $120,000 just before the housing crash [of 2007–2009]," Virginia Mayou wrote to an advice column at *USA Today*. "She now owes [the bank] about $89,000. She needs to sell, but the home is valued at less than she owes. She is a single mom with teenage children and doesn't have funds to pay off the mortgage. What are her options?"

Mayou's daughter did not know she was taking a risk in purchasing her home. She had every reason to expect that the value of her home would continue increasing. House prices in the US had doubled over the decade before she borrowed from the bank to purchase the home.[1] The home loan, called a mortgage, did not seem risky to the bank either. The reason is that Mayou's daughter had given the home itself as the collateral, meaning, that if for some reason she failed to pay back the loan the bank could take ownership of the house.

Figure 13.1 Abandoned homes in Portugal after the crash in house prices that was part of the global financial crisis of 2008.

Photo © Nelson Garrido "HOME LESS" Exhibition, Venice (June 2016).

But that collateral would no longer be enough to offset the unpaid portion of the mortgage. For many people, the best they could do after the mortgage crisis began was to give the bank the keys to the house and walk away, leaving the bank with the loss.

US house prices on average would fall by more than one-quarter between the peak of the housing market in the summer of 2006 and the low in early 2012. Mayou's daughter's loss was even greater.[2]

! REMINDER Remember from Chapter 12 that wealth is the stock of things (or the value of that stock) owned by a person or household that contribute to a flow of income or other benefits. Income is the largest amount that a household or person could consume over a given period without borrowing or reducing the value of the household's wealth (their stock of assets such as home or car). The two differ because income is a *flow*—that is, it is some amount *per period of time* while wealth is a *stock*—that is, a set of things or their value with no time dimension.

She was an unfortunate and common casualty of the global financial crisis in which she was a small and unwilling player. The median wealth—the wealth of the family for which half of households are wealthier and half are poorer—fell from $107,000 to $57,800, a drop of almost one-half. Median wealth losses of Hispanic and African-American households were far greater.[3]

Households like the Mayou's were betting and perhaps even expecting—on the grounds of recent experience—the value their home to increase. For most families the value of their home was most of their wealth. Their home served as a potential source of funds—through sale or additional borrowing with the house value as collateral—in case of some health, job loss, or other emergency. For any family, to find out that their home was worth half as much as they had thought was an unimaginable disaster.

The ripple effects from the downturn in housing prices in the summer of 2006 quickly turned into a financial tsunami. Households sought to restore their declining wealth, cutting back on purchases. The car industry in Detroit was the first major sector to be devastated as people decided they could not afford a new car. Sales of cars and small trucks collapsed, from 16 million in 2007 to nine million in 2009. Experienced auto workers who were counting on their jobs until retirement were on the street looking for work. General Motors and Chrysler were headed for bankruptcy.[4]

Few were spared, and the effects went beyond losses in the value of a person's home, or GM stock, or another asset, or losing a job. When the

leading French bank BNP Paribas on August 9, 2007 announced that it could no longer pay back its loan holders, their report pointed to an entirely new dimension of uncertainty: it had become "impossible to value certain assets fairly regardless of their quality or credit rating." As the global financial crisis unfolded it dawned on people that they had no idea what their assets were worth.

What happened to the value of Virginia Mayou's house, or the value of GM stock, or the GM workers jobs, or the Paribas creditors is called a shock, an unexpected difference between what might have been expected and what actually occurred. Shocks can also be welcome, as when the price of one's house rises more rapidly than expected. This occurred year after year during the housing price bubble that led to the crash of 2007–2009. Another positive shock: the tripling of the value of GM stocks between March 2020 and June 2021, during the COVID-19 pandemic.

The fact that expected outcomes are often not realized (do not actually occur) is the foundation of the study of risk. We have already seen its importance in the credit market model of the previous chapter. Here we deepen the analysis by studying people's preferences about risk and how both individual decisions and public policy can mitigate risk.[5]

! REMINDER In Chapter 12 two very different realized outcomes occurred depending on whether the machine failed or not.

! REMINDER As we saw in Chapter 2, an *expected value* or *expected payoff* is not some particular outcome that you expect to occur. It is the weighted average of the payoffs that could occur multiplied by the probability that each payoff (or outcome) will occur.

! REMINDER The term *risk* is conventionally used to describe situations where the probabilities of the possible outcomes are known. The term *uncertainty* describes situations where the decision maker does not know and *cannot learn* these probabilities.

13.2 CHOOSING RISK: GENDER DIFFERENCES

People are *not* natural-born gamblers: if offered the choice between receiving \$100 for sure, or flipping a coin to determine whether you get \$1,200 or have to pay \$1,000, most people would take the sure \$100. But the expected payoff of the bet (the good outcome multiplied the probability it occurs plus the bad outcome multiplied by the probability it occurs) has the same value as the sure thing: $\$1200 \cdot (0.5) - \$1,000 \cdot (0.5) = \$100$. People go for the sure thing because they prefer to avoid risks.

Choosing a level of risk

The coin flip just mentioned is an example of many choices we make in which for one or more of the options before us there is a good and a bad outcome. Think about the Assurance Game (from Chapter 1) representing the problem of planting early or planting late in Palanpur shown again in Figure 13.2. For each of the two strategies that Aram or Bina could choose (Plant Early or Plant Late) there is a good outcome (resulting in a higher payoff) and a bad outcome (resulting in a lower payoff). If Aram plants early, the good outcome occurs when Bina also plants early and he receives a payoff of $y^G = 4$ and a bad outcome (when Bina plants Late) namely $y^B = 0$.

Therefore, for planting early, Aram's **risk exposure** is $\Delta_E = y^G - y^B = 4 - 0 = 4$. Planting late also has a good outcome $y^G = 3$ and a bad outcome

RISK EXPOSURE The difference between the better and worse outcome when the two are equally likely. We also term this the level of risk.

Figure 13.2 Planting in Palanpur: an Assurance Game. The *degree of risk* is the difference between the good and bad payoffs. So in Palanpur, Plant Late has a degree of risk of 1 ($3-2=1$), whereas Plant Early has a degree of risk of 4 ($4-0=4$).

		Bina	
		Early	Late
Aram	Early	4, 4	0, 3
	Late	3, 0	2, 2

$y^B = 2$. Therefore, for planting late, Aram's risk exposure is $\Delta_L = y^G - y^B = 3 - 2 = 1$.

If we assume that both Aram and Bina think that the other is equally likely to Plant Early or Plant Late, then playing Plant Late is less risky because the difference (the Δ) between the good and bad outcomes is smaller. Then Δ is a measure of the extent of the risk of the two strategies. His degree of risk is four times as great when Aram plays Plant Early as when Aram plays Plant Early. The same is true of Bina.

We treat the level of risk Δ as something that people choose, ranging from everyday actions like whether to carry an umbrella or not, to life-setting decisions such as whether to emigrate to another country.

The choice of Δ may refer to any of the following:

- A student choosing to specialize in nuclear engineering (where salaries are high but opportunities few outside the nuclear power industry) rather than in liberal arts where expected salaries are modest, but your training may equip you for a large variety of jobs, exposing you to less risk than the nuclear engineer.
- Relocating to a booming region of your country in search of work (substantial returns if you land a job, losses if you do not) rather than taking one of the available low-wage jobs in in your home town (low risk with lower potential returns).
- Going into business as a self-employed person—electrician, software engineer—rather than taking a salaried job with a predictable course of future raises.
- Producing and selling a single product (high risk, with potential high returns if it is very good) rather than a range of products (low risk, spread across many products).

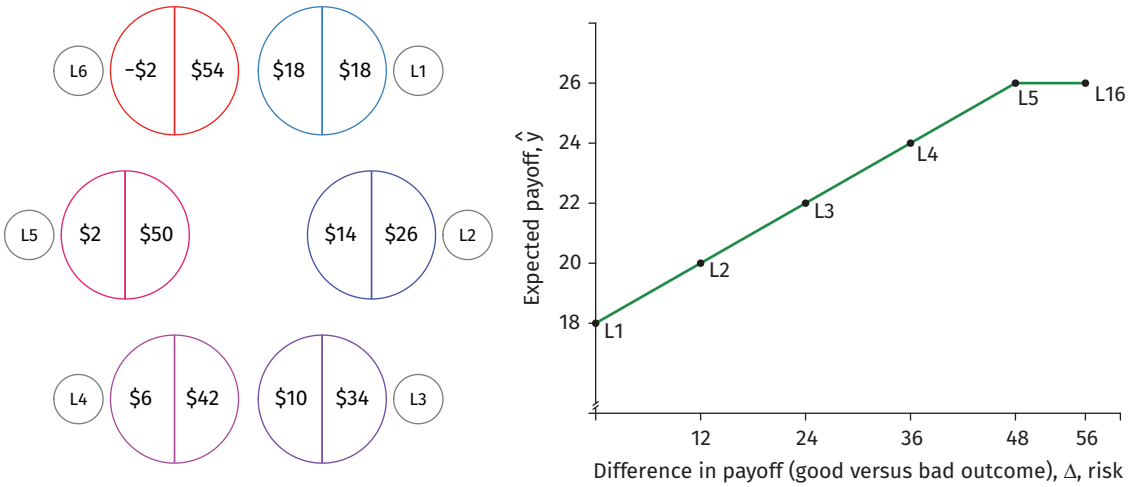
A lottery

We can describe the choices open to a person by listing what are termed the lotteries in which they may engage. A **lottery** is a set of outcomes and the probabilities that each will occur. In ordinary language a lottery is a gambling game played in casinos or online, such as Powerball in the US, EuroMillions in Europe, or the National Lottery in the UK. In game theory a lottery is a way of modeling risky choices involving two (or more) outcomes with given probabilities of occurring.

Aram, Bina, and the farmers of Palanpur faced two lotteries: Plant Early and Plant Late. In each case there was a risk—the difference between the good and the bad outcome of the lottery. And we assumed that in order

LOTTERY In game theory a lottery is a set of uncertain outcomes and the probabilities that each will occur.

Figure 13.3 Six lotteries: expected payoff and risk. In panel (a), we show the set of lotteries a person could choose in an experiment. The two numbers in the circles indicate for that particular lottery, the equally likely bad and good outcome (in that order). Panel (b) shows the expected payoff of each of the lotteries along with the degree of risk.



(a) Experimental choice of lotteries

(b) Expected payoffs and degree of risk

to make a decision each of the players assigned some probability that the good and bad outcomes would occur, depending on what the other player did, if they chose plant early or plant late.

Some examples of lotteries are shown in the left-hand panel of Figure 13.3. Lottery (L6) in the upper left is the riskiest: it describes a 50–50 chance of winning \$54 and losing \$2. The expected payoff is just the average of these two numbers, or \$26, and the degree of risk is $\Delta = \$54 - (-\$2) = \$56$. By contrast, L1 is not risky at all: if you play that lottery you get \$18 with certainty. L1 is not really a lottery as there is no uncertainty.

In the Figure 13.3 (b) comparing lotteries 1 through 6, the size of the difference between the good and bad outcomes (that is, the level of risk or Δ) increases and the expected payoff increases (up to L5).

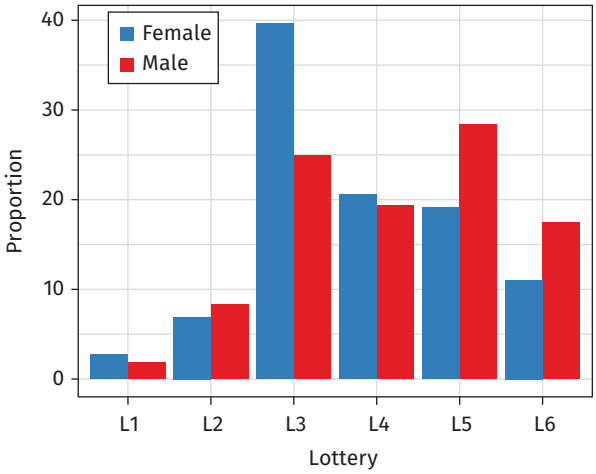
Gender differences in risky choices

Economist Sheryl Ball and her coauthors studied gender differences among US university students in choices among the six lotteries. We present the results of their experiments in Figure 13.4.

All subjects confronted the same set of expected payoffs and risks given by the six lotteries (Figure 13.3). The most common choices differed by gender. You can see from the figure that both men and women are among those choosing the riskiest lottery (L6) and also choosing the no-risk lottery (L1). But the most common choice (the modal choice) for men was lottery 5. The modal choice for women was lottery 3. On average men were risk-takers and women were risk-avoiders.

Figure 13.4 Choices of the level of risk by male and female university students. Lottery 1 (L1) is risk free, while Lottery 6 (L6) is the riskiest lottery. The height of each bar is the percentage of all women (blue) or men (red) who chose the lottery indicated. The means of the women's and men's choices, respectively were 3.79 and 4.16.

Source: Ball et al. (2010).



CHECKPOINT 13.1 Risk and expected payoffs Which lottery would you choose? Why do you think some of the subjects in the experiment chose L6?

13.3 RISK PREFERENCES OVER LOTTERIES

Why do some people choose riskier lotteries than others? For example, what accounts for the differences in Figure 13.4 between those (both men and women) who choose the risky L5 and the sure thing, L1? The risk-takers must have evaluated the benefits of a greater expected income in L5 highly and not been too concerned about the fact that they might end up with just $2 (when they could have had $18 for sure, if they had chosen L1). Placing a negative value on being exposed to risks or uncertainty is called **risk aversion**. It is a common reaction of people and it could be due to a combination of:

- Anxiety about not knowing what will occur and a related personality trait that psychologists term harm avoidance.

RISK AVERSION A risk-averse person dislikes uncertainty about outcomes and will choose a certain outcome valued at $x over some lottery whose expected value is greater than $x.

- A desire to avoid regret about having made a bad decision.
- Diminishing marginal utility of wealth, income, or whatever the currency of the lottery is.

To see why the last bullet is true think about a person with just enough money to purchase one adequate meal in a day. He is then offered a lottery that instead of his one meal for sure, he will with equal probability have no meal or two meals. Think about what you would choose.

You would likely choose the sure thing because one meal is much better than no meal, and while two meals would be nice, the "good" outcome of the risky lottery (two meals) is not good enough to run the risk of having nothing to eat. Where the difference in one's wealth, health, or income associated with risk is substantial–having a job or not, developing extraordinary vision capacities, or losing one's sight–then, like no meal at all, the "bad outcome" may be sufficiently catastrophic to motivate avoiding it at almost any cost.

M-CHECK Risk and diminishing marginal utility. Risk can be analyzed using a utility function $u(y)$ in which marginal utility is diminishing in the level of realized income, y. In this framework people do not have feelings of anxiety or excitement about risk itself; instead attitudes toward risk are explained entirely by the marginal utility of income. We do not use this approach here because we prefer a more general model that can include all of the reasons—including those studied by psychologists—why people might want to avoid taking risks or enjoy risk-taking. Another advantage of the approach presented here is that it makes use of analytical tools already taught: indifference curves and feasible sets.

Risk-averse indifference curves

To understand how people make risky choices we describe their evaluations of different outcomes using a utility function in which

- *Expected income is a good:* something the decision maker prefers and wants more of.
- *Risk is not a good:* something the decision maker would like to avoid or is possibly indifferent to, but does not prefer.

Here is the function:

$$\begin{aligned} \text{person's utility function} \quad & u = u(\Delta, \hat{y}) \qquad (13.1) \\ \text{with marginal utility of risk} \quad & u_\Delta \leq 0 \\ \text{marginal utility of expected income} \quad & u_{\hat{y}} > 0 \end{aligned}$$

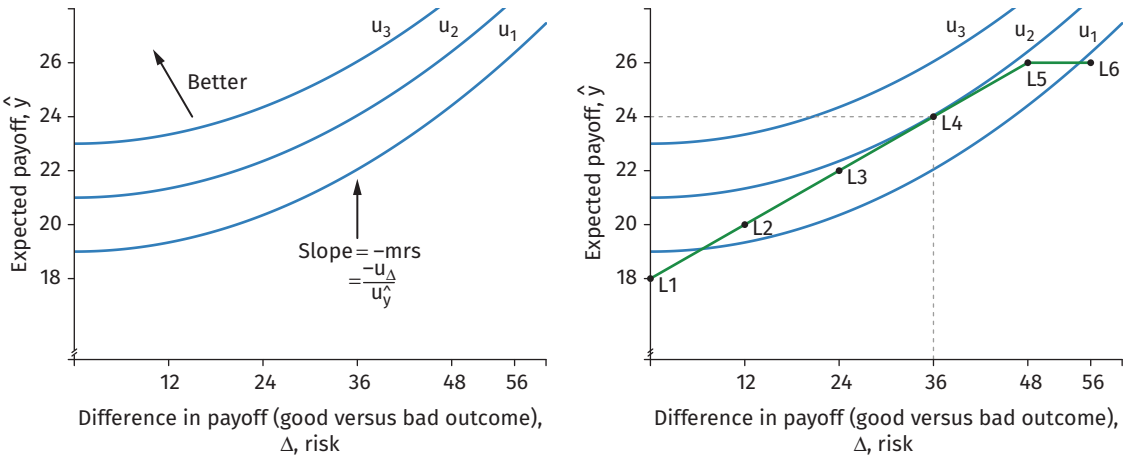
We often refer to the marginal *dis*utility of risk and this is just the negative of the marginal utility of risk or $-u_\Delta \geq 0$.

!REMINDER This is similar to what we did in Chapters 10 and 11 where we referred to the disutility of providing quality or effort as the negative of the marginal utility of providing quality or effort.

We allow for people with a given level of risk exposure and expected income $(\Delta, \hat{y})$ whose marginal utility of risk $u_\Delta = 0$, so that the person is termed **risk neutral**, that is, indifferent to the level of risk. If the level of risk in a lottery were very small (e.g. a payoff of \$10.01 versus \$9.99) a person might not place any negative value on the risk involved. But for risks of any significant magnitude we assume that $u_\Delta < 0$ so the marginal utility of risk is negative, the person is risk averse.

RISK NEUTRAL A risk-neutral person is indifferent between receiving \$x with certainty and playing an uncertain lottery with the same expected value. A risk-neutral person is not risk averse.

Figure 13.5 Feasible lotteries and indifference curves for risk-averse preferences. Because expected payoff is a "good" and risk is a "bad," preferred outcomes are above and to the left. As a result the indifference curves slope upward: if two outcomes are associated with the same level of utility, then one must have both higher risk and higher expected payoff than the other.



(a) Indifference curves of risk and expected income

(b) Feasible outcomes and indifference curves

M-CHECK As before (e.g. in M-Notes 10.1 and 12.2) to find the slope of an indifference curve we totally differentiate the function in question and set the result equal to zero. In this case we set $d\hat{y}(u_{\hat{y}}) + d\Delta(u_{\Delta}) = 0$ and then solve for the slope of an indifference curve: $\frac{d\hat{y}}{d\Delta} = \frac{-u_{\Delta}}{u_{\hat{y}}}$. See also M-Note 13.1.

! REMINDER The points making up a particular indifference curve are bundles of goods consumed or other results of actions taken associated with identical levels of utility. On a given indifference curve in Figure 13.5, therefore, utility is the same for differing levels of risk and expected income.

To see if a utility function like Equation 13.1 could explain gender differences in the modal choice of a particular lottery, look at the indifference curves in the Figure 13.5 (a). The indifference curves are upward-sloping because risk is a bad. How steep they are at a particular point is a measure of how risk averse the person is under the conditions given by the particular level of Δ and $\hat{y}$ at that point. The slope of the indifference curve (as we explain in the M-Check) is:

$$-mrs(\Delta, \hat{y}) = \frac{\text{Marginal disutility of } \Delta}{\text{Marginal utility of } \hat{y}} = \frac{\overbrace{-u_{\Delta}}^{>0}}{\underbrace{u_{\hat{y}}}_{>0}} \geq 0 \tag{13.2}$$

This expression gives the answer to the question: Suppose you were required to take on a little more risk (an increase in Δ, the bad, so a move to the right in the figure); how much additional expected income (a move up) would you need to be no worse off than before (that is, to get you back to the same indifference curve)? The answer depends on how bad the risk is (u_{Δ}) compared to how much you value expected income ($u_{\hat{y}}$). So, the slope of the indifference curve is a measure of how risk averse the person is at each point along the curve.

Doing the best you can in a risky situation

To take account of the constraints that limit what the decision maker can do, we define a feasible set to include all of the combinations of the good (expected income, $\hat{y}$) and the bad (the level of risk, Δ) that the decision maker

can implement by their actions. For the experiment just described these are the points in the right panel of Figure 13.5 representing the six lotteries. You can see that for the person whose indifference map is shown, the best they can do is to select L4, which will expose them to a substantial amount of risk ($36) but also with a substantial expected payoff $24.

How might this way of understanding risky choices explain the gender differences found in the choice of lotteries?

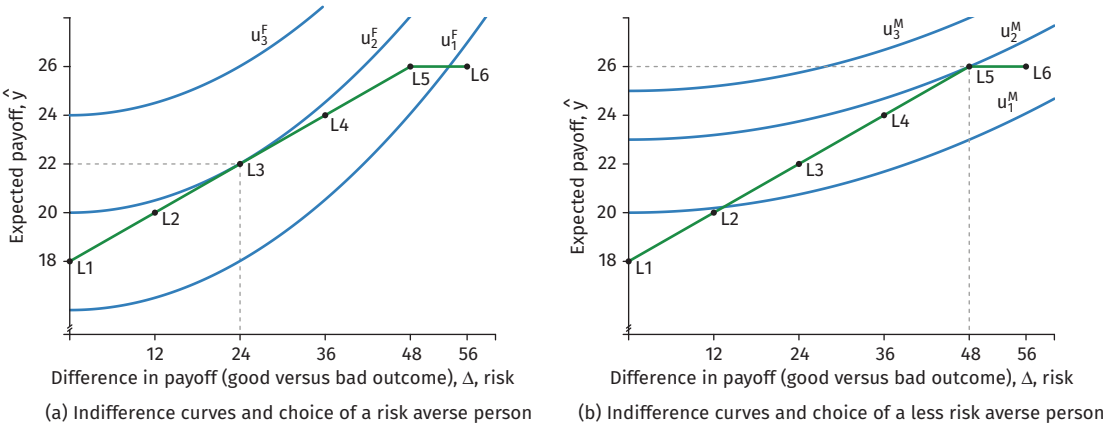
✓ **FACT CHECK Gender differences in risk preferences** There is no widely accepted explanation for gender differences in risk preferences. A study of two different cultural groups in China suggests that differences in upbringing and the social interactions that one experiences may be important.[6]

Gender differences in risk aversion?

In Figure 13.6 we provide a possible answer: the two panels show possible indifference maps that would lead to a different choice of lottery. Those in the left panel are steeper, and thus illustrate higher levels of risk aversion that than those on the right, which are flatter, meaning less risk averse. The person depicted on the left would choose L3, the modal choice of women in the experiment, while the person on the right would choose L5, the modal choice of men. Thus a gender difference in risk preferences—women more like the person on the left, men more like the person on the right—could explain the experimental results.

We do not know, of course, that the differences between men and women in the experiment are explained by the kind of risk-averse indifference curves that we have introduced. Look carefully at their choices over all of the lotteries in Figure 13.4. The expected income of L6 is no greater than L5, and it is riskier. But a substantial number (both men and women) chose L6. They may have placed a positive value on the stakes of the game being

Figure 13.6 More and less risk averse indifference maps motivating the modal choices of women and men. In panel (a) we show a relatively steep—that is, risk averse—indifference map that could have resulted in the less risky modal choice for women. Panel (b) shows a flatter set of indifference curves that could have motivated the riskier modal choice of men.



(a) Indifference curves and choice of a risk averse person

(b) Indifference curves and choice of a less risk averse person

✓ **FACT CHECK Novelty seeking** Psychologists have identified a personality trait called *novelty seeking* that is associated with risky behaviors and inversely correlated with another trait mentioned above, *harm avoidance*.

large, even if the expected income from taking the chance was no higher. Explaining *their* actions would require a different approach, which could also describe the preferences of those who take part in sky-diving and other "thrill-seeking" activities.

Differences in risk preferences between men and women or among other groups can be important because they may explain real-world decisions about risky choices. For example, other things equal, the evidence suggests that men are more likely to choose riskier assets with higher expected returns like stocks and mutual funds for their retirement portfolios than are women. This difference will result in greater incomes on average when men reach retirement.

Risk preferences can affect academic performance (risk-takers are more willing to guess answers) and subsequent choices about careers. Also affected are job choice (e.g. going into business on your own versus becoming a teacher), and structure of one's compensation package (e.g. share in a firm's profit vs a fixed salary).

Taking risks in all of these cases may raise a person's lifetime expected earnings. If women are on average more risk averse than men, then this will contribute to women on average having lower incomes than men. Differences in risk aversion may be part of the explanation of the persistence of inequalities not only between men and women, but more generally between wealthy and wealth-poor people.

✓ **FACT CHECK** Allison Booth and Patrick Nolen found that in experiments similar to that reported in Figure 13.4 when women are competing with other women in groups of only women, the risk levels they choose do not differ from men. So the context in which the choice takes place makes a difference. Taking account of the effect of context would require representing women's preferences by a different utility function when competing in an all-women group, with fatter, less risk-averse indifference curves than when competing with both men and women.[7]

CHECKPOINT 13.2 Risk neutrality Draw an indifference map (like Figure 13.5) of a risk neutral person.

13.4 DECREASING RISK AVERSION: THE PERSON AND THE SITUATION

People differ in their degree of risk aversion, as the results of the experiment shown in Figure 13.4 suggest. These differences may be the result of:

- *Their type of person*: Some people feel anxiety or distress about uncertainty; others value "surprises." When we distinguish between different types of people we mean people with different utility functions and as a result different entire indifference maps, like the two people contrasted in Figure 13.6.
- *Their situation*: People already facing significant risks and unable to afford any serious loss will be very averse to additional risks. If the same person were wealthier and less exposed to risk, they might be less averse to risk. By the person's situation we mean their particular bundle of a level of risk exposure Δ and expected income $\hat{y}$ indicated by a point in the figure. Differences in the situation faced by a person could

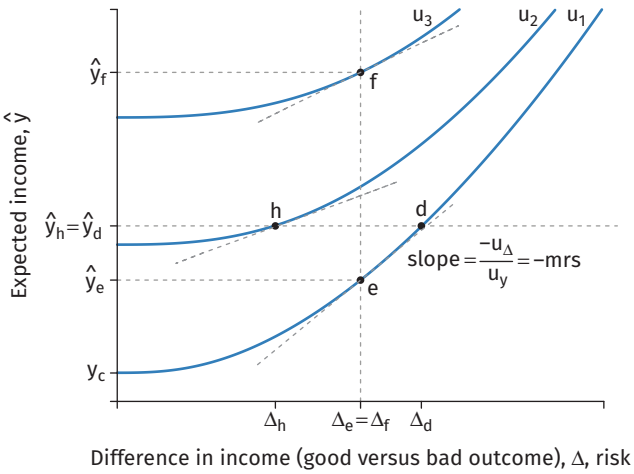
result in differing levels of risk aversion, indicated by the slopes of their indifference curves at different points in the space of expected income and risk. The same person will have differing levels of risk aversion depending on their situation.

In Figure 13.7 we illustrate two important influences of the situation on one particular person's level of risk aversion:

- *Risk exposure*: the degree of risk exposure (how far to the right the person's situation is in the figure). People exposed to substantial risks experience a large marginal disutility of additional risk. This can be seen by comparing the slope of the indifference curve at points **h** and **d**. For a given level of expected income, where risk exposure is greater (at **d**), risk aversion (the slope of the indifference curve) is also greater.
- *Income or wealth*: the person's expected income (how far up in the figure the person is). Those with little income also experience a large marginal disutility of additional risk. This is illustrated by comparing points **e** and **f**. For a given level of risk exposure, when the person's expected income is greater (**f**) the person is less risk averse (the indifference curve is flatter) than where their expected income is less (**e**).

This is why the indifference curves of any given person are steeper as you move horizontally to the right, and flatter as you move vertically upward.

Figure 13.7 Decreasing risk aversion as shown in three indifference curves for a person. The slope of the indifference curve at some point $(\Delta, \hat{y})$ is the ratio of the person's marginal disutility of risk to their marginal utility of expected income, that is, a measure of risk aversion.



EXAMPLE A 20,000 euro loss in income—say, due to an injury or illness—will inflict a much greater catastrophe on a person who is poor than it would if they were well-off. What this means is that for the poor person u_Δ is much greater (in absolute value) so (by Equation 13.2) the indifference curve will be steeper.

REMINDER Reminder Insurance was defined in Chapter 12 as any costly action one can take that reduces the level of risk to which one is exposed.

The tendency of a person to be less risk averse if she has more income than if she has less income is called **decreasing risk aversion** (risk aversion is less if income or wealth is greater). The most important reason for decreasing risk aversion is that a negative shock of a given size is likely to be a much greater loss in well being for a poor person than it would for the same person were they to have greater wealth.

Two additional features of this indifference map are important.

- **The certainty equivalent (y_c)**: The intercept of the indifference curve u_1 and the vertical axis, y_c, has the same utility as points **e, d**, and all of the other combinations of risk and expected income that make up the indifference curve. What is unique about y_c is that there is *no risk* (it is like Lottery 1 in the experiment, a sure thing). So what y_c tells us is the level of *certain* income that would be valued by the person equally to each of the other combinations on the same indifference curve (involving more risk and more expected income).
- *Risk neutrality*: We have said that the indifference curves are upward-sloping, but for a person sufficiently wealthy they might be flat (not shown in the figure). Because risk aversion is measured by the slope of the indifference curve and a flat indifference curve has a slope of zero, in this (very wealthy) situation the person would be risk neutral: she would care only about expected income, not about risk. You can also see that at the vertical intercept–that is, very small levels of risk, the indifference curves are approximately flat, the person exposed to virtually no risk at all, would not be risk averse.

Recall from Chapter 12 that people of limited wealth may be unable to borrow at all unless at usurious payday loan interest rates. We also saw that access to credit provides a kind of insurance, because the lender bears some of the loss in the case of project failure.

These facts along with risk averse indifference curves can help us explain why poor people may choose not to make risky investments–including investments in their own ability to earn higher incomes such as further training or moving to a distant part of the country. As a result poor people may end up poorer on average than they would have been had they taken the risk.

DECREASING RISK AVERSION The tendency of a person to be less risk averse if she has more income (or wealth) than if she has less.

CERTAINTY EQUIVALENT The level of *certain* income that would be valued by the person equally to each of the other combinations on the same indifference curve (involving more risk and more expected income).

The result can be a vicious circle or self-reinforcing poverty, contributing to economic inequality and its tenacious persistence. In the next section we provide an example.

CHECKPOINT 13.3 **Risk aversion** In Figure 13.7 explain how the indifference curves show that at point **f** a person is more risk averse than they would be at **e**, and at **d** more risk averse than they would be at **h**. Explain why each would be true.

13.5 APPLICATION: RISK, WEALTH, AND THE CHOICE OF TECHNOLOGY

Farming is one of the riskiest occupations. This is because the farmer's income depends on three things that vary substantially and are out of the farmer's control:

- weather and other environmental conditions affecting crop growth;
- susceptibility of crops and livestock to disease; and
- the prices at which inputs are purchased and (especially) outputs are sold.

Partly for these reasons, researchers have studied farmers to better understand behavior in risky situations. Research on farming families in Indian villages recorded both the types of risk to which the farmers are exposed—uncertainty about the date that the dry season would end and the rains start—and the differing ways that farmers coped with the resulting uncertainty about their incomes. The researchers also recorded the total wealth of the farmers and the forms that the wealth took: land, irrigation equipment, tools, stocks of grain, and draft animals (neutered bulls, that is, bullocks).

There were substantial differences in wealth among the farmers studied: the richest one-fifth of the farmers owned 54 percent of the total wealth. Rich and poor villagers also differed in *how* they farmed: those with a substantial amount of total wealth favored investments in pumps and other irrigation equipment, while the less wealthy rarely purchased pumps and invested their limited wealth in bullocks.[8]

The researchers also found that few of the farmers had access to credit in times of need. Instead, to meet their needs they resorted to selling one or more of their bullocks, in which there was a market to sell them.

By not investing in irrigation equipment the less well-off farmers missed an opportunity to make substantially more profits: taking account of the costs, an installed pump would have raised profits by 72 percent on average. The reason the farmers avoided this profitable investment is simple: the poor farmers owned bullocks in order to have something they could sell to get them through times of need. The bullocks were a kind of combination savings account and insurance policy! In the villages under study there was no second-hand market in irrigation pumps and other equipment, so, unlike

Figure 13.8 **Bullocks tilling the ground in Karnataka, India.**

EXAMPLE Farmers in the US can purchase insurance to cover losses due to adverse weather. Insurance against damage by hail has been available in France and Germany for almost two centuries. But for most farmers in poor countries today insurance is not available. Think about the farmers in India. If they could have purchased crop insurance their resulting lesser degree of risk exposure would have allowed them to invest in irrigation pumps rather than bullocks, greatly increasing their income, and perhaps lifting them out of poverty.

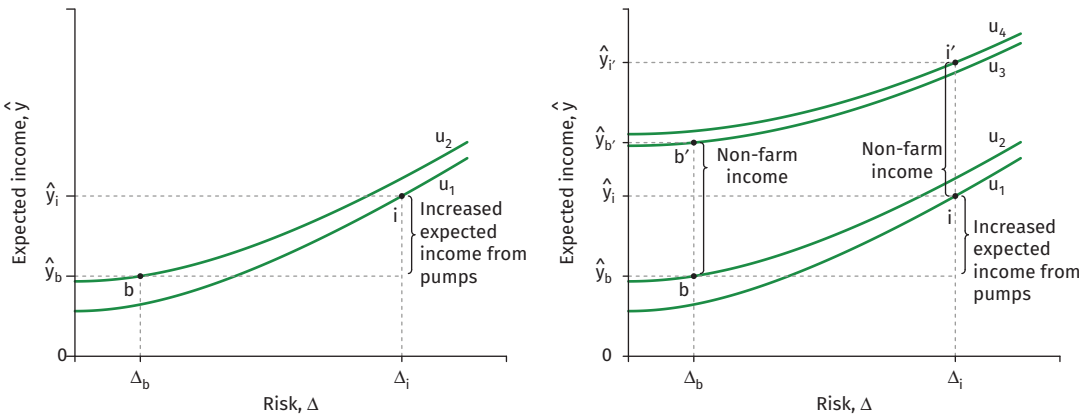
a bullock, owning a pump did not provide a buffer against risk (adverse weather or pests destroying the farmers-crops).

The fact that the wealthier farmers invested in riskier and more profitable assets (pumps rather than bullocks) had the effect of perpetuating or widening the income differences between them and the other farmers whose low income and lack of access to borrowing made them more risk averse.

Figure 13.9 illustrates this process and how it works for a poor farmer, Anil. On the horizontal axis is the risk undertaken by his choice of investments, with two levels shown: a less risky one with substantial investments in bullocks, the other riskier option with a greater investment in pumps. Thus, Anil must choose between investing in irrigation pumps with risk Δ_i and investing in bullocks with risk Δ_b.

An indifference curve through Anil's choice (point **b**) shows all combinations of risk and expected income that are equally preferred by him (as before points higher and to the left of this curve are preferred). In panel (a), we can see that Anil is better off with the bullocks (point **b**) than with the pumps: point **i** on indifference curve u_1 is below and to the right of the indifference curve through point **b** on indifference curve u_2 (remember $u_2 > u_1$).

Figure 13.9 Risk, wealth, and choice of farm technology. We consider Anil in two different states of the world. In the first state, shown in panel (a), he does not have much wealth and he must choose between investing in irrigation pumps with risk Δ_i and investing in bullocks with risk Δ_b. He chooses the less risky point **b** on indifference curve u_2 rather than at point **i** on indifference curve u_1 ($u_2 > u_1$). In the second state of the world (shown in panel (b)), Anil is rich because he has non-farm income. In Anil's wealthy state of the world, he chooses the riskier point **i'** on indifference curve u_4 rather than at point **b'** on indifference curve u_3 ($u_4 > u_3$). Remember: the *vertical* distance between points **b'** and **b**—his non-farm income—is the same as the distance between points **i'** and **i**.



(a) Anil's choice when he is poor

(b) Anil's choice when he has non-farm income that increases his wealth

Now suppose that Anil wins the lottery or somehow obtains substantial wealth, from which he will receive an amount of *non-farm income*. His expected income would be his farm income *plus* his non-farm income (from the newly acquired wealth).

In the second state of the world (shown in panel (b)), Anil is rich and has non-farm income as indicated by the equal difference between points **b** and **b'** or between **i** and **i'** (his difference in income is $\hat{y}_{b'} - \hat{y}_b = \hat{y}_{i'} - \hat{y}_i$). In the state of the world where Anil has non-farm income, he chooses point **i'** on indifference curve u_4 rather than at point **b'** on indifference curve u_3. When he is wealthy he invests in the riskier asset of irrigation rather than bullocks and has higher expected income as a result with expected income $\hat{y}_{i'}$ rather than $\hat{y}_{b'}$ if he had invested in bullocks.

What explains the difference in his choice of assets when he is wealthy? As in Figure 13.7, the indifference curve through point **i'** is flatter (less risk averse) because people with more income are less risk averse. Because of this, when **b** and **i** both increased by the same amount of non-farm income, **i'** ended up on a higher indifference curve than **b'**. In this case, Anil, the same farmer who was once poor, became less risk averse because he became rich.

We can summarize what happened. In the first state of the world, Anil was caught in what is called a poverty trap in which his low income and lack of access to borrowing (as protection against risk) led him to choose the less risky but lower profit option: bullocks.

He did not make a mistake; Anil was doing the best that he could given his situation. The choice made for good reasons of prudence made him safer, but it also kept him poor. If he had had higher income to start with, he would have maintained a high income and invested in the higher return investment: irrigation.

A key part of the story was the level of risk that Anil chose. He had just two options. But in general we have a range of choices about the level of risk we undertake. We can analyze these choices using the same tools of constrained optimization that you have used since Chapter 3.

! REMINDER From Chapter 1, a *poverty trap* is a self-reinforcing set of processes (a vicious circle) that perpetuates low income for a person or group that, were the trap broken, could enjoy self-reinforcing prosperity.

CHECKPOINT 13.4 **You Win!** Recall the You Win! competition in Nigeria introduced in Chapter 12 where businesses were randomly selected to get the equivalent of $50,000 to invest in their businesses. People invested in minibus taxis, a factory to manufacture paint, and many other opportunities. Using the tools in Figure 13.9 explain why this policy might have worked.

› EXAMPLE Look up Episode 702 of the NPR (National Public Radio, USA) Planet Money Podcast if you find the You Win! program interesting.[9]

13.6 DOING THE BEST YOU CAN IN A RISKY WORLD

In the previous chapter we modeled the risk-taking choices of the operator of a "machine," choosing the speed at which it is run, f. In that model the speed of the machine determined both:

- *the difference between the good and bad outcome*: namely, the revenues possible by selling the goods produced by the machine if it does not fail, minus the revenues possible if it does, that is zero; and
- *the likelihood of failure*: The probability that the bad outcome would occur, which increased the faster the machine was run.

In this chapter, we study the choice of the extent of the difference between the good and bad outcome, Δ, but not the probability that each will occur. And for simplicity we will assume that the good and bad outcomes are equally likely: they each occur with probability one-half, independently of any actions that the decision maker takes. Recall that we call the difference between the good and the bad outcome the level of risk, or just risk.

Feasible choices of risk and return

The terms "returns to risk" or just "**returns**" are the realized income or expected income resulting from having made an investment or some other risky choice. It is what you "get back" (hence the term "returns."). Here we let the expected income resulting from a risky choice represent the returns.

Of course most people would like to choose a course of action with high returns and low risk. But not all combinations of risks and returns to risk in terms of expected income are *feasible*. The feasible combinations of expected income ($\hat{y}$) and risk (Δ) are bounded in Figure 13.10 by the risk–return schedule, $\hat{y} = \hat{y}(\Delta)$. Similar to other feasible frontiers you can see that it divides the space of risk and expected returns into combinations that could possibly occur (in light green) and those which will not under any circumstances occur (in blue).

! REMINDER The *marginal rate of transformation* is the negative of the slope of the feasible frontier, here given by the risk–return schedule. It is a measure of the *opportunity cost* of the one in terms of the other, that is, the opportunity cost of less risk is lower expected income.

Like the risk and expected income curve in Chapter 12, with the risk–return schedule expected income first rises with risk-taking—posing a trade-off to the decision maker—then expected income reaches a peak and thereafter falls. The slope of the risk–return schedule is the (negative of the) marginal rate of transformation of risk into expected income.

$$\text{Slope of the risk–return schedule} = -mrt(\Delta, \hat{y}) = \frac{d\hat{y}}{d\Delta} \equiv \hat{y}_{\Delta} \tag{13.3}$$

The choice of a risk level by a risk-averse person

We can use the risk–return schedule along with indifference curves that capture the decision maker's risk preferences to understand the choice of a risk level. To do this, we introduce two people: Arjun (A) and Nicolas (N). Arjun is risk *a*verse and Nicolas is risk *n*eutral, we explain in the next section how these risk preferences relate to their levels of wealth.

RETURNS The term "returns to risk" (or just returns) is the realized income or expected income resulting from an investment or some other risky choice.

Figure 13.10 Feasible combinations of risk and expected income. Feasible combinations of risk (Δ) and expected income ($\hat{y}$) are shown using the risk-return schedule, $\hat{y} = \hat{y}(\Delta)$. A combination within the feasible set, for example point **e**, is feasible, but would not be selected by a decision maker because if it is not on the feasible frontier then there must be some other point **d** with the same expected income and less risk and another point **a** with the same risk and greater expected income. Both of these points dominate point **e**. Points outside the set are infeasible. Point **m** shows the choice of risk (Δ_m) that maximizes expected income at $\hat{y}_m$—you can see that the slope of the risk-return schedule is zero as shown by the tangent line.

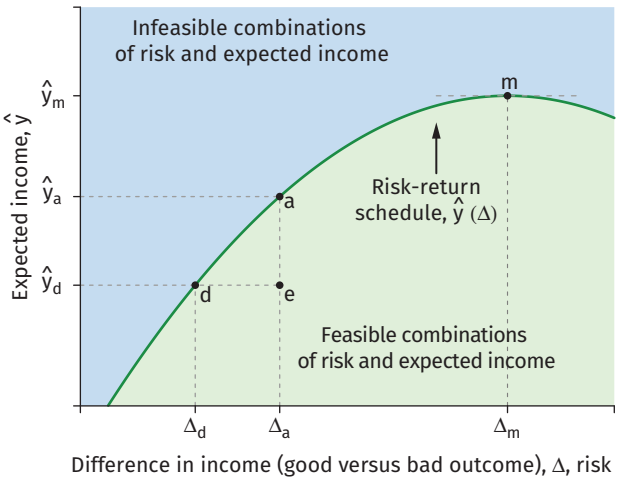
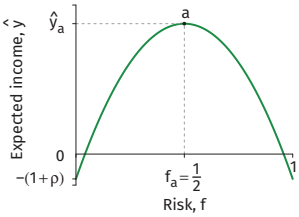


Figure 13.11 Risk and expected income. You have already seen a risk-return schedule in Chapter 12, reproduced here. In Chapter 12 we modeled risky choices using the relationship between the level of risk taken f and the expected profit of a project. "Schedule" is just another word for "function."



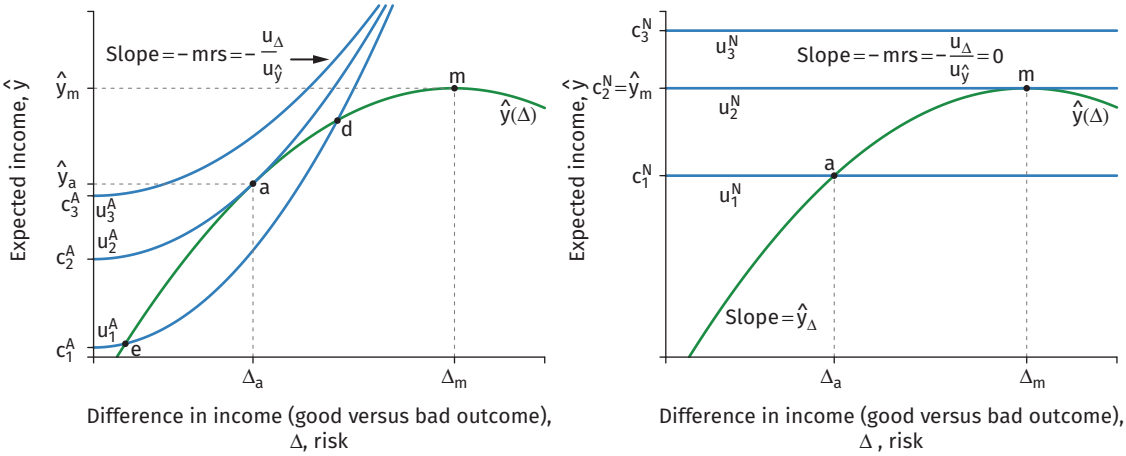
The decision maker will vary Δ to maximize $u(\Delta, \hat{y})$ subject to the risk–return schedule $\hat{y} = \hat{y}(\Delta)$. We show in M-Note 13.1 that this requires choosing the Δ that equates the:

$$\text{Slope of the indifference curve} = \text{Slope of the risk–return schedule}$$
$$-mrs(\Delta, \hat{y}) = \frac{-u_\Delta}{u_{\hat{y}}} = \hat{y}_\Delta = -mrt(\Delta, \hat{y}) \tag{13.4}$$

Arjun, the risk-averse person (that is, $u_\Delta < 0$) shown in Figure 13.12 (a) could pick any point on the risk–return schedule (including **a, d, e**, or **m**). Arjun's indifference curves will be upward-sloping because, as a risk-averse person, expected income is a good and risk is a bad. Finding the level of risk, Δ, at which there is a tangency between his risk–return schedule and his highest indifference curve, he will select a level of risk Δ_a, with an expected return of $\hat{y}_a$, at point **a**.

What level of risk will Nicolas, the risk-neutral decision maker, choose when restricted to feasible combinations of risk and expected income? He will have horizontal indifference curves like those shown in Figure 13.12 (b),

Figure 13.12 Indifference curves and risk choices of a risk-averse (panel (a)) and risk-neutral (panel (b)) person. A risk-averse person chooses a point like **a** where his indifference curve (u_2^A) is tangent to his risk–return schedule; as a result he obtains a bundle of expected income and risk ($\Delta_a, \hat{y}_a$). The risk choice of a risk-neutral person Δ_m corresponds to expected income $\hat{y}_m$ on u_2^N.



(a) Risk-averse person

(b) Risk-neutral person

M-CHECK Remember, we use the symbol $\hat{y}_\Delta$ to mean $\frac{d\hat{y}}{d\Delta}$, the derivative of expected income with respect to the choice of risk. This is the slope of the risk–return schedule or feasible frontier, or the (negative of) the marginal rate of transformation.

! REMINDER The slope of an indifference curve is $\frac{-u_\Delta}{u_{\hat{y}}}$. A risk-neutral person does *not* care about risk so $u_\Delta = 0$, therefore the indifference curves of a risk-neutral person are horizontal (zero slope). A risk-averse person does care about risk as a bad, so $u_\Delta < 0$, therefore their indifference curves are upward-sloping (have a positive slope).

and so will select point **m**, implementing the level of risk that maximizes his expected income with bundle ($\hat{y}_m, \Delta_m$).

Another way to compare the actions of risk-averse and risk-neutral people is to compare how they value different risky choices against a *certain* level of income. The certain amount of money that has the same utility as a set of risky choices is shown by the vertical intercept of the indifference curves in Figure 13.12 or the "risk-free return." We call this the certainty equivalent as it is the certain amount of money that is equivalent in utility terms to a lottery or risky choice. For example, the quantity c_2^A is the certainty equivalent of every point—that is every combination of risk and return—making up Arjun's indifference curve u_2^A (remember every point along the indifference curve corresponds to a lottery). That is, the certainty equivalent c_2^A is equivalent in utility terms to the risky choice indicated by point **a** with risk and expected income ($\Delta_a, \hat{y}_a$) shown in Figure 13.12 (a).

Now, consider the certainty equivalent for Nicolas for the risky choice given by point **a** in Figure 13.12 (b). Nicolas's indifference curves are flat and therefore his certainty equivalent for point **a** is given by c_1^N. Nicolas's certainty equivalent for the risky choice given by point **a** is higher than Arjun's certainty equivalent for point **a** ($c_1^N > c_2^A$). This demonstrates a common pattern: for the same risky choice, a risk-averse person has a certainty equivalent that is *lower* than the certainty equivalent for a risk-neutral person.

A contrast with the credit market model in Chapter 12

Recall that in Chapter 12 when the lender extended a loan to a borrower under an incomplete contract, the borrower chose a level of risk greater

than the expected income-maximizing level, which is $\Delta = \Delta_m$ here and was $f = \frac{1}{2}$ in Chapter 12. Here, unless she is risk neutral, the decision maker chooses a level of risk *less* than Δ_m. Two differences in the models of the two chapters explain the difference in the choice of risk:

- In Chapter 12 we had not yet introduced risk aversion, so borrowers and lenders alike (and the owner-operator too) were risk neutral. This is why the owner-operator (with horizontal indifference curves) chose the expected income maximizing level of risk, $f = \frac{1}{2}$.
- In Chapter 12 we studied loan contracts in a legal setting (bankruptcy law and limited liability) such that the lender bore the entire risk of nonrepayment of the loan if the project failed. We explained that lending under these circumstances is equivalent to also providing insurance to the borrower. This fact—that the lender shared the risk with the borrower—is the second reason why borrowers took more risk than would have maximized expected income.

In the setting for this chapter—risk-averse actors who are not engaged in loan contracts—above Δ_m, additional risk-taking is a lose–lose proposition: it incurs more of the "bad" while reducing the "good."

M-NOTE 13.1 **Choosing a level of risk to maximize utility**

The decision maker selects a level of risk so as to maximize her expected utility, subject to the feasible set of risk and return, captured by the risk–return schedule:

$$\text{Vary } \Delta \text{ and } \hat{y} \text{ to maximize} \qquad u = u(\Delta, \hat{y}) \tag{13.5}$$

$$\text{subject to} \qquad \hat{y} = \hat{y}(\Delta) \tag{13.6}$$

Substituting Equation 13.6 into Equation 13.5, we have a maximization problem in one variable, Δ:

$$\text{Vary } \Delta \text{ to maximize} \qquad u = u\left(\Delta, \hat{y}(\Delta)\right)$$

To find the first-order condition for a maximum, we differentiate this equation and set the result equal to zero:

$$\frac{du}{d\Delta} = u_\Delta + u_{\hat{y}}\hat{y}_\Delta = 0$$

Which rearranged is:

$$\text{Slope of indifference curve} = \text{Slope of risk–return schedule}$$

$$-mrs(\Delta, \hat{y}) = \frac{-u_\Delta}{u_{\hat{y}}} = \hat{y}_\Delta = -mrt(\Delta, \hat{y}) \tag{13.7}$$

This is the condition stated in Equation 13.4.

M-NOTE 13.2 **Choosing risk: Numerical example**

In M-Note 13.1 we analyzed the general case. Now, we will give explicit functional forms to her expected utility and risk–return schedule, illustrate

continued

them with numerical values for the parameters and show how the utility-maximizing level of risk taking is determined.

Let the expected utility be $u(\hat{y}, \Delta) = \hat{y} - 0.5\Delta^2$. Let us assume that the risk–return schedule can be characterized as $\hat{y}(\Delta) = a\Delta - b\Delta^2$. Therefore, the maximization problem is:

$$\text{Vary } \hat{y} \text{ and } \Delta \text{ to maximize} \quad u = \hat{y} - 0.5\Delta^2 \tag{13.8}$$

$$\text{subject to} \quad \hat{y} = a\Delta - b\Delta^2 \tag{13.9}$$

Plugging Equation 13.9 into Equation 13.8, the problem becomes:

$$\text{Vary } \Delta \text{ to maximize} \quad u = (a\Delta - b\Delta^2) - 0.5\Delta^2$$

As before, we differentiate this equation with respect to the single variable Δ and set the result equal to zero to find the first-order condition:

$$\text{First-order condition:} \quad \frac{du}{d\Delta} = a - 2b\Delta - \Delta = 0$$

Which rearranged is:

$$\begin{aligned} \text{Slope of indifference curve} &= \text{Slope of risk–return schedule} \\ \underbrace{\Delta}_{-\text{ mrs}} &= \underbrace{(a - 2b\Delta)}_{-\text{ mrt}} \\ -\Delta(1+2b) &= -a \\ \Delta^{eq} &= \frac{a}{1+2b} \end{aligned} \tag{13.10}$$

Plugging Equation 13.10 into Equation 13.9, the expected income is:

$$\hat{y} = a\left(\frac{a}{1+2b}\right) - b\left(\frac{a}{1+2b}\right)^2$$

If we set $a = 200$ and $b = 2$ the risk is $\Delta_b = 40$ and the expected income $\hat{y}_b = 4{,}800$ where "b" (as in Chapter 3) indicates doing the best you can.

CHECKPOINT 13.5 Doing the best you can in a risky world

a. Use Equation 13.4 to explain with words why points **e** and **d** in Figure 13.12 panel (a) are not the best Arjun can do. (Hint: compare the slopes of the indifference curve and the risk–return schedule).

b. Use that reasoning (involving marginal rates of substitution and transformation) to explain why a risk-neutral person will, in general, choose higher levels of risk.

13.7 HOW RISK AVERSION CAN PERPETUATE ECONOMIC INEQUALITY

EXAMPLE Among the farmers in India described above, an irrigation pump is a specific asset (because it is difficult to sell once initially installed) while bullocks are more general assets (because there is a ready market for pre-owned or "second-hand" bullocks.)

Investments are risky because they involve a fundamental transformation. Before making an investment the decision maker has money that can be used to buy a broad array of goods or financial or other assets that can be easily sold for money. After the investment the decision maker owns specific assets—buildings and machinery or other assets such as patents or trademarks—dedicated to the production of particular goods and services. Specific assets are harder to sell than general assets. An investment is

therefore a gamble that the specific assets—and the goods or services that they can produce—will be a source of profit for their owner.

The future profitability and hence the value of these goods and services may change dramatically due to unforeseen future events and so the choice to invest is a type of decision-making under uncertainty. A person with income to invest typically has a choice to invest in more or less risky assets. Less risky financial assets include cash or US Treasury Bills (called "T-bills" for short) or UK government "gilts." These are promises to pay a given amount at a future date, an IOU ("I owe you") from the government. A person could also choose to hold moderately risky stocks in well-established firms or highly risky venture capital investments in unknown startup firms or government-issued bonds (similar to T-Bills) issued by unstable governments that might not honor the promise to pay the IOU.

✓ **FACT CHECK** The name gilt comes from "gilt-edge bond" a reference to the low or zero risk associated with them; the British government has never defaulted on a gilt.

None of these is truly without risk. Cash or the government's obligated payment to the owner of a T-bill or a gilt may change in its value—what it can buy—due to an increase or (less likely, decrease) in prices.

Arjun's and Nicolas's investment options are indicated by the risk–return schedule $\hat{y}(\Delta)$. We know that Arjun will find point **a** and choose Δ_a with return $\hat{y}_a$ (a for risk **a**verse). Nicolas would select point **m** with a level of risk Δ_m and expected income $\hat{y}_m$ (m, for the maximum of the risk–return schedule which he chooses as risk-neutral actor). Assume for each of them they have a certain amount of money c_1^A and c_1^N respectively, which are their fallback positions, or, for each of them, a fallback indifference curve of $u_z^A = u_1^A$ for Arjun and $u_z^N = u_1^N$ for Nicolas.

We see two things from this model given their investment choices:

- Arjun did well by investing, increasing his certainty equivalent from c_1^A to the certainty equivalent income of his risky investment, namely c_2^A (which corresponds to what he receives on indifference curve u_2^A through point **a**).
- Wealthy and risk-neutral Nicolas does better than Arjun, with an expected income of $\hat{y}_m$ ($\hat{y}_m > \hat{y}_a$). Because he is risk neutral and his indifference curves are horizontal, $\hat{y}_m$ is also his certainty equivalent, $c_2^N = \hat{y}_m$ ($c_2^N > c_2^A$). Because he is risk neutral he is indifferent between taking the risk Δ_m or holding an asset with no risk at all, as long as the expected income of his risky investment is the same as the certain income of the asset.
- Lacking wealth, Arjun had a totally different view of the options: had he made Nicolas's very risky investment (Δ_m) this would have made him even *worse* off than at his fallback option (point **m** lies *below* Arjun's fallback indifference curve u_1^A).

What this means is that over many investment decisions made by the two, or for a population made up of many Arjuns and many Nicolases, the average return to investment by Arjuns (or the class of Arjuns) will fall short of the average return to Nicolases (or his class of wealthy Nicolases). This is one

of the ways that income differences and the societal inequalities associated with them are self-perpetuating:

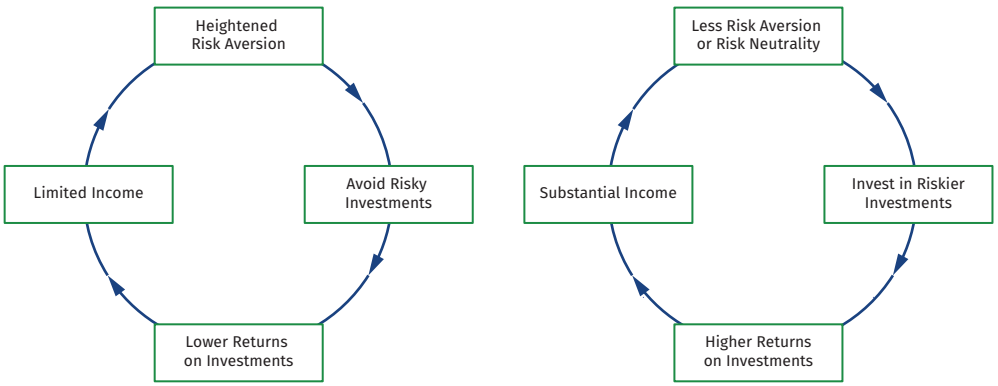
- *vicious circle*: A *vicious* circle of low income, risk aversion, avoiding risky investments, and low expected income; or
- *virtuous circle*: A *virtuous* circle of substantial income, risk neutrality (or close to risk neutrality) investing in risky assets, and substantial expected returns.

Figure 13.13 illustrates these two circles.

The different tales of these two circles need not have anything to do with Nicolas's or Arjun's basic psychology. We illustrated the case by the contrasting two indifference maps. But they could have had the same utility function, so that if Arjun were as rich as Nicolas his indifference curves would be horizontal too. Or if Nicolas had been as poor as Arjun, like Nicolas he would have been risk averse. How they differed could have been only their *initial* income. Had Arjun been the one with high income and Nicolas the one with low income, Arjun's circle would have been virtuous and Nicolas's vicious.

CHECKPOINT 13.6 **Wealth and risk aversion** Why would unexpectedly inheriting a large sum make a person less risk averse (locate illustrative before and after points in Figure 13.7)?

Figure 13.13 **How risk aversion can perpetuate economic inequality: vicious and virtuous circles.** Start with the left box "Limited income" in panel (a). The figure depicts a *vicious* circle (or cycle) where limited income leads to heightened risk aversion, which results in avoiding risky investments, and means lower expected returns on average and lower levels of income. Panel (b) depicts a *virtuous* circle where substantial income leads to lower risk aversion or risk neutrality, which means investment in riskier assets, and higher expected income on average.



(a) Vicious circle (b) Virtuous circle

13.8 HOW INSURANCE CAN MITIGATE RISK AND REDUCE INEQUALITY

What can be done to break or mitigate these cycles, and to allow more people to benefit from undertaking risky investments? One answer is: **insurance**.

Insurance can take many forms other than the familiar car and house insurance. Learning how to code in a widely used programming language–like Python or R–means that you will have job opportunities in many sectors of the economy should your current job end. Acquiring this or some other general skill is a form of *insurance* against risky outcomes. In many countries learning English is also a form of insurance as it expands the range of jobs to which one can apply and even countries in which one could seek employment.

Insurance reduces risk exposure

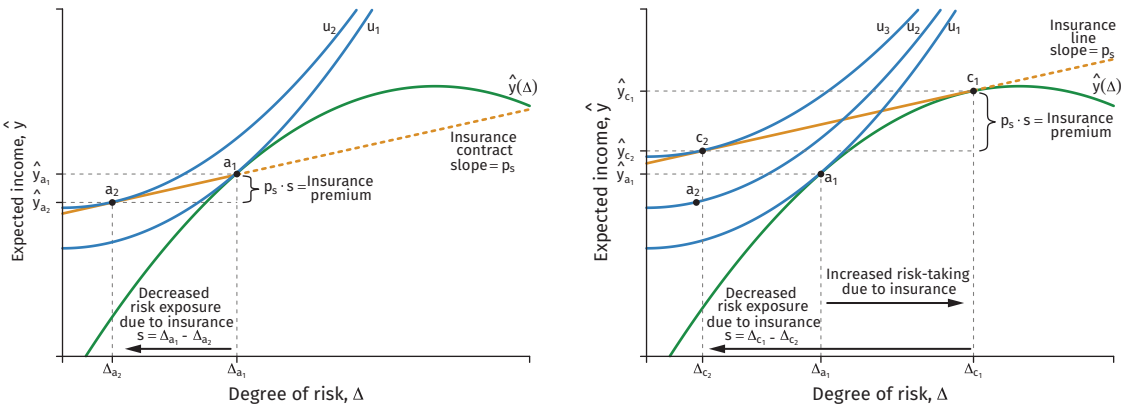
We will now see that if insurance is available, risk-averse people will purchase it and as a result be willing to take more risks and to benefit from the higher expected returns associated with riskier investments. The reason is that for any given investment project or other decision, insurance reduces the difference between the good and the bad outcomes that the person will experience. Insurance makes the bad outcome not as bad because the person who purchases insurance is compensated for the realization of the bad outcome. Insurance makes the good outcome less good, also: whichever outcome is realized, the payment of the cost of insurance means that there will be less income left over for other expenditures by the insured.

To see why this is so, consider a person, Juliana, who is making a decision involving risk. In the absence of insurance, as Arjun did in the previous example, Juliana maximized her utility by choosing risk level Δ_{a_1} (point $\mathbf{a_1}$ is the utility-maximizing point for risk-averse Juliana) with corresponding expected income $\hat{y}_{a_1}$. This is shown in Figure 13.14 (a).

Now introduce an insurance contract, which allows Juliana to "buy" less risk, by paying an amount–called the insurance premium–in reduced expected income in exchange for a reduction in her degree of risk, Δ. (What is reduced is her income left over for other purchases after paying the insurance premium; for simplicity we call this a reduction in expected income.) If she has chosen point $\mathbf{a_1}$, then the opportunity to purchase insurance is shown by what is called the *insurance contract line* through that point, the orange upward-sloping line in the figure.

From point $\mathbf{a_1}$ she can move to any point on that line. She is interested in reducing risk, so she will consider moving to the *left* on the insurance contract line (which we will call the insurance line, for short). This is also why the orange line to right of point $\mathbf{a_1}$ is dashed. If she chose a point on that portion of the line she would be *selling* insurance, not buying it, that is, taking on extra risk in return for a higher expected income.

Figure 13.14 Effect of insurance on risk-taking and utility: two tangencies. When insurance is *unavailable* (shown by point $\mathbf{a_1}$) the person takes a limited amount of risk and as a result can expect a limited amount of income. The availability of insurance is indicated by the orange "insurance contract" line: an amount of insurance s (meaning a reduction in the risk) can be purchased by paying the amount $p_s \cdot s$. The slope of the insurance contract line is p_s so a steeper line means *more costly* insurance. If she were to choose an investment with the risk level indicated by point $\mathbf{a_1}$, as she did without insurance, then she could purchase sufficient insurance to be at point $\mathbf{a_2}$. But her utility maximum is achieved by choosing the risk level Δ_{c_1} (point $\mathbf{c_1}$) and then purchasing insurance to be at point $\mathbf{c_2}$ with risk exposure Δ_{c_2}. The first tangency is point $\mathbf{c_1}$ and the second tangency is $\mathbf{c_2}$.



(a) Insurance and risk taking that does not maximize utility

(b) Utility-maximizing risk taking and insurance

To see how this could work let $\mathbf{a_2}$ be some point on the insurance line to the left of $\mathbf{a_1}$. Then instead of her current risk exposure and expected income, by purchasing an amount of insurance s at a price per unit of risk reduction p_s she can have:

$$\text{Less risk} \qquad \Delta_{a_2} = \Delta_{a_1} - s$$
$$\text{and less expected income} \qquad \hat{y}_{a_2} = \hat{y}_{a_1} - p_s s$$

We can rearrange these equations to find the slope of the insurance line:

$$\text{Extent of reduction in } \Delta \qquad \Delta_{a_1} - \Delta_{a_2} = s$$
$$\text{Extent of reduction in } \hat{y} \qquad \hat{y}_{a_1} - \hat{y}_{a_2} = p_s s$$

From which we see that the slope of the insurance line is:

$$\text{Slope of insurance line} \quad \underbrace{\frac{\overbrace{\hat{y}_{a_1} - \hat{y}_{a_2}}^{\text{Extent of reduction in } \hat{y}}}{\Delta_{a_1} - \Delta_{a_2}}}_{\text{Extent of reduction in } \Delta} = \frac{p_s s}{s} = p_s$$

The negative of the slope of the insurance line is the marginal rate of transformation of reduced expected income into reduced risk by purchasing insurance.

How will insurance affect her choices? You can see from Figure 13.14 (a) that at her previous choice of a risk level Δ_{a_1} Juliana would be better off buying insurance. This is because at point $\mathbf{a_1}$:

$$\text{Slope of indifference curve} = \frac{-u_\Delta}{u_{\hat{y}}} > p_s = \text{Slope of insurance line}$$

$$mrs > mrt$$

Remembering that $-u_\Delta$ is the marginal disutility of risk which is the same thing as how much she would benefit from risk reduction, we can rearrange the above to be:

$$-u_\Delta > u_{\hat{y}} p_s$$

$$\text{Marginal benefits of risk reduction} > \text{Marginal costs of risk reduction}$$

As before she could choose an investment with risk and expected income of $(\Delta_{a_1}, \hat{y}_{a_1})$, but then also buy insurance. As a result, after taking account of both the insurance premium and the risk mitigation afforded by the insurance, she would reduce both her expected income and the risk exposure that she experiences.

As a result, Juliana moves down and to the left along the insurance line, mitigating the risk of the investment she chose, and reducing her expected income by an amount equal to her insurance premium. How much insurance will she purchase? She should buy an amount such that the marginal benefit of further risk reduction is neither greater than nor less than the marginal cost. In other words she should buy the amount of insurance such that

$$\text{Slope of indifference curve} = \frac{-u_\Delta}{u_{\hat{y}}} = p_s = \text{Slope of insurance line}$$

$$mrs = mrt$$

From this rule, she will find that the bundle $(\Delta_{a_2}, \hat{y}_{a_2})$ is the best she can do if she chooses the investment at point $\mathbf{a_2}$, bringing her to indifference curve u_2.

! REMINDER Recall the price lines in Chapters 3 and 6; this line is similar except only risk has a "price" on the market (p_s). Alternatively, we can think of the price of insurance being the amount of the other good (expected income) that the decision maker has to give up in order to reduce risk exposure, as we compare points along the price line—the same understanding we already have of a marginal rate of transformation.

M-NOTE 13.3 **Choosing insurance given some initial** $(\Delta_{a_1}, \hat{y}_{a_1}(\Delta_{a_1}))$

Suppose the bundle $(\Delta_{a_1}, \hat{y}_{a_1}(\Delta_{a_1}))$ corresponds to a given level of risk and resulting expected income that Juliana has chosen (such as point $\mathbf{a_1}$ in Figure 13.14), with utility $u_1 = u(\Delta_{a_1}, \hat{y}_{a_1})$.

When she is given the opportunity to buy insurance, Juliana will choose the level of insurance s that will maximize her utility. Choosing this level s will reduce both her risk (by the amount of insurance she bought, s) and her expected income (by the cost per unit of risk reduction p_s multiplied by the amount of risk reduction s). We index her situation after purchasing the insurance by the subscript 2:

$$\text{Experienced risk with insurance} \quad \Delta_{a_2} = \Delta_{a_1} - s \tag{13.11}$$

continued

$$\text{Expected income after paying for insurance} \quad \hat{y}_{a_2} = \hat{y}_{a_1}(\Delta) - p_s s \quad (13.12)$$

Thus, Juliana's maximization problem is as follows:

$$\text{Vary } s \text{ to maximize} \qquad u = u(\Delta_{a_1} - s, \quad \hat{y}_{a_1} - p_s s)$$

For a given Δ_{a_1} and $\hat{y}_{a_1}$, we find her first-order condition for a maximum by differentiating her utility with respect to the single variable s and setting the result equal to zero:

$$\frac{du}{ds} = u_\Delta(-1) - u_{\hat{y}} p_s = 0$$

which, rearranged, means that

$$\text{slope of indifference curve} = \text{slope of insurance line}$$

$$-mrs = -mrt$$

$$\frac{-u_\Delta}{u_y} = p_s$$

The slope of the indifference curve tells us the income she is willing to give up to reduce risk. The slope of the insurance line is the income she has to give up to reduce risk. Because at point $\mathbf{a_1}$ the indifference curve is steeper than the insurance line, she can increase her utility by buying more insurance.

CHECKPOINT 13.7 **Buying risk** Use panel (a) of Figure 13.14 to explain why Juliana would not prefer some outcome to the right of point $\mathbf{a_1}$ along the dashed portion of the insurance line, rather than to the left on the solid portion of the line.

Insurance encourages risk-taking

But given the introduction of the insurance option, she could do even better than point $\mathbf{a_2}$ if she reconsidered her initial choice of Δ_{a_1} jointly buying insurance and choosing more risk than Δ_{a_1}.

How much additional risk should she take, and how much insurance should she buy? To answer this question it is important that the insurance line is not unique to point $\mathbf{a_1}$: from any point that Juliana chooses on the risk–return schedule, this line shows her opportunities to move to less risky states by buying insurance.

! REMINDER This two-step optimization process is exactly what we did in Chapter 6 (Section 6.5) to model the choice of specialization followed by trade.

So think about the insurance line as something she can drag along the risk–return schedule, never changing its slope (which is the price of insurance p_s), and including as one of the points on the line the point indicating her investment choice. The insurance line will be the constraint facing her when, *after* making her investment decision, she considers *reducing* her risk level.

Because it is a constraint on her choices, and because she prefers choices that are higher up and to the left, she wants to drag the line to a point at which it is as high and to the left as possible. Her decision to make an even riskier choice and buy more insurance is shown in Figure 13.14 (b). The answer is given in two steps:

- *Step* 1: point $\mathbf{c_1}$ is her choice of a risk level and expected income (*before* paying for insurance) that results from that choice; and

- *Step 2*: point $\mathbf{c_2}$ is where she will be as the result of her decision to take the risk Δ_{c_1} and to purchase an amount s of insurance at a price p_s for a total insurance cost of $p_s \cdot s$ and move to Δ_{c_2} after purchasing insurance (and therefore increasing her utility from u_1 at the no-insurance outcome to u_3 taking additional risk and buying insurance).

In this two-step process, "doing the best she can" requires finding two tangencies, not just one as in our usual case so far.

- *First tangency*: To determine her *choice of risk* (point $\mathbf{c_1}$) she equates the marginal rate of transformation of risk taken into expected income (the slope of the risk–return schedule) to the marginal rate of substitution of premium paid (reduced expected income after paying for insurance) and reduced risk exposure. That is, she finds where the insurance contract line is tangent to the risk–return schedule.
- *Second tangency*: To determine *how much insurance to buy* (point $\mathbf{c_2}$) she equates the marginal rate of substitution between risk and expected income (the slope of the indifference curve) to the marginal rate of transformation of insurance premium paid into reduced risk exposure. That is, she finds where her indifference curve is tangent to the insurance contract line.

M-CHECK When she is choosing a risk level she wants to be on the insurance contract line that is highest and to the left. Representing her objectives, the negative of its slope ($-p_i$) is the marginal rate of substitution. When, having chosen a risk level, she is then deciding how much insurance to buy, the insurance contract line is her constraint. So the negative of its slope (same $-p_i$) is the marginal rate of transformation.

The availability of insurance has two effects: it allows Juliana to reduce her total risk exposure, and because of that it also encourages Juliana to invest in a riskier option, raising her expected income from $\hat{y}_{a_1}$ to $\hat{y}_{c_2}$.

If insurance were more expensive—higher p_s—the insurance contract line would be steeper and she would choose a lower risk level and purchase less insurance. If the price of insurance were so high that the insurance contract line was as steep as the risk–return schedule at the point that she chose, then she would buy *no* insurance at all.

This case illustrates a broader point about the economy as a whole, which we return to later in this chapter: when insurance is available, people are able to take more risks and enjoy greater expected income. This win–win outcome is possible because Juliana made a risky and high expected income choice, but she could also transfer some part s of her resulting risk exposure to the insurer.

But in order for Juliana to 'sell' her risk, there must be somebody willing to buy it.

✓ **FACT CHECK** Using a seat belt while driving in a car is another form of insurance: it reduces the difference between consequences of the good and bad outcome (no crash, crash). Consistent with the main lesson of this section, there is some evidence (see Chapter 16) that drivers wearing seat belts drive faster, insurance supporting higher levels of risk taking, just as the model predicts.

M-NOTE 13.4 **Jointly choosing the level of risk and of insurance**

Here we show how the decision maker jointly chooses the type of investment and the associated risk level given the risk–return schedule and the price of insurance, and the level of insurance to purchase, given the investment chosen and the price of insurance. Using notation similar to M-Note 13.3, subscripts 1 and 2 refer to before and after buying insurance, respectively.

We can define the utility that she maximizes as a function of two variables that she chooses: the risk before buying insurance (Δ_{c_1}) and the amount of insurance (s):

continued

Vary Δ and s to maximize $\quad u = u(\Delta_{c_2}, \hat{y}_{c_2})$

which, using Equations 13.11 and 13.12, is $\quad u = u(\Delta_{c_1} - s, \hat{y}(\Delta_{c_1}) - p_s s)$

As utility is now a function of the two variables Δ and s, to find the solution to this constrained maximization problem we partially differentiate the utility function with respect to Δ and s, and set the results equal to zero:

First-order condition #1 $\quad \dfrac{\partial u}{\partial \Delta} = u_\Delta + u_{\hat{y}} \hat{y}_\Delta = 0$

Rearranging $\quad \dfrac{-u_\Delta}{u_{\hat{y}}} = \hat{y}_\Delta \qquad (13.13)$

First-order condition #2 $\quad \dfrac{\partial u}{\partial s} = -u_\Delta - u_{\hat{y}} p_s = 0$

Rearranging $\quad \dfrac{-u_\Delta}{u_{\hat{y}}} = p_s \qquad (13.14)$

Notice that the left-hand side of the two conditions is identical. Thus, combining Equations 13.13 and 13.14 we have:

$$\hat{y}_\Delta = p_s = \frac{-u_\Delta}{u_{\hat{y}}}$$

where

- $\frac{d\hat{y}}{d\Delta} = p_s$ is the slope of the risk–return schedule equal to the slope of the insurance line and
- $p_s = \frac{-u_\Delta}{u_{\hat{y}}}$ is the slope of the insurance line equal to the slope of the indifference curve.

This tells us, referring to Figure 13.14, that Juliana will choose:

- *First tangency, point* c_1: the risk level Δ such that the marginal rate of transformation of risk taken into expected income equals the marginal rate of transformation of premium paid into reduced risk exposure.
- *Second tangency, point* c_2: the amount of insurance s such that the marginal rate of substitution between risk and expected income is equal to marginal rate of transformation of reduced expected income into reduced risk.

CHECKPOINT 13.8 **Insurance and risk–taking** In Figure 13.14 explain why in the absence of insurance the best Juliana can do is indicated by point $\mathbf{a}_1$ in panel (a); but once insurance is available, she does better by taking more risk.

13.9 BUYING AND SELLING RISK: TWO SIDES OF AN INSURANCE MARKET

Who is buying the risk that Juliana is selling? Could the buyer be someone like her? Would that person be willing to accept greater risk exposure for herself in return for an increase in her expected income? You can see from Figure 13.14 that at the price p_s shown as the slope of the insurance contract line, someone in Juliana's situation would have no interest in selling insurance. The reason is that selling insurance would mean moving

to the right *up* the insurance contract line and reaching even *lower* indifference curves.

If not Juliana, then who?

Enter: Weikai (WAKE-eye). He is exposed to zero risk and is wealthy (maybe the only form of his substantial wealth is US T-bills). He is therefore not very risk averse; he is close to risk neutral. Juliana is poor and exposed to risk, with a good and a bad state affecting her realized income and occurring with equal probability. As a result of both limited income and a high level of risk exposure Juliana is highly risk averse.

Insurance: Buying and selling risk

The two may engage in an exchange to alter the distribution of expected income and risk between them. We can represent their interaction using the Edgeworth box that you encountered in Chapter 4. But there is a difference: in the Edgeworth boxes you previously studied, what was being allocated was two goods—coffee and data. In the case we now consider two people who are exchanging a "good"—expected income—and a "bad"—risk. In Figure 13.15 we present the Edgeworth box with Juliana's and Weikai's indifference curves.

The dimensions of the box are the total amount of expected income that the two will jointly experience ($\bar{\hat{y}} = \hat{y}^J + \hat{y}^W$) and the total amount of risk to which the two will be exposed ($\bar{\Delta} = \Delta^J + \Delta^W$ which is $\bar{\Delta} = \Delta_z^J$ at Juliana's endowment (at point **z**) because prior to their exchange she bears all the risk and Weikai bears none).

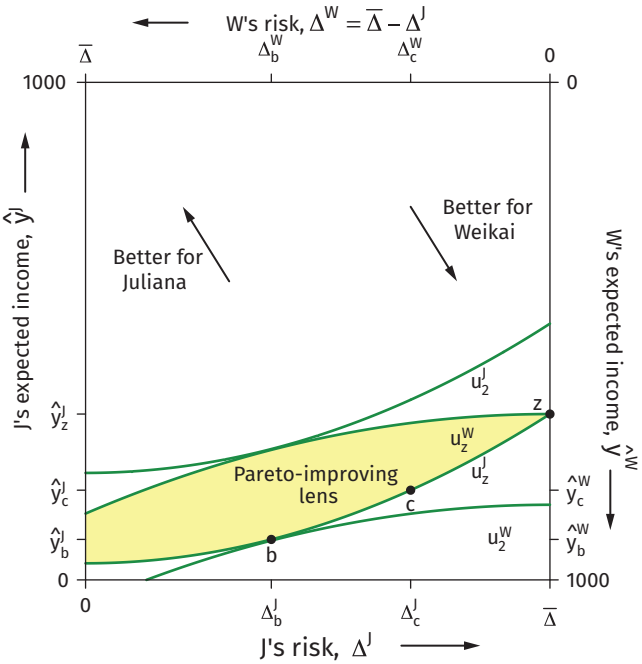
In the initial state with no exchange between the two (point **z**) the (expected and realized) income of Weikai is $\hat{y}_z^W = y_z^W$ (he faces no risk, so his expected income and his realized income are equal). Juliana's expected income is $\hat{y}_z^J$ and her realized income is $y^J = y_z^J + 0.5\Delta_z$ in the good state and $y^J = y_z^J - 0.5\Delta_z$ in the bad state. Recall, that the good and bad states are equally likely, so her expected income is y_z^J.

Because as before risk is a bad and expected income is a good, the indifference curves slope upward. For example Juliana is indifferent between two possible allocations indicated by point **b**, namely, exposed to a lower level of risk (Δ_b^J) along with a lower level of expected income ($\hat{y}_b^J$) and point **c**, being exposed to more risk (Δ_c^J) and a higher level of expected income ($\hat{y}_c^J$).

The slope of each of her indifference curves is as before a measure of her degree of risk aversion. The steeper the indifference curve, the more expected income she is willing to give up to reduce the amount of risk to which she is exposed. She would prefer any point above and to the left of point **z**—less risk and more expected income—but that would make Weikai worse off (less utility than his participation constraint u_z^W), so she will not have that opportunity.

! REMINDER Remember from Chapters 4 and 5 that an allocation is Pareto efficient (and therefore will be a point on the Pareto-efficient curve) if at that point the two participants' marginal rates of substitution are equal, meaning that their indifference curves are tangent. We call the tangency condition the $mrs^A = mrs^B$ rule where A and B are two people engaged in some interaction.

Figure 13.15 Feasible allocations of expected income and risk and indifference curves over these allocations. Every point in the Edgeworth box is a possible allocation that divides the total amount of expected income ($\bar{\hat{y}}$) and risk ($\bar{\Delta}$) between the two people. Weikai (blue indifference curves) prefers allocations lower and to the right; Juliana (green indifference curves) prefers allocations higher and to the left. So there are conflicts of interest in comparing higher points on the left with lower points on the right. But there are some mutually beneficial reallocations, starting at point **z** and comparing points down and to the left as indicated by the shaded Pareto-improving lens.



! REMINDER To determine the price he would set, we could use the models of first-mover advantage (price-setting) and "take-it-or-leave-it power" studied in Chapter 4 or the Cournot model of competition among those wishing to sell insurance to Juliana studied in Chapter 9.

A Pareto-improving insurance contract

There are, however, opportunities for a mutually beneficial bargain. In Figure 13.15 we can see that the endowment point **z** is not Pareto efficient because at that point the indifference curves of the two are not tangent, they intersect. Therefore, there are other allocations of risk and expected income that both would prefer to point **z**. The yellow-shaded area indicates all of these win-win allocations. In them Weikai takes over some of the risk exposure in return for Juliana transferring to him some of her expected income. In other words Weikai sells Juliana some insurance.

The agreement between the two to move to an allocation in the shaded area would take the form of an insurance contract: Juliana would give up some of her income and Weikai would take on some of her risk exposure. Figure 13.16 provides the game tree illustrating the interaction.

The game begins with Weikai offering a price for an amount of risk reduction that Juliana wishes to purchase ($p_s(\bar{\Delta} - \Delta^J)$) (we do not analyze why he offers this particular price). Prior to the good or bad state having been realized, Juliana either rejects the offer (she takes the right branch of the tree) or accepts (she takes the left branch).

In the latter case she buys her chosen amount of insurance $\bar{\Delta} - \Delta^J$, where Δ^J is the risk exposure that Juliana will be subjected to under this contract, paying a total of $p_s(\bar{\Delta} - \Delta^J)$. Because the total amount of risk exposure of the two is $\bar{\Delta}$, it follows that Weikai's risk exposure at the post-exchange allocation is $\Delta^W = \bar{\Delta} - \Delta^J$, meaning that Δ^W is the amount of insurance that Juliana receives from Weikai–it is risk to which she initially was (but is no longer) exposed. That part of her initial risk is now his.

After the state is revealed, Juliana pays $0.5(\bar{\Delta} - \Delta^J)$ to Weikai if the good state has occurred while Weikai pays $0.5(\bar{\Delta} - \Delta^J)$ to Juliana if instead the bad state has occurred.

Table 13.1 summarizes the realized income in the good and bad states for the two people when the insurance contract is implemented. Table 13.1 gives the expected income and risk exposure in the initial situation and the situation following the implementation of the insurance contract. Figure 13.17 captures the interaction. The the orange insurance contract line illustrates Weikai's price offer. Juliana is constrained to a point somewhere along the orange insurance contract line, including point **z**, meaning reject

Figure 13.16 A game tree explaining the sequence of the insurance contract.
Weikai offers a price for insurance, p_s. Juliana can accept or reject that price of insurance. In either case of purchasing the insurance or not, she will face a good or bad state. The expressions at the bottom nodes of the tree are the realized outcomes in the bad and good states for Weikai (top row in red) and Juliana (bottom row in blue).

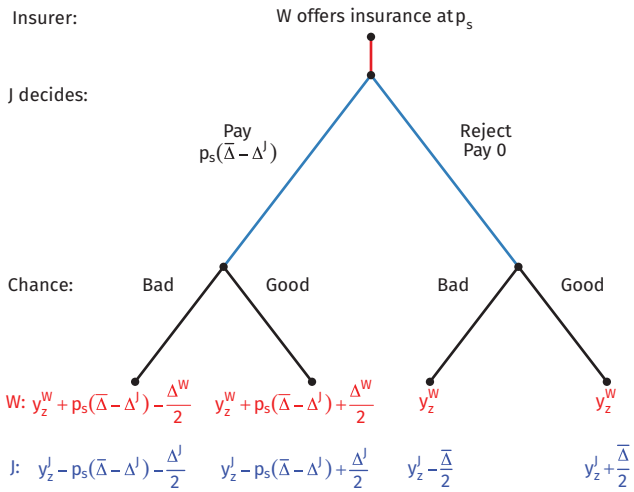


Table 13.1 Risk exposure and expected income in the initial state and following implementation of the insurance contract.

Person	Before insurance: risk and expected income	After insurance: risk and expected income
Weikai: rich not risk exposed, less risk averse	Risk: $\Delta_z^W = 0$	$\Delta_s^W = \overline{\Delta} - \Delta_s^J$
	Expected income: y_z^W	$\hat{y}_z^W + p_s\Delta_s^W$
Juliana: poor risk exposed, more risk averse	Risk: $\Delta_z^J = \overline{\Delta}$	Δ_s^J
	Expected income: $\overline{y}_z^J$	$\hat{y}_z^J - p_s\Delta_s^W$

the offer. The point **s** is on the highest indifference curve that is available to her at this price. Her indifference curve at point **s** is tangent to the insurance line, meaning that her marginal rate of substitution is equal to the marginal rate of transformation (the slope of the insurance line, or p_s). She picks point **s**, pays $p_s(\overline{\Delta} - \Delta_s^J)$ and receives $\Delta_s^W = \overline{\Delta} - \Delta_s^J$ insurance.

Allocation **s** represents a Pareto improvement over initial endowment **z**. But, while both Juliana and Weikai are better of as a result off exchange, the exchange has increased inequality of expected income. On the other hand, the exchange reduced inequality between the two in risk exposure.

CHECKPOINT 13.9 Buying and selling risk In Figure 13.15 explain why at the initial point (**z**) Weikai would be better off taking on some of Juliana's risk in return for a payment from her, and she would be better off paying him to do that.

HISTORY Bill Gates, who is said to have risked it all by dropping out of Harvard to found Microsoft, explained that he actually did not quit Harvard but took a leave of absence so that he could return to Harvard if necessary. His family also had enough money so it would not be a problem if Microsoft had failed. Think how much innovation we might get if everyone had such a good fallback position.

13.10 APPLICATION: FREE TUITION WITH AN INCOME-CONTINGENT TAX ON GRADUATES

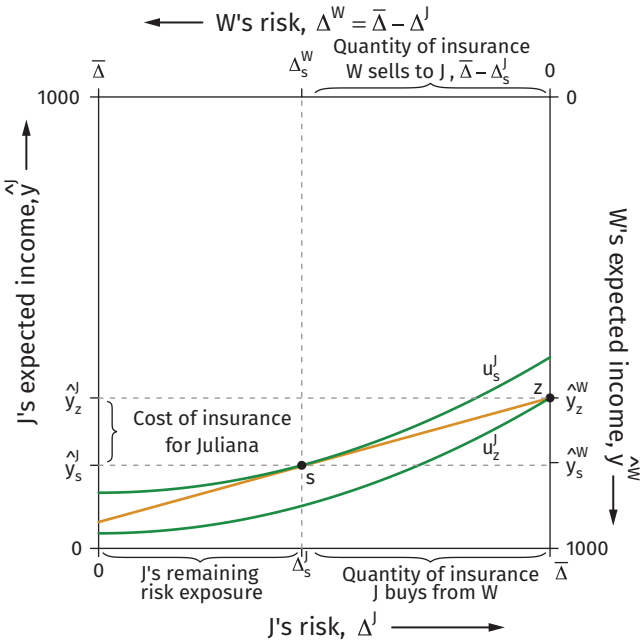
The term "risky investment" is associated with headline-grabbing disasters like the insurance giant AIG's (American International Group) business model prior to its crash in 2008.[10] But one of the riskiest investments of all is investing in yourself.

Yourself: A risky investment

The decision to attend a particular university and to study a particular subject is made while uncertain about:

- Do I have the talent and discipline to do well studying this subject?
- Will I find it interesting enough to pursue a career in a related area?
- Will instructors at the university I attend teach me well, and certify that I am competent in my field?
- Will there be a demand for the skills (and credentials of them) I will acquire as a result of my decisions?

Figure 13.17 Buying and selling risk. Weikai has price-setting power, sets a price equal to the slope of the line p_s. Juliana, who may choose any point on that line, picks **s**, which is on the most preferred of her indifference curves that is feasible: it is the highest indifference curve (u_2) she can obtain given the price line which constrains her given the price Weikai stipulates (p_s). Juliana has given up some expected income to purchase insurance in the form of reduced risk exposure.



Then there is the cost, both directly to pay tuition and accommodation at the university and also the opportunity cost of being a student, that is, forgoing the wage or other income that one would have gained if you had not attended university. Most people are not rich enough to be even close to risk neutral, and as we saw in Chapter 12, a great many families, even in high-income countries, are unable to borrow the substantial sums required to pay for higher education privately. So unless higher education were provided or subsidized by government or private philanthropic contributions few people would attend university.

To understand why, consider Sofia, who has the opportunity to attend university without cost. Sofia has completed a two-year degree as a medical technician and is considering two options:

- *no risk, certain job, and salary*: taking a job with a certain salary, which would be supplemented by the annual interest on a modest financial asset that she has, giving her a total income of $\hat{y}_1 = y_a$ (**a** for the certain

option available to her now and it does not have a hat because it is certain); or

- *continue education, bear risk, higher expected income*: continuing her education and as a result increasing her expected income to $\hat{y}_e$ while also being exposed to risk, the amount of which is Δ, the difference between her income in the good and the bad state.

If she chooses to get more education, the good state might be described by a positive answer to the four bulleted questions about yourself as a risky investment. The bad state could be a negative answer to the questions. As before, we let the good and the bad state be equally likely to occur. We also assume that the realized income in the bad state is less than the certain income she would have were she to decide not to continue her education (otherwise she might just ignore the risk, as she could be certain to have a higher income—even in the bad state—by going to university).

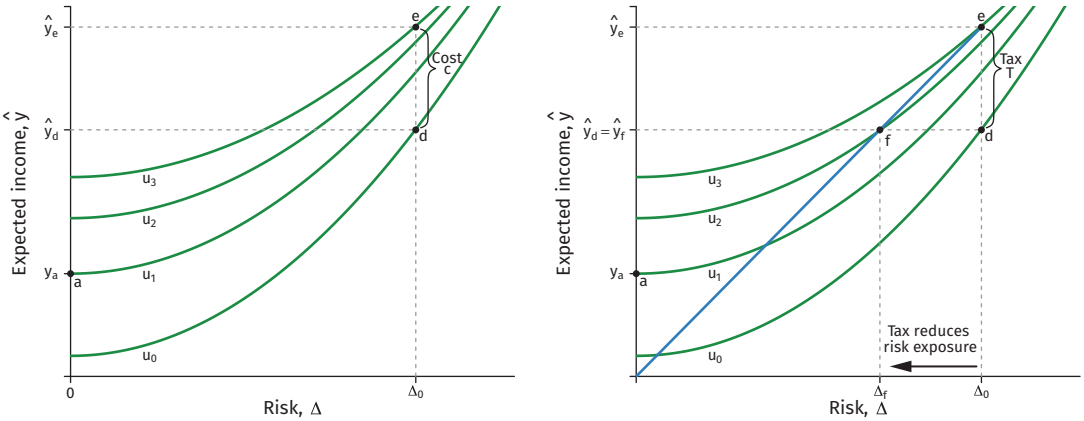
In Figure 13.18, point **e** shows her expected income and risk exposure if she continues her education when education is free (paid for by the government from general tax revenues), point **d** corresponds to when she pays for her education privately. If instead she chooses *not* to continue her education, she would receive a certain income of $y_1 = y_a$ and be exposed to no risk, indicated by point **a** in the figure. Her choices are mutually exclusive. If she chooses to continue her education, then point **a**—the job that she has currently been offered—is no longer available to her. (Her options as just described are summarized in Table 13.2.)

How would Sofia compare the outcomes available to her? This is where the indifference curves come in. If Sofia were wealthy, she might not be concerned about risk exposure—like Bill Gates deciding to leave Harvard—but having a modest income and being limited in how much she can borrow, she is risk averse, as the upward slope of her indifference curves indicate.

Her indifference curve u_1 through point **a** gives the combinations of risk exposure and expected income that Sofia prefers equally to the certain income of y_a. Therefore, y_a is the certainty equivalent of the combination of expected income and risk given by every point on the indifference curve labeled u_1. Comparing points **a** and **e** and the indifference curves on which they appear, we can see that if tuition is free she will choose to undertake further education ($u_3 > u_1$).

But other citizens might object: it was they who paid the taxes that allowed the government to provide Sofia's education for free. Why should taxpayers subsidize Sofia's investment in herself? As a result of higher education, Sofia would have a higher expected income, but taxpayers who had not attended university would have a lower income (having subsidized Sofia's education). It does not appear to be fair to those not as fortunate as Sofia who has already attended 14 years of schooling.

Figure 13.18 Effect of income-contingent taxation of graduates on choices concerning further education. In panel (a) we show Sofia's indifference curves and her possible choices: taking a zero-risk job now (no further education) at income $\hat{y}_1 = y_a$, point **a** or undertaking additional education along with additional risk exposure, shown if she pays no costs by point **e**, and by point **d** if she has to pay the cost of her education. From the indifference curves you can see that she prefers taking the job now (no further education) if she has to pay the cost, that is, $u_1 > u_0$. But she would prefer to continue with her education if it were free because $u_3 > u_1$. In panel (b), we show the case in which tuition is free, but following graduation she will pay a tax proportional to her income. Under the income-contingent tax her expected income and risk exposure are reduced in the same proportion (along the blue line through point **f**). So instead of being at the simple free tuition (no tax) outcome **e**, she is now at point **f** on indifference curve u_2. Because $u_2 > u_1$ she will continue her education.



Let us therefore consider the case in which Sofia pays the cost of her education. As a result her expected income is now: expected income with free tuition minus the costs of her education. The difference between her incomes in the good and bad state is not affected–her realized income in the two states is just reduced by the private cost of her education, as shown by point **d**. So her exposure to risk is unaffected: it is still $\Delta_e = \Delta_0 = \Delta_d$. Point **d**, however, lies on indifference curve u_0. If she has to pay the cost of her tuition she would be better off not pursuing further education because u_0, her utility if she pays for her education, is less than u_1, her level of utility when she takes the risk-free job.

Table 13.2 Comparison of the free, privately paid, and income-contingent tax policies for education. Point **a** is not shown in the table as the risk (Δ_a) is zero and the only state is the certain income y_a. We use the y-subscript T to indicate the results under the income-contingent tax when she has continued her education. ^{G} and ^{B} refer to the good and bad states respectively. The tax rate is τ.

Outcome (point in Figure 13.18)	Free tuition (e)	Private cost (d)	Free tuition and income-contingent tax on graduates (f)
Good outcome	$y_e^G = \hat{y}_e + 0.5\Delta_0$	$y_d^G = \hat{y}_e + 0.5\Delta_0 - c$ $= \hat{y}_d + 0.5\Delta_0$	$y_T^G = (\hat{y}_e + 0.5\Delta_0)(1-\tau)$
Bad outcome	$y_e^B = \hat{y}_e - 0.5\Delta_0$	$y_d^B = \hat{y}_e - 0.5\Delta_0 - c$ $= \hat{y}_d - 0.5\Delta_0$	$y_T^B = (\hat{y}_e - 0.5\Delta_0)(1-\tau)$
Expected income	$\hat{y}_e = \frac{y_e^G + y_e^B}{2}$ $= \frac{\hat{y}_e + 0.5\Delta_0 + \hat{y}_e - 0.5\Delta_0}{2}$	$\hat{y}_d = \frac{y_d^G + y_d^B}{2}$ $= \frac{\hat{y}_d + 0.5\Delta_0 + \hat{y}_d - 0.5\Delta_0}{2}$	$\hat{y}_f = \frac{y_T^G - y_T^B}{2}$ $= \frac{(\hat{y}_e + 0.5\Delta_0)(1-\tau) + (\hat{y}_e - 0.5\Delta_0)(1-\tau)}{2}$ $= \hat{y}_e(1-\tau)$
Experienced risk exposure	$y_e^G - y_e^B = \hat{y}_e + 0.5\Delta_0$ $- (\hat{y}_e - 0.5\Delta_0)$ $\Delta_0 = \Delta_e$	$y_d^G - y_d^B = \hat{y}_d + 0.5\Delta_0$ $- (\hat{y}_e - 0.5\Delta_0)$ $\Delta_0 = \Delta_d$	$y_f^G - y_f^B = \hat{y}_e + 0.5\Delta_0(1-\tau)$ $- (\hat{y}_e - 0.5\Delta_0)(1-\tau)$ $= \Delta_0(1-\tau)$ $= \Delta_f$

EXAMPLE In many countries students can finance the expenses of higher education by borrowing money with the unusual feature that the extent of repayment will be less the lower is the income of the graduate, so that only those with substantial incomes fully repay the cost of education. The shortfall due to less than full repayment by lower-income graduates is made up by tax revenues.

Income-contingent taxation of graduates

An alternative way of funding higher education that would reduce the risk exposure of those pursuing further studies, while at the same time addressing the unfairness of general tax payers subsidizing the advancement and higher income prospects of university graduates has been proposed. Here is the idea: let tuition be free when the student attends university, but then impose a tax on graduates that is based on the income that they actually receive (their *realized* income not their *expected* income) after they graduate.

One version of this idea would have the tax be proportional to income and set at a rate such that the taxes collected from graduates would on average pay for their education. This would require that graduates who experienced a bad state (meaning a low income) would pay in taxes an amount less than the cost of their education, while those who experienced a good state would pay more than what their education cost.

Adding the taxes paid by the unlucky and the lucky graduates, the total revenue collected would totally cover the cost of their education. We call this kind of policy *income-contingent taxation of graduates*.

Under the income-contingent tax, her income in each state, and her expected income would be as indicated in the third column of Table 13.2. The policy allows Sofia to trade away some of her gain in expected income were she to continue her education in order to reduce the degree of risk to which she would be exposed if she chose to invest in herself.

This is shown in Figure 13.18 (b). The government would set the level of the tax to collect total revenues sufficient to cover the full costs of the education that people had decided to pursue. The tax rate as a fraction of income, τ (the Greek letter "tau"), that would accomplish this is equal to the costs of her education as a fraction of her before-tax expected income as shown in M-Note 13.5.

The blue line through point **e** shows the effects of differing tax rates on reducing both the after tax expected income and risk exposure of the graduates. For any tax rate we know from Table 13.2 that

$$\text{Risk exposure with tax:} \quad \Delta_0(1-\tau)$$
$$\text{Expected income with tax:} \quad \hat{y}_e(1-\tau)$$

Because the tax reduces both expected income and risk by the same proportion this means that increasing the tax rate will move the resulting allocation downward to the left of point **e** along the blue line. The tax rate the government will implement is the one that collects in taxes an amount equal to the cost of education. This is indicated by point **f**.

With the income-contingent taxation of graduates, Sofia now has a choice between point **a** and point **f**. Because **f** is on a higher indifference curve than point **a** Sofia will continue her education ($u_2 > u_1$). Under the income-contingent tax plan, Sofia "purchases" a bundle that includes two years additional of higher education, expected income before taxes of $\hat{y}_e$, an amount $\tau\hat{y}_e = c$ in expected taxes (meaning what she would pay averaged over the bad and good state) and a reduced level of risk exposure $\Delta_0(1-\tau) = \Delta_f$. She is better off than had she taken her risk-free job instead.

M-CHECK Points like **e** and **f** on the blue line give the final bundle of expected income and risk exposure associated with differing tax rates, ranging from $\tau = 0$ at point **e** to $\tau = 100$ percent at the origin. Point **f** is selected as the tax rate that covers the cost of the graduates' additional education.

M-NOTE 13.5 **The tax rate and after-tax income in the good and bad state**

The tax collected from graduates on average will be $\tau\hat{y}_e$ and this amount will have to cover the cost of their education, c. So $\tau\hat{y}_e = c$ which means that $\tau = c/\hat{y}_e$.

Sofia gets tuition-free education and in the good state she obtains income $\hat{y}_e + 0.5\Delta$ minus the tax she must pay on her income which is $\tau(\hat{y}_e + 0.5\Delta)$ as in Table 13.2:

$$y_T^G = \overbrace{\hat{y}_e + 0.5\Delta}^{\text{Income in the good state}} \overbrace{-\tau(\hat{y}_e + 0.5\Delta)}^{\text{Tax payment in good state}} \tag{13.15}$$
$$y_T^G = (\hat{y}_e + 0.5\Delta)(1-\tau)$$

Similarly, if she experiences the bad state, then her income will be the following:

$$y_T^B = \overbrace{\hat{y}_e - 0.5\Delta}^{\text{Income in the bad state}} \overbrace{-\tau(\hat{y}_e - 0.5\Delta)}^{\text{Tax payment in bad state}} \tag{13.16}$$
$$y_T^B = (\hat{y}_e - 0.5\Delta)(1-\tau)$$

CHECKPOINT 13.10 **Investing in yourself** Explain why, in the absence of the income-contingent graduates' tax, Sofia's choice was between points **a** and **d**, but with the tax it is between **a** and **f**. How does the tax alter her decision about continuing her education?

13.11 ANOTHER FORM OF INSURANCE: A LINEAR TAX AND LUMP-SUM TRANSFER

The income-contingent graduates' tax mitigated the riskiness that Sofia faced in choosing to continue her education. Like deciding whether to attend university, and if so what subject to major in, or to invest in bullocks rather than irrigation pumps, deciding on a policy of taxation and spending by a government involves citizens' trade-offs between expected income and risk. So, what you have learned about risk aversion and insurance gives you an insight into the sometime controversial economics and politics of taxation and redistribution.

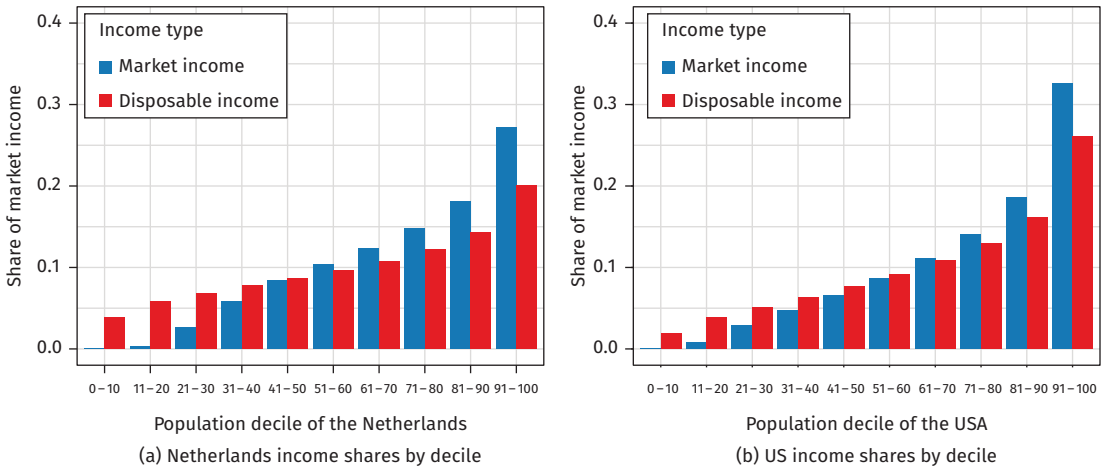
Taxes, transfers, and redistribution: A look at the data

Countries differ in the extent to which government taxes and payments to citizens reduce the degree of inequality in what families and people are able to spend. Figure 13.19 provides an illustration, using data from the Netherlands and the US. The horizontal axis in both panels refer to deciles of the population from the poorest 10 percent (on the left) through ever richer segments of the population to the richest 10 percent.

The height of the bars shows what fraction of the total income of the country is received by each of these groups of 10 percent (called deciles).

Figure 13.19 **Inequality of disposable income and market income in the Netherlands and the US.**

Source: CORE, *The Economy*.[11]



The blue bars show the share of each decile in what is termed **market income**, that is, income before taxes are paid or payments from the government received. The red bars show the income shares of the ten deciles in what is termed **disposable income**, that is, income after the payment of taxes and receipt of government payments.

The blue bar on the far right of panel (b), for example, shows that the top 10 percent of income recipients in the US receive about one-third of all of the market income; the farthest left blue bar in the same panel shows that the poorest 10 percent in the US receives about 2 percent of all market income.

Studying the two panels you can see:

- *Redistribution*: The effect of taxes and transfers is to reduce the disposable income of the rich and to raise the disposable income of the poor: the red bars are shorter than the blue bars in the right of the figures, and the opposite is true on the left.
- *Differences in the extent of redistribution*: This redistribution to the less well-off is greater in the Netherlands than in the US. For example in market income the poorest two deciles in the Netherlands are poorer than the bottom two deciles in the US (compare the blue bars); but they receive a much larger share of disposable income (the red bars).
- *Differences in inequality of disposable income*: Inequality in disposable income is less in the Netherlands than in the US. The ratio of the disposable incomes of the top decile to the bottom deciles is 6.95 in the Netherlands and 16.8 in the US.

If the distribution of disposable income after taxes and transfers is more equal than the distribution of market income, then the tax and transfer policy is termed progressive. If disposable income is more unequal than market income, the tax and transfer policy is termed regressive. The tax and transfer systems in the Netherlands and the US are progressive, but the US is less progressive than the Netherlands.

MARKET INCOME Market income is income before the payment of taxes or the receipt of transfers from the government; it includes earnings (wages and salaries from employment) as well as income from self-employment and from the ownership of assets (interest, rents, or dividends).

DISPOSABLE INCOME Disposable income is the maximum a household can spend ('dispose of') without borrowing, after paying tax and receiving transfers (such as unemployment insurance and pensions) from the government.

HISTORY Arthur Okun, who served on US President John F. Kennedy's Council of Economic Advisors wrote this about tax and transfer policies: "The money must be carried from the rich to the poor in a leaky bucket. Some of it will simply disappear in transit, so the poor will not receive all the money that is taken from the rich." The leaks he had in mind included not only the administrative costs of the policies but also the possible effects of the taxes on incentives. Corruption by government officials would also be part of the leakage.[12]

The degree of risk exposure with linear taxes and lump-sum distribution

To understand tax-financed redistribution, let's consider what is called a **linear tax and lump-sum transfer** policy.

- *Linear tax*: The *linear tax* part is that each family or person pays a fixed fraction of their income in taxes, so the taxes paid are a linear function of the pretax income a person has.
- *Lump sum*: The *lump-sum* part is that the total amount of taxes collected, net of the costs of collecting taxes and distributing the transfers, is divided equally among all of the citizens.

Because we will eventually turn to the politics of taxes and transfers, we call the taxpaying (and transfer-receiving) person or family "the citizen." The citizen is exposed to some risk: they can experience either a "good state" or a "bad state," occurring as before with equal probability. To introduce a tax and transfer policy we define τ as the tax rate, that is the percentage of market income paid in taxes. The cost of administering the program is ϕ percent of the taxes collected. Let us now consider the citizen's after-tax and transfer disposable income.

In Table 13.3 we show the taxes paid and transfer received in the good and bad states for two citizens of differing income levels. Taxes paid depend on which of the two equally likely states occur, so just averaging across these states we have:

$$\text{Individual's expected taxes} \quad \hat{T} = \tau\left(\frac{y^G + y^B}{2}\right) = \tau\hat{y} \tag{13.17}$$

A citizen's expected income after taxes but before receiving the transfer is expected income ($\hat{y}$) minus expected taxes or ($\hat{T}$): $\hat{y} - \hat{T} = \hat{y} - \tau\hat{y} = \hat{y}(1-\tau)$. Because the tax is proportional to pretax income, just as with the income-contingent graduates' tax, citizens pay more if they experience the good state than if the bad state occurs.

But what they receive as a transfer does not depend on the state that they experience.

$$\text{Transfer received} = \frac{\text{total taxes collected - cost of admin}}{\text{number of citizens}}$$

The total taxes collected divided by the number of citizens is just the tax rate τ times the average pretax income of citizens ($\underline{y}$) and so is equal to $\tau\underline{y}$. The cost of administering the tax and transfer policy is ϕ times that amount. So we have:

LINEAR TAX AND LUMP-SUM TRANSFER A tax that is proportional to income (a linear tax), the proceeds of which are divided equally and transferred to citizens (a lump sum).

Table 13.3 A comparison of the before and after tax experiences of low-income and high-income citizens. To make this table we set a mean income of the entire economy (not just the two citizens chosen here as examples) equal to 80,000, a degree of risk pretax of $\Delta = 40{,}000$, a tax rate, $\tau = 0.3$, and cost of the tax, $\phi = 0.05$. The table shows what is also illustrated in Figure 13.20, for example, that the citizen with lower expected income receives more in transfers (22,800) than she can expect to pay in taxes (15,000) and that both higher- and lower-income citizens experience a lesser degree of risk (28,000, the difference between the good and bad outcomes after the tax) than they did before the tax (40,000).

Outcome	General result	Low-income citizen	High-income citizen
	Before tax		
Good outcome before tax	$y^G = \hat{y} + \frac{\Delta}{2}$	70,000	140,000
Bad outcome before tax	$y^B = \hat{y} - \frac{\Delta}{2}$	30,000	100,000
Experienced risk before tax	$y^G - y^B = \Delta$	40,000	40,000
Expected income before tax	$\hat{y} = \frac{y^G + y^B}{2}$	50,000	120,000
	After tax		
Transfers received	$= \tau\underline{y} - \phi\tau\underline{y}$ $= \tau\underline{y}(1-\phi)$	22,800	22,800
Good outcome after tax	$y_T^G = (\hat{y} + \frac{\Delta}{2})(1-\tau) + \tau\underline{y}(1-\phi)$	71,800	120,800
Bad outcome after tax	$y_T^B = (\hat{y} - \frac{\Delta}{2})(1-\tau) + \tau\underline{y}(1-\phi)$	43,800	92,800
Experienced risk after tax	$y_T^G - y_T^G = \Delta(1-\tau)$	28,000	28,000
Expected income after tax	$\hat{y}(1-\tau) + \tau\underline{y}(1-\phi)$	57,800	106,800
Expected taxes	$\hat{T} = \tau\left(\frac{y^G + y^B}{2}\right) = \tau\hat{y}$	15,000	36,000

$$\begin{aligned}\text{Transfer received} &= \tau\underline{y} - \tau\underline{y}\phi \\ &= \tau\underline{y}(1-\phi)\end{aligned} \tag{13.18}$$

And the citizen's expected income after taxes and transfers is therefore:

$$\hat{y}_T = \underbrace{\hat{y}(1-\tau)}_{\text{Expected after-tax income}} + \underbrace{\tau\underline{y}(1-\phi)}_{\text{Expected transfers}} \tag{13.19}$$

In Figure 13.20 we show the expected taxes and transfer levels for citizens of all income levels. The first term in Equation 13.19 is the upward-sloping line (it's a line not a curve because the tax is linear). The horizontal blue line is the second term in Equation 13.19, the transfers received by all citizens.

The vertical distance between the blue and the green lines shows the difference between the expected taxes paid and transfer received, positive for higher income citizens and negative for lower income citizens. For a person whose income is $y = \underline{y}(1-\phi)$, the expected taxes paid equal the

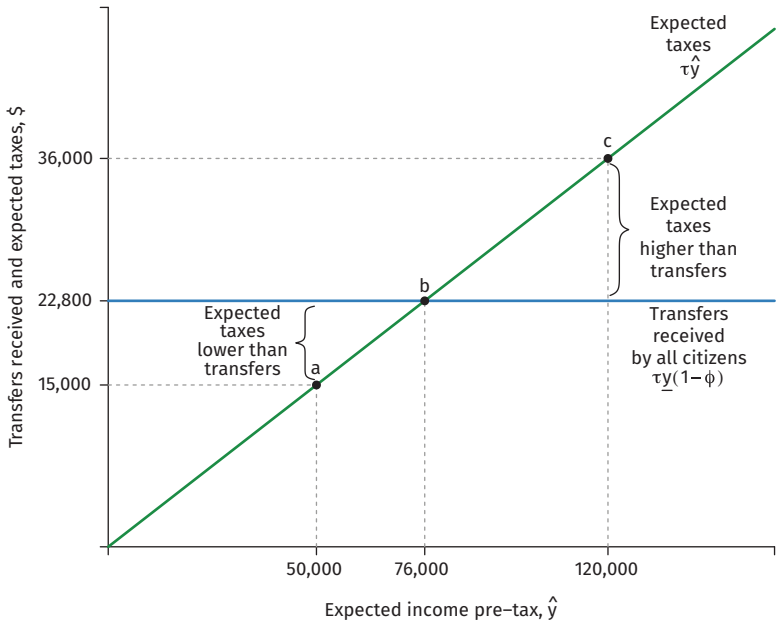
transfer received. This can be seen from the equation above rewritten as follows:

$$\hat{y}_T = \hat{y}(1-\tau) + \tau\underline{y}(1-\phi) = \hat{y} - \tau(\hat{y} - \underline{y}(1-\phi)) \tag{13.20}$$

We can conclude two things about the linear tax and lump-sum transfer policy.

- the policy is *progressive*: it redistributes income from higher- to lower-income citizens. Expected disposable income–that is $\hat{y}_T$–is more equal than before tax and transfer income. This is because (from Figure 13.20) expected taxes exceed the transfer for citizens with higher expected income and the reverse is true for people with lower expected income.
- the policy is a form of *insurance*: it redistributes income from the lucky to the unlucky. This is because independently of whether their expected incomes are high or low, people experiencing the good state pay more in expected taxes than do people experiencing the bad state.

Figure 13.20 The relationship between income and taxes paid and transfers received. Someone who has low pretax expected income ($y_L = 50{,}000$) receives *more* in lump-sum transfers (22,800) than the taxes they pay (15,000). Someone who has high pretax expected income ($y_H = 120{,}000$) receives *less* in lump sum (22,800) transfers than the taxes they pay (30,000). To make this figure we have used $\tau = 0.3$, $\phi = 0.05$, and $\underline{y} = 80{,}000$ as in Table 13.3. Remember that $\underline{y}$ is the mean income of the *entire* economy, *not* the average income of the two citizens included in the example.



Concerning the second bullet, from Equation 13.20 (repeated below) we can see that the policy *reduces the risk exposure* of both high income and low income citizens. Expected disposable income is:

$$\begin{aligned}\hat{y}_T &= \hat{y}(1-\tau) + \tau\underline{y}(1-\phi)\\ &= (1-\tau)(\text{the risky part}) + \tau(\text{the certain part})\end{aligned}$$

The reason why the tax and transfer policy serves as a kind of insurance is that the transfer received—that is, $\tau\underline{y}(1-\phi)$—is certain: unlike before tax income, it is independent of the state that the individual experiences, whether good or bad. This is because there are a very large number of citizens paying taxes, some experiencing good states some experiencing bad states. Equation 13.21 says that the larger is the tax, the less weight does the *risky* part of her income have in determining in citizen's the expected disposable income. And correspondingly, the larger is the weight of the *certain* part.

The effect, as was the case with the income-contingent tax on graduates, is to reduce the difference between disposable income in the good and the bad state from $\underline{\Delta}$ to $\underline{\Delta}(1-\tau)$. (This was illustrated for the two citizens in Table 13.3.)

CHECKPOINT 13.11 The citizen with income equal to $\underline{y}(1-\phi)$ Consider the values in Table 13.3. The mean income is \$80,000 with $\tau = 0.3$ and $\phi = 0.05$:

a. Calculate the expected income of the citizen with income $\underline{y}(1-\phi)$.

b. Show that the citizen with this income pays an amount in taxes exactly equal to the transfers they receive as a lump sum.

✓ **FACT CHECK** If families are limited in how much they can borrow and do not have any other way of insuring against income shocks, the equation for expected income with taxes in Table 13.3 tells us that, where taxes are more redistributive from the higher to lower income families, we expect to see also smaller differences in consumption between families (some lucky, some unlucky). Using the fact that tax rates and transfer policies differ greatly from one state to the other in the US, economists have determined that this is exactly the case. Consumption is more equally distributed in states with more progressive tax and transfer systems.[13]

13.12 A CITIZEN'S PREFERRED LEVEL OF TAX AND TRANSFERS

Even though most people place some value on the well-being of people other than themselves, citizens will obviously differ in the level of the linear tax that they would prefer unless they are perfect altruists (value others' income as much as their own). Of course people do not get to pick their preferred tax rate—the same rate applies to all. But political parties propose differing levels of taxation, and citizens do pick which party to support.

Think about a particular citizen, Helmut, who in the absence of the tax and transfer policy would experience risk level Δ_f and expected income $\hat{y}_f$, as indicated by point **f** in Figure 13.21. Helmut's expected income is substantially above $\underline{y}(1-\phi)$ and so, like the person indicated by point **c** in

! REMINDER The tax and transfer line is similar to the insurance contract line (insurance line) introduced earlier (Figure 13.14): it describes the opportunities for reducing risk exposure at the cost of a reduction in expected income. A difference, however, is that the slope of the tax and transfer line—the opportunity cost of risk reduction in terms of reduced expected income—is not a constant but depends on the level of before tax expected income.

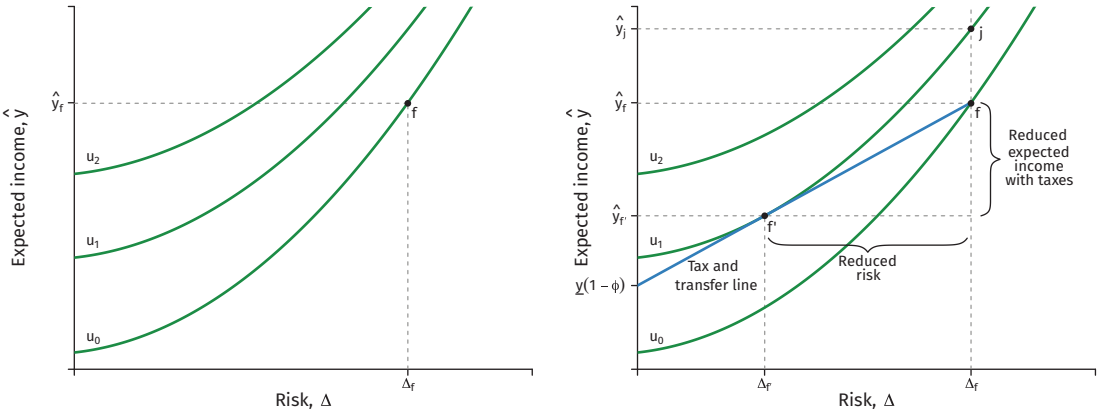
Figure 13.20, his expected income after taxation will be less than in the absence of taxation.

But even if he is entirely self-regarding (cares only about his own income and risk, not that of others) he may favor a tax on his income along with a lump-sum transfer of the average tax revenues. The reason is that, like the free tuition and graduates' income-contingent tax policy that made it advantageous for Sofia to continue her education, the linear tax and lump-sum transfer will reduce the difference between his realized income in the good and bad state. A reduction in his expected after-tax income may be a cost he is willing to pay for the risk reduction that the tax and transfer policy implements.

In Figure 13.21(b), the blue line through point **f**—called the tax and transfer line—shows how various levels of taxation could transform his combination of after tax and transfer expected income and risk exposure. The options range from $\tau = 0$ in which case he would remain at point **f** to $\tau = 1$. In this case, that is, with a 100 percent tax rate, the only after-tax income a person received would be the transfer, so he and everyone else would have an income after taxes and transfers of $\underline{y}(1-\phi)$. Points on the line closer to $\underline{y}(1-\phi)$ on the vertical axis represent greater risk reduction by means of higher taxes and transfers.

Even though Helmut can expect to pay more in taxes than he will receive in the transfer, will he nonetheless prefer some positive level of taxes? To answer this we need to identify the benefits and costs to Helmut of increasing the tax.

Figure 13.21 A citizen's preferred tax and transfer policy. Helmut's potential choices about risk and expected income. His bundle of risk and expected income puts him on indifference curve u_0 at point **f** in panel (a). Taxes that he pays and which are distributed to him (and others) as a lump sum could move him to u_1, with decreased risk exposure and lower average income at point **f′**.



(a) Indifference curves and allocation without taxes & transfers

(b) Post-tax and transfer allocation with reduced income and reduced risk exposure

As you can see from Figure 13.21 (and as is shown in M-Note 13.6) the slope of the tax and transfer line is:

$$\text{Slope of the tax and transfer line} = \frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} \tag{13.21}$$

We can see from Equation 13.20 that the numerator of Equation 13.21 is the reduction in disposable income associated with the tax and transfer policy. This is the cost of the policy. We know from the fact that the experienced level of risk under the tax and transfer program is $\Delta(1-\tau)$ that the denominator is the effect of the reduction in risk resulting from the tax and transfer policy. This is the benefit of the policy. The ratio of the two—how much the benefit of risk reduction costs in terms of the expected disposable income forgone—is the opportunity cost of risk reduction by this policy.

As in the previous cases studied, the slope of his indifference curves, a measure of his risk aversion, is $-u_\Delta/u_{\hat{y}}$, his marginal disutility of risk divided by the marginal utility of expected income when experiencing the indicated level of risk and expected income. The slope of an indifference curve indicates how much he values risk reduction relative to disposable income. So we have:

$$\text{Slope of an indifference curve} = \frac{-u_\Delta}{u_{\hat{y}}} = \frac{\text{marginal disutility of risk}}{\text{marginal utility of } \hat{y}}$$

We can see that at point **f**—the status quo that Helmut would experience in the absence of any tax—the following inequality holds:

$$\text{Slope of tax and transfer line} < \text{Slope of the indifference curve}$$

$$\frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} < \frac{-u_\Delta}{u_{\hat{y}}} \tag{13.22}$$

which, rearranged, reads

$$u_{\hat{y}} \cdot (\hat{y} - \underline{y}(1-\phi)) < -u_\Delta \cdot \Delta \tag{13.23}$$

$$\text{Marginal costs of risk reduction} < \text{Marginal benefits of risk reduction}$$

This means that at Helmut's status quo (point **f**, that is, $\tau = 0$) the reduction in income associated with a tax times the marginal utility of expected income is less than the reduction in risk times the disutility of risk. Because the costs are less than the benefits, Helmut would prefer some positive level of taxation.

To find the level of τ that he would prefer, we use the analytical framework above, but we require that the marginal costs and benefits of risk reduction be equal, which means that the slopes of the tax and transfer line and the indifference curve are equal. The level of taxation that implements this rule is shown by point **f′** in Figure 13.21.

His utility at that point $(u(\hat{y}_T, \Delta_T) = u_1)$ is the same as he would have experienced without the tax and transfer policy if he had been substantially richer, shown as point **j** also on the indifference curve u_1 in the figure. At that point he has an expected before tax and transfer income of $\hat{y}_j$ with his

initial level of risk exposure (Δ_f). The difference between this hypothetical income at point **j** in the figure and his actual expected before tax and transfer income ($\hat{y}_j - \hat{y}_f$) is a measure in income units of how valuable the reduction in risk exposure accomplished by the tax and transfer policy is to him.

You think of point **f** as the citizen's fallback option: no tax and transfer policy. Then the amount $\hat{y}_j - \hat{y}_f$ is the rent he would receive if the tax and transfer policy indicated by the point **f**′ were implemented. We call these political rents because they are the result of public policies that are implemented as the result of a political process.

M-NOTE 13.6 **The opportunity cost of reduced risk exposure**

Summarizing the results so far, the citizen's expected disposable income (after taxes and transfers) is (see Table 13.3):

$$\hat{y}_T = \hat{y}(1-\tau) + \tau\underline{y}(1-\phi) = \hat{y} - \tau\left(\hat{y} - \underline{y}(1-\phi)\right) \tag{13.24}$$

And her exposure to risk under the tax and transfer policy is:

$$y_T^G - y_T^B = \Delta_T = \Delta(1-\tau)$$

We can use these two equations to find the equation for the tax and transfer line, expressing $\hat{y}_T$ as a function of Δ_T. First we rearrange the equation immediately above to get an expression for the tax rate, or

$$\tau = 1 - \frac{\Delta_T}{\Delta}$$

Then we substitute this expression into Equation 13.24 for $\hat{y}^T$:

$$\hat{y}_T = \hat{y} - \left(1 - \frac{\Delta_T}{\Delta}\right)\left(\hat{y} - \underline{y}(1-\phi)\right) \tag{13.25}$$

This is the equation for the tax and transfer line for a person with market expected income of $\hat{y}$ and market risk exposure of Δ. A tax and transfer policy transforms a reduction in expected income (a cost) into a reduction if risk exposure (a benefit). We find the slope of the tax and transfer line by differentiating $\hat{y}_T$ with respect to Δ_T or:

$$\frac{d\hat{y}_T}{d\Delta_T} = \frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} \tag{13.26}$$

This is the negative of the marginal rate of transformation of reduced expected income into reduced risk, which is the opportunity cost of reducing risk, or the slope of the tax and transfer line. The numerator is the marginal cost of higher taxes in terms of expected income foregone. The denominator is the marginal benefit of higher taxes in terms of reduced risk exposure.

Equation 13.26 shows that the tax and transfer line will be flat (mrt = slope = zero, meaning no cost) for the citizen with an income equal to $\underline{y}(1-\phi)$. For those with incomes less than $\underline{y}(1-\phi)$, the tax and transfer line is downward-sloping, meaning that the tax and transfer policy raises their expected disposable income.

Equation 13.25 says that a tax and transfer policy that reduces risk exposure to zero ($\Delta_T = 0$) means that everyone will have the same expected disposable

continued

income. To see this substitute in $\Delta_T = 0$ so that we then have:

$$\begin{aligned}\hat{y}_T &= \hat{y} - \left(1 - \frac{\Delta_T}{\Delta}\right)\left(\hat{y} - \underline{y}(1-\phi)\right)\\ &= \hat{y} - \left(\hat{y} - \underline{y}(1-\phi)\right)\\ \Rightarrow \hat{y}_T &= \underline{y}(1-\phi),\end{aligned}$$

which is the citizen's after tax and transfer income irrespective of the person's level of market expected income.

M-NOTE 13.7 **Helmut identifies his preferred tax rate**

Helmut wants to determine which tax rate (including zero) would be the best for him. We first write his utility function as follows:

$$u(\hat{y}_T, \Delta_T) = u(\overbrace{\hat{y}(1-\tau) + \tau\underline{y}(1-\phi)}^{\text{After-tax expected income}}, \overbrace{\Delta(1-\tau)}^{\text{After-tax risk exposure}}) \tag{13.27}$$

Equation 13.27 says that utility depends on a person's *after tax* and transfer income and risk exposure. Now, we want to see the tax rate under which his utility is maximized. Differentiating with respect to τ and equating the result to 0 so as to find the maximum we have:

$$\frac{\partial u}{\partial \tau} = u_{\hat{y}} \cdot \left(-\hat{y} + \underline{y}(1-\phi)\right) + u_{\Delta} \cdot (-\Delta) = 0$$

Simplifying:

$$\frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} = -\frac{u_{\Delta}}{u_{\hat{y}}} \tag{13.28}$$

You are familiar with the right side of the equation. It is the marginal rate of substitution. The left side of the equation you also know (from Equation 13.26) is the opportunity cost of risk reduction, that is, the marginal rate of transformation of reduced expected income into reduced risk. In other words, the left side of the equation is the slope of the tax and transfer line in Figure 13.21.

Helmut's utility is maximized at the point where the slope of the tax and transfer line is equal to the slope of the highest indifference curve that is feasible for Helmut, that is, when:

$$mrt(\Delta, \hat{y}) = \frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} = -\frac{u_{\Delta}}{u_{\hat{y}}} = mrs(\Delta, \hat{y})$$

CHECKPOINT 13.12 **The "leaky bucket" problem** The cost of administering the tax and transfer policy represented by ϕ is sometimes referred to as the fraction of tax revenues that "leaks away" before making it back to citizens as transfers. Redraw the figure of the tax and transfer lines with a value of ϕ that is:

a. *Higher* than that shown; and

b. *lower* than that shown.

✓ **FACT CHECK** Political scientists Torben Iversen and David Soskice explored the relationship between support for taxation and redistribution policies on the one hand and the degree of risk exposure associated with a person's skills on the other. For example, a skilled operator of a specific kind of machine would be more risk exposed than a person with coding skills. They found that taking account of other influences on political preferences (income, sex, employment status, party affiliation, and age) the degree of a person's skill risk exposure predicts support for tax and transfer policies. In their study of 11 high-income democracies Iversen and Soskice found that the risk exposure effect was equal in size to the effect of income on their preferences about taxes and redistribution.[14]

! REMINDER Everyone pays the same tax rate τ as a percent of their income; the total taxes paid differ because people's incomes differ.

13.13 POLITICAL RENTS: CONFLICTS OF INTEREST OVER TAXES AND TRANSFERS

So far the tax and transfer policy resembled the insurance policy that Juliana purchased. Helmut was willing to trade a reduction in expected income (the taxes he paid minus the transfer he received) for a reduction in risk exposure, just as Juliana paid an insurance premium so as to reduce her risk exposure. Their risk mitigation opportunities were represented by a tax and transfer line, and a similar insurance contract line.

But there is a big difference. Unlike Juliana who was free to choose any point along the insurance contract line, citizens cannot choose *individually* how much risk reduction they will "purchase": the amount will be determined by a tax and transfer policy adopted by the government applied to all citizens. And this means that citizens will *disagree* about the policy to adopt, due to their differing levels of expected market income and risk exposure.

Those with higher expected incomes will typically favor lower taxes than those with lower expected incomes. There are two reasons for this:

- As Equation 13.21 makes clear, the opportunity cost of risk reduction by means of taxes and transfers is greater for those with higher expected incomes. They pay more of the taxes simply because averaging across the good and bad outcomes, they have higher incomes.
- The risk reduction implemented by the tax and transfer policy is worth less to those with higher expected incomes because of diminishing risk aversion as Figure 13.7 shows.

To see why this leads to conflicts about the level of taxation think about a person with the same level of risk exposure as Helmut, shown by point **e** in Figure 13.22. Taking account of the fact that not all of the taxes collected will be distributed (the "leaky bucket problem") her expected income is such that the expected amount she will pay in taxes is exactly offset by the amount she will receive in transfers.

So for her there is no cost to risk reduction. This is why her tax and transfer line is horizontal. Her utility will be maximized if $\tau = 100$ per cent. She and Helmut have a conflict of interest, as Helmut's preferred level of taxation is much less than 100 percent because his utility-maximizing point is at $\mathbf{f}'$.

Now consider another citizen much richer than Helmut and exposed to the same level of risk, shown as point **i** in Figure 13.22. You can see that all the points on the rich person's tax and transfer line are on indifference curves with a lower level than u_4, which is what he will experience in the absence of any tax and transfer policy. The tax that Helmut prefers (which would bring Helmut to point **f'**) would inflict a substantial loss in utility on this richer individual. So the political rent that Helmut would receive under

his preferred tax (the quantity $\hat{y}_j - \hat{y}_f$ in Figure 13.21) has a counterpart in the higher-income individual's loss if Helmut's plan were adopted.

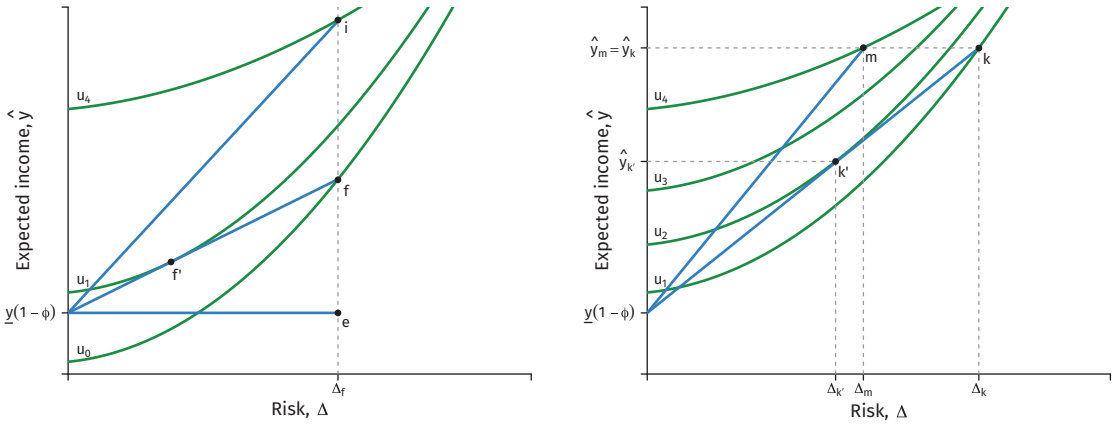
These differences in expected income and risk exposure levels are the basis of conflicts of interest, in which the stakes for some of the citizens are substantial. You can think of the political rent that Helmut would receive were his preferred tax implemented as the maximum amount he would be willing to pay in order to ensure the election of a party that would implement that plan, if the alternative were no tax and transfer policy at all.

In Figure 13.22 (b) we can see that even citizens with the same level of expected income will differ in their preferred level of taxation if they have differing levels of risk exposure. To see this think about the preferred tax level for Citizen M and Citizen K whose expected income and levels of risk exposure are shown by points **m** and **k** in the figure. Citizen K prefers a substantial reduction in risk (shown by point **k**′) even thought it would cost her a major loss in expected income. By contrast, from his status quo point, the tax and transfer line provides Citizen M with no opportunities for a gain in utility (if he moved down and the left along his tax and transfer line he would be crossing ever *lower* indifference curves).

The differences between what citizens would choose show the conflicts at the heart of many policy debates in contemporary societies, with citizens'

M-CHECK The slope of the tax and transfer line is the opportunity cost of risk reduction in terms of reduced expected income. For the low income person at point **e**, this is zero (the line is flat) because the transfer they receive exactly offsets the modest tax that they pay, while for the person at point **i** the opportunity cost is high (the line is steep) because the tax paid greatly exceeds the transfer that they will receive.

Figure 13.22 Conflicts of interest over the level of taxation. In panel (a), citizens with pretax and transfer expected incomes and risk exposure given by points **f, e,** and **i** differ in the level of taxes they would prefer because they differ in their level of expected income. The person with the bundle **i** prefers no tax, **e** prefers a 100 percent tax, and **f** prefers a substantial (but not 100 percent) tax that would result in the bundle **f**′. In panel (b), citizens with risk exposure and expected incomes indicated by points **m** and **k** differ in their preferred level of taxation because they differ in the extent of their risk exposure. The person with the bundle **m** prefers no tax, while the person at **k** is more risk exposed and prefers a substantial tax that would move them to point **k**′.



(a) Conflicts of interest between rich and poor

(b) Conflicts of interest between more and less risk exposed

preferences mirroring their choices of the political outcomes they would prefer. The model we have presented can help us understand these debates and conflicts over the extent of tax-based redistribution from the lucky to the unlucky and from the well-off to the less well-off.

13.14 APPLICATION: CHOOSING JUSTICE, A QUESTION OF ETHICS

So far we have asked questions about the level of taxes and redistribution that people would prefer if they cared only about themselves—their own expected income and risk exposure, not that of their fellow citizens. But we know from Chapter 2 that people do care about others and sometimes dislike inequality, even if their own income is higher as a result.

The same model about risky decisions that we have been using, with some important additions, can clarify how people might think, not about their own self-interest, but about an ethical question: How much inequality is just and how much redistribution a government should do?

The common element in questions about risk on the one hand and economic injustice on the other is that both concern differences in income. Risky decisions are about actions that result in differences in one's own realized income in good and bad states. Judgment about the justice or injustice of income inequality is also about differences, but in this case it is differences between particular people in their income (averaged over good and bad states if they are risk exposed).

! REMINDER *Inequality aversion* is a preference for more equal outcomes and a dislike for either *disadvantageous inequality* that occurs when others have more than the actor and (to a lesser extent typically) also *advantageous inequality* that occurs when the actor has more than others, or both.

Just as people have preferences about risk—ranging from strong risk aversion to risk neutrality—people have preferences about inequality: in behavioral experiments, for example, many subjects are *inequality averse*. People can be both risk averse and inequality averse. Or risk neutral and inequality averse.

Feasible choices of the extent of inequality and the level of average income

! REMINDER Adam Smith in his *Theory of Moral Sentiments* wrote "We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it."[15]

We now ask you to imagine that you are Adam Smith's "Impartial Spectator" and you are asked to design your ideal society. In the society you're considering, it has already been determined that there will be two groups of equal size: the first called "richer" and the second called "poorer." Your job is to answer the following question: *How much* richer than the poorer people should the richer people be?

Maybe you would question why there should be any rich and poor at all. But you are also told that perfect equality in the society would mean that there were insufficient incentives for people to work hard, take risks, and

study. As a result some inequality would be better for everyone, *even for the poor*, because the poor would have more in a world in which most people work hard, study and are willing to take some risks than in a world of no inequality.

But too much inequality can also be a problem, lowering average income. For example where high-income people are much richer than low-income people, social conflicts–ranging from labor strikes to property theft may be more common. In all countries a substantial fraction of the labor force work as police, private security workers, and others whose job it is to maintain order and protect property. But the ranks of what is called "**guard labor**" are substantially larger in highly unequal countries. People guarding things are not producing things, so inequality may reduce average incomes by diverting a nation's productive potential from producing to guarding.

✓ **FACT CHECK** We provide more evidence about inequality and guard labor and illustrate how policies that reduce inequality could increase average incomes in Chapter 16. There is some evidence that the most unequal of the high-income economies (the UK and the US) are to the right of point **m** in the figure, meaning that there exist policies that could both reduce inequality and raise average incomes.

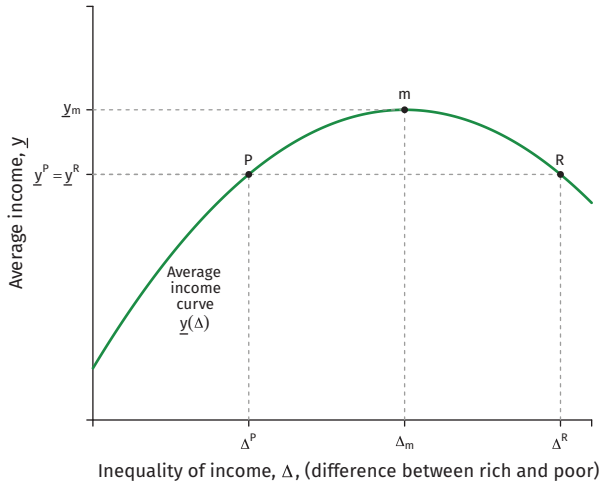
The comparison of the Netherlands and the US in Figure 13.19 provides an illustration. Disposable income is much more equally distributed in the Netherlands than in the US. And the fraction of the Dutch labor engaged in what are termed "protective services" (literally guards) is less than a third of the fraction of the US labor force engaged as guards. Countries even more equal than the Netherlands–Denmark and Sweden–employ even fewer guards.[16]

Figure 13.23 illustrates an economy in which average income first rises, reaches a maximum, and then falls as inequality increases, as depicted by $\underline{y}(\Delta)$. The $\underline{y}(\Delta)$ function is similar to the risk–return schedule. It gives the feasible combinations of average income and inequality. The negative of the slope of $\underline{y}(\Delta)$ is *the marginal rate of transformation of inequality into average income*, or the opportunity cost of greater average income in terms of greater inequality. For low levels of inequality the opportunity cost of greater inequality is negative because greater inequality results in higher average income.

Because the classes are of equal size, average income is the midpoint between the income of the rich and the income of the poor. So the rich get average income plus one-half of Δ and the poor get average income minus one half of Δ. Recall that the average income, $\underline{y}$, is determined by the level of inequality using the function $\underline{y}(\Delta)$. Therefore:

GUARD LABOR Those employed as police, private security personnel, the armed forces, and others whose job is enforcing and perpetuating the rules of the game.

Figure 13.23 Average income and income inequality. How much equality is too much, too little, just right? Point **m** indicates the level of inequality at which average income is the greatest. **P** and **R** respectively are the inequality levels that maximize the income of the poor and the rich.



$$\text{Income of the poor} \qquad y^P = \underline{y}(\Delta) - \frac{1}{2}\Delta \tag{13.29}$$

$$\text{Income of the rich} \qquad y^R = \underline{y}(\Delta) + \frac{1}{2}\Delta \tag{13.30}$$

Between 0 inequality and point **P** in the Figure 13.23, not only the rich but even the poor would benefit from greater inequality. (We show in M-Note 13.8 that **P** is the point on the average income function where the slope is equal to $\frac{1}{2}$.) So increased inequality would result in Pareto improvements over an economy in which average income and inequality are both very low.

Inequality beyond **P** would benefit the rich and hurt the poor, but average incomes would still rise: the rich would be receiving a larger piece of a larger pie. Beyond point **m**, however, with greater inequality, average income would fall, and the rich would be getting a larger slice of a smaller pie. There would eventually be some level of inequality so extreme that for inequality greater than this, even the income of the rich would suffer. This is indicated by point **R**.

M-NOTE 13.8 **How much inequality would maximize the income of the poor?**

To answer the question we have to vary Δ to maximize Equation 13.29.

$$y^P = \underline{y}(\Delta) - \frac{1}{2}\Delta$$

continued

So we differentiate Equation 13.29 with respect to the single variable Δ and set the result equal to zero:

$$\frac{dy^P}{d\Delta} = \underline{y}_\Delta(\Delta) - \frac{1}{2} = 0$$

$$\underline{y}_\Delta = \frac{1}{2}$$

So, the income of the poor is maximized when level of inequality is such that the slope of the average income function equals one-half, shown by Δ^P in Figure 13.23.

An Impartial Spectator's preferences about inequality

What level of inequality would an Impartial Spectator choose?

We assume that the Impartial Spectator is not so inequality averse that they would choose a level of inequality less than Δ^P, because point **P** would be a Pareto improvement over any outcome for which $\Delta < \Delta^P$. Both the rich and the poor would benefit if the economy were organized at point **P** rather any less unequal outcome. Similarly, the Spectator would be unlikely to select a level of inequality greater than Δ^R because such extreme levels of inequality would be Pareto inefficient: both the rich and the poor would be better off by a reduction in equality to Δ^R.

To consider the possible levels of inequality between points **P** and **R** we need to know more about the Impartial Spectator's preferences. Remember indifference maps and the utility functions on which they are based do not give us information about "how happy" each bundle of outcomes would make the decision maker, or what they "get." The indifference curves tell us what the decision maker will choose (selecting points on higher indifference curves over points on lower indifference curves). In the case of the Impartial Spectator the decision will be made on the basis of how just she judges each outcome to be.

Figure 13.24 shows average income as a function of the degree of inequality as in the previous figure, along with two sets of indifference curves that the Impartial Spectator might have, an inequality-averse Spectator on the right and a inequality-neutral Spectator on the left. The slope of these indifference curves is a measure of inequality aversion.

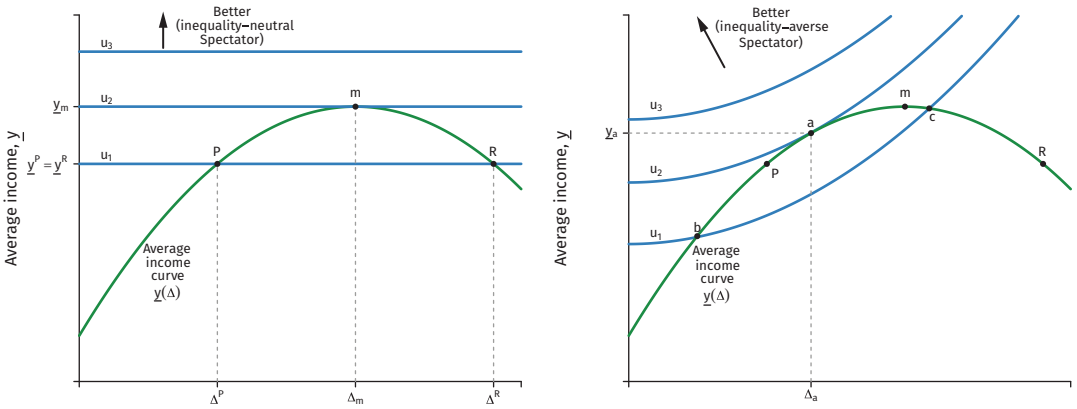
$$\text{Slope of an indifference curve} = \text{A measure of inequality aversion}$$

$$\frac{-u_\Delta}{u_{\underline{y}}} = \frac{\text{Marginal disutility of inequality}}{\text{Marginal utility of average income}}$$

With preferences indicated by these indifference curves, the Spectator will seek a point on the highest possible indifference curve that is also feasible, which is where the feasible frontier is tangent to an indifference curve.

Figure 13.23 shows that an inequality-averse Spectator would choose a lower level of inequality.

Figure 13.24 The degree of income inequality chosen by an inequality-neutral and an inequality-averse Spectator. The curved green line $\underline{y}(\Delta)$ is the set of feasible average incomes the Impartial Spectator faces when choosing a just level of inequality. The inequality-averse Spectator chooses $(\Delta_a, \underline{y}_a)$. An inequality-neutral Spectator would choose point **m** with a combination of average income and inequality of $(\Delta_m, \underline{y}_m)$ (as shown in panel (a)).



(a) Indifference curves of an inequality-neutral Spectator

(b) Indifference curves of an inequality-averse Spectator

CHECKPOINT 13.13 The trade-off between average income and inequality. Using Figure 13.24 consider the following:

a. Explain why average income falls as inequality increases after Δ_m.

b. Can you think of any reason why the Impartial Spectator might choose a point to the left of Δ^P?

13.15 RISK, UNCERTAINTY, AND LOSS AVERSION: EVALUATION OF THE MODEL

In evaluating the model of doing the best you can in risky situations remember two things:

- *The map is not the territory*: We have deliberately left out aspects of the problem that are not essential to answering the question at hand.
- *This is not a model of how people think*: The model is intended to understand and predict what people do, not to describe the literal thought processes in which they engage in choosing an action.

The model has expanded your analytical took kit, clarifying important aspects of economic behavior: investment in risky assets, what methods of production to use, the value of continuing one's education, how much insurance to purchase, citizens' evaluations of alternative tax and transfer policies. You have also seen that with some amendments, the same model can be used to consider an important ethical question: injustice. A feature of

the model is that it bases the study of risky decisions on familiar analytical tools: indifference curves and feasible sets.

But the model is limited in a number of ways.

Uncertainty: Not knowing the probabilities of the relevant contingencies

We have assumed that our decision makers—the student considering higher education, the citizens, the Indian farmers—knew the probability that the good and bad state would occur (we assumed that the two states are equally probable). But in many situations the probabilities of the various contingencies affecting the outcomes of people's actions are not known. So people face not the problem of risk, but the much more difficult problem of *uncertainty*. People typically have a pretty good idea about the risk of rain (you can get an estimate of that online, so this is a risky situation), not an uncertain one but no clue about the likelihood of an earthquake or the outbreak of a new pandemic (so this is a case of uncertainty).

In our utility function we have let Δ, the difference between the equally probable good and bad outcomes, represent what the risk-averse person considers to be the "bad"—namely risk. But not knowing the probabilities of the relevant contingencies there is no way to assign a particular level of risk to a choice, even if we know the difference in expected incomes between the good and bad state. To see this suppose that in the good state you gain 10,000 euros and in the bad state you lose 10,000 euros. How risky is your decision?

It is obviously a lot more risky if the probability of the bad outcome is one-half than if it is, say one-in-a hundred. The same would be true if the probability of the bad outcome were 99 in 100: the expected income associated with this choice would be very low and most decision makers whether risk averse or not would want to avoid it. But the reason would not be its riskiness: the choice would not be as risky as the fifty-fifty probability.

So not knowing the probabilities we cannot say how risky a choice is, making it impossible to use our risk-averse utility function. The same problem arises using another way that economists treat risk: by assuming that decision makers make choices to maximize their expected utility. If the probabilities of the contingencies are not known, expected utility cannot be computed.

Often the best we can do is to identify the consequences of the one or more "bad states" and adopt policies to reduce the likelihood of their occurring. The term used to describe these policies is commonly "prudence" (meaning, roughly "caution" in the face of uncertainty). But prudence does not deliver the prescriptions for action such as those that are possible when the probabilities of contingent events are known.

The limits of the model are especially clear when applied to the problem of climate change. The reason is that while many of the relevant facts are reasonably well established—that human activity contributes to climate change, for example—we really are not able to assign well-informed prob-

abilities or even guesses to some critical contingencies. For example we cannot know or even intelligently guess how probable human extinction is over the range of relevant earth surface temperatures.

Loss aversion as an alternative reason for avoiding risks or uncertainty

Loss aversion is a well-documented aspect of human behavior according to which the loss of some given amount—say, a euro or a favorite coffee cup—reduces our utility by more than a gain of the same amount or object would have raised our utility. A loss-averse person would refuse a coin flip in which they stood to gain 1 euro if the coin came up heads, and lose 1 euro if it came up tails.

Loss aversion differs from risk aversion because it captures the fact that we treat losses and gains differently, even if they are very small. This differs from risk aversion in that a person who was highly risk averse in a situation involving a substantial risk (a large difference between the payoff at the good and bad state) would be approximately risk neutral for very small risks.

Suppose you have an income of $1,000 per month. At the beginning of one month, you receive the news that you will receive a positive income shock, but you don't know exactly the amount. It could be either $100 or $300, each with a probability of 0.5. Now, suppose that you have an income of $1,200 per month, but you know that you will face either a positive income shock of $100 or a negative shock of $100, each equally likely. The two situations are analogous. The expected income is $1,200, the good and bad states exactly the same. However, a loss-averse person treats gains and losses differently: preferring the first scenario to the second because in the second there are losses despite the expected income being the same. We prefer a gain of $100 to achieve an income of $1,100 than a loss of $100 that results in the same income.

No single model is entirely adequate to cover all of the relevant cases, ranging from problems in finance, to loss-averse behaviors, to the absence of information about the probabilities of the relevant contingencies concerning climate change. A combination of models chosen on a case-by-case basis, with contributions from psychology, biology and other sciences seems the best way of understanding how we behave in situations in which we lack information on the outcomes of our actions.

CHECKPOINT 13.14 **Limitations of the model** Explain why the model of decision making with risky situations that you have learned may be more useful to study the risks arising from unexpected changes in the value of your home (and the demand for insurance) than the uncertainties of climate change or risky behaviors such as driving a car too fast.

LOSS AVERSION Loss aversion is present when the loss of some given amount reduces a person's utility by more than a gain of the same amount would have raised their utility.

13.16 CONCLUSION

When it comes to how we live our ordinary lives and the decisions we have to take, risk and uncertainty are the rule, not the exception. Richard Feynman, Nobel Laureate in physics–a field widely known as an 'exact science'–had this to say: "I have approximate answers and possible beliefs and different degrees of uncertainty about different things, but I am not absolutely sure of anything."[17]

In this chapter we have adapted and extended the model of constrained optimization to analyze risky decisions. The main new element is risk aversion, which builds on the idea that just as we may have likes and dislikes when it comes to different foods, or engaging in prolonged hard work, we also have preferences about the degree of risk to which we are exposed. Indifference curves based on risk preferences allow us to study the trade-offs that we face when the opportunities we have to increase our expected income also expose us to more risk.

We will see in Chapter 15 that the model of decision-making in the presence of risk will provide important insights into why capitalism is such a dynamic economic system. And we will continue a theme introduced in this chapter: how risk aversion may also contribute to a second attribute of capitalism, namely elevated levels of economic inequality.

MAKING CONNECTIONS

Constrained optimization: Treating expected income as a "good" and risk as a "bad" allows us to extend the familiar analytical tools—indifference curves and feasible sets—for use in understanding decision-making over risky options.

Mutual gains and conflicts over their distribution: Insurance allows a valuable risk reduction to the insured and profits to insurance providers, thus implementing a Pareto improvement over the 'no insurance' default option. How the mutual benefits made possible by insurance are divided between the insurance provider and the insured is a matter of conflict.

Heterogeneity: Wealth differences. Differences in risk aversion reflect not only personality differences among people but also differing situations in which people find themselves. Having substantial wealth and (as a result) having access to credit reduces risk exposure and hence risk aversion. Lesser income and wealth limits a person's ability to borrow and increases people's risk exposure and risk aversion.

Inequality and poverty traps: The risk aversion of people without access to credit will motivate them to avoid making high-risk choices with high expected returns (such as changing one's occupation, starting a business, or relocating), thereby perpetuating their limited income.

Policies: Tax and transfer policies by governments are a kind of insurance that distributes income not simply from higher-income to lower-income citizens, but also from the lucky to the unlucky; this can reduce the extent of risks to which a person is exposed and thereby allow people with limited wealth to make less risk-averse choices with higher expected incomes.

IMPORTANT IDEAS

contingency	progressive/regressive taxation	risk
expected utility	uncertainty	risk aversion
risk neutrality	certainty equivalent	insurance
insurance premium	loss aversion	risk-loving
risk–return schedule	investment	degree of risk
inequality aversion	market (pretax) income	tax and transfer line
disposable income	specific/general asset	"leaky bucket"
lump sum transfer	linear tax	guard labor
lottery	expected value/payoff	

MATHEMATICAL NOTATION

Notation	Interpretation
$u(\Delta, \hat{y})$	a utility function (for the study of risk as a "bad")
Δ	risk, the difference between the good and (equally probable) bad outcome
p	probability of an event
y	realized income
$\hat{y}$	expected income
$\underline{y}$	average income
$\hat{y}(\Delta)$	risk–return schedule
$\underline{y}(\Delta)$	inequality and average income schedule
s	insurance (amount of risk reduction)
p_s	price of insurance
τ	tax rate
ϕ	cost of taxes ("leaky bucket")
$\hat{T}$	expected taxes paid

Note: We use the superscripts G and B to indicate realized values in the good and bad states respectively. The subscript T refers to an after-tax and transfer value. A variable with a "hat" (such as $\hat{y}$) means "expected" and an underlined variable (such as $\underline{y}$) means average.

CHAPTER

14

PERFECT COMPETITION AND THE INVISIBLE HAND

[The investor] intends only his own gain, and he is in this, as in many other cases, led by an *invisible hand* to promote an end which was no part of his intention. Nor is it always worse for the society that it was not part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.

Adam Smith
Wealth of Nations (1776)[1]

DOING ECONOMICS

This chapter will enable you to:

- Explain how markets are a means of processing information; that prices can provide information about the relative scarcity of goods and their value to people.
- Show that if all aspects of exchanges are covered by complete contracts, then a perfectly competitive equilibrium allocation will be Pareto efficient, and also give reasons why the complete contracts assumption is unlikely to hold.
- Explain why, in this model, market exchanges at equilibrium prices do not affect the distribution of wealth.
- Explain how an imaginary "auctioneer" with complete knowledge of all of the relevant facts and the power to prohibit out-of-equilibrium trades could set prices under which markets would clear.
- Describe a decentralized process in which buyers and sellers with limited information could bargain their way to a Pareto-efficient allocation as long as there were no impediments to efficient bargaining.
- Explain why in this decentralized process, the resulting distribution of wealth will differ from the distribution prior to market exchange.
- Understand the conditions under which bargaining among private economic actors can implement Pareto improvements or even Pareto-efficient outcomes, providing an alternative to governmental remedies such as taxes, subsidies and regulation.

14.1 INTRODUCTION: KITCHEN TALK IN MOSCOW

Figure 14.1 Nikita Kruschev and Richard Nixon at the so-called Kitchen debate, July 24, 1959.[2]

Photo by Howard Sochurek/The LIFE Picture Collection via Getty Images.

At the height of the Cold War, in a model kitchen installed as an exhibit at the US embassy in Moscow to impress Russians with the living standards of the American people, two aggressive debaters faced off: Nikita Khrushchev (KRUS-chef), Premier and First Secretary of the Communist Party of the Soviet Union, and Richard Nixon, Vice President of the United States (and later to be President). The date was July 24, 1959.

> Nixon: I want to show you this kitchen. It is like those of our houses in California.
>
> Khrushchev: We have such things.
>
> Nixon: This is our newest model . . . Our steel workers as you know are now on strike. But any steel worker could buy this house.
>
> Khrushchev: In Russia, all you have to do to get a house is to be born in the Soviet Union. You are entitled to housing . . . In America, if you don't have a dollar you have a right to choose . . . sleeping on the pavement.
>
> Nixon: We have 1,000 builders building 1,000 different houses. We don't have one decision made at the top by one government official. This is the difference.
>
> Khrushchev: . . . in another 7 years, we'll be at the level of America, and after that we'll go farther. As we pass you by, we'll wave "hi" to you.

The "kitchen debate" was a surprisingly genial episode in the competition between the Soviet Union and its allies on the one hand and the US and other capitalist nations on the other. The conflict would bring humanity to the brink of global nuclear war on more than one occasion. At issue was the superiority, even survival, of two competing economic systems: central economic planning and capitalism.

Four decades before Khrushchev and Nixon met in the model kitchen in Moscow, economists had debated whether or not a *government* could do a better job of allocating society's resources than the *market*. At issue was whether the economy should be guided by the "one decision . . . by one government official" that Nixon mentioned (a **centralized economy**), or instead by the countless decisions of the buyers and sellers, investors and workers, and others whose actions are coordinated by markets in a **decentralized economy**.

A *centralized* economy (also called a "centrally planned" economy) is one in which the government (the central planner) decides what should

CENTRALIZED ECONOMY A centralized economy is one in which the government decides what should be produced, where, by whom, and when, and how the resulting goods should be distributed among the population.

be produced, where, by whom, and when, and how the resulting goods should be distributed among the population. In a decentralized or "market" economy, on the other hand, individual firms, people, and other economic actors make choices about production and consumption with only limited coordination by a government. No economy is entirely centralized, as private decisions of people are never entirely controlled by a government. Similarly no economy is entirely decentralized, as government policies limit the feasible actions that people may take (policies to protect public health, for example), and alter the benefits and costs of particular actions (through taxes, for example).

! REMINDER In Chapter 1 what we called an Invisible Hand Game was one in which entirely self-regarding players achieve a Pareto-efficient outcome even though they interact noncooperatively, that is without coordinating their actions in any way.

The sometimes academic debate (which we return to at the end of this chapter) took a decidedly practical turn when, in 1928, the Soviet Union became the first-ever country to centralize economic decision-making. With their first five-year plan the government of the Soviet Union replaced markets as the main mechanism for determining the priorities and functioning of the economy. They substituted instead the decisions of government officials.

For the half century following the introduction of the five-year plans, the Soviet economy did well compared to some capitalist countries—Brazil and South Korea, for example—despite the devastation of the German invasion and occupation during World War II.

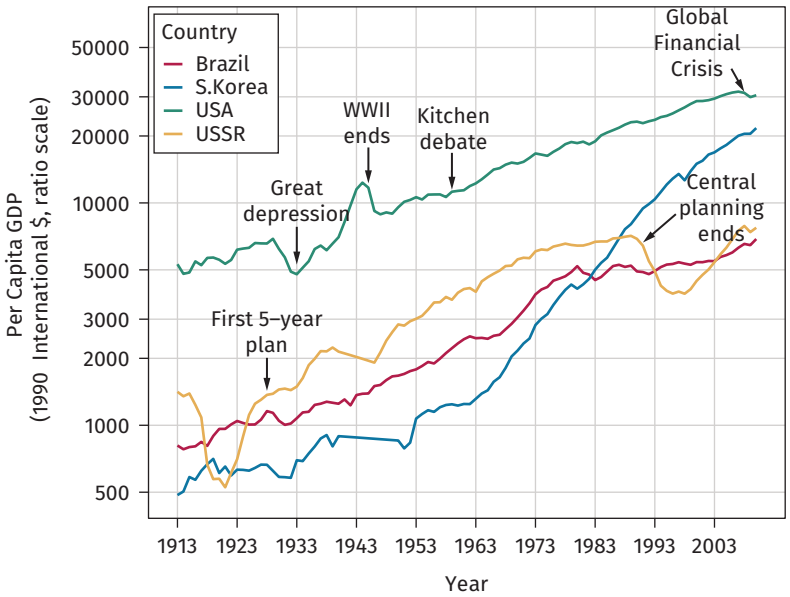
Just a year after the Soviet Union's first five-year plan was launched, the capitalist world plunged into the Great Depression, a cataclysm of economic insecurity and unemployment that barely affected the Soviet Union. Shortly thereafter, Adolf Hitler's rise to power in Germany—in part a result of the high levels of unemployment—raised the stakes of the ongoing debate. Germany's economy under fascist rule for the most part avoided the Great Depression. These events and the idea that the centrally planned economy could avoid the boom-and-bust dynamic of the capitalist economies buttressed the case against capitalism.

Opinions differ on who won the kitchen debate, but Figure 14.2—depicting the GDP per capita over time of several countries, including the US and USSR—shows that Khrushchev's claim that the Soviet Union would overtake the US simply did not happen. Notice from Figure 14.2, too, that at the height of the debate on planning versus the market the Soviet economy was growing rapidly, under the first and second five-year plans (1928–1938), while the US economy was struggling to recover from the crash of 1929. But in the long run, the centrally planned economy about which Khrushchev had boasted was outstripped by the US especially in the three decades following World War II, termed the "golden age of capitalism."

DECENTRALIZED ECONOMY In a decentralized economy, who produces what, when, how, and for whom is determined by the uncoordinated decisions of owners of individual firms, employees, and other private economic actors.

Figure 14.2 The natural logarithm of real per-capita GDP of the USA, the USSR, Brazil, and South Korea over the twentieth century. When the natural logarithm is plotted against time, the slope of the curve is the rate of growth, so aside from the Great Depression and World War II, the rate of growth of per-capital GDP in the US has been fairly constant over the period. The rate of growth in the USSR exceeded that of the US during the 1920s and 1930s but then declined to less than the US in the decade prior to the end of Communist Party rule there. South Korea's growth dramatically increased in the mid-1960s and has remained very high since.

Source: Inklaar et al.(2018).



HISTORY In the century after Adam Smith wrote, the idea of the invisible hand made its way into popular culture, as this snip from Lewis Carroll's 1865 *Alice in Wonderland* shows: "The game seems to be going on rather better now," Alice said. "'Tis so," said the Duchess: "and the moral of it is—'Oh, 'tis love, 'tis love, that makes the world go round.'" "Somebody said," whispered Alice, "that it's done by everyone minding their own business."[3]

14.2 A GENERAL COMPETITIVE EQUILIBRIUM

Long before the 1930s debate on centralization vs decentralization, economists have studied what we call a perfectly competitive general equilibrium system. A key objective has been to determine the conditions under which a set of interrelated markets (that is, a market system) that allow unimpeded entry and exit to a large number of entirely self-regarding buyers and producers could implement a Pareto-efficient allocation.

The invisible hand and general equilibrium analysis

The question dates back to Adam Smith and his, at the time, shocking idea that the uncoordinated individual pursuit of self-interest in a competitive economy would be guided by an "invisible hand" to serve the public interest.

To better understand how Smith's invisible hand might work (and why it might not) economists have modeled the economy as a whole (called

general equilibrium), rather than focusing on a single market (partial equilibrium).

- *General equilibrium:* When we analyze the equilibria of all markets simultaneously, then we engage in **general equilibrium** analysis, to understand the ways in which markets affect each other.
- *Partial equilibrium:* When we analyze the equilibrium of a single market, say the market for fish populated by fish buyers and fish sellers, then we are engaged in **partial equilibrium** analysis.

Our analysis of the market for Benetton's clothing items in Chapter 10 and credit in Chapter 12 and buying and selling risk in Chapter 13 are examples of partial equilibrium economics. Our "whole economy" model in Chapter 11 is an example of general equilibrium analysis because it brings together the equilibrium number of firms and prices in the goods market with the equilibrium wages and employment levels in the labor market. Our models using the Edgeworth box–buying and selling coffee and data for example in Chapter 4–are also simple examples of general equilibrium analysis (and we will use that method here). Attempts to clarify the conditions under which Adam Smith's surprising claims for the invisible hand might be true have occupied some of the best minds in economics since the origin of our discipline.

What they found out is of some interest for that reason alone. Kenneth Arrow and Frank Hahn put it this way:

> There is by now a long and . . . imposing line of economists from Adam Smith to the present who have sought to show that a decentralized economy motivated by self-interest and guided by price signals would be compatible with a coherent disposition of economic resources that could be regarded in a well-defined sense as superior to a large class of possible alternative dispositions.[5]

They were not doing this as advocates of the invisible hand but more out of scientific curiosity, they explained:

> That [this claim] has permeated the economic thinking of a large number of people who are in no way economists is itself sufficient grounds for investigating it seriously . . . it is important to know

GENERAL EQUILIBRIUM General equilibrium analysis is a study of two or more markets and their interactions.

PARTIAL EQUILIBRIUM Partial equilibrium analysis is the study of a single market.

Figure 14.3 Friedrich Hayek (1899–1992). An Austrian-born philosopher and economist, Hayek was a critic of the centrally planned economy and an advocate of limited government. But he had little time for the economic models often deployed in opposition to government interventions in the economy. His dissatisfaction with the model of perfect competition is clear in the head quote of Chapter 9. He won a Nobel Prize in economics for his work demonstrating (in the words of the prize committee) "how prices as such are the carriers of essential information on cost and demand conditions, how the price system is a mechanism for communication of . . . information."[4] Hayek's own words on this topic are the head quote for this final part of our book.

Photo from From LSE Library, IMAGELIBRARY/1025.

! REMINDER Remember the *classical institutional challenge* from Chapter 1? Consider how remarkable it is if a set of institutions as simple as those regulating a competitive market can address the challenge, even if only in special circumstances.

not only whether it is true but whether it *could be true*. (original emphasis)

Almost two and a half centuries after Smith, we now know that there are indeed conditions under which Adam Smith's remarkable conjecture could be true. But we will also see that, except in very special cases, market competition even if "perfect" does *not* lead to a Pareto-efficient outcome.

We begin by presenting the model of the general equilibrium of a perfectly competitive economy. The model of the equilibrium is:

- *general*: because it concerns buying and selling of two goods rather than a single good; and
- *perfectly competitive*: because at the equilibrium of this model players are price-takers—they take prices as given—and markets clear.

CHECKPOINT 14.1 Allocations in the Edgeworth box Consider Figure 14.5 and do the following:

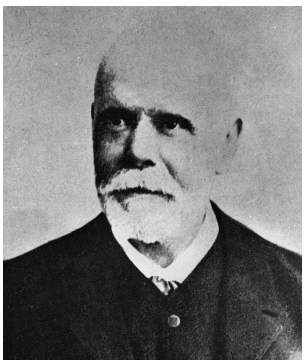
a. Reading from the figure say who gets what amount of the two goods at points **z** and t^B.

b. Find the points which are the opposite of Pareto improvements over allocation **z**, namely allocations which are worse for one of the players and not better for any.

c. How does the figure show that allocations t^B and t^A are Pareto improvements over allocation **z**?

Figure 14.4 Léon Walras (1834–1910). The idea of a perfectly competitive general equilibrium was developed by Walras, a French economist with a passion for social justice and for mathematics. His father had been a schoolmate of Augustin Cournot, whose model of competition among firms you studied in Chapter 9. Along with Alfred Marshal, Walras (pronounced val-RA) is considered the founder of the "neoclassical school" of economics that in most countries was the predominant approach in economics during the twentieth century, substantially extended by the introduction of Keynesian macroeconomics following the Great Depression.[6]

Photo by Album/Alamy Stock Photo.



A model of perfectly competitive general equilibrium

We return to the model that you studied in Chapter 4. There we showed how goods might be exchanged in an Edgeworth box under a variety of rules of the game, including cases in which one of the two players has first-mover (price-setting) advantage or "take-it-or-leave-it" (TIOLI) power (price and quantity setting).

Here we study a symmetric interaction among players, neither of whom has any special advantage in bargaining. They will exchange goods at prices that they take as given and that clear the market. We will not ask yet what determines the price, but will instead be interested in the properties of the resulting equilibrium, and particularly in the conditions under which it will be Pareto efficient.

The model is illustrated by the Edgeworth boxes shown in Figure 14.5 for two traders, Adamo and Beatriz, who are allocating some fixed total number of goods, similar to our illustration with Ayanda and Biko in Chapter 4 but under different rules of the game. The total amount of the two goods available ($\bar{x}$ and $\bar{y}$) determine the dimensions of the box: $\bar{x} \times \bar{y}$ (in this case, 10×15). The initial allocation of the two goods is indicated by point **z**: Adamo

has a lot of coffee and Beatriz has a lot of data. These two bundles of coffee and data are referred to as Adamo's and Beatriz's endowments.

We term the endowment the fallback option for two traders because, as in our previous use of the Edgeworth box, it is the outcome that they will experience if they do *not* trade. The participation constraints of the two are the indifference curves u_z^A and u_z^B which show all the allocations resulting from a trade that each values as highly as not trading at all.

They both can be better off if Adamo exchanges some of his coffee for some of Beatriz's data. This is shown by the Pareto-improving lens bounded by their participation constraints. The lens is the set of all allocations that are Pareto superior to their endowments at **z**.

The Pareto-improving lens narrows down the set of possible trades that A and B might make. Remembering that because their exchange must be voluntary, points not in the Pareto-improving lens cannot occur. The reason is that at least one of the two would be worse off as a result, and so would refuse to trade.

Every point in the Pareto-improving lens is an allocation that both A and B could agree to if the only alternative was no trade at all. The reason is that moving from point **z** to that point would make both better off (or at least one of them better off and the other not worse off).

To understand the trades that each might actually make, we need to consider the prices at which the two exchange their goods. To do this we adopt an indirect strategy. We *do not* ask where the prices come from, as we did in Chapter 4.

Instead we work backward from what is required for an outcome of the traders' interactions to be a Nash equilibrium, specifically, the properties that a price must have if it is to be an equilibrium consistent with the rules of the game of a perfectly competitive market. So we:

- consider all possible hypothetical prices, with the same prices determining the rate at which the two will exchange goods;
- suppose that the traders take each price as given (we don't ask why they would do this);
- analyze the trades that they would like to make at each of these prices using the price-offer curve that you studied in Chapter 3;
- find one of these hypothetical prices for which the pair of price-offer curves is mutually consistent (so that, for example, the amount of coffee that A wishes to sell at that price is the same amount as B wishes to buy).
- this is a general equilibrium because the trades indicated by the price–offer curves are the best that each can do at any given price, so neither trader could benefit by offering to change the quantity of goods that they are transacting.

! REMINDER Perfectly competitive means that the Nash equilibrium of a market is a situation in which supply equals demand and neither buyers nor sellers can benefit by altering either of the two strategies available to them, namely, the prices at which they offer to transact or the amounts they seek to transact.

! REMINDER An A is Pareto superior to allocation B if at allocation A at least one person is better off than at allocation B and no person is worse off than at A. The terms "Pareto improvement over" or "Pareto dominates" are equivalent to "Pareto superior to."

HISTORY The perfectly competitive general equilibrium model is sometimes termed the Walrasian (val-RAY-sian) model (after one of its originators), or in its more mathematical variants, the Arrow—Debreu model (after two economists who contributed to its twentieth-century development).

› EXAMPLE In this video (tinyurl.com/y6eaqpgn) Kathryn Graddy explains how real markets work, illustrated by a fish market (from the CORE project. www.core-econ.org).

! REMINDER A trader's price-offer curve shows the greatest amount of one good that a person will give up in return for each amount of another good (or money). In the Edgeworth box, the price means the price of the x good in terms of how many y goods are required to purchase one unit of the x good. This means that a "price" is the price of coffee. So a price like $p_j = 1$ in Figure 14.6 means that 1 kilogram of coffee costs 1 gb of data. A higher relative price of coffee would be 2 gb of data for 1 kg of coffee (a steeper line), a lower price would be 0.5 gb of data for 1 kg of coffee (a flatter line). Points on the price–offer curve are thus the maximal willingness to pay (in gb of data) for 1 kg of coffee.

! REMINDER If you would like to review the derivation of a price-offer curve for a price-taking actor return to Chapter 3 (section 3.9).

Because price-offer curves play a central role in this line of reasoning Figure 14.6 shows A's price-offer curve. The difference between Adamo's endowment **z** and any point on his price-offer curve represents a rearrangement of his ownership of goods (giving up some coffee, getting more data) that he would be *willing* to implement through voluntary exchange. For example point **j** indicates that in order to get 4 gb of data (so he would then have 5 gb rather than just one in his endowment), he would at most be willing to give up 4 kilograms of coffee (he would as a result have 5 kilograms of coffee rather than the 9 kilos at his endowment). This information tells us how much Adamo will offer to sell (and be willing to buy) for each of the hypothetical prices. There is no money in this model so a price is a ratio of one good to the other, such as how many gb of data Adamo will receive in return for 1 kg of coffee is given by the (negative of the) slope of a line passing through point **z**. One of these possible prices for the two goods is indicated by the price line p_j in Figure 14.6 at which Adamo gets 1 gb of data in exchange for 1 kg of coffee, so the "price of coffee" is 1 (gb of data) and the price line p_j has a slope of –1.

Figure 14.5 The Edgeworth box, participation constraints, and the Pareto-improving lens. Remember, the dimensions of the box are given by the total amounts of the two goods to be divided among the two players, that is 10 kg of coffee and 15 gb of data. Each trader—Adamo and Beatriz—has an endowment shown by point **z** with corresponding participation constraints u_z^A and u_z^B. The yellow-shaded area between these two curves is made up of all of the allocations that are better for both than the endowment allocation, **z**.

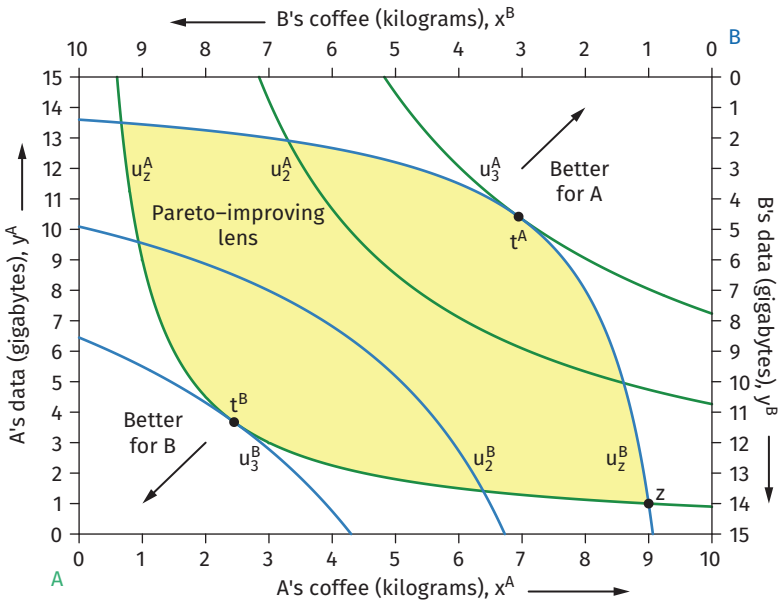
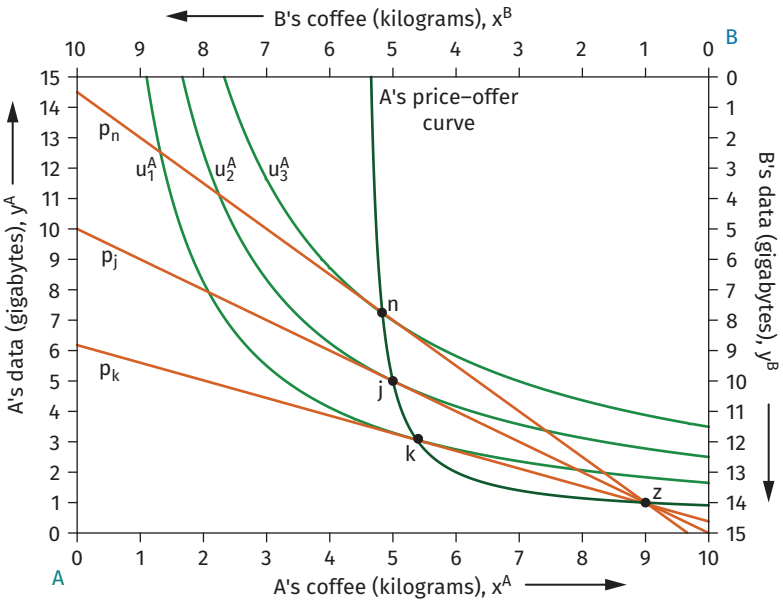


Figure 14.6 A's price-offer curve with three different prices. The price of coffee (gb of data required to purchase one kg of coffee) is the (negative of) the slope of each of the three price lines p_k, p_j, and p_n. Given some low price of coffee such as $p_k = 0.58$, A will sell less coffee (meaning that he will buy fewer gb of data indicated by point **k**) than at a higher price $p_j = 1$ (point **j**).



! REMINDER The orange price lines are a kind of budget constraint for Adamo. The negative of its slope is the marginal rate of transformation of kilograms of coffee into gigabytes of data. So at points **k**, **j**, and **n** where the price line is tangent to the indifference curves Adamo has implemented the individual utility-maximizing rule, equating $mrs = mrt$.

Recall from Chapter 4 a price-offer curve is constructed by treating the price line as a *constraint* on the trader's utility-maximizing process. So in Figure 14.6 think of each of the three price lines p_k, p_j, and p_n as alternative budget constraints corresponding to the three prices. We then find a point of tangency between each of the given price lines and one of Adamo's indifference curves. This means that each point on his price-offer curve—such as **n, j,** and **k**—is a tangency between the price line through that point and an indifference curve.

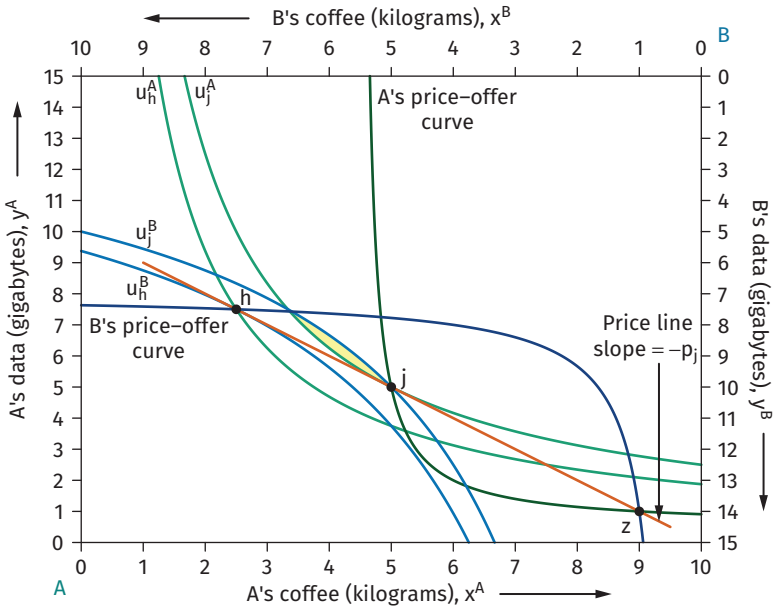
14.3 MARKET CLEARING AND PARETO EFFICIENCY

How will the two traders respond to some hypothetical price?

A non-clearing market

To answer the question we need to refer to Beatriz's price–offer curve. In Figure 14.7 we show the desired trades of the two when the price is p_j. Here is what we find out:

Figure 14.7 A non-clearing market in an Edgeworth box. The blue and green price-offer curves show that at the relative prices of the two goods given by the slope of the orange price line each trader will choose a corresponding allocation given by the intersection of the price line and their price–offer curve. The price line (p_j) results in excess demand for coffee and excess supply of data. So if the price, for some reason, were p_j, the transaction indicated by allocation **j** (and not **h**) would be implemented because the amount transacted is determined by the short side of the market, that is, the person whose desired level of transactions is least, in this case Adamo. There would remain unrealized mutual gains shown by the small yellow-shaded Pareto-improving lens.



- Adamo would not want to sell much coffee (indicated by point **j**, the intersection of the price line and his price–offer curve); while
- Beatriz would want to purchase more coffee (indicated by point **h**, the intersection of the same price line and her utility-maximizing choice).

This means that at price p_j the markets *do not clear*:

- At the price p_j, the demand for coffee (by Beatriz) exceeds the supply (Adamo's offer); or, what is the same thing,
- Beatriz's *supply* of data exceeds Adamo's *demand* for data.

At the price p_j, Adamo is on the short side of a non-clearing market, the side for which the number of desired transactions is least. Beatriz, who wants to exchange more goods than Adamo, is on the long side. As you saw in Chapter 9 and as you know from the way the labor market works,

the transaction that will occur is determined by the amount of trade the short-sider desires.

This is because the exchange is voluntary. The long-sider, Beatriz, who would like to trade more at the Beatriz-favorable price p_j, cannot require Adamo to trade more than he would like. In this case Adamo is like the employer, the short side in the labor market: the number of people hired is determined by the employers not by those wishing to find a job. And the amount of coffee and data that will be exchanged is determined by Adamo.

Could point **j** be a Nash equilibrium? To see that it cannot, imagine that you are Beatriz. At the price p_j you would like to buy more of Adamo's coffee but he is not willing to sell more. What would you do? You would offer him a somewhat higher price of coffee and he would agree to sell (as you can see from Adamo's price–offer curve, raising the price, e.g. to p_n will get Adamo to sell more). Both would then be better off. Or Adamo, thinking along the same lines, could have just offered to sell more coffee if Beatriz would pay a bit more of her data per kg that she gives up.

So allocation **j** resulting from the price p_j cannot be a Nash equilibrium. Exchanging goods at that price cannot be an equilibrium. The fact that at point **j** Beatriz could offer Adamo terms under which exchanging more of the goods would be mutually beneficial means that the allocation at point **j** could not be Pareto efficient. The following reasoning confirms this:

- At point **j**, Adamo's indifference curve u_j^A is tangent to the price line (that's why Adamo wanted to sell that amount).
- At point **j**, Beatriz's indifference curve u_j^B is *not* tangent to the price line (that's why Beatriz wanted to buy more coffee than Adamo wanted to sell).
- Therefore the two indifference curves have different slopes at point **j**, which means that they intersect and so they cannot be tangent.
- But this means that the allocation at point **j** cannot be Pareto efficient: Above and to the left of point **j** you can see a small Pareto-improving lens representing allocations with more goods being exchanged.

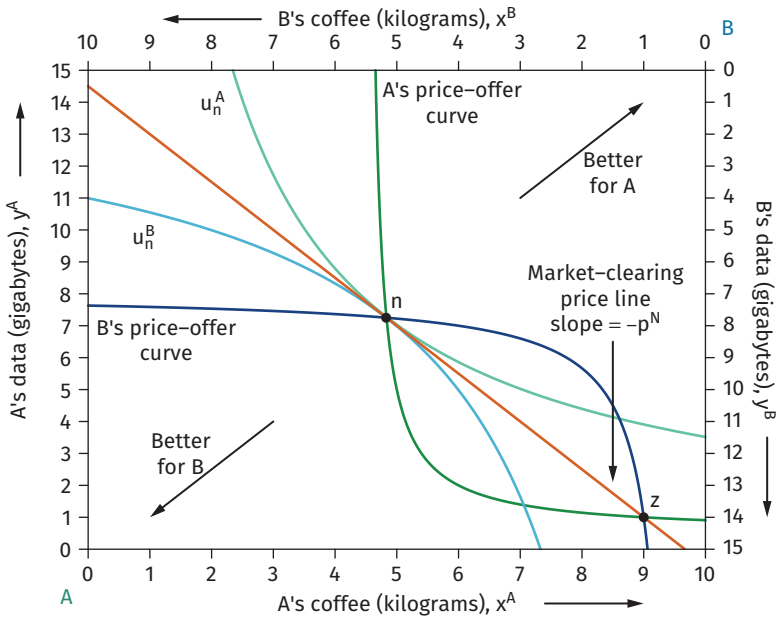
We have already shown that the allocation **j** and the price p_j cannot be a Nash equilibrium; we can now also conclude that it is also not Pareto efficient.

CHECKPOINT 14.2 **Non-clearing markets** Using Figure 14.7 explain what exchange would occur if the price was p_j: who trades how much of what to whom? What trade would Beatriz have preferred to make at that price?

A Pareto-efficient Nash equilibrium

Is there a Nash equilibrium allocation that is also Pareto efficient? To find the Nash equilibrium price, think about the definition: a Nash equilibrium is a mutual best response. The Nash equilibrium must be a point on each of

Figure 14.8 A general competitive equilibrium. The blue and green price-offer curves show that at the relative prices of the two goods given by the slope of the orange price line each trader will choose a corresponding allocation given by the intersection of the price line and their price–offer curve. The price line, p^N, shows how Adamo's desired sales of coffee to Beatriz are equal to Beatriz's desired purchases of coffee. At the same price Beatriz's desired sales of data to Adamo are equal to Adamo's desired purchases of data. The market clears at p^N and there is neither excess demand nor excess supply of either good.



! REMINDER The $mrs^A = mrs^B$ rule. An allocation among *two or more people* is *Pareto efficient* if their two indifference curves are tangent at that allocation. The $mrs = mrt$ rule applies to *an individual* requiring that the slope of her indifference curve be tangent to the feasible frontier that is the constraint on her *individual utility maximization* process.

their price–offer curves, in other words the intersection of their price-offer curves.

To see this, turn to Figure 14.8, in which we show the two price-offer curves. You can see that there is a price of the two goods–shown by the price line p^N–under which three conditions hold: *market clearing*, *Pareto efficiency*, and *Nash equilibrium*.

First, *markets clear*: at point **n** Adamo's supply of coffee is equal to Beatriz's demand for coffee and Beatriz's supply of data is equal to Adamo's demand for data.

Second, the allocation is *Pareto efficient*. Here is why.

- At point **n**, Beatriz's indifference curve is tangent to the price line; it is on Beatriz's price-offer curve which are all points of tangency between a price line and the person's indifference curve ($mrs^A = mrt^A$).
- By the same reasoning, at point **n**, Adamo's indifference curve is tangent to the same price line ($mrs^B = mrt^B$).

- If both people's indifference curves are tangent to the same price line at point **n** (the law of one price), they must have the same slope.
- Therefore the two indifference curves must be tangent to each other, meaning that the allocation at point **n** is Pareto efficient because their marginal rates of substitution are equal (the $mrs^A = mrs^B$ rule).

! REMINDER The law of one price states that in equilibrium identical goods or services will transact at the same price.

Third, the allocation is a *Nash equilibrium*: as a mutual best response, Adamo is doing the best he can given the quantities of data and coffee Beatriz is willing to transact, and Beatriz is doing the best she can given the transactions that Adamo is willing to implement. Neither can benefit by offering to trade at some price other than p^N, so being a price-taker is also the best they can do.

! REMINDER This is another case of an allocation that implements the $mrs^A = mrs^B$ rule as the unintended result of the the two actors using the $mrs = mrt$ rule to maximize their own utiltiy.

The model we have used is very simple, but extensions of the model to cases with many buyers and sellers of many goods in which they are not simply "endowed" an initial allocation but *produce* the goods they exchange yields the same result.

Because the buyers and sellers are price-takers and we made use of the law of one price, the resulting allocation is sometimes referred to as a perfectly competitive general equilibrium. But the model does not address the question: *How* did competition determine the price p^N?

What was shown instead is that there *exists* a price like p^N resulting in an allocation with the three properties: *market clearing*, *Pareto efficiency*, and *Nash equilibrium*. The gap in the model—the lack of an explanation of how the equilibrium prices come about—has proven difficult to repair without dropping the assumption that people act as price-takers. We will suggest a way that this can be done in section 14.10 by adopting a more realistic model of how competition works.

14.4 PRICES AS MESSAGES, MARKETS AS INFORMATION PROCESSORS

Even keeping this caveat in mind, the general equilibrium model supports a surprising result: there exists a set of prices such that people independently maximizing their own utility would implement a Pareto-efficient allocation *if these prices were somehow the ones at which they were constrained to transact*.

The italicized text is important, and we will later ask: Can we expect markets (even if perfectly competitive) to produce prices like p^N? But for now we assume that set of relative prices like p^N is (somehow) known and are the prices at which goods transact.

To see the advantages of the decentralized coordination of the allocation of goods to people by means of prices, imagine that markets were outlawed, and that you were tasked with allocating two goods, x and y, between Adamo (A) and Beatriz (B) in a Pareto-efficient way. We will use this two-

HISTORY Looking ahead, the main theme of Friedrich Hayek's 1945 critique of the centralized economy was that it would be impossible to collect all of the information necessary to plan a centralized economy. People buying and selling on competitive markets unknowingly produce the information that they need (prices) not intentionally, but as an unintended by-product of their privately motivated actions.[6]

person case as a lens for studying the relevant case, in which there are thousands of Adamos and Beatrizes.

Consumption: Unintentionally equating marginal rates of substitution

You know that Pareto efficiency requires allocations that satisfy the $mrs^A = mrs^B$ rule. That is, Pareto efficiency requires that people's indifference curves are tangent or:

$$\text{Slope of A's indifference curve} \quad = \quad \text{Slope of B's indifference curve}$$

$$mrs^A(x^A, y^A) = \frac{u_x^A}{u_y^A} \quad = \quad \frac{u_x^B}{u_y^B} = mrs^B(x^B, y^B) \tag{14.1}$$

To do this you would have to know the indifference curves of the two, a challenging task in the case of just two people, and virtually impossible for an entire economy. You would have to know (which means: devise ways of finding out) the utility functions of each person.

But if the perfectly competitive general equilibrium prices were in force the problem would be a lot simpler. Here is why.

- Each person maximizes utility *constrained by* a budget constraint.
- To do this they use the $mrs = mrt$ rule, equating the slope of their indifference curve to the slope of their budget constraint, which is given by the (negative of) the price ratio $\frac{p_x}{p_y}$.
- But that price ratio is the same for all buyers and sellers (the law of one price, again).
- This means that they are equating their marginal rates of substitution.

! REMINDER Remember that a *relative price* can be thought of as the *ratio of prices* of two goods, e.g. p_x/p_y, which shows the price of one good (x) *relative to* the price of another good (y). We have also thought of relative prices as the *marginal rate of transformation* or the *opportunity cost* of choosing one good over another.

Because all traders are doing the same thing—constrained by the same ratio of prices—each unknowingly equates their own marginal rate of substitution to the marginal rates of substitution of all the other traders. In other words:

$$mrs^A(x^A, y^A) = \frac{u_x^A}{u_y^A} = mrt^A = \frac{p_x}{p_y} = mrt^B = \frac{u_x^B}{u_y^B} = mrs^B(x^B, y^B) \tag{14.2}$$

Equation 14.2 says that even though they did not intend it, by pursuing their own interest—maximizing their utility subject to a budget constraint by implementing the $mrs = mrt$ rule—they implement an allocation in which their marginal rates of substitution are equal, and which, therefore, is Pareto efficient.

What information did Adamo and Beatriz need to have to unintentionally implement an efficient outcome? Not much. Each had to know their own preferences (not the preferences of the other) and the prices of the goods.

Production: Unintentionally equating marginal rates of transformation

We now extend the model to include production. To see what Pareto efficiency requires return to Figure 6.7. There we showed that a person optimizes her utility by producing and consuming such that:

$$\begin{aligned} \text{marginal rate of substitution} &= \text{marginal rate of transformation} \\ \text{which requires:} & \\ \text{slope of indifference curve} &= \text{slope of feasible frontier} \end{aligned}$$

Each actor—Adamo and Beatriz—not only consumes the goods x and y, they *produce* the goods x and y. To produce the goods, they make choices about which technologies to use, the scale of output, the mix of inputs resulting in the marginal costs of producing the goods $c_x^A, c_y^A, c_x^B, c_y^B$.

Remember from Chapter 8 that in a perfectly competitive equilibrium firms act as price-takers, and as a result they maximize profits by producing up to the point that marginal cost equals the given price of the good. So, considering each Adamo and Beatriz as producers of the goods, the ratio of prices will equal the ratio of marginal costs for both. That is:

$$\text{A's mrt} = \frac{c_x^A}{c_y^A} = \frac{p_x}{p_y} = \frac{c_x^B}{c_y^B} = \text{B's mrt} \tag{14.3}$$

!REMINDER In Chapter 3 you saw how the person, doing the best she can, equated the slope of her indifference curve (the marginal rate of substitution) to the slope of the feasible frontier (the marginal rate of transformation or opportunity costs).

Bringing together production and consumption

Having added production by using each trader's costs of production, efficiency now requires that the marginal rate of substitution in people's indifference curves be equal to the marginal rate of transformation in the production of the two goods. We now have:

$$\frac{u_x^A}{u_y^A} = \frac{c_x^A}{c_y^A} = \frac{p_x}{p_y} = \frac{c_x^B}{c_y^B} = \frac{u_x^B}{u_y^B}. \tag{14.4}$$

All traders optimize with respect to the same relative prices. They therefore equate their own marginal rate of substitution in consumption as well as their marginal rate of transformation in production (the ratio of marginal costs) to the other trader's marginal rates of substitution and transformation. They thereby implement a Pareto-efficient allocation.

Equation 14.4 shows that prices convey two kinds of information about goods:

- *People's subjective value*: How *valuable* it is to people who consume or use it (measured by their marginal rates of substitution, which is their willingness to pay); and
- *Cost*: How *costly* it is to produce it (measured by the ratio of their marginal costs, which is the same thing as the marginal rate of transformation).

M-CHECK In Equation 14.3 the ratio of marginal costs is equal to the price ratio and the ratio of marginal costs is the same for the two producers, A and B. The marginal cost ratio is the opportunity cost of good x in terms of good y—how much of y each would have to give up to get produce an additional unit of good x—or, the marginal rate of transformation.

Equation 14.4 shows that the perfectly competitive equilibrium equates these two aspects of **scarcity**: the marginal rate of transformation (the ratio of marginal costs, the relative costliness of the goods) and the marginal rate of substitution (the ratio of marginal utilities, the relative value of the goods to users) are both equal to the ratio of prices and therefore are equal to each other.

EXAMPLE Example If (contrary to Equation 14.4) the opportunity cost of producing x were greater for A than for B, then this would mean that B has a comparative advantage in producing x, and that they both could do better if B would produce more of x and A produced more of y. This would be a case of specialization according to comparative advantage studied in Chapter 6.

As a result, perfectly competitive equilibrium prices send messages to people who buy and produce goods. If a drought in the American Midwest has decimated the wheat crop, then the price of bread will rise, and the message to the person is "maybe put potatoes on the table tonight instead of bread." If the price of tin has risen due to dwindling reserves of the metal, the owner of the firm hears: "consider redesigning your product using plastic."

Prices do more than convey messages about scarcity: they also provide the motivation to act on the information. If bread or tin is more expensive, the person who buys it or the producer using tin as an input will save money by shifting to an alternative.

To summarize:

Price = Message about scarcity + Motivation to act on the message

REMINDER The fishermen in Chapter 5 incurred a private cost of fishing, but they also imposed costs on the other fishermen due to depletion of fish stocks. The reason for the coordination failure in that case—overfishing—was that the fishermen took account of *only* the private cost of their own choice to fish more. This did not accurately measure the social cost of fishing because it did not include the external cost that one fisherman fishing more imposed on the other.

But there is a hitch: the prices have to be right. Imagine that instead of rising due to a drought, the price of bread had *fallen* because farmers are now using a new fertilizer, which, when it runs off into nearby rivers and streams, *destroys* the aquatic environment and the tourism or commercial fisheries that depends on it. The message sent by the lower price of bread would be "let's have bread tonight rather than pasta." But the message would be mistaken: the lower price does not measure the full cost of putting bread on the table.

CHECKPOINT 14.3 **Law of one price** Explain why the Law of one price is essential to showing that individuals will consume bundles of goods such that their marginal rates of substitution are equated (as in Equation 14.2).

14.5 PARETO EFFICIENCY AND THE INVISIBLE HAND: THE FIRST WELFARE THEOREM

We can generalize from the bread example by returning to the distinction between **private costs** and benefits and **social costs** and benefits first introduced in Chapter 5. The private cost (marginal or average) is the cost that the decision maker bears as a result of some action that he or she

SCARCE A good is scarce if it is valued by people (they would prefer to have more of it) and there is an opportunity cost of acquiring more of it.

takes. The social cost is the private cost plus any costs imposed on others as negative external effects. That is:

$$\text{Marginal social cost} = \text{Marginal private cost} + \text{External cost} \quad (14.5)$$

For prices to send the right messages they must:

- measure how much goods contribute to people's satisfaction (utility); and
- how much they cost society to produce (the marginal social costs including the negative external effects).

We can use Equation 14.5 to understand whether an outcome is Pareto efficient or not. For the perfectly competitive general equilibrium allocation to be Pareto efficient Equation 14.4 is *not* sufficient. The reason is that what perfect competition (along with the law of one price) ensures is that the ratio of prices is equal to the ratio of *private marginal costs*. This is what Equation 14.4 says.

But what is required for Pareto efficiency is that the price ratio equals the ratio of marginal *social* costs. That is, Pareto efficiency requires that the price equals the *entire* costs the producer's marginal private costs plus the external costs the producer imposes on others.

In the example of bread prices and wheat farming above, this means that the price of bread would have to include the costs that the new fertilizer imposes on fishing and tourism. In the example of "conspicuous consumption" in Chapter 7 building a luxury home in an otherwise modest neighborhood imposes a disutility on the neighbors. This disutility is an uncompensated negative external effect, so it would have to be included in the price of the new home for home sales to result in Pareto–efficient outcomes.

In this model (that is, assuming perfect competition) for marginal private costs to equal marginal social costs it must be the case that:

- *No missing markets*: there are markets for any aspect of an exchange that people value (pro or con), so that everything that matters has a price.
- *No uncompensated external effects*: when people exchange goods and services any aspect of the production and use of the good that affects anyone's well-being (including those not party to the transaction) is measured in the price.

When these two conditions are both met, we say contracts are complete. In discussions of general equilibrium this is sometimes called the *market completeness assumption*. Where markets are entirely missing we do not have complete contracts because *there are no contracts*.

Where these two conditions are met, the equilibrium of a perfectly competitive economy will be Pareto efficient. This is expressed by what

! REMINDER A Pareto-efficient outcome can be implemented if policies are adopted so that actors bear the otherwise external costs that their actions impose on others. The fishermen in Chapter 5 reduced their fishing hours when they were forced to pay (in taxes) for the external costs of their actions on others: they internalized the external costs and therefore reduced their hours spent fishing. They therefore achieved a Pareto-efficient outcome.

! REMINDER If contracts are complete, there are:

- *No missing markets*: there is a market in every good and service that people value, so that everything that matters has a price and
- *No uncompensated external effects*: when people exchange goods and services any aspect of the production and use of the good that affects anyone's well-being (including those not party to the transaction) is measured in the price.

HISTORY This result was proven independently by Kenneth Arrow and Gerard Debreu in 1951 and published in a paper they coauthored three years later.[7]

is called the first theorem of welfare economics but which might better be called, honoring Adam Smith, the Invisible Hand theorem.

First theorem of welfare economics: A *perfectly competitive equilibrium of an economy with complete contracts is Pareto efficient.*

When contracts are complete, the competitive general equilibrium of the economy is Pareto efficient because prices send the right message: the marginal private cost of a good to the firm or consumer is exactly equal to the marginal social cost to society of having another unit of that good available. We can summarize the two requirements that if met ensure that an equilibrium be Pareto efficient:

- *Perfect competition:* All markets must be *perfectly competitive*, so that in equilibrium each buyer and seller take prices as given, and the law of one price holds. This insures that *prices will be equal to the marginal private cost* of their production.
- *Contracts must be complete*: so that there are no uncompensated external effects. This insures that *marginal private cost equals marginal social cost.*

If both conditions hold, then the ratio of the prices of any two goods will be equal to the ratio of their marginal social costs.

Using the superscripts P to refer to marginal private costs and S to refer to marginal social costs, we can extend Equation 14.4 to take account of the requirement that marginal private costs = marginal social costs:

$$mrs(x,y) = \overbrace{\frac{u_x}{u_y} = \frac{p_x}{p_y} = \underbrace{\frac{c_x^p}{c_y^p} = \frac{c_x^s}{c_y^s}}_{\text{Complete contracts}}}^{\text{Perfect competition}} = msrt(x,y) \tag{14.6}$$

Equation 14.6 is the familiar condition for Pareto efficiency stating that the marginal rate of substitution in consumption (*mrs*) must be equal to the marginal rate of transformation in production. But here we use *msrt* to mean the marginal *social* rate of transformation, meaning the ratio of marginal social costs, that is, taking account of both private costs and the costs imposed on others by the uncompensated external effects of the production of the good.

Equation 14.6 also makes it clear that the equality of the *mrs* to the ratio of prices and the ratio of marginal private costs and marginal private rate of transformation *mrt* requires perfect competition while the private and marginal social rate of transformation requires contracts to be complete.

FIRST WELFARE THEOREM A perfectly competitive equilibrium of an economy with complete contracts is Pareto efficient.

We will see in the next section that these requirements are not likely to be met.

CHECKPOINT 14.4 **First Theorem of Welfare Economics** Explain why both perfect competition and complete contracts are required by the theorem (Equation 14.6).

HISTORY Gerard Debreu, who along with Kenneth Arrow proved what we have called the Invisible Hand theorem, in 1984 told the French journal *Le Figaro* "The superiority of the liberal economy [meaning substantially unregulated competitive markets] is incontestable and can be demonstrated mathematically…" He was referring to his own justly famous theorem.[8]

14.6 MARKET FAILURES DUE TO UNCOMPENSATED EXTERNAL EFFECTS

At first glance, the first welfare theorem appears to vindicate Adam Smith's conjecture that competitive markets would ensure that traders be "led by an invisible hand to promote an end which was not part of" the participants' intentions.

But few economists take the theorem as demonstrating that real-world market institutions in fact implement Pareto-efficient outcomes. The reason is that the above two conditions required by the theorem—perfect competition and complete markets—are not even remotely descriptive of the world's economies today.

To develop more empirically grounded models we have introduced price-*making* actors—owners of firms in markets with limited competition, as well as principals interacting under incomplete contracts with agents in non-clearing markets for credit, labor, and goods of variable quality—in Chapters 8, 9, 10, and 11. Here we draw general conclusions from these examples.

Uncompensated external effects

The two fishermen overfishing a lake in Chapter 1 were our first example of a coordination failure—Nash equilibria that are Pareto inefficient due to the lack of coordination among the economic actors involved. In Chapter 5 we returned to the problem, showing how the two fishermen determine the hours they devote to fishing by balancing their marginal private costs (their own disutility of effort in fishing) and the marginal benefits (how much each additional hour added to the catch). The coordination failure among the fishermen was due to the fact that each fisherman—in deciding how much to fish—did not include in the costs of their own fishing the negative effect of their fishing on how much fish the other fisherman caught. These additional costs are an uncompensated external effect.

Another example: when fuel costs are low, more people decide to drive to work rather than taking public transport or choosing to ride their bicycles. The information conveyed by the low price does not include the carbon emissions and other environmental costs of deciding to drive. The effects on the decision maker are termed *private* costs and benefits, while the total effects, including costs inflicted or benefits enjoyed by others, are *social* costs and benefits.

! REMINDER Costs inflicted on others (pollution and congestion that are worse because you drive to work) are termed negative external effects; while uncompensated benefits conferred on others are positive external effects.

! REMINDER In section 1.16 we showed how an application of tort law in a prisoners' dilemma game—requiring the defector to pay the cost that they impose on a cooperator—could convert the interaction to an invisible hand game.

We can understand why these and other market failures are common by thinking about how they could be avoided. How could the cost of driving to work accurately reflect all of the costs incurred by anyone, not just the private costs incurred by the decision maker? The most obvious (if impractical) way would be to require the driver to pay everyone affected by the resulting environmental damage (or traffic congestion) an amount equal to the damage inflicted.

This is impossible to do, but it sets a standard of what has to be done or approximated if the "price of driving to work" is to send the correct message. A similar approach applies if you drive recklessly on the way to work, skid off the road, and crash into somebody's house. Tort law (the law of negligence) in most countries would require you to pay for the damage to the house. You are held liable for the damages so that you would pay the cost you had inflicted on another.

Knowing this, you might think twice about driving to work (or at least slow down a bit when you are late). But while tort law in most countries covers some kinds of harm inflicted on others (reckless driving), other important external effects are not be covered by tort law (adding to air pollution of congestion by driving your car).

Table 14.1 gives examples of important types of market failure. In all of these examples there are *uncompensated external effects*. For each of them, it is the case that either:

- the *private costs* to the decision maker of an activity differ from the *social costs* including both the private costs and the negative external effects on others, in this case the private costs are less than the social costs; or
- the *private benefits* to the decision maker differ from the social benefits, in this case the social benefits exceeding the private benefits;

or both of the above.

An example we have already mentioned illustrates how Table 14.1 is constructed. A farm uses pesticides that contaminate the local water supply. The farm's private costs for its pesticide use do not include the external cost the farm's owners impose on fishermen and private citizens who use the river water or other water sources contaminated by the use of the pesticides. Because the farmers do not pay the full cost of the pesticides (social cost = private cost + external cost), they *overuse* pesticides.

Now consider an example not from the table, but describing the kind of neighborhood you might live in. A homeowner plants a beautiful garden that her neighbors can see. She enjoys the flowers she has planted at her house, but the social benefits include not only her enjoyment of the flowers but also the enjoyment of neighbors who pass by. Because homeowners do not reap the full benefits of maintaining a beautiful home (social benefits = private benefits + external benefit), many will under-provide home maintenance and garden cultivation.

Table 14.1 Examples of uncompensated external effects and market failures.

Type of problem	The decision	Uncompensated external cost or benefit	Market failure
Public good	A firm invests in R&D	Other firms can use the innovation	Too little R&D
Public bad	You take an international flight	Resulting increased carbon emissions imposes climate change costs on others	Overuse of air travel
Negative external effect	A farm uses pesticides that contaminates water	Damage to fishing industry	Overuse of pesticide
Positive external effect	A firm trains a worker	Another firm benefits if the worker quits	Too little worker training
Common property resource	You travel to work by car	Congestion for other road users	Overuse of roads and highways
Moral hazard Hidden action (labor)	An employee on a fixed wage decides how hard to work	Hard work increases her employer's profits	On the job effort is too low
Moral hazard Hidden action (credit)	A borrower decides how much risk to take	Lender is exposed to default risk	Excessive risk
Adverse selection (insurance)	People with serious health problems are more likely to buy insurance	Insurer's profits fall insurance prices rise	Insurance costs rise so only the ill seek to purchase insurance
Veblen effects (conspicuous consumption as a public bad)	A wealthy person engages in conspicuous consumption	Other people feel that their income is inadequate	Excessive working and competitive consumption to catch up

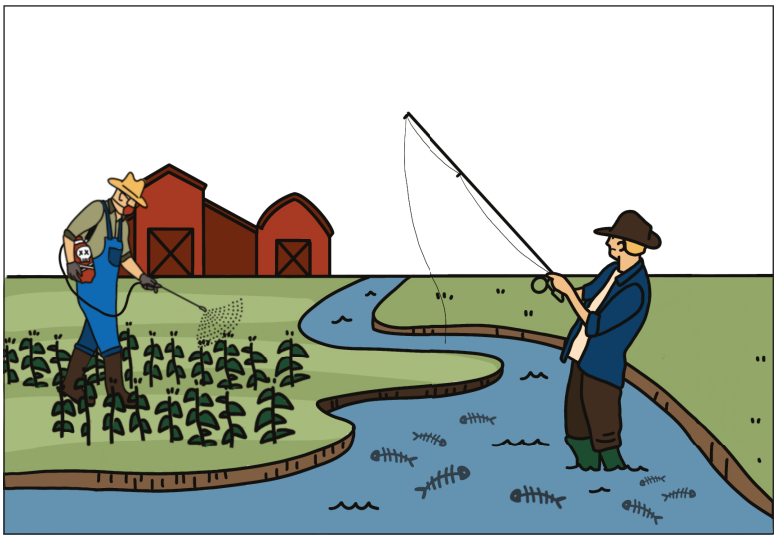
Incomplete contracts and missing markets

If the gardener somehow owned the sight of her flowers which she could sell or rent to people walking by, or the neighbors of the polluting farmers owned the clean water around their house which they could sell to the polluting farmers, the market failure could be avoided. Market failures occur because the external benefits and costs of a person's actions are not owned by anyone.

Think about waste: if you redecorate your house and you tear up the floor or knock down a wall, you own the debris and you have to dispose of it, even if you need to pay someone to take it away. But this is not the case with pollutants from a farm or loud music you might play at night. You do

Figure 14.9 Market failures and other coordination failures resulting from uncompensated external costs or benefits due to incomplete contracts. The farmer using pesticides or herbicides imposes a negative external cost on the person who depends on fishing for their livelihood.

Image credit: Anmei Zhi.



not have a contract with the farm company specifying at what price you are willing to accept contaminated water. You cannot bargain with your neighbor about the price you will charge to give him the right to play music after 10 p.m. (though strange as this last case may seem, we will return to it later in the chapter).

If the problem is uncompensated external effects, why don't countries just rewrite their laws so that benefits conferred on others must be rewarded, and costs inflicted on others be paid by the decision maker? In Chapter 11 we reviewed the reasons why the kinds of contracts that would enforce these objectives are incomplete or unenforceable: the necessary information is either not available or not verifiable, the external effects are too complex or difficult to measure to be written into an enforceable contract, or there may be no legal system to enforce the contract (as in pollution that crosses national borders).

For these and other reasons, in most cases it is impractical to use tort law or any other body of law to make people liable for the costs they inflict on others. It is equally infeasible to use the legal system to compensate people for the beneficial effects they have on others, for example, to pay those who keep beautiful gardens an amount equal to the pleasure this confers on those who pass their house, because a court would have to know how much that pleasure was worth to each person who walked by.

In the earlier examples of uncompensated external effects, the reason why uncompensated external costs and benefits occur is the same:

- Some information about an aspect of an exchange that is of concern to someone other than the decision maker is *asymmetric* or *non-verifiable*.
- Therefore there can be no contract or property rights ensuring that external effects will be compensated.
- As a result, some of the social costs or benefits of the decision maker's actions will be *excluded* from (or will not be sufficiently valued in) the decision-making process.

! REMINDER When information is *non-verifiable* it means that it cannot be taken to a court or other third party to adjudicate on a contract, and the contract is therefore incomplete even if information is not asymmetrical.

The result is a market failure even in a perfectly competitive equilibrium. Market failures are a kind of *coordination failure*, where the broader term is sometimes used to describe a Pareto-inefficient outcome in problems like traffic congestion where the relevant social interactions are not on a market.

CHECKPOINT 14.5 Uncompensated external effects Review: Make sure you can explain the uncompensated external effects shown in each row of the third column of Table 14.1, and why these result in the market failures shown in the final column.

14.7 PERFECT COMPETITION AND INEQUALITY: DISTRIBUTIONAL NEUTRALITY

Economists celebrate the first welfare theorem because it explains what is required for the prices to be right. Doing this clarifies the (very demanding) conditions under which a perfectly competitive market would allow buyers and sellers to realize all of the possible mutual gains from exchange.

The same model also has a lot to say about how these gains from mutually beneficial exchanges will be shared. A remarkable feature of the model that we explain below is that the total value of the goods that each player holds in the competitive equilibrium after exchange is the same as the value of each actors' endowment bundle when both are valued at the competitive equilibrium prices.

This means that the distribution of wealth is unchanged by the process of exchange, a feature termed the *distributional neutrality of the market*. Both are better off after the exchange—they have moved to higher indifference curves—it is the value of their bundles that has not changed. We will now see how this comes about.

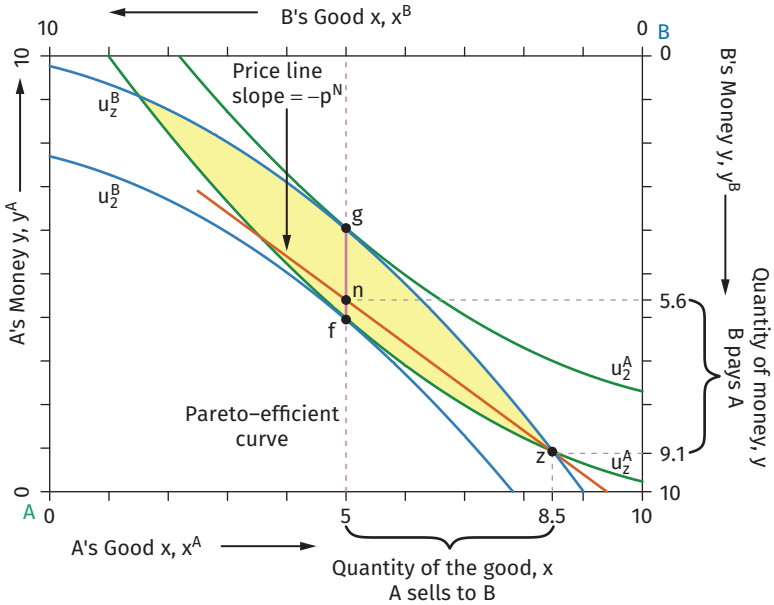
Distributional neutrality: A graphical illustration

As before, the two traders in Figure 14.10 each have a positive but different *initial allocation* or *endowment* of both goods (x_z^A, y_z^A) and (x_z^B, y_z^B) indicated in Figure 14.10 by point **z** (as in the previous figures, and in Chapter 4). The endowment is an exogenous initial distribution of wealth, the determination of which is beyond the model. The participation constraints of the two traders (u_z^A and u_z^B) are the indifference curves passing through

! REMINDER Recall that with a quasi-linear utility function that is linear in y, the person's marginal rate of substitution does not depend on how much y she has. A difference in the amount of y that the person has shifts the indifference curves without changing the slope. We call this a 'vertical displacement' or 'vertical shift.'

M-CHECK In Figure 14.10 and later figures we use quasi-linear utility functions (linear in y) so the marginal utility of y is a constant as y increases. The result is that the Pareto-efficient curve is a vertical line because changing the amount of y that the two have does not alter their marginal rates of substitution, so Pareto efficiency requires a particular distribution of the x good that is *independent* of the distribution of the y good. This allows us to treat distribution separately from the Pareto efficiency.

Figure 14.10 Distributional neutrality of exchange in the perfectly competitive equilibrium. At their endowment point **z**, Adamo has 8.5 units of x and 0.9 units of y. He trades 3.5 units of x, to obtain 3.5 units of y, resulting in his competitive equilibrium allocation of 5 units of x and 4.4 units of y, at point **n**. At the equilibrium, Beatriz has 5 units of x and 5.6 units of y. The market-clearing price is shown by the slope of the market-clearing price line, $-p^N = 1$ (this is the ratio of the number of units of good y exchanged for units of good x, therefore a *relative* price). The value of the traders' wealth at the equilibrium outcome (**n**) is the same as the value of their wealth at their endowments (**z**). So with the Nash equilibrium price, $p^N = 1$, for Adamo, $p^N \cdot x_n^A + y_n^A = p^N \cdot x_z^A + y_z^A = 9.4$ and for Beatriz $p^N \cdot x_n^B + y_n^B = p^N \cdot x_z^B + y_z^B = 10.6$.



their initial endowments (**z**). The yellow-shaded lens between them is the set of Pareto improvements over z.

Points making up the Pareto-efficient curve in Figure 14.10—both the dashed and solid portions of the curve—are the allocations of the two goods that are at the tangencies of the two traders' indifference curves, where their marginal rates of substitution are equal. We know:

M-CHECK The two traders have identical utility functions of the following form: $u^A(x^A, y^A) = y^A + 2x^A - \frac{1}{12}(x^A)^2$ and $u^B(x^B, y^B) = y^B + 2x^B - \frac{1}{12}(x^B)^2$.

- *from the first welfare theorem* that the final allocation is Pareto efficient so it will be somewhere along the Pareto-efficient curve in the figure; and
- *from the fact that the exchange will be voluntary* and the two traders have *endowments given by point* **z** that the perfectly competitive equilibrium allocation will be at some point between **f** and **g**.

The law of one price means that all transactions are made at the same market equilibrium price. The law has two far-reaching implications.

- *Equal treatment*: An implication of the assumption that all transactions take place at equilibrium prices is that two people who have the same preferences and initial endowment will end up in equilibrium with the same consumption bundle, because they face the same equilibrium prices and the same budget constraint. This is the *equal treatment property* of the perfectly competitive general equilibrium model. No matter which Adamo type we consider, all of their transactions will be at the same price, so all of the Adamos will end up at the same final allocation. (The same will be true of all the Beatrizes.)
- **Distributional neutrality**: The final allocation reached by all the Adamos will have the same value (as total wealth) at the market equilibrium price, p^N, as the initial endowment ($\mathbf{z}$). This follows because all points on the price line—including the endowment and the equilibrium allocation—have the same wealth value (at the price given by the slope of the price line). The value of what they get at the market equilibrium prices is the same as the value of the initial endowment at those prices. The same is true of all the Beatrizes.

Thus, in addition to Pareto efficiency we have a second important result about the equilibrium of a perfectly competitive economy: the degree of inequality that results after the trading process is identical to the inequality in the value of the traders' initial endowments, when the endowment and post-exchange allocations are valued at the equilibrium prices.

M-CHECK Both equal treatment and distributional neutrality are general results of the model, independent of the quasi-linear utility functions we have chosen as a simplification to illustrate the model.

! REMINDER The price line is the budget constraint facing each of the traders, and the value of expenditure along any budget constraint is the same, namely, it is the total expenditure that constrains the maximum the buyer can spend.

M-NOTE 14.1 Wealth and distributional neutrality

For either trader A or trader B, wealth at their endowment is given by m_z:

Endowment wealth: $m_z = p_x x_z + p_y y_z$

We can let $p_y = 1$, so: $m_z = p_x x_z + y_z$

Let $p = \frac{p_x}{p_y}$. As $p_y = 1$, $p = p_x$. We can then define their post-exchange wealth as:

Wealth after exchange $m_n = p x_n + y_n$

y after exchange $y_n = y_z + \Delta y$

x after exchange $x_n = x_z + \Delta x$

Wealth after exchange $m_n = p(x_z + \Delta x) + (y_z + \Delta y)$

continued

DISTRIBUTIONAL NEUTRALITY Distributional neutrality is a characteristic of the perfectly competitive general equilibrium model in which the distribution of wealth following competitive exchange is identical to the distribution of wealth prior to exchange.

In every exchange whatever the price, it must be the case that the value bought is equal to the value sold or, what is the same thing, the value of the purchase minus the value of what was bought sum to zero. In reading the equations below, remember that if the person exchanged some of the y-good for some of the x-good, then $\Delta x > 0$ and $\Delta y < 0$:

Value purchased equals value sold $\quad p\Delta x + \Delta y = 0$

from which we see that $$p = \frac{-\Delta y}{\Delta x} \tag{14.7}$$

therefore $$m_n = px_z + y_z + \underbrace{p\Delta x + \Delta y}_{=0} = px_z + y_z = m_z \tag{14.8}$$

Equation 14.8 says that when wealth is valued at the equilibrium prices, the trader's wealth at their endowment (x_z, y_z) is equal to their wealth at their post-exchange allocation (x_n, y_n).

Because this is true of both players—no difference in their wealth before and after exchange—the distribution of wealth *must* be the same at point **z** and point **n** when wealth is valued at the price at which the exchanges took place.

We can use our numerical example in Figure 14.10 to see that equation 14.8 holds. We will show this holds for Adamo, but you can similarly see that it holds for Beatriz. Recalling that $p = 1$ we have:

$$m_n^A = 1 \times 8.5 + 0.9 + \underbrace{1 \times 3.5 + (-3.5)}_{0} = 1 \times 8.5 + 0.9 = m_z^A$$

CHECKPOINT 14.6 **Distributional neutrality** Using Figure 14.10 explain why the distribution of wealth (the goods held by the two individuals valued at the relative prices in equilibrium) is the same after the exchange as it was at the endowment.

! REMINDER It is confusing that indifference curves with the usual *concave to the origin* (bowed in) shape are part of the *convexity assumption*.

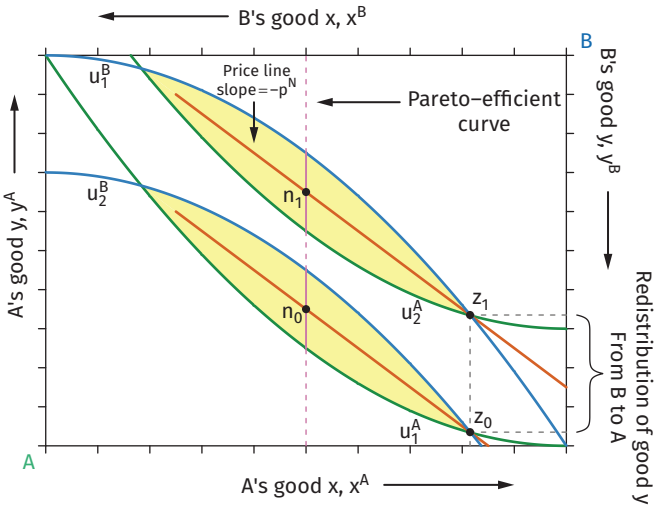
14.8 EFFICIENCY, FAIRNESS, AND WEALTH REDISTRIBUTION: THE SECOND WELFARE THEOREM

The distributional neutrality of the perfectly competitive market combined with the first welfare theorem is the basis for a kind of "division of labor" among policies which can be stated "let market competition address objective of efficiency and let wealth redistribution address the objective of fairness."

Stated mathematically, what is called the second welfare theorem addresses how an outcome that is both fair and efficient might be implemented by the combination of a redistribution of wealth and the process of perfectly competitive market exchange. The conditions under which the second theorem of welfare economics holds are somewhat more demanding than the first (or Invisible Hand) welfare theorem, for which the assumption of complete contracts is sufficient.

A second key condition for the second theorem to hold is called the *convexity assumption*. This is satisfied if:

Figure 14.11 The second theorem of welfare economics. Initially, each Adamo has lots of the good x, x^A, but very little good y, y^A, as shown by point z_0. With this initial allocation and price p^N the traders would end up at n_0. But, if the citizens through their government chose to redistribute the initial endowments to z_1 by taking some of good y from each Beatriz and giving it to each Adamo, then trading would start at point endowment z_1. At the resulting Pareto-efficient outcome n_1 the Adamos are better off, as intended by the policy. (The Beatrizes are clearly worse off).



- the indifference curves for two goods are "bowed inward" toward the origin (as is true here, due to diminishing marginal utility of the x-good); and
- the production possibility frontier (the frontier of the feasible production set) is "bowed outward" from the origin as it is in Figure 6.7 in Chapter 6.

The assumptions of the second theorem are more limiting than those of the first theorem. The reason is that in addition to ruling out missing markets and other kinds of incomplete contracts, also ruled out is any significant economies of scale such as those suggested by our survey of the evidence on cost curves in section 8.5 of Chapter 8. Also inconsistent with the sufficient conditions for the convexity assumption are the "bowed inward" feasible frontiers illustrated in Figures 6.4 and 6.5 of Chapter 6, where we showed how a country might get poverty trapped in an inferior Nash equilibrium.

Here is the theorem.

Second theorem of welfare economics: *Given both complete contracts and convexity, any Pareto-efficient allocation can be implemented by an*

HISTORY Nobel Laureate Amartya Sen wrote that the second welfare theorem "belongs to the revolutionists' handbook" because it suggested to many readers that an efficient way to reduce inequality is by redistributing wealth.[9]

assignment (a redistribution) of endowments among parties, followed by a perfectly competitive exchange process.

To see the theorem's importance, suppose that the citizens of an economy wish to redistribute goods to the less well-off members of society and they select a particular Pareto-efficient allocation as their preferred outcome.

The **second welfare theorem** says that an alternative, fair outcome can be implemented by reassigning property rights among the citizens (changing who has what at the initial endowments) followed by a perfectly competitive exchange process.

Why are the results of the second welfare theorem interesting to economists?

HISTORY Shortly after South Africa's first democratic election in 1994 that ended the system of racial inequality called apartheid, President Nelson Mandela asked one of the authors of this book to take part in a commission advising him about economic policy. Bowles proposed that the land and other wealth that had been confiscated from the black African population should be returned to their former owners (or their families) and that the market exchange process would then implement a more just allocation. He was simply repeating the logic of the second welfare theorem.

Often, governments adopt policies to alter final allocations of goods and services by changing *prices*, for example, by placing price controls on certain goods, or taxing others, or minimum wage laws that put a floor under wages, and rent control laws that put a ceiling on rents. The second welfare theorem tells us that instead of changing prices, governments seeking to address economic inequalities could achieve the same result by altering the initial distribution of endowments and then letting markets operate, rather than by changing prices. The contrast of the two approaches is illustrated by two types of policies that helped poor farmers.

- *Changing prices*: The government of the Indian state of West Bengal, elected with support of less well-off farmers, placed a ceiling on the share of the farmers' crops that could be claimed by the landlord.[10] This is a policy to affect *prices*, in this case, the rent that the tenant farmers have to pay to their landlords (the owners of the farms).
- *Redistributing wealth*: Half a century ago the governments of South Korea and Taiwan adopted policies that limited the amount of land that large landlords could own and distributed land to landless farmers. This is how policy can alter *initial endowments*, in this case the *ownership* of the land through redistributing wealth.[11]

We can illustrate these alternative approaches by returning to Figure 14.11. Consider what would happen if the Adamos and Beatrizes making up the population decided that the Adamos really deserve more. They propose to pursue this end by changing the *price*. Because the Adamos are the sellers of the good, this could be done by passing a law setting a

SECOND WELFARE THEOREM Given complete contracts and the convexity assumption about production and preferences, any Pareto-efficient allocation can be implemented by some assignment of the endowments of all parties, followed by a perfectly competitive market exchange process.

price of the good p_1 that is higher than the competitive equilibrium price p^N and then letting the two groups exchange goods in any way they wished. But, at a higher price, markets would not clear (p^N is the market-clearing price) and there would be excess supply of the good. We also know that the result would not be Pareto efficient.

Suppose instead that they decided to redistribute the *wealth*: in Figure 14.11 the endowments are redistributed from $\mathbf{z_0}$ to $\mathbf{z_1}$, after which perfectly competitive exchange at equilibrium prices takes the traders to $\mathbf{n_1}$ rather than the original post-exchange allocation $\mathbf{n_0}$. So under the assumptions of the second welfare theorem, particularly the assumption that all exchange takes place at the eventual equilibrium prices, wealth redistribution followed by perfectly competitive exchange represents a **mechanism** that can implement *any* equal-treatment Pareto-efficient allocation.

The two welfare theorems suggest that under the assumed conditions the second approach—wealth distribution—is preferred. The distributional neutrality of the market means that a redistribution of wealth translates directly into a redistribution of the value of final asset ownership. This is the basis of what we called the division of labor for public policies mentioned above: let prices do the work of achieving an efficient allocation, and let reassignments of initial endowments (the *z*s) do the work of achieving fairness.

Pareto efficiency and distributional neutrality: Why the theorems are important

When the two welfare theorems are taken together they appear to leave little room for ethical concerns about the operation of a competitive market system. The distribution of well-being is determined not by markets themselves but rather the distribution of initial endowments because "markets" are distribution neutral.

Kenneth Arrow pointed out that under the conditions specified by the theorems that he first proved:

> Any complaints about [the market system's] operation can be reduced to complaints about the distribution of income . . . the price system itself determines the distribution of income only in the sense of preserving the status quo.[13]

MECHANISM A mechanism is a set of rules of the game—possibly designed deliberately—that provide the incentives, constraints, and information that will result in an allocation—often a preferred allocation—being implemented.

› EXAMPLE John Roemer's treatment of the Marxian theory of exploitation is based on the same distributional neutrality of the market: "If the exploitation of the worker seems unfair, it is because one thinks the initial distribution of capital stock, which gives rise to it is unfair."[12]

Figure 14.12 Kenneth Arrow (1921–2017) was an American economist who by the age of 30 had proved three theorems that were to shape the development of economics and the other social sciences, the two welfare theorems presented here and his so-called impossibility theorem. The latter theorem showed that if citizens' preferences are ordinal (they rank outcomes rather than assigning cardinal numbers to them) and are not comparable across individuals, then there is no system of voting that can meet a set of criteria broadly capturing our idea of how a democracy should work. He also made major contributions to understanding the process of learning-by-doing introduced in Chapter 6.

Photo by Linda A. Cicero/Stanford News Service. Wikimedia Commons.

Others, including the philosopher David Gauthier, have drawn even more general conclusions from the model:

> The operation of a market cannot in itself raise any evaluative issues. Market outcomes are fair if . . . they result from fair initial distributions . . . the presumption of free activity ensures that no one is subject to any form of compulsion, or to any type of limitation not already affecting her actions as a solitary individual. . . . [Thus] morality has no application to market interaction under the conditions of perfect competition.[14]

In other words, at the perfectly competitive general equilibrium, there is nothing morally wrong that was not already there at the original endowment.

HISTORY Arrow's and Roemer's observations had been anticipated by the US Supreme Court in its 1915 decision *Coppage v. State of Kansas*, "wherever the right of private property exists, there must and will be inequalities of fortune; . . . it is impossible to uphold the freedom of contract and the right of private property without at the same time recognizing as legitimate these inequalities of fortune that are the necessary result of the exercise of those rights."[15]

CHECKPOINT 14.7 **From the "revolutionist's handbook"** Use Figures 14.8 and 14.11 to contrast the two strategies for redistributing income from B to A: changing prices or redistributing wealth. In Figure 14.8 draw a new price line which if imposed (by the government) would raise A's income and utility and lower B's.

14.9 MARKET DYNAMICS: GETTING TO AN EQUILIBRIUM AND STAYING THERE

But does the price-taking competitive equilibrium help us to understand real economies? For us to answer that question, we need to know, in addition:

- Would an economy that got to a price-taking equilibrium stay there?
- How could an economy get to a price-taking competitive equilibrium to begin with?

Staying there: Is a price-taking "equilibrium" a Nash equilibrium?

In the real world, the markets that most closely approximate price-taking are also those where the commodity being traded has been standardized and certified by some third party (like the market for #2 red winter wheat we introduced in Chapter 10). The standardization of commodities increases the effective number of competitive suppliers and greatly reduces the influence any one competitor can have on the market price. Large numbers of competitors make it more difficult for firms to collude. If an entire economy were like the market for #2 red winter wheat with goods defined by complete contracts as enforced by the Chicago Board of Trade and found itself in a price-taking competitive equilibrium, would it stay there? We explored in Chapters 8 and 9 how buyers and sellers would have

an incentive to change their behavior in ways that would result in a different kind of equilibrium.

If, in a competitive price-taking equilibrium, there is something a firm or a family could do to raise its profits or utility—for example acting like a price-maker rather than a price-taker—given what everyone else is doing (that is, price-taking), then the price-taking "equilibrium" is not a Nash equilibrium and it will not persist.

The first thing that at least some firms would explore is finding a way to escape from the competition of other firms by making their product distinct in some way so that other firms' products are no longer perfect substitutes for the firm's differentiated output. Product redesign or advertising might accomplish this. If successful, the firm would then face a downward-sloping demand curve, and be able to restrict output, charge a price greater than marginal cost, and make profits above the opportunity cost of capital.

The important conclusion is that price-taking may not be profit-maximizing behavior even in an initially competitive market because there are opportunities to make more profits by deliberately altering the nature of the competition that a firm faces.

Even without product differentiation, in real-world economies firms may face downward-sloping demand curves because the number of effective competitors they face is limited. One example of these limits is geography; think of restaurants—How many competitors does a mid-priced Italian restaurant have in a small city?

CHECKPOINT 14.8 **Adam Smith on competition** In *The Wealth of Nations*, Adam Smith wrote that "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices." Less often cited is his next sentence: "It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty." On the basis of this passage how would Smith respond to the following question: Is price-taking behavior among competitors a Nash equilibrium?

Getting there: The parable of the auctioneer

Walras—whose model of perfectly competition general equilibrium we have so far employed—did not explain the *process* by which an economy of people and firms would reach the price-taking competitive equilibrium that his equations describe. In one edition of his book, Walras proposed the idea that the market-clearing equilibrium prices could be found by adding to the economy adding a fictional character, the *Auctioneer*.

The Auctioneer, Walras wrote, "cries out a system of prices at random" and asks that all the people and firms report their price-taking profit- and

REMINDER In Chapters 8 and 9 we discussed how firms would *differentiate* their goods to ensure that they could maintain market power.

REMINDER In the Cournot model of Chapter 9 the typical firm sees itself as sharing the market with a limited number of competitors, and therefore as having some power to influence market price.

REMINDER Remember (from Chapter 9) that price-taking is not an *assumption* about how people behave in general. Price-taking is the *result* of a perfectly competitive market equilibrium. Price-takers at the market equilibrium will do better if they become *price-makers* when the market is out of equilibrium. Competitive markets reach an equilibrium through some of the buyers or sellers changing their prices in order to capture the rents that exist when the market is not in equilibrium. Price-taking does not mean that buyers or sellers cannot set a price different from what others are setting; they can set any price they wish. Price-taking means they *cannot benefit* from doing so.

HISTORY All that Kenneth Arrow and Frank Hahn claimed for the second welfare theorem is that "in a certain sense any desired efficient allocation can be achieved by redistribution of initial assets followed by the achievement of an equilibrium." They are careful not to suggest that the equilibrium can be achieved in a decentralized manner. They illustrate the second welfare theorem with an example of "an omniscient state" that "computes a price vector… satisfying the hypotheses of the theorem."[16]

utility-maximizing supplies and demands of all the commodities at those random prices. Given that the Auctioneer is just randomly choosing prices, for any given set of prices suppliers, the result would typically be:

- *excess demand*: too little of some commodities (so that demand exceeds supply); and
- *excess supply*: too much of others (supply exceeding demand).

But, excess supply and demand are not a problem for the Auctioneer because no one actually produces or exchanges at these random prices in Walras's parable: the Auctioneer simply does not allow it. Sooner or later the Auctioneer by chance would hit upon the market-clearing price-taking competitive equilibrium prices. Having confirmed that markets would clear, once the equilibrium prices were known to all traders and everyone instructed to restrict their transactions to the equilibrium prices the Auctioneer would then allow production, exchange, and consumption to take place.

Neither Walras nor anyone since has imagined such experimentation by an Auctioneer is how prices and market clearing come about. Walras created the parable as a device to demonstrate that market-clearing prices could exist and could somehow be discovered. The lack of a convincing account of how an economy might get from an initial endowment to a price-taking competitive equilibrium challenges a common interpretation of the second welfare theorem: namely that redistribution of endowments followed by market exchange can implement any Pareto-efficient allocation.

How do buyers and sellers act out of equilibrium? Can buyers and sellers move to a Pareto-efficient equilibrium from any initial endowment? The perfectly competitive general equilibrium model does not provide answers to these questions. As a result, we cannot say *how* buyers and sellers can achieve any Pareto-efficient equilibrium from a given initial endowment without exploring alternatives to Walras's Auctioneer. This is what we do in the next section.

14.10 BARGAINING AND RENT-SEEKING: A MORE REALISTIC MODEL OF MARKET DYNAMICS

Traders might get to an equilibrium without the services of the Auctioneer. How? Traders can reach market equilibrium through a decentralized process involving actual trades at disequilibrium prices. To do this we return to thinking about traders as perfect competitors (introduced in Chapter 9), not as price-takers as in the perfectly competitive general equilibrium. A perfect competitor is an economic actor who seeks out each and every opportunity for gains through exchange, driving the economy, eventually to an equilibrium, by a process of rent-seeking that continues until there are no rents left to seek.

More realistic assumptions for exchange

What does an adequate account of such a system require? Let us start with basic facts about the *people trading*: They are empirically plausible in what they are attempting to do and in what they know.

- People differ in their preferences and, because of specialization and the division of labor, people differ too in their endowments.
- People refuse exchanges that make them worse off, so the prices at which they exchange are mutually agreed upon.
- People know their own preferences but not those of other people.

Next are the facts about the institutions that provide the rules of the game for trading: the process of exchange is truly *decentralized*.

- *Decentralized trade*: trading is arranged by traders not by a government or some other centralized body.
- *No auctioneer*: there is no imaginary Auctioneer saying which trades are allowed.
- *Trade is voluntary*: a trade takes place if the trade is mutually beneficial to the two traders involved, and does not take place if it is not mutually beneficial.

Figure 14.13 illustrates the setting for our model, showing the endowment allocation (point **z**) and indifference curves of one of many As and one of many Bs. Because the As have none of the y good and most of the x good you can think of the As as sellers of the x good and buyers of the y good. In Chapter 4 (section 4.7) we showed how, without the help of an Auctioneer, two traders could bargain into the yellow Pareto-improving lens.

To remind you how this works, let's start in Figure 14.10 at the endowment allocation point **z**, before any trades have taken place. Here:

$$mrs_z^A = u_x^A/u_y^A \quad < \quad u_x^B/u_y^B = mrs_z^B \tag{14.9}$$

The marginal rate of substitution of Beatriz types exceeds the Adamo types marginal rate of substitution. Her maximum willingness to pay for some of Adamo's x-good is greater than the least amount of her y-good that he would *accept* in return for giving up some his x-good. So each Adamo might wish to exchange some of his x for some of a Beatriz's y and each Beatriz would conversely wish to trade some of her y for some of an Adamo's x, so mutually advantageous trades will be possible.

The market process can be thought of as a large collection of bargains over potential surpluses like the interactions we have analyzed in Chapters 4 and 9. In section 4.7 we illustrated a simple bargaining rule: the price at which they exchange should be midway between their willingness to pay of one and the corresponding willingness to pay of the other. We do not

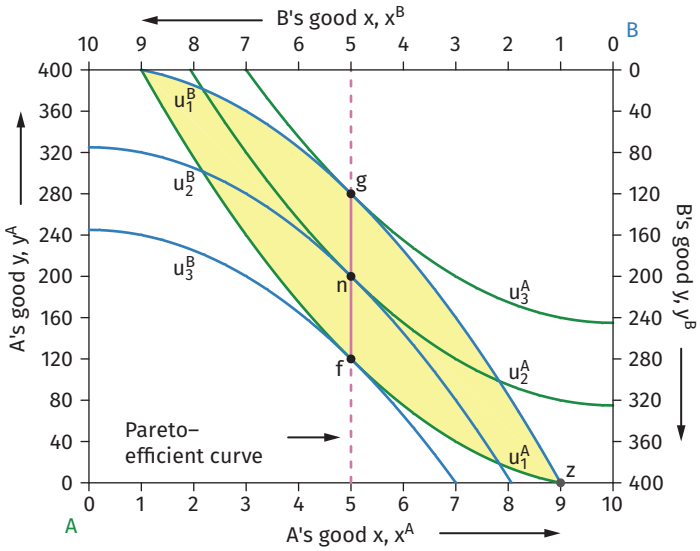
HISTORY Duncan Foley (1943–) pioneered the decentralized bargaining approach to general equilibrium theory. In 1994 he described a setup like this: "[traders] enter the market knowing only the transactions they view as improving their condition given their endowments, preferences, technology, and expectations [when they] encounter other [traders]; and make mutually advantageous transactions with them in a disorderly and random fashion."[17] He also (in 1967) introduced the idea that a fair allocation should be "envy free" so that each person's bundle was such that they would not want to exchange bundles with any other person if they were given the opportunity.

! REMINDER Remember that (given the *current* amount of x and y he has) a person's marginal rate of substitution (his mrs^A) is the maximum amount he is willing to pay in units of y for an additional unit of x.

M-CHECK The two traders have identical quadratic, quasi-linear utility functions of the form: $u^A(x^A, y^A) = y^A + 100x^A - \frac{1}{20}(x^A)^2$ and $u^B(x^B, y^B) = y^B + 100x^B - \frac{1}{20}(x^B)^2$.

Figure 14.13 The Edgeworth box that two traders confront in a computer simulation. All of the Type A traders—the Adamos—have several items of the good ($x^A = 9$) to sell, but no good y ($y^A = 0$). We could think of them as the *producers* or *sellers* in this economy. All of the Type B traders—the Beatrizes—have very little of the good ($x^B = 1$), but a significant endowment of cash to purchase the good ($y^B = 400$). We could think of them as the *buyers* in the economy. These simulation models are based on similar simulations from Foley (1994). The solid portions of the Pareto-efficient curve are allocations that could be implemented by bargaining between the two (the dashed portions are allocations that violate the participation constraints of one or the other). Point **g** is the best that A could possibly do by bargaining, while point **f** is the best that B could do. Point **n** is the Nash equilibrium from Figure 14.8.

HISTORY Stephen Smale, a mathematician, introduced an element of market realism by abandoning the Auctioneer and allowing transactions take place at nonequilibrium prices. Similar to Duncan Foley's model, Smale comments: "The exact equilibrium depends on factors such as which agents first encounter each other."[18] As in Foley's model, final wealth and utility therefore come down to an individual trader's endowment, trading history, and luck.



! REMINDER We spoke about path dependence in previous chapters, starting in Chapter 1. In basic terms, path dependence means that "history matters," that is, the sequence in which trades occur matters for which traders outperform other traders, or which equilibrium of a game is more likely to occur.

have to specify exactly how they trade or at what prices other than that the exchange is voluntary, so that at least one of them benefits and neither loses as the result of an exchange.

What is important is that *any* exchange process in which agents seek out and execute mutually advantageous trades as long as any exist has to lead them eventually to some Pareto-efficient allocation. So our decentralized bargaining framework replicates the result of the first welfare theorem but without bringing in the Auctioneer.

Each trader might engage in a series of trades with different partners, always implementing Pareto improvements (remember trade is voluntary!). This process will continue until *no one* can find a partner with whom a mutually advantageous trade was possible. When this is the case, we know two things:

- *Pareto efficiency*: The outcome is Pareto efficient for the same reason that it is a Nash equilibrium: there are no mutually beneficial exchanges that could be carried out.
- *Nash equilibrium*: The outcome must be a Nash equilibrium because given the strategy sets available to the players–buy or sell offers to others–there are no actions that any player can make that will result in further positive payoffs. The reason for this is that, to be executed, an exchange must confer mutual benefits, and from the fact the outcome is Pareto efficient, we know there are no such opportunities remaining. So no further trading is a best response for each player.

Without knowing more about the details of the exchange process, such as the exact order in which the agents meet and the exact trades they might make, we cannot say where along the Pareto-efficient curve the As and Bs will end up. But we do know two additional things about the result:

- *The market is no longer distributionally neutral*; *nor is equal treatment observed*. One Adamo may trade consistently at more favorable prices than another on some particular trading path, and get a larger share of the consumer surplus. The same is true of the Beatrizes. The result is that after the exchange process there are inequalities among the Adamos and among the Beatrizes even thought they had the same initial wealth endowments.
- *Exchanges take place at disequilibrium prices*. None of the traders know what the Nash equilibrium price of the perfectly competitive model is, they simply trade at whatever mutually beneficial price they can agree on.

We can illustrate how such a decentralized market exchange process might work using a computer simulation.

CHECKPOINT 14.9 **Bargaining to better outcomes** Using Figure 14.8 explain how the two types—Adamos and Beatrizes—might bargain so that starting at an initial endowment **z**, the final allocation could be **f**, **n**, or **b**. What will determine which of these final allocations (or others that are also possible) will occur?

14.11 COMPUTATIONAL GENERAL EQUILIBRIUM: MARKETS, EFFICIENCY, AND INEQUALITY

We simulate disequilibrium trade using two goods (x and y) and two types of traders, A and B. Each trader has their own utility function and initial endowment of the two goods. There is an equal number (1,000) of each type of trader. The "traders" are just lines of code in the computer program that simulates the market in our model.

In each round of trading, the traders are paired randomly, so over the course of many rounds of the simulation they interact with many different traders of the other type rather than a single other as we have so far assumed in our one-on-one version of the Edgeworth box. For each pair in a given round, the program computes offer prices given their current holdings of the two goods, and chooses a random price in the interval between the two offer prices.

! REMINDER Remember that "offer price" is another name for marginal rate of substitution because it is the price at which a trader is willing to give up some y in order to get an extra unit of x. Therefore when the offer prices are equal, the marginal rates of substitutions are equal and the allocation is Pareto efficient.

We limit trade at the random price between the traders' offer prices by the smaller of their two utility-maximizing offers. This is consistent with the short side of a non-clearing market determining the quality of goods traded as you saw in Chapter 9 and Figure 14.18. This involves calculating the demand function for each trader and ensures that the trades are feasible and do not decrease a trader's utility. Each trader's holding of the two goods is updated by the trade, and carried forward to the next round of trading.

This process continues until there are no available mutually advantageous trades, which means that the final allocation is Pareto efficient.

! REMINDER Remember that a trader's participation constraint is their fallback utility or the utility that they would receive at their endowment, **z**.

Because both traders have quasi-linear utility functions, their offer prices depend only on their holding of the non-linear good. In this case, all Pareto-efficient allocations give the same amount of the nonlinear good to all the traders of the same type, and the price ratio in the final trade is the same as the perfectly competitive general equilibrium.

We illustrate the basics of the simulation in Figure 14.13 where we will let the y-good represent "money available for other purchases" or just "money."

The Pareto-efficient allocation of the good is for each of the As and Bs to have five units of the good. The Pareto-efficient allocations are shown in both Figures 14.13 and 14.14. The traders will go on trading until they reach that allocation. Remember trading continues until there are no mutually beneficial exchanges possible, until a Pareto-efficient outcome has been reached.

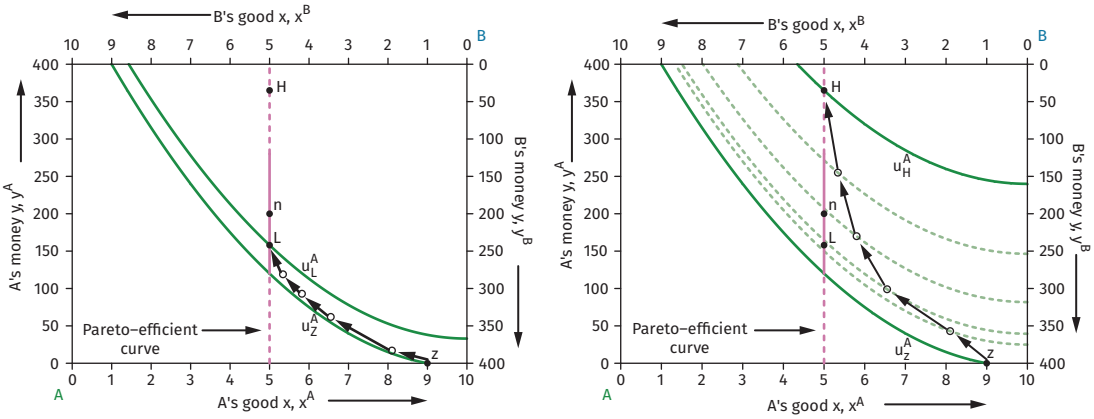
Disequilibrium trading creates inequality

What is undetermined is how much money each will have left over for other purchases when the trading ends.

Figure 14.14 shows trading paths for two Type A traders: one who did poorly and another who did well. The difference between them is purely a matter of luck. Though the traders did not perform equally well, their trading in every round implemented Pareto-efficient outcome. The same inequalities are also evident among the Bs in the simulation.

Why did one do so much better than the other? Talent? Bargaining power? Mistakes? No, there are no mistakes: each trade made by both traders made them better off as the panel (b) in the figure shows, where the trader moves to ever higher indifference curves.

Figure 14.14 Bargaining to a Pareto-efficient and unequal outcome. In both panels, **n** is the perfectly competitive general equilibrium as would be implemented for all traders by the fictional Auctioneer. Panel (a) shows a simulated path for a Type A trader who we selected to illustrate the case of a trader who does not obtain a significant increase in utility as a consequence of the trades that they make in the simulation. He traded several times, but his trades resulted in only small increases in utility, with him eventually arriving at a final allocation (**L** for low) on the Pareto-efficient curve. Panel (b) show a simulated path for a Type A trader who obtains a significant increase in his utility, resulting in a final allocation (**H** for high) on the Pareto-efficient curve. Each intermediate trade is shown by a hollow circle. The endowments are shown by point **z**, and the final allocations are shown by the black dots **L** and **H**. The figures are generated with agents who have the same preferences and the starting wealth shown at point **z**. Their trades take them above the utility they received at **z**.



(a) A low-performing Type A trader

(b) A high-performing Type A trader

Moreover, the model does not advantage any one player over any other. The pairing to trade is random, and so is the selection of prices (from the mutually beneficial range at which they both wish to trade). To understand the results shown in Figure 14.14 remember that a particular A interacts with many *different* Bs (not–except by chance–multiple times with the same B).

In addition to tracing their gains in utility we can also track their wealth and its distribution. Because the wealth of each player is made up of two goods we need some price in order to sum the total wealth of a player. We know that following exchange, the bundles of all of the players are arrayed somewhere along the Pareto-efficient curve. Because the equated marginal rates of substitution along that curve are all equal we know the prices at which trade *concluded* is given by that common slope. We use that price as the relative value of the two goods making up their wealth in our calculations of the wealth of each player.

Referring back to Figure 14.13 you know that one of the points on the Pareto-efficient curve, point **n**, is on the price line that includes the endowment allocation **z**. Remember this is the trade that *would have been* executed by *all* of the players had the trading been organized by the Auctioneer. And we know from the distributional neutrality of that process that the distribution of wealth at points **n** and **z** are identical.

But then it follows that the As that ended up at points below **n**, for example at point **f** in Figure 14.13 have less wealth than at their endowment. An example in Figure 14.14 is point **L** the post exchange bundle of the low-performing A trader. They *lost* wealth in the trading process. The Bs at point **f** gained wealth. Other As—for example those near point **g** in Figure 14.13—gained wealth, as did the Bs who ended up near point **f**.

How much inequality did the trading process create? In Figure 14.15 we show the distribution of wealth following the exchange process for both the As and the Bs. Remember there were no endowment differences among the As or among the Bs, so all of the wealth inequality shown in the figure is the result of the trading process itself.

Here is what we have learned from our decentralized market model and its simulations.

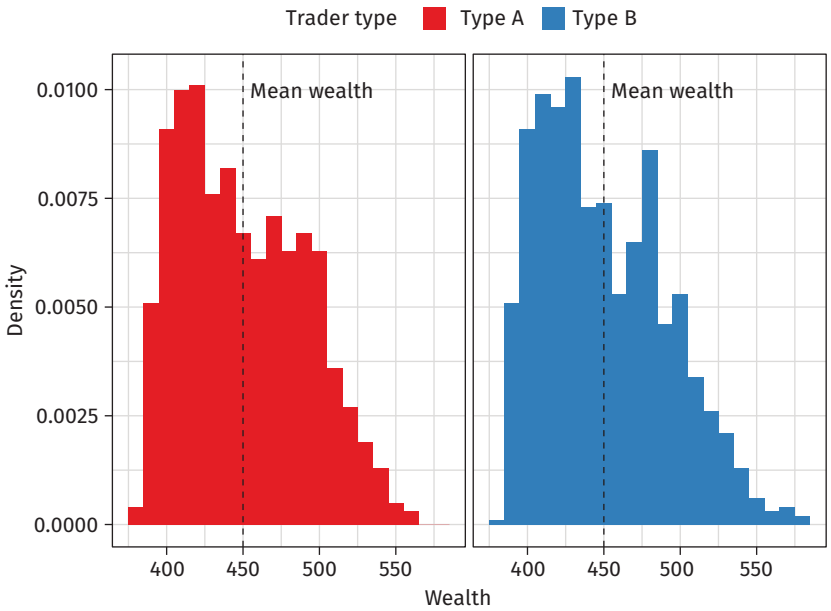
- *Pareto efficiency*: All of the traders end up on the Pareto-efficient curve.
- *Gains from trade*: Both As and Bs benefited from trading (no A is below his participation constraint, and no B is below hers).
- *Wealth inequality*: The two types of traders started off with identical wealth within their types, but they ended up with very different levels of wealth; many losing modest amounts of wealth and some capturing very significant wealth increases.

! REMINDER You have seen that unequal treatment is a result of a competitive labor market in Chapter 11. Among a group of identical workers (with identical skills and other endowments) some workers will be employed and others not employed. In other words, some are renting their endowments for a wage w while others are not transacting at all (they are *labor market-excluded* workers).

A more realistic market environment would have generated greater inequalities:

- *Bargaining power*: If we also introduced some differences in bargaining power, then the price at which a trader settles may be more or less favorable than some other trader; so inequality would no longer just be a matter of luck.
- *Discrimination*: Or, if some traders were the targets of racial, religious, gender, or other discrimination, then they would tend to face less favorable prices, introducing yet more inequality.
- *Non-random matching*: Finally perfect competitors in the real world would not settle for being randomly matched: those with more information would seek out and find trading counterparts holding very different quantities of the goods than they had, and with whom the range for mutually beneficial trades was therefore especially large.

Figure 14.15 The wealth distributions of the traders from the simulations. The figure shows the distribution of wealth at the final allocations for all traders. Their wealth at the endowment point for both As and Bs is 450, using the prices that would have occurred had the trades been organized by an Auctioneer, and that were the final trading prices. As the distributions demonstrate, within the types of traders, wealth is unequally distributed, with some traders achieving much higher levels of wealth than other, less fortunate, traders.



Our model has affirmed that competitive exchange can implement Pareto-efficient outcomes, even without the Auctioneer as long as there are no impediments to bargaining. But there has been some collateral damage to the other results of the perfectly competitive general equilibrium model.

We have to give up Arrows's insistence on the distributional neutrality of markets, namely that markets merely preserve the status quo distribution of wealth at the initial endowment. Whether the inequalities emerging in the trading process among identical people are of significant magnitude remains an open question. Our simulation cannot speak to that question. But as a theoretical proposition, the distribution neutrality of the perfectly competitive market based on the law of one price, like the Auctioneer called in to enforce the law, is a fiction.

CHECKPOINT 14.10 Markets generate inequality: Beyond distributional neutrality Explain why, after bargaining with the other type, both As and Bs in Figure 14.15 end up with different levels of wealth even though all As started off with identical endowments, and the same is true of all the Bs.

14.12 BARGAINING TO AN EFFICIENT OUTCOME: THE COASE THEOREM

Our model of traders with limited information, buying and selling at whatever prices they can agree upon until there are no more mutually beneficial trades to be made suggests that bilateral bargaining may play a key role in achieving a Pareto-efficient outcome for an economy.

Ronald Coase observed that even if some external effect cannot be covered in a contract, the economy can still reach a Pareto-efficient outcome through bilateral exchanges as long as traders can bargain efficiently, including bargaining over the uncompensated external effects. Our model of competitive exchange is an example of Coase's reasoning, as we have assumed that any possible exchange with the potential to benefit both traders will be implemented.

Figure 14.16 Ronald Coase (1910–2013). You encountered Coase's contributions to understanding firms in Chapter 11. In accepting the Nobel Prize in 1992 he reminisced that as a young man he had wondered: "How did one reconcile the views expressed by economists on the role of the pricing system and the impossibility of successful central economic planning with the existence . . . of these apparently planned societies, firms, operating within our own society."[19] Both legal practice and economic theory have been shaped by his theory of bargaining.[20]

Image courtesy of Coase-Sandor Institute for Law and Economics, University of Chicago Law School.

Coase versus standard approaches in economics and law

In the field called **welfare economics**, the standard approach to coordination failures of the type illustrated in Table 14.1 is that the government should impose taxes or provide subsidies designed so that private economic actors internalize the external benefits and costs of their actions on others. The tax or subsidy transforms each person's objective function and hence, their utility- or profit-maximizing first-order conditions, so that each will act as if they were taking account of the effects of their actions on others. We saw an example of how taxes do this in Chapter 5 when each fisherman was taxed by the government for each hour that they fished. The tax reduced their fishing time to the Pareto-efficient level.

Compelling arguments for "green taxes" and government-provided schooling are routinely made on these grounds, invoking reasoning originating with Alfred Marshall and A. C. Pigou (1877–1959) early in the twentieth century. In legal theory, standard approaches to activities generating external costs are to prohibit the activities or make those generating the external costs legally responsible for ("liable for") compensating those harmed for the external costs inflicted on them. An example of a prohibition would be requiring the replacement of incandescent light bulbs with LED lighting to reduce greenhouse gas emissions. We used a liability rule in Chapter 1 to get the fishermen to a Pareto-efficient outcome by making each liable for the damage they caused the other.

WELFARE ECONOMICS A branch of economics that studies the effect of economic policies and institutions on individual and societal well-being ("welfare").

Table 14.2 Three approaches to internalizing external costs and benefits.

Approach	Ways to internalize or limit external effects	Examples
Welfare economics	(Pigouvian) taxes and subsidies by governments	Green taxes, public education
Legal theory	Prohibitions and liability (tort) law by courts	Prohibition of incandescent bulbs
Coase	Bargaining by private parties	Cap-and-trade environmental policies

Coase provided a private alternative to governmental intervention through taxes, subsidies, or prohibitions to address the problem of external costs and benefits. He demonstrated that, under specific conditions, entirely *private* transactions motivated by self-interest could accomplish the same objectives. Table 14.2 summarizes approaches to coordination failures including Coase's contribution.

Coase reconsidered Pigou's theory and ideas. In 1920, Pigou had explained the role of costly external effects imposed on others by the example of a railroad whose engines caused sparks, lighting fires in the farmland through which the train passed. The fires damage the crops and cost the farmers lost profits.

Pigou had written that to internalize the external cost imposed on the farmers the owners of the railroad should be liable for the damage caused by the trains. If the owners of the railroad anticipated the liability, this would induce them to account for the costs of their actions on others.

Coase responded that "if the railroad could make a bargain with everyone having property adjoining the railway line and there were no costs involved in making such bargains, it would not matter whether the railway was liable for damages caused by fires or not."[22] We call this key assumption "no costs involved in making such bargains" the Coase's proviso.

This surprising conclusion is motivated by the following observation. If the costs of the fires exceeded the cost of preventing the sparks (say, by redesigning the engines), then even if the railroad was under no obligation to prevent the fires, those who were harmed could simply pay the railroad a sufficiently large sum to induce them to agree to eliminate the fire damage, and still be better off than before. Conversely, supposing that the farmers had the right be be compensated for the fires, if the cost to the railroad of controlling sparks is very high, the railroad could pay the neighboring farmers not to cultivate the land near the tracks.

› EXAMPLE Pigou did not dream up his example of trains igniting fires. In 1918, the Cloquet-Moose Lake Fire in Minnesota, US, was started by sparks from a railroad; 453 people died and over 50,000 were displaced or injured. More than 250,000 acres of land burned and the fire caused over $1.2 billion (in 2020 US dollars) in damage to properties.[21]

You have already encountered Coase's proviso–*costless bargaining*–in the simulation model of competitive exchange where we assumed that traders exchanged goods if there was a mutual benefit to be had. If executing a trade involved, for example, substantial legal expenses–that is **transaction costs**–then Coase's proviso is violated.

! REMINDER As the head quote to Chapter 10 suggests, Coase was scornful of economic theories such as the perfectly competitive general equilibrium model or our decentralized bargaining model that assume the absence of transactions costs.

Bargaining is *costless* when the parties to the bargain, such as traders in an Edgeworth box interaction, do not incur costs in executing a trade other than the price of the good exchanged. **Costless bargaining**–or as Coase sometimes put it the lack of "impediments to bargaining"–is important, and, unlike many who have invoked Coase against governmental regulation, Coase himself stressed it: "if market transactions were costless all that matters (questions of equity aside) is that the rights of the various parties should be well defined and the results of legal actions easy to forecast. But . . . the situation is quite different when market transactions are so costly as to make it difficult to change the arrangement of rights established by the law."[23] What came to be called the *Coase theorem* achieves a seemingly dramatic extension of the theorems of welfare economics. Even where contracts are incomplete and, as a result, uncompensated external effects like those illustrated in Table 14.1 occur, efficient allocations can result from bargains. The **Coase theorem** states that Pareto-efficient allocations occur if the people affected can bargain efficiently over the rights governing the actions that result in the external costs or benefits.

HISTORY Because there is some controversy about what the "theorem" means (as Coase himself said, there is no explicit "Coase theorem"), it may be useful to consult its author. In his 1992 Nobel lecture, Coase wrote: "What I showed . . . was that in a regime of zero transactions costs, an assumption of standard economic theory, negotiations between the parties would lead to those arrangements being made which would maximize wealth, and this irrespective of the initial assignment of rights."[24]

CHECKPOINT 14.11 The Coase theorem Explain Coase's reasoning that in the case of railway sparks igniting farmers' fields, the external effect could be internalized by private bargaining between the railroad owners and the farmers as long as there were no transaction costs, irrespective of whether the railroad initially had the right to emit the sparks or not.

TRANSACTION COSTS Costs that impede the bargaining process when contracts are incomplete, including costs of acquiring information about the good to be traded, and costs of enforcing a contract.

COSTLESS BARGAINING Bargaining is costless when the parties to the bargain do not incur transaction costs in executing a trade other than the price of the good exchanged. Efficient bargaining is often used interchangeably with the term "costless bargaining."

COASE THEOREM The Coase theorem states that Pareto-efficient allocations will be implemented as long as transactions costs are absent, so that those affected are able to bargain costlessly over the rights governing an exchange or other interaction, independently of the initial assignment of these rights.

14.13 AN EXAMPLE: HOW COASEAN BARGAINING WORKS

To understand the conditions under which Coase's argument would be true, think about a situation similar to one of the examples that Coase used when he first introduced the idea.

Let's consider two people: Anders and Bianca. Bianca operates a small metalworking shop next to Anders's yoga studio and when Bianca cuts metal the loud noise disturbs the people practicing yoga next door. The loud noise is a negative external effect (a cost) that Bianca imposes on Anders.

To explain what occurs we shall describe a set of payoffs for Anders and Bianca depending on the actions they take. This will allow us to construct game trees to explain what the Nash equilibrium outcomes of the interaction will be. We consider two cases:

- *Bianca has the initial rights*: Bianca can *use* her loud machinery—imposing a cost on Anders—as much as she likes; he can bargain with her to get her to restrict her usage.
- *Anders has the initial rights*: Anders has the right to *restrict* Bianca's usage of the machinery; she can bargain with him to allow her to use it.

Let's say that if there is no bargaining, and if Bianca uses her metal-cutting saw during the hours of operation of Anders's yoga studio, then Bianca receives a payoff of 5 and Anders 1. If she is restricted to not use the machine when the yoga studio is open, then both get payoffs of 4.

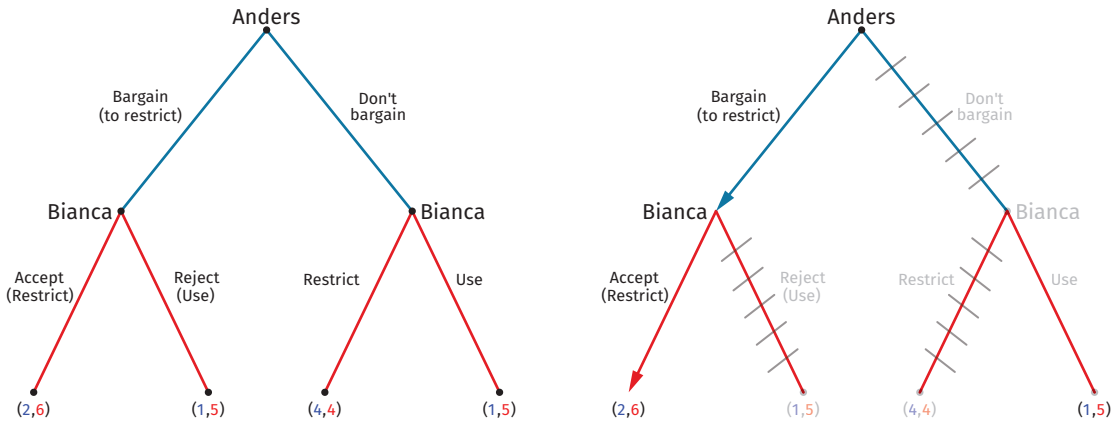
Bargaining when Bianca has initial rights

Problems like this are sometimes addressed by what are called zoning laws, which for example *exclude* "nuisance activities" such as garbage incinerators or pig farms from residential neighborhoods or professional office locations. Coase's idea is that private bargaining rather than government policies could get Bianca to *internalize* the external cost she imposes on Anders, giving her the private incentive to operate her shop in a way that does not disturb Anders's yoga sessions.

The initial assignment of rights in this case is whether or not Bianca has the right to generate as much noise as she wants, whenever she wants. The possible bargains between her and Anders include a payment from one to the other along with the transfer of the initial rights from one to the other. For example, Bianca initially has the right to operate her machinery in an unrestricted way, but she might be willing to give up that right—giving Anders the right to restrict her use of the saw—if Anders paid her enough.

Remember, like any other exchange, for a bargain to be implemented it must result in both parties being better off (or at least one of them being better off and the other not worse off). This is because exchange is *voluntary*. Both parties must accept the deal. So the result of the bargain must be a Pareto improvement over the result that would occur without the bargain.

Figure 14.17 The case of bargaining between Anders and Bianca when Bianca initially has the right to use her machine unrestrictedly. Anders is the first mover, with his decision branches and payoffs in blue. Bianca is the second mover, with her decision branches and payoffs in red. Anders can Bargain (to restrict) or not. If Anders bargains, Bianca can accept or reject the bargain. If Anders does not bargain, Bianca can Restrict voluntarily or choose to Use unrestricted. Panel (a) shows the game without the solution; panel (b) shows the solution to the game with the pruned branches faded out and the Nash equilibrium shown as (Bargain (to restrict), Accept).



(a) Bargain to restrict (b) Solved game tree

The game tree in Figure 14.17 illustrates the case in which Bianca has the right to unrestricted use of her metal saw. At the top of the tree Anders can choose either:

- *Don't bargain* with Bianca to get her to restrict her usage of the machinery; or
- *Bargain* (*to restrict*) and offer her 2 to restrict her use of the machinery.

If he offers this bargain, then, moving down the game tree, Bianca has the choice between:

- *Reject* Anders's offer and exercise her right to unrestricted use of her machinery; or
- *Accept* the bargain and Restrict.

On the right branch of the tree, Anders does not bargain, and Bianca then chooses between:

- *Restrict*: to limit her usage of the metal-cutting saw; or
- *Use unrestricted*: meaning use her metal-cutting saw whenever she pleases.

The numbers at the end of the branches are the payoffs to the two people: Anders's in blue, Bianca's in red. You already know that if Anders does not bargain and Bianca chooses Use then Anders gets 1 and Bianca gets 5, while if Bianca restricts they both get 4.

On the left branch of the tree, that is, if Anders chooses Bargain (to restrict) offering Bianca 2 to *not* use her saw the outcomes are as follows:

- If Bianca rejects the offer and retains unrestricted rights to use her machine, then she gets 5 and Anders gets 1.
- But if she accepts his offer, then before being paid by Anders Bianca gets 4 and then with the payment from Anders she gets $4 + 2 = 6$; as a result, Anders gets $4 - 2 = 2$.

How would this game be played if both Anders and Bianca cared only about their own payoffs? To find the outcome of the game, look for the Nash equilibrium, that is, an outcome resulting from Bianca's choice being a best response to Anders's choice and Anders's choice also being a best response to what he can anticipate will be Bianca's choice.

Anders is the first mover. To decide what to do, Anders anticipates what Bianca will do in response. He goes directly to the payoffs. The best he could do is to get 4, that is, if he did not bargain and Bianca chooses Restrict. But, then, looking at Bianca's payoffs, he can see that if he does not bargain, she will choose Use, because she will then get 5 rather than 4. And he will get 1.

So Anders now considers the bargaining option. The best he can do is 2, that is if Bianca accepts his bargain. To figure out if she will, he needs to consider how Bianca will make her choice. Bianca would think that she can get 5 if she chooses Reject and can use her machine in an unrestricted way. But she could get 6 if Anders offers her 2 to secure an agreement to restrict her use (that is, her payoff of 4 if she restricts plus the 2 that Anders offers in this case).

Anders will be better off bargaining with Bianca, giving her 2, to Restrict and ending up with the 4 he gets if Bianca chooses Restrict, minus the 2 he pays to Bianca, or 2 for himself, which is better than the 1 he would get if Bianca chose to Reject his offer. So the outcome of the game will be (Bargain (to restrict), Accept) with the total payoffs of the two being $6 + 2 = 8$, which is larger than the total payoffs of the two, or $5 + 1 = 6$, *without* the bargain.

Coase's point, "negotiations between the parties would lead to those arrangements being made which would maximize wealth," applied here, means that the noise problem would be solved by the bargaining process providing Bianca's incentives for Accept a bargain and restrict the use of the machine that imposes the negative external effect. This is true even though she has the right to use her machine in an unrestricted way (she was free to Reject Anders's offer).

Bargaining when Anders has initial rights

Now consider the alternative case in which Bianca initially does *not* have the unrestricted right to use her metal saw. Could she bargain with Anders to get that right?

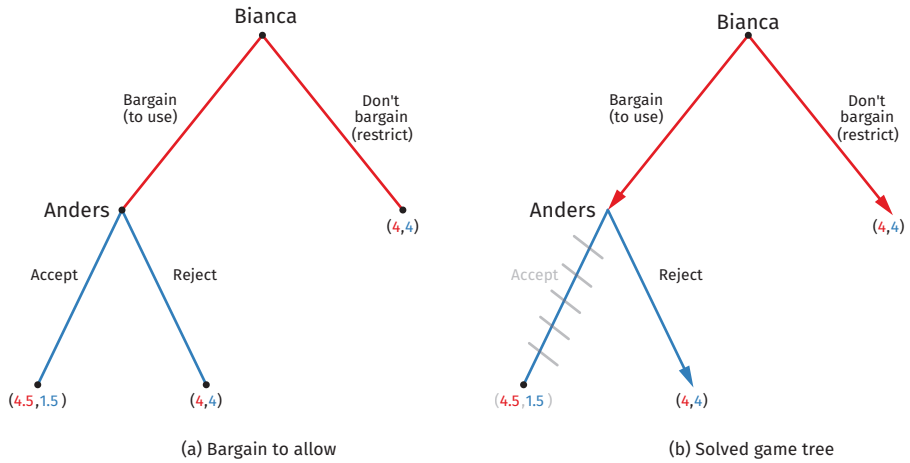
Figure 14.18 shows that the answer is "no." Alternatively, Bianca could propose to Anders to transfer the right to unrestricted use to herself, in return for a payment from Bianca to Anders. Recall that if Bianca has the right to unrestricted use, she receives a payoff of 5 and Anders receives a payoff of 1. Suppose Bianca offers Anders half of what she would gain by unrestricted use of her machines. This would be 0.5, as, without the bargain, Bianca was receiving a payoff of 4.

In this case, Bianca would receive a payoff of 4.5, which is better off than she was without the bargain. Anders, however, would only receive a payoff of 1.5, which is worse off than he was without the bargain, where he received a payoff of 4. He would thus reject Bianca's offer.

Is there any offer he wouldn't reject? Consider Anders's fallback option: he can get a payoff of 4 if he rejects the offer, lets the initial assignment of rights do its work, and Bianca has to Restrict. He therefore needs a payoff of at least 4 when he transfers the right to her.

We saw in Figure 14.17 that if Bianca Rejects when Anders offered a Bargain, Anders would get a payoff of 1. The difference between 1 and 4 tells

Figure 14.18 Bargaining to *allow* when Anders has the initial rights to limit Bianca's use of her machine. Bianca is the first-mover with her decision branches in red. Anders is the second mover with his decision branches in blue. Bianca can Bargain (to allow) or not. If Bianca bargains with Anders, Anders can accept or reject the bargain. Panel (a) shows the game without the solution; panel (b) shows the solution to the game with the pruned branches faded out. Bianca will restrict at either solution.



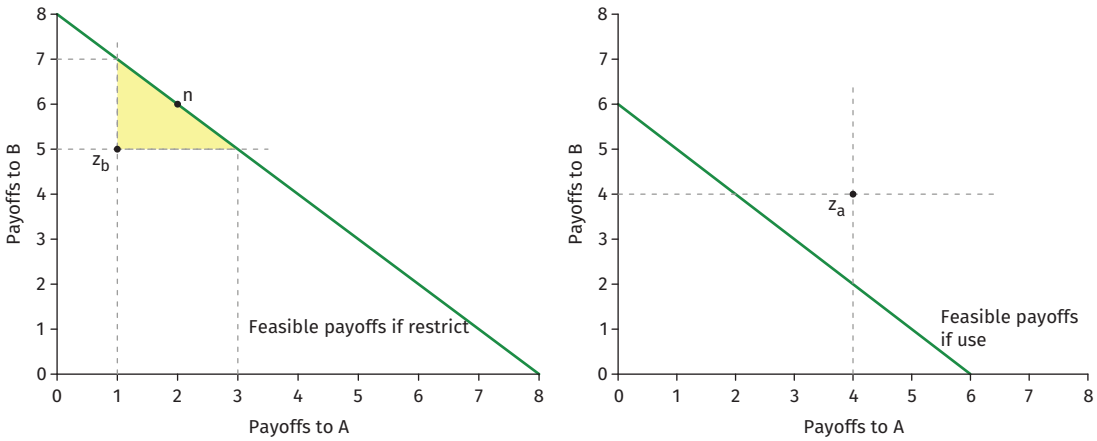
us how much Bianca would have to pay him: $4-1=3$ is the *least* amount Bianca could pay.

But if Bianca paid Anders 3 (allowing her unrestricted use of her noisy machines), she would end up with $5-3=2$. A payoff of 2 is *less* than if she had voluntarily chosen to restrict, where she gets 4 and Anders gets 4. Bianca would not offer this bargain, and will Restrict with a payoff (of (4,4)).

Figure 14.19 allows us to visualize our main result in a different way: that, whether Bianca or Anders has the initial right, bargaining (or not bargaining) will lead to the same outcome–Bianca restricting her use of machines, with a total combined wealth of 8. This is what Coase meant by writing that the external effect will be internalized "irrespective of the initial assignment of rights."

The initial distribution of rights, however, *does* matter in terms of who gets what. Points **n** in panel (a) and $\mathbf{z_a}$ in panel (b) of Figure 14.19 are our two solutions, where Bianca has the initial right and Anders has the initial right respectively. We can see that in both cases total combined wealth will be 8. However, the *distribution* of wealth differs. When Bianca has the

Figure 14.19 The distribution of payoffs depends on the initial assignment of rights. In panel (a), we show the case where Bianca has the initial right to use her machines and Anders seeks to bargain with her. $\mathbf{z_b}$ shows the fallback position with no bargain, and the yellow-shaded area shows all possible Pareto-improving bargains, **n** being the particular solution to our game. The green line shows the feasible frontier for payoffs to Bianca and Anders if they arrive at an agreement to restrict Bianca's use of the machine. In panel (b), Anders has the initial right to restrict Bianca's use of the machine, and Bianca seeks to bargain with him. $\mathbf{z_a}$ shows the fallback position with no bargain, which is more than any of the points on the feasible frontier if they arrive at an agreement to allow Bianca to use the machine without restriction. Therefore, they will *not* enter into this agreement.



(a) Bianca has initial right, Anders seeks to bargain (b) Anders has initial right, Bianca seeks to bargain

initial right, she gets a payoff of 6, while Anders gets a payoff of 2. When Anders has the initial right, they both receive a payoff of 4.

CHECKPOINT 14.12 Explaining the payoffs

a. What is the smallest payment that Anders could have offered to Bianca that would have led her to agree to the bargain and choose Restrict?

b. Suppose that Bianca has the unrestricted right to operate her machine, that Anders plans to operate his yoga studio for just one year, and the payoffs shown are for the year. Use Figure 14.19 to explain how much Anders would be willing to pay if he could bribe the city government to amend the zoning law so as to withdraw Bianca's unrestricted right.

Coasean bargaining: Why it works and why it might not

Looking at the two cases in which Bianca had or did not have unrestricted rights, the outcome–Restrict–occurs independently of who had the initial rights. Notice that the outcome with bargaining when she *did* have the rights, with payoffs (2,6) is Pareto superior to the outcome in that same case without bargaining with payoffs (1,5). But the payoffs when Bianca has the right to Accept and use her machines in an unrestricted way (2,6) are *not* Pareto superior to the outcome in which she did *not* have that right (4,4).

The Pareto improvement shows the Coase theorem in action. As long as the two parties can bargain costlessly:

- *The assignment of rights does* ***not*** *affect efficiency*: Bargaining will internalize the external costs and the inefficiency will be addressed independently of the initial assignment of rights.
- *The assignment of rights* ***does*** *affect distribution*: who has the rights will affect the *distribution* of payoffs. Bianca does better and Anders does worse when Bianca has the right to the unrestricted use of the machine.

As Coase said, the reason why bargaining will implement the Restrict outcome is that the total payoffs under restrict (8) exceed the payoffs when Bianca plays Unrestricted (6). The difference between the two is the sum of two effects:

- the *external cost* to Anders of unrestricted use of Bianca's machine, that is, the difference in Anders's payoffs under Restrict and Unrestricted ($4 - 1 = 3$); and
- the *opportunity cost* of Bianca restricting his use of the machine, that is, the difference between Bianca's payoffs under Unrestricted and Restrict ($5 - 4 = 1$).

Restricting Bianca's use of the machine–eliminating the external cost in the first bullet above–will benefit Anders by 3 but cost Bianca just 1. Whoever

HISTORY In a paper published in 1960 (one of the two that won him the Nobel Prize in economics three decades later) Coase wrote that the idea on which his theorem is based, namely, that there were no costs of bargaining, was "a very unrealistic assumption."[25] To carry out a market transaction people need to discover who it is that one wishes to deal with, to inform people that one wishes to deal and on what terms, to conduct negotiations leading up to a bargain, to draw up the contract, to undertake the inspection needed to make sure that the terms of the contract are being observed, and so on.

has the initial rights, there will be some bargain that will result in the use being restricted.

But in stating his "theorem" Coase was careful to assume that "there are no costs involved in making such bargains." Letting the example be a little bit more realistic makes it clear why.

- Bargaining is costly.
- People affected are many and diverse.
- People adversely affected face a coordination problem.

Bargaining is costly. Imagine that Anders and Bianca are not neighbors able to bargain informally with one another at virtually zero cost but instead hire teams of lawyers to bargain on their behalf.

Return to the left panel of the figure: assume that legal fees are three times any amount transferred, and ask if bargaining will succeed in implementing the outcome (Bargain, Restrict). Remember this occurred because Anders paid Bianca 2 to secure her agreement to Restrict. Studying the payoffs, you will see that he could have offered Bianca just a little more than 1 (so that Bianca would get a bit more than 5 if she chose Restrict). But even paying this lesser amount (just 1) to Bianca, along with the legal fees would cost Anders a total of 4, leaving him with a payoff of zero, which is worse than what he would get if Bianca chose Unrestricted. So there would be no agreement that Anders could have offered that would benefit both and so be both proposed and accepted.

People adversely affected are many and diverse. A second step in the direction of realism provides another reason why Coase's zero bargaining costs assumption is not generally applicable. Pigou's example of the railroad company and the farmers along the train's route is a good one because at least on one side of the interaction there are a large number of people (in this case the farmers).

Before bargaining with the railroad company, the farmers would have to agree with one another, perhaps bargaining about which farmers should pay more or less of the costs of securing the railroad's agreement to redesign their engines, if that is the solution. Even if the external costs imposed on the farmers exceeded the opportunity cost to the railroad of redesigning the engines to avoid the fires, the costs in legal fees and the farmers' own time could be large enough so that no bargain could be struck, and the railroad would continue causing the farmers' crops to burn.

People adversely affected face a coordination problem. Related to the previous problem, as the number of people bearing an external cost increases, the likelihood that they can coordinate to bargain with the party imposing the costs decreases. This problem is sometimes called the problem of concentrated benefits and diffuse costs. The person imposing the cost on others—Bianca in this case—received concentrated benefits from being able

EXAMPLE Imagine Bianca's metalworking studio had two more neighbors: say Caroline's Hair Salon and Deepal's Law Offices. Caroline plays music and so isn't as badly affected as Anders. Deepal really suffers from the noisy machines and it affects her ability to keep law clients. So there are differential harms which would make it difficult for all three of Bianca's neighbors to agree on a common bargain.

to run her machine (the profit she makes from doing so). But, if instead of affecting Anders alone, Bianca's actions affected Anders, Caroline, Deepal, Erkan, Friederike, and many others each of whom bore some costs, but not that great a cost individually, then the costs are *diffused* among many people.

If one or some of them were able to reach a bargain with Bianca to restrict her activities, the result of the bargain itself would be a public good (non-rival and non-excludable, as discussed in Chapters 1 and 5). Each person would therefore have an incentive to free ride on the efforts of others to bargain with Bianca: obtaining the benefits without paying the costs. This means bargaining itself is a coordination problem. We explore this dynamic further in Chapter 16 and explain why it is hard to design policies to overcome such problems of coordinating collective action among many people who bear costs they'd rather others incur.

CHECKPOINT 14.13 **Bargaining to *allow* external effects** Think about a case different from that shown in Figure 14.18 in which the payoff to Bianca from Unrestricted is much larger, 10, which means that the opportunity cost of restricting (that is 6, or 10 minus the 4 she gets if she chooses Restrict) is much greater. The payoff to Anders if Bianca is unrestricted is, as before, 1. Using the reasoning above but with these different payoffs:

a. Show that if the two can bargain, then Bianca will make unrestricted use of her machinery whether she initially has the right to do so or not.

b. If Bianca did not have the unrestricted right to use the machine and if Anders could make a take-it-or-leave-it (TIOLI) offer to Bianca to grant her permission to unrestricted use of the machine, what bargain would he offer?

c. If Bianca did have the unrestricted right to use her machine explain why Anders cannot bargain with Bianca to secure her agreement to Restrict.

14.14 APPLICATION: BARGAINING OVER A CURFEW

Anders's and Bianca's simple game tree clarified the basic idea of Coase's bargaining model, but most bargaining problems involve strategy sets with a much greater range of choices than Restrict, Use unrestricted, Accept, Reject, and so on. Most bargaining models involve strategies like we have considered when studying firms' choices of output, how much time people spend exploiting the fishing stock, and similar choices in which the action of a player can vary over a wide range.

To see how we can use Coasean reasoning in cases like this we introduce a new setting.[26]

The neighbors' utilities and indifference curves

Anna (A) and Bertolt (B) are neighbors. Bertolt enjoys playing loud music late into the night, while Anna worships the rising sun, and hence wants to go to sleep early. The coordination problem they face arises from the negative external effect that Bertolt's music imposes on Anna, similar to the noise of Bianca's metal saw disturbing Anders's yoga sessions.

A government-imposed curfew is proposed specifying the time of night, T, after which no music is to be played. If A could determine the curfew she would set $T = T^A = 9$ p.m., while B would select $T = T^B = 3$ a.m.

M-CHECK So as to be able to write $T^B > T^A$, and to construct the figures below we let $T^A = 9$ p.m. take a value of 1 and T^B, that is 3 a.m. or six hours later as $1 + 6 = 7$.

The time at which the curfew is set is the initial assignment of rights. For example if $T = T^B$, then Bertolt has the unrestricted right to play music whenever he would like. Or, if $T = T^A$, then Anna has the initial rights and can have Bertolt charged with a violation of the curfew if his music disturbs her sleep.

The curfew law has a proviso that the curfew can be reset if the two neighbors can agree on the time. This means that supposing the curfew was set at $T = T^A$ Bertolt could offer Anna a sum of money to secure her agreement to a later curfew.

Here is the game (we do not specify the rules under which they may bargain as that would not add any insights to the model):

- The city's mayor sets a curfew $T^A \leq T \leq T^B$ and informs A and B.
- If $T < T^B$, then B can offer to transfer a sum of money to A in return for her agreeing to some later curfew (reducing his income left over).
- If $T > T^A$, then A can offer to transfer a sum of money to B in return for his agreeing to some earlier curfew (reducing her income left over).
- If one of these offered bargains is accepted by the other, then the mayor is informed of the new agreed upon curfew time, the transfer of funds (if any) occurs, and the two neighbors enjoy the utilities shown below
- This ends the game.

Letting α and β be positive constants indicating the importance of the curfew time relative to income in the well-being of each (the relative intensity of their preferences) here are their utility functions.

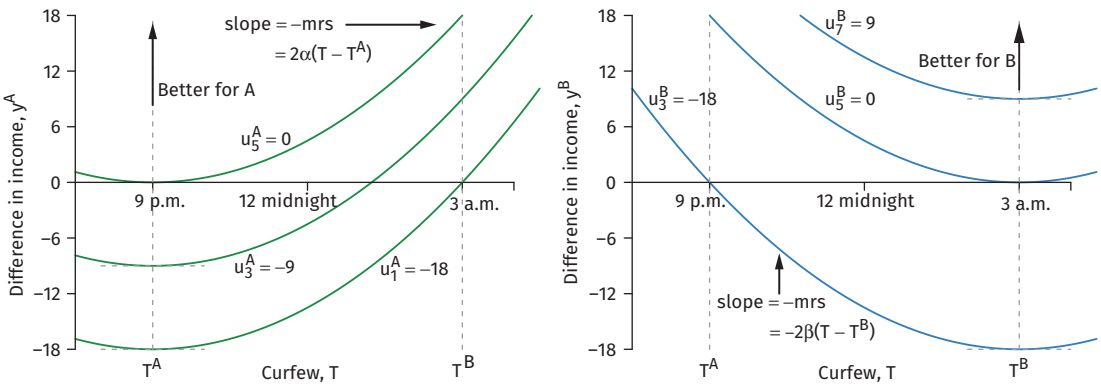
$$\text{Anna's utility: } u^A(T, y) = y - \alpha(T^A - T)^2 \tag{14.10}$$

$$\text{Bertolt's utility: } u^B(T, y) = -y - \beta(T^B - T)^2 \tag{14.11}$$

M-CHECK You can see from the utility functions and the indifference curves that for both of them the curfew could be either too early or too late: Anna would not like to be prohibited from playing music prior to 9 p.m. and Bertholt would not like to be bothered by loud music after 3 a.m.

The term y is a transfer from Bertolt to Anna. Where y takes a negative value it is a transfer from in the other direction, from Anna to Bertolt. For simplicity, we let $\alpha + \beta = 1$. The two utility functions are quasi-linear, so that transfers of y are equivalent to transfers of utility, and the marginal utility of income is a constant equal to one. The neighbors' indifference curves are shown in Figures 14.20 (a) and (b). The vertical axis in both cases is

Figure 14.20 Indifference curves for Anna and Bertolt over payments and the curfew time. The difference in income $y^A = y$ in panel (a) is Bertolt's payment to Anna. In panel (b) the difference in income $y^B = -y$ is Anna's payment to Bertolt. Anna dislikes any curfew later than 9 p.m., so for a later curfew to be equivalent to the earlier one for her it would have to come along with a payment from Bertolt. This is why her indifference curves slope upwards in the after 9 p.m. portion of the figure. A similar reasoning explains why the indifference curves for Bertolt, who dislikes curfews earlier than 3 a.m., slope downward in the portion of the figure before 3 a.m.



(a) A's indifference curves (b) B's indifference curves

the change in income. A positive y-coordinate indicates that the person is receiving income from the other; a negative coordinate means that they are paying the other.

We can see that for Anna, after 9 p.m:

- Later curfew times are a *bad*.
- Payments from Bertolt (positive values on the vertical axis) are a good (alternatively, paying Bertolt ($y < 0$)) reduces her utility.
- Her indifference curves therefore slope upward.

For Bertolt, before 3 a.m:

- Later curfew times are a good.
- A payment from Anna (that is, some $y < 0$) is also a good (alternatively, *paying* Anna, ($y > 0$)) is a bad.
- His indifference curves therefore slope *downward*.

Given the conflict of interest between the neighbors about the curfew time, the best an Impartial Spectator such as the mayor of the town in which Anna and Bertolt live, can do is to minimize the level of disutility each incurs, as we now show.

CHECKPOINT 14.14 Conflicts of interest

a. Why are Anna's indifference curves downward-sloping for times earlier than 9 p.m.?

b. Why are Bertolt's indifference curves upward-sloping for times later than 3 a.m.?

c. Explain why (for both of them) the indifference curves are flat for their preferred curfew time.

d. Explain why the utility associated with Anna's indifference curves is less for curves farther to the right after 9 p.m. and is also less for curves farther down.

A mayor proposes a curfew between two conflicting neighbors

!REMINDER In Chapter 5, the Impartial Spectator had a social welfare function, W, that was the sum of the players' utilities as it is here. In Chapter 4 it was a Cobb-Douglas function.

The mayor of the town faced with this conflict of interest among the (only) two citizens consults Adam Smith's Impartial Spectator. As you recall from meeting her in earlier chapters, the Impartial Spectator places an equal value on the well-being of all citizens. As a result she suggests that the mayor set T to maximize total social utility, $W = u^A + u^B$.

Accepting this idea and assuming that he knows the utility functions of the two citizens, we now explain why the mayor will choose the curfew, T^i (the i superscript is for impartial):

$$\text{Socially optimal curfew} \qquad T^i = \alpha T^A + \beta T^B \tag{14.12}$$

This is a weighted sum of the two preferred curfew times. If $\alpha = \beta$ (meaning the time of the curfew is equally important to the two citizens), the socially optimal curfew is halfway between the two neighbors' preferred times.

Figure 14.21 shows how T^i is determined. To see why midnight is the socially optimal curfew, imagine that an earlier curfew were imposed, for example, $T' = 10{:}30$ p.m. Then comparing points **g** and **h** in Figure 14.21 we see that the marginal disutility to A of a later curfew (that is, $-u_T^A(T')$ shown as point **h**) would be less than the marginal utility to B of a later curfew (that is, $u_T^B(T')$, point **g**) or

$$-u_T^A(T') \quad < \quad u_T^B(T') \tag{14.13}$$

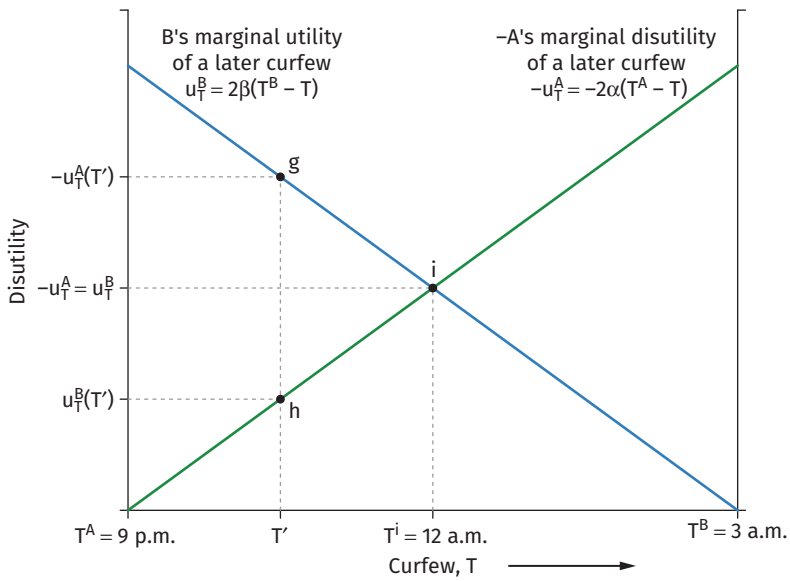
A's marginal disutility of a later curfew < B's marginal utility of a later curfew

Marginal cost of a later curfew < Marginal benefit of a later curfew

We can also interpret Equation 14.13 to say that B's willingness to pay for a slightly later curfew (the right side of the equation) is greater than A's willingness to accept a later curfew, that is, the least amount that she would accept in return for agreeing to a later curfew.

So the benefits of extending the curfew to a later hour exceed the costs, and some later curfew would be socially optimal. Similar reasoning

Figure 14.21 The socially optimal curfew. The horizontal axis is the time of the curfew. The upward-rising line is A's marginal disutility of a *later* curfew, which is zero if the curfew is 9 p.m. and it rises the later is the curfew. B's marginal utility of a later curfew is substantial when the curfew is early and it declines to zero when the curfew is B's ideal time, 3 a.m. The socially optimal curfew T^i is the time that equates these two quantities.



applies to the reverse case—a curfew later than is socially optimal so that the marginal benefits of an earlier curfew exceed the marginal costs. The conclusion is that the socially optimal curfew is the value (T^i) such that:

$$-u_T^A(T^i) \quad = \quad u_T^B(T^i) \tag{14.14}$$

or, the marginal disutility (to A) of a later curfew equals the marginal utility (to B) of a later curfew.

M-NOTE 14.2 The socially optimal curfew

The mayor wishes to impose the socially optimal curfew, taking account of both Anna's and Bertolt's preferences in her social welfare function, W.

$$\begin{aligned} W &= u^A + u^B \\ W &= y - \alpha(T^A - T)^2 - y - \beta(T^B - T)^2 \\ W &= -\alpha(T^A - T)^2 - \beta(T^B - T)^2 \end{aligned} \tag{14.15}$$

To find the maximum total social welfare, we differentiate Equation 14.15 with respect to T and set the result equal to zero.

$$\frac{\partial W}{\partial T} \quad = \quad 2\alpha(T^A - T) + 2\beta(T^B - T) = 0 \tag{14.16}$$

continued

We can rearrange this to read

$$\underbrace{-2\alpha(T^A - T)}_{-u_T^A} = \underbrace{2\beta(T^B - T)}_{u_T^B} \quad (14.17)$$

A's marginal disutility of a later curfew = B's marginal utility of a later curfew

The left-hand side of Equation 14.17 is the green upward sloped line in Figure 14.21; the right-hand side of the equation is the downward sloped blue line in the figure.

To find the socially optimal curfew (T^i) that is implied by this procedure, divide Equation 14.16 by 2 and isolate the T terms:

$$\alpha T + \beta T = \alpha T^A + \beta T^B$$

$$T(\alpha + \beta) = \alpha T^A + \beta T^B$$

$$T^i = \frac{\alpha T^A + \beta T^B}{\alpha + \beta} \quad (14.18)$$

We have assumed $\alpha + \beta = 1$, so

$$T^i = \alpha T^A + \beta T^B \quad (14.19)$$

CHECKPOINT 14.15 The socially optimal curfew

a. Use the text and Equation 14.13 to explain why a curfew later than T^i would not be socially optimal.

b. Suppose that Anna cared less about uninterrupted sleep than Bertolt cared about uninterrupted music, so that $\alpha < \beta$. Redraw Figure 14.21 to reflect this new situation and explain how the socially optimal curfew is affected.

! REMINDER Bargaining is efficient if the allocation resulting from the bargain is Pareto efficient. Coase described efficient bargaining as "costless" or the absence of impediments to bargaining.

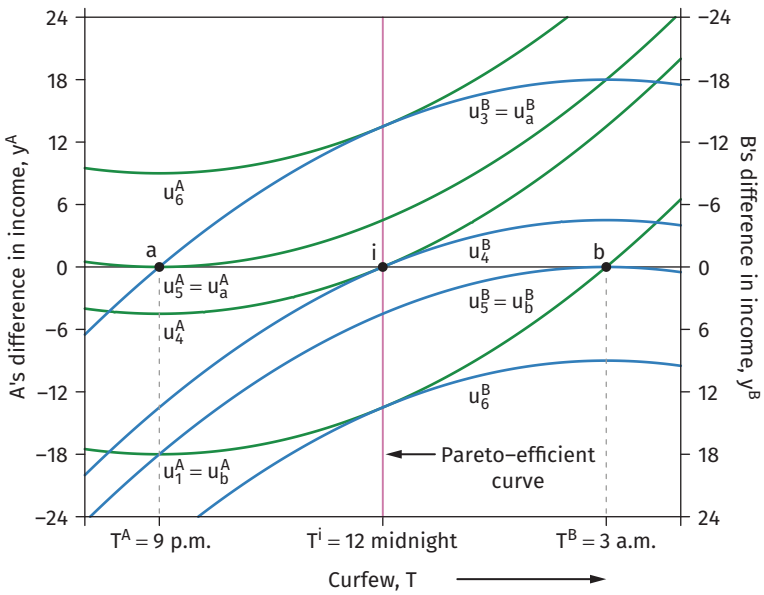
Private bargaining

The Coase theorem says that it doesn't matter for Pareto efficiency which of the two determines the curfew or even if some third party determines it as long as the two can efficiently bargain to rearrange the relevant property rights, meaning in this case, the curfew itself. Suppose the bargaining takes the form of either neighbor paying the other an amount of money. For example B pays y to A in return for A agreeing to a later curfew than whatever is initially announced (with $y > 0$ being a payment from B to A, $y < 0$ would be a payment from A to B for an earlier curfew). Would private bargaining achieve the same result as the curfew chosen by the impartial spectator?

Figure 14.22 illustrates the possibilities of a bargain between the two. The solid horizontal curfew-time line at $y = 0$ is the case in which no money changes hands. Points **a** and **b** correspond to each neighbor's preferred curfew time: Anna's at 9 p.m. (with indifference curves u_a^A for her and u_a^B for Bertolt) and Bertolt's at 3 a.m. (with indifference curves u_b^B for him and u_b^A for Anna). The vertical dimension as before represents a possible

! REMINDER As we did in constructing the earlier Edgeworth boxes, Figure 14.22 is based on Figure 14.20 but with Bertolt's vertical axis scale and indifference curves inverted.

Figure 14.22 Indifference curves at different endowments and at the allocation chosen by the mayor (impartial spectator). The horizontal axis is the time of the curfew (T), with T^A and T^B indicating A's and B's preferred curfews. The vertical axis shows the transfers between A and B in terms of money (y). T^i is the social optimum. As with any Edgeworth box, each point in the figure is an allocation experienced by both A and B, that is, a curfew T and a transfer from B to A, y (which if $y < 0$ is a transfer from A to B.)



transfer between the two, points above the horizontal curfew-time line being allocations with transfers from B to A, and those below the line, transfers from A to B. Any point in the figure is an allocation, that is a bundle of (T, y) experienced by each of the citizens. Figure 14.22 is another Edgeworth box of the kind that you have already seen in this and earlier chapters. But here what is being traded is not two goods, or a good and money, but instead a legal right—to play music or to have uninterrupted sleep—and money.

M-CHECK The Pareto-efficient curve is vertical here for the same reason that it was in Figure 14.11 and other figures earlier in the chapter. Because with the quasi-linear utility function, the marginal utility of y is $u_y^A = 1$ and $u_y^B = -1$, the marginal rates of substitution based on the ratios $\frac{u_T}{u_y}$ are simply u_T. This is why in Figure 14.20 the expressions for the slopes of the indifference curves do not include y.

The impartial spectator's choice—a midnight curfew—is shown by point **i** on the curfew-time line where neither transfers money to the other and each has a utility $u_4^A = u_4^B = 4.5$. The Pareto efficient curve is all of the allocations at which the curfew is at midnight and where the neighbors' indifference curves are tangent or $mrs^A(T, y) = mrs^B(T, y)$. As in Figure 14.20, because the marginal utility of income is constant for both people (due to their quasi-linear utility functions) the indifference curves are vertical displacements of one another.

So, other tangencies can be found along a vertical line through midnight on the curfew-time line, giving the Pareto-efficient curve. Pareto-efficient

outcomes will set the curfew at midnight, but differ in the payments between the neighbors.

We introduced the Impartial Spectator as a thought experiment, to determine what the socially optimal curfew would be. But returning to the real lives of our citizens, consider what would appear to be the worst case, no curfew at all, which means that in the absence of any bargaining between the two, B will impose loud music on A until T^B o'clock every night (3 a.m.). That is, Bertolt has the initial rights just like Bianca did in the previous example to run her machine whenever she wanted to unless Anders struck a bargain with her.

To see if Bertolt and Anna might strike a bargain, consider the interaction between the two as illustrated in Figure 14.23. As in the previous figure, the time of the curfew (T) is on the horizontal axis and the difference in incomes for A and B are measured vertically.

If B has the right to play music as late as he wants, then the initial endowment is point **b** where there is no curfew and no transfer. The values of the fallback options of the two—their utility if no bargain is struck—are shown by the indifference curves u_1^A and u_5^B (these are equivalent to u_b^A and u_b^B in 14.22). These are the participation constraints limiting the bargains that they might voluntarily agree to. The area between them shaded in yellow is the Pareto-improving lens: each point in the lens is an allocation different from the no-bargain outcome at which both would be better off.

Using Equations 14.10 and 14.11 we see that if the status quo curfew is T^B, B gets utility 0 while A gets –18. Both would prefer any point in the Pareto-improving lens. We do not know what bargain the two traders will agree to. The bargain will depend on the institutions and norms governing the bargaining process. Here is an example. Suppose B can make a take it-or-leave-it (TIOLI) offer to A.

! REMINDER A negative y means A transfers to B, whereas a positive y would mean B transfers to A.

Referring to Figure 14.23:

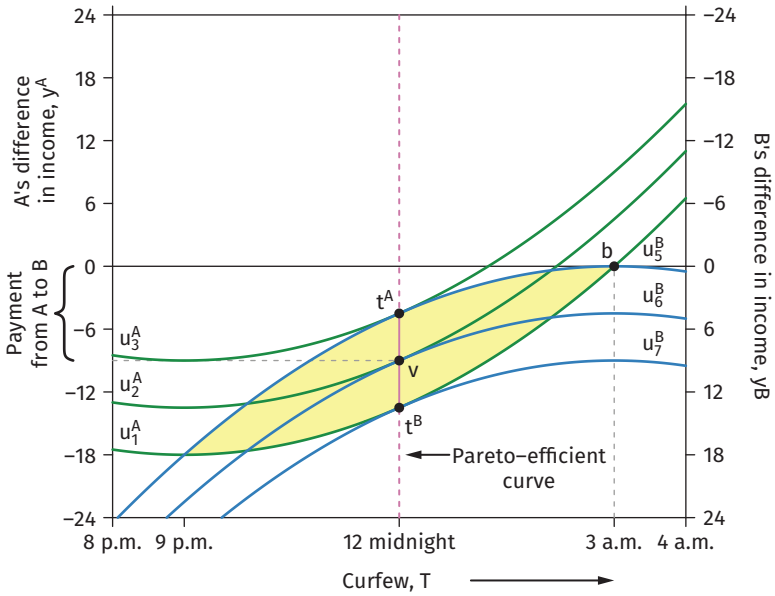
- B will require A pay B the amount $y = -15$ (A has a negative income difference $y^A = -15$ and B a positive with $y^B = 15$).
- B will agree to a curfew set at $T^i = 4$ (midnight).
- The outcome will be $u_{t^B}^A = -18, u_{t^B}^B = 9$.

The result is that:

- the curfew is set at the Pareto efficient time because the two bargained efficiently; and
- all of the gains from the bargaining between A and B have been captured by B because the bargaining rules allowed B to make a TIOLI offer.

From this we can see that Bertolt had two distinct advantages: *the initial assignment of rights* (he was free to play music as late as he wished) and *superior bargaining power* (he could make a TIOLI offer).

Figure 14.23 Optimal Coasean bargaining. The horizontal axis is the time of the curfew (T), with T^A and T^B indicating A's and B's preferred curfews. The vertical axis shows the money transfers from B to A (y). Midnight is the social optimum and the Pareto-efficient curfew. The initial endowment is point **b**, which means that B has the right to play music as late as he wants (3 a.m.) If they arrive at a negotiated solution to split the rents 50-50, then they will arrive at point **v**. We assumed $T^A = 1$ (9 p.m.), $T^B = 7$ (3 a.m.), $\alpha = \frac{1}{2}, \beta = \frac{1}{2}$.



M-CHECK The Pareto-improving lens must exist because at $T = T^B$, $du^B/dT = 0$ (T^B is *B*'s preferred curfew time) while $du^A/dT < 0$ (A would benefit from an earlier curfew). So the marginal cost to B of a slightly earlier curfew is virtually zero, and as a result there will exist some combination of an earlier curfew $dT < 0$ and some payment from *A* to *B* $-y$ that will make both better off.

Even without TIOLI power, however, Bertolt would have done very well due to the initial assignment of rights. Had the two split the rents from their bargaining equally, both gaining 4.5 over their fallback option—shown by point **v** in Figure 14.24—the utilities at the bargained allocation would have been $u^A = -13.5, u^B = 4.5$.

CHECKPOINT 14.16 Switching the initial rights Instead of Bertolt having initial rights as in Figure 14.21, assume that Anna had initial rights to impose a curfew at 9 p.m. and that Bertolt could offer her a bargain to extend the curfew to midnight.

a. Show on Figure 14.22 what the Pareto-improving lens would be.

b. Show the TIOLI offers each would make to the other were a bargain to be made.

c. Indicate an equal sharing of the rents bargain when neither of them has TIOLI power.

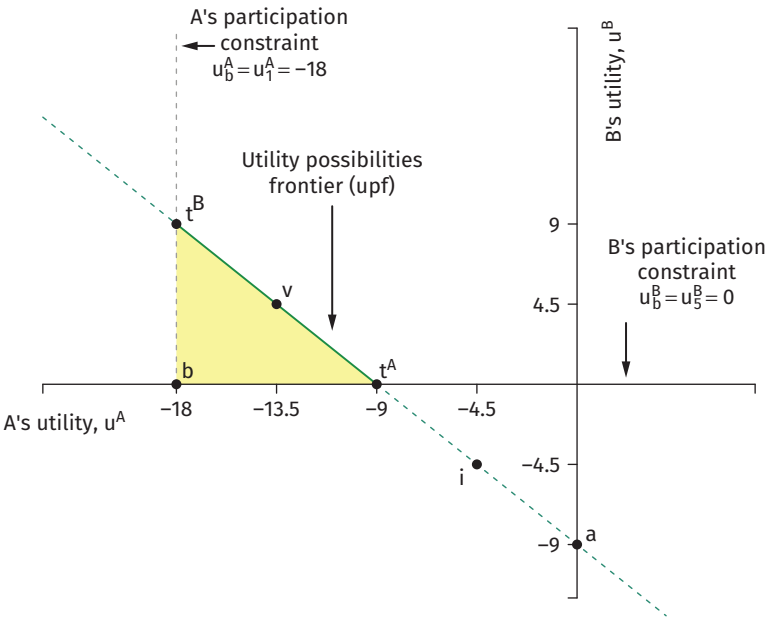
Efficient Coasean bargaining and the bargaining set

What we do know—here is the Coasean condition for an agreement—is that if the *institutions* and *norms* governing the bargaining process allow Pareto-efficient bargains, then the outcome will be Pareto efficient. That is, the allocation of goods and income resulting from traders bargaining with each other will be an allocation along the Pareto-efficient curve within the Pareto-improving lens.

Figure 14.24 shows the utility possibilities frontier and the bargaining set for the allocations of time and money. Because Figure 14.24 is in terms of utility it depicts the surplus that the players can obtain as a consequence of trading: the gains from exchange. The Pareto-improving lens in Figure 14.23 in terms of allocations of the curfew and money corresponds to the utilities in the yellow-shaded Pareto-improving triangle in Figure 14.24.

! REMINDER In Chapter 4 we used the utility possibilities frontier when discussing the choices the Impartial Spectator might make about which Pareto-efficient outcome to choose depending on their social welfare function. Those outcomes that were Pareto efficient lay on the utility possibilities frontier. If bargaining is efficient, the result will be one of the points on the utility possibilities frontier.

Figure 14.24 The utility possibilities frontier and bargaining set that results from the Coasean bargaining. The triangle in yellow is the bargaining set and corresponds to the Pareto-improving lens in Figure 14.23, similar to the Pareto-improving set and utility possibilities frontier you've seen in Figure 4.6 in Chapter 4. Points **a**, **b**, $\mathbf{t^A}$, $\mathbf{t^B}$, **i**, and **v** correspond to the same points in Figures 14.22 and 14.23, but now represent the utilities at each allocation rather than the curfew time (T) and differences in income (y^A and y^B). Point **b** is the no-bargain fallback option. Point $\mathbf{t^B}$ is the outcome of the bargaining when B has TIOLI power. Point **v** corresponds to the implementation of the Pareto-optimal curfew after A makes a payment to B.



To see why this has to be true, imagine that the two had settled on some curfew other than T^i, say T', in Figure 14.21. A would be willing to pay B to further reduce T, the maximum payment offered by A ($u^A(T')$) exceeding the minimum acceptable to B ($u^B(T')$). The outcomes consistent with efficient Coasean bargaining differ from the standpoint of distribution, but all are Pareto efficient as Coase put it, "irrespective of the initial assignment of rights."[27]

Wealth and credit constraints on Coasean bargaining

In our model so far, from the standpoint of Pareto efficiency Coase is right: who holds the property rights does not matter ("questions of equity aside"). But our result illustrating the Coase theorem—showing that the Pareto-efficient allocation results irrespective of the initial assignment of rights—depends on a critical assumption. We assumed that Anna had the funds necessary to pay Bertolt 15. But we know from Chapter 12 that many people are *credit-constrained* or *credit-market excluded*. So it could be that A is not wealthy and does not have (and cannot borrow) the funds necessary to compensate B.

If Anna's exclusion from credit markets had constrained her to pay Bertolt no more than 1, while Bertolt was wealthy enough to be unconstrained, then the initial assignment of rights would have mattered.

In this case if the status quo—no bargain—fallback position had been a 9 p.m. curfew instead of 3 a.m. curfew, then the two could have bargained to the Pareto-efficient midnight curfew, Bertolt compensating Anna 15 for agreeing to the postponement of the curfew. But if Bertolt initially had the right to play music as late as he pleased, Anna would not have had the funds necessary to compensate him sufficiently to secure his agreement to turn off the music at midnight. In this case the initial assignment of rights does affect the Pareto efficiency of the Nash equilibrium.

CHECKPOINT 14.17 **Wealth constraints on efficient bargaining** Use Figure 14.23 to do the following:

a. Show the feasible set of Pareto-improving bargains if A is unable to transfer more than 1 to B, and if the initial assignment of rights is a 3 a.m. curfew.

b. Explain why this means that the result of their bargaining will not be a Pareto-efficient allocation.

14.15 BARGAINING, MARKETS, AND PUBLIC POLICY

The Coase theorem shares with the second welfare theorem the idea that a Pareto-efficient Nash equilibrium can be implemented irrespective of the distribution of wealth (whether that takes the form of an endowment allocation of goods or the initial curfew). The mechanisms ensuring this

result differ—perfect competition and complete contracts for the second theorem and efficient bargaining for the Coase theorem. But the effect is the same: to separate the questions of efficiency and fairness.

Economists have for the most part welcomed the resulting opportunity to focus on questions of efficiency, while setting aside fairness. But the domain in real economies for which the Coase theorem applies may be quite limited. This is due not only to the limited wealth and credit market exclusion or quantity constraints of some actors, but also to the fact that efficient bargaining requires that whatever is agreed upon be part of a complete and enforceable contract.

Among the more surprising claims said to be based on Coase's reasoning is that the assignment of property rights is efficient in actual economies. However, when the Coase theorem is presented sufficiently precisely to be correct, all it says is that *if* there are no impediments for traders to bargain efficiently, *then* outcomes will be Pareto efficient. This seems disappointingly similar to the first welfare theorem.[29] The conditions under which the Coase theorem holds—no impediments to efficient bargaining—include complete contracting in whatever the bargain is about. This has led to a concern that:

- Where the Coase theorem works, the welfare theorems also hold, and the Coase theorem is unnecessary.
- Where the welfare theorems fail (due to contractual incompleteness), the zero-bargaining costs assumed by the Coase theorem will also not hold.

Some have concluded on this basis that when the Coase theorem is needed it fails, and is therefore of little relevance. But this interpretation misunderstands the significance of Coase's analysis. What he pointed out is that trading on competitive markets is *not* the only way to get from an inefficient initial endowment to a point on the Pareto-efficient curve or at least closer to it: people can bargain bilaterally *without* an Auctioneer and still obtain Pareto-efficient outcomes.

We therefore need to understand the Coase theorem not as a case *against* the Pigouvian tax-and-subsidy welfare economics tradition. Instead, where neither markets nor governments succeed, we can understand the Coase theorem as a specification of the conditions under which private rearrangements of property rights may overcome—or lessen the effects of—coordination failures.

By indicating what is required—efficient bargaining—the Coase theorem makes clear just how improbable it is in many situations that private decentralized allocations will be Pareto efficient. In this respect, it may resemble the first welfare theorem: it neither advocates nor opposes decentralized solutions, rather it clarifies what is required for the results to be Pareto efficient.

The Coase theorem also underlines the value of distinguishing between efficiency arguments and fairness arguments concerning policies for cop-

HISTORY Coase's contribution proved controversial, in large measure because it was used to advocate a limited role for government in addressing market failures. James Buchanan and Gordon Tullock wrote: "If the costs of organizing decisions should be zero, all externalities would be eliminated by voluntary private behavior regardless of the initial structure of property rights. There would, in this case, be no rational basis for state or collective action beyond the initial minimum delineation of the power of individual disposition over resources."[28] Advocates of a lesser economic role of government often favored the conclusion and often forgot the "if" clause on which it is based.

✓ **FACT CHECK** The World Health Organization reported in 2018 that about three billion people worldwide cook using polluting open fires or stoves with kerosene, dung, wood, and other biomass. The external cost? About four million people a year die of respiratory diseases related to air pollution caused by air pollutants from such cooking and heating.[30]

ing with coordination failures. The Pigouvian position—for example, that polluters should pay for the harm they do—is often voiced by environmentalists. But does the "make the polluters pay" principle apply if the polluters in question were not carbon-based energy companies but instead people who live in cities in poor nations who cook and heat by means of fires because they lack the income to purchase more environmentally-friendly stoves? Many people would object to this application of the principle on grounds of fairness.

While Coase is helpful in clarifying challenges of this nature, it is not clear how a Coase-inspired solution to this problem—those who care about air quality bargaining with those cooking over wood fires—could work, given the extraordinary transaction costs of arranging a private bargain among all of those involved. A tax-financed subsidy for environment-friendly stoves along with a prohibition of wood fires might be a more effective solution.

CHECKPOINT 14.18 Pigou and external effects caused by people living in poverty

a. Explain why the example above—asking that a person living in poverty who causes pollution should bear the costs of that pollution—is consistent with the Pigouvian position outlined earlier in the chapter.

b. What would a Coasean bargain potentially look like in this case between another party and the poor person who cooks with their polluting stove?

HISTORY In the lecture commemorating his Nobel Prize, Coase commented on the role of economic advisers from the US and the costly transition from a centrally planned to a capitalist economy in the 1990s (see Figure 14.2): "Without the appropriate institutions no market economy of any significance is possible. If we knew more about our own economy, we would be in a better position to advise them."[31]

14.16 APPLICATION: PLANNING VERSUS THE MARKET IN THE HISTORY OF ECONOMICS

The fierce debates over capitalism vs. the centrally planned economy in the 1920s and 1930s—known as the *socialist calculation debate*—were about ideal systems, that is ways of organizing an economy in which troublesome details like incomplete contracts and asymmetric information were simply ignored.

Debating the possibility of central planning

Theoretical economists including Enrico Barone (1859–1924), a student of Pareto, had explored exactly these questions in depth before World War I and the Russian Revolution. Barone concluded that the job of the "ministry of production" of a socialist regime would be to *mimic* the conditions of price-taking competitive equilibrium by equating marginal rates of transformation and marginal rates of substitution across the different sectors of the economy. Because any profits or other surpluses or rents that might accrue to socialist enterprises would be controlled by the socialist state, Barone envisioned the state as redistributing the economic surplus in line with its political preferences over income distribution.[32]

After the Bolshevik Revolution in 1917 that brought the Communist Party to power in Russia and the rest of the Soviet Union, the Austrian philosopher Ludwig von Mises (1881–1973) predicted the failure of the Soviet attempt to institute a "planned" economy. He reasoned that to implement an efficient government plan it would be necessary to solve thousands (or tens of thousands) of simultaneous equations (like Equation 14.6) describing the equality of marginal rates of transformation and marginal rates of substitution for each of the huge number of produced goods and services that make up a modern economy.

Barone and later advocates of what came to be called "market socialism" including Oskar Lange (1904–1965) and Abba Lerner (1903–1982) noted that even if the direct solution of the numerous equations describing the conditions for Pareto efficiency proved to be beyond the capacity of existing mathematical methods, a socialist economy could in principle instruct its managers to act *as if* they were price-taking profit-maximizing capitalists to achieve the same efficient outcomes.

HISTORY Just about the same time that the socialist calculation debate erupted, Alan Turing (1912–1954), John von Neumann (1903–1957), and other pioneers of the computer age were laying the foundations for the construction of digital computers that would come to have the power to solve very large systems of equations of the type that a central planner would use. A bit later the Russian mathematician and economist Leonid Kantorovich (1912–1986) was developing computational methods for solving very large constrained maximization problems of the type that central planners face.

The backdrop to the debate during the 1920s and 1930s was the Great Depression in the capitalist countries (except in fascist Germany) and rapid economic growth in the Soviet Union under their initial five-year plans (see Figure 14.2). By the 1940s the debate was all but over, and surprisingly the "pro-planning" side appeared to have won. Even the arch opponent of socialism, Joseph Schumpeter, conceded: "Can socialism work? Of course it can. . . . There is nothing wrong with the pure theory of socialism."[33] He was echoing another opponent of socialism, Pareto, who much earlier had affirmed the feasibility of rational economic calculation in what he called "an argument in favor of collectivist production," concluding that "pure economics does not give us a truly decisive criterion for choosing between the organization of society based on private property and a socialist organization."[34]

HISTORY Long before the debate on central planning, John Stuart Mill had provided a critique of the problems of a hypothetical nonmarket economy—worker motivation, and reduced innovation—far more carefully than any produced prior to Hayek in 1945.

What then *was* wrong with centralized planning? And what was wrong with the economic theory that so inadequately captured the economic shortcomings of centralized allocations and had seemingly vindicated the planned economy?

The problem of information and price signals

A striking feature of the calculation debate is that both sides had deployed the *same* economic model on behalf of their *opposing* arguments. This was the perfectly competitive general equilibrium model along with the often implicit assumption that all actors had access to information about prices and trading throughout the economy.

Hayek soon appreciated the error and counterattacked on stronger grounds. In his 1945 paper "The Use of Knowledge in Society" he reframed the debate in terms of the costs and limited availability of information, ideas the perfectly competitive general equilibrium model left out.

HISTORY One of the leaders of the Russian Revolution, Leon Trotsky, explained the essentials of Hayek's idea well over a decade prior to Hayek's celebrated paper: "If a universal mind existed, such a mind, of course, could *a priori* draw up a faultless and exhaustive economic plan, beginning with the number of acres of wheat down to the last button for a vest. The bureaucracy often imagines that just such a mind is at its disposal . . ."[35]

› EXAMPLE In this video (tinyurl.com/y3pcuxa8), Joseph Stiglitz comments on why the financial crisis was a market failure (from the CORE project. www.core-econ.org).

The problem with centralized economics planning, according to Hayek, is that the information needed by the planner is *privately* held by millions of economic actors. The actors lack the will and the means to transfer their information to a central authority. By contrast, according to Hayek, *decentralized* markets make effective use of dispersed information. People know their own preferences and respond to prices. Under ideal conditions, people observe prices and behave as if the prices reflect the scarcity of the goods in question.

By focusing attention on which institutions more effectively use the available information, Hayek's paper counts as a landmark work in the theory of economic institutions. In identifying a major shortcoming of centralized planning, Hayek also pointed to the deficiencies of the perfectly competitive general equilibrium model, namely the assumption that all actors have sufficient information so that complete contracts can regulate the exchange process.

For some economists, the general equilibrium model of perfect competition had provided a strong justification to leave the allocation of resources to the market and to limit government activities to providing institutions ensuring competition and private property rights. But most economists, including Arrow, held that the empirical implausibility of the first welfare theorem's assumptions—especially complete contracts—disqualified it as a defense of limited government.

CHECKPOINT 14.19 The perfectly competitive model and central planning Economics Nobel Prize winner Joseph Stiglitz wryly observed that "if the neoclassical model of the economy [meaning, the perfectly competitive complete contracts model] were correct, . . . centrally planned socialism would have run into far fewer problems."[36] Use Hayek's ideas presented here and in the head quote to Part IV of this book to explain why this seemingly contradictory observation by Stiglitz could be true.

14.17 PERFECT COMPETITION, MARKETS, AND CAPITALISM

The perfectly competitive general equilibrium model is not really about capitalism, or any other market system. Nor does it capture even the idealized logic of a system of decentralized allocation among people with limited information of the type that Hayek described. The all-knowing and all-powerful Auctioneer that Walras invented to provide the equilibrium prices of the model bears a striking resemblance to the equally idealized central planner.

Markets play no real role in this model, nor is the model consistent with any plausible *process* of how equilibrium is reached. The reason is that buyers and sellers do not set prices (they are "price-takers").

Fortunately economics has a lot more to offer about how markets work than the perfectly competitive market equilibrium, as the contrast in Table 14.3 between prefect competition and the model of bargaining with zero transactions costs that we introduced and simulated in sections 14.10 and 14.11 illustrates.

It is not surprising then that the claimed Pareto efficiency of the perfectly competitive equilibrium outcomes now plays virtually no role in scholarly discussions of economic policy and institutions. Instead people have refocused on the practical question of choices among feasible institutions and policies supporting real-world outcomes that improve welfare, but may not be Pareto efficient.

The fact that the conditions under which even highly competitive markets will implement Pareto-efficient outcomes are unlikely to be observed in real economies is not a reason to abandon or limit the use of markets as a component of our system of economic institutions. It is simply a reminder that discovering an appropriate balance of markets, governments, and other institutions requires comparing imperfect systems, not idealized ones. And in this comparison a key element must be, as Hayek wrote in the head quote of this part of our book, which actors—whether it be the owners of firm, their employees, government officials, or consumers—have information adequate to implement improved outcomes.

In this practical task, however, the lessons of the welfare theorems remain important. Under the right conditions, people acting autonomously in pursuit of their own interests may implement socially desirable outcomes. Enhancing the capacity of private actions—either buying and selling on markets, or Coasean bargaining—to accomplish these ends remains an important aim of policy.

HISTORY Arrow and Hahn drew attention to this gap: "If we did not stipulate . . . an auctioneer, we would have to describe how it comes about that at any moment of time two goods exchange on the same terms wherever such an exchange takes place and how these terms come to change under market pressure."[37]

CHECKPOINT 14.20 **Rent seeking and perfect competition** Make sure you can explain in your own words the contrasts in each row of Table 14.3.

Table 14.3 A comparison of rent-seeking competition with zero transaction costs and perfect competition.

	Rent-seeking competition (Sections 14.10 and 14.11)	Perfect competition (Sections 14.3 to 14.9)
Actors	Price-makers	Price-takers
Law of one price	No	Yes
Equal treatment	No	Yes
Distribution	Markets affect inequality	Distribution neutrality
Pareto efficiency	Yes	Yes
How prices change	Rent-seeking, bargaining	Fictional Auctioneer

14.18 CONCLUSION: IDEAL SYSTEMS IN AN IMPERFECT WORLD

The classical institutional challenge that we introduced in Chapter 1—how can we design a set of institutions that will improve the conditions of humanity—continues to be an important theme in economics. In the last half century or so—often assisted by mathematical reasoning—we have been able to contribute some light to the often heated debates about the invisible hand and its policy implications.

Three conclusions stand out:

- The perfectly competitive general equilibrium approach on which the welfare theorems are based does *not* provide a model of how the decentralized process of competition among firm owners and other people would reach and remain at some particular allocation.
- The assumptions under which a perfectly competitive equilibrium—if reached and sustained—would be Pareto efficient are unlikely to hold in any real economy.
- The main value of the bargaining approach proposed by Coase is that it expands the menu of institutional choices beyond perfect competition versus centralized government allocations. But the conditions under which bargaining would achieve Pareto-efficient outcomes are, like those underlying the first and second welfare theorems, unlikely to hold as a general rule, even approximately in real economies.

These somewhat disappointing results do not detract from the contribution of the invisible hand and the debates surrounding it. They clarify the economics of the world we live in. In this world, a process approximating Smith's invisible hand reasoning can sometimes actually operate—bargaining among private actors—providing a set of institutions that along with government policies to address coordination failures is superior in many respects to some of the alternatives, including a highly centralized economy, like the one that Nikita Khrushchev advocated in the kitchen debate with Richard Nixon.

The lessons of the perfectly competitive general equilibrium model, efficient bargaining, and the invisible hand motivate the final two topics we will address.

First, in the next chapter we turn away from the abstract and idealized world of the perfectly competitive equilibrium and efficient bargaining to a real historical and current entity: capitalism and its economic constitution. The inhabitants of this world are not the undifferentiated price-taking "traders" who we have considered here but instead a much more interesting and lifelike cast of characters: employers and employees, lenders and borrowers, the wealthy and the property-less, the included and the excluded, the price-making first movers, and the second movers.

Second, in the final chapter we will study governmental policies and how they might improve the functioning of a capitalist economy. The

observation that the conditions under which Adam Smith's invisible hand will work are not realized in any economy is not a sufficient basis for concluding that government interventions in the economy will improve economic outcomes. While well-designed policies can play an essential role in sustaining both more fair and more efficient outcomes, idealized models of the government (like the idealized model of the perfectly competitive economy or the process of bargaining studied here) are inadequate.

MAKING CONNECTIONS

Rules of the game: Perfect competition and the perfect competitor: Two views of competition—the perfectly competitive general equilibrium exchange process governed by complete contracts and coordinated by an "auctioneer" and decentralized bargaining among rent-seeking buyers and sellers—differ in how they represent the game in which buyers and sellers are engaged.

Mutual gains from trade (rents): People choose exchanges and other economic interactions over their fallback endowments because, through buying and selling or bargaining, they share in the mutual gains from trade that these interactions allow.

The distribution of the resulting rents: In the perfectly competitive general equilibrium model the process of exchange among price-takers does not affect the distribution of wealth among the parties to the exchange. This is not the case under alternative rules of the exchange game such as decentralized bargaining.

Wealth redistribution and fairness: Both models—bargaining and perfectly competitive general equilibrium—clarify the relationship between the initial distribution of wealth (endowments) or rights and the resulting post-exchange differences in utility or wealth of the players. This provides the basis for considering policies that redistribute endowments and rights in order to promote fairness.

Pareto efficiency: There are conditions under which the Nash equilibria of both models are Pareto efficient: complete contracts in the perfectly competitive general equilibrium model and efficient bargaining in the bargaining model.

Optimization by the $mrs^A = mrs^B$ and $mrs = mrt$ rules: We used both rules: the first for identifying points on the Pareto-efficiency curve, the second for understanding buyers, and sellers, individual optimization underlying their price–offer curve.

Dynamics: In contrast to the bargaining model, the perfectly competitive general equilibrium model provides just a snapshot of an equilibrium outcome but not a "film" of a process by which that outcome could be reached or sustained.

History: Over the last century economics has evolved both through the application of mathematical reasoning to our studies and through an engagement with dramatic changes in economy and society, including the emergence and collapse of an alternative economic system—centralized planning—and the Great Depression. Today we anticipate that engagement with current realities—climate change, pandemics, mounting inequalities, and the information revolution—will continue the process of change.

IMPORTANT IDEAS

decentralized exchange	convexity assumption	coordination failure
Coase theorem	marginal rate of substitution	Edgeworth box
Pareto-improving lens	utility possibilities frontier	Pareto-efficient curve
first welfare theorem	Auctioneer	endowment allocation
post-exchange allocation	market clearing	excess supply
second welfare theorem	gains from trade	institutions
excess demand	market failure	wealth inequality
social costs and benefits	incomplete contract	uncompensated external effect
equal treatment	distributional neutrality	invisible hand
efficient bargaining	non-verifiable information	asymmetric information
Nash equilibrium	initial assignment of rights	path dependence
rent seeking	costless bargaining	property rights
transaction costs	marginal social rate of transformation ($msrt$)	perfect competitor
dynamics		incomplete contract
socialist calculation debate	ideal systems	
central planning	computer simulation	

MATHEMATICAL NOTATION

Notation	Interpretation
x, y	goods to be allocated among people
p	relative price of a good, i.e., $\frac{p_x}{p_y}$
$u(\ )$	utility function
c	marginal cost of production
m	wealth
T	time of curfew
α and β	disutility parameter for a curfew different from A's and B's ideal time
W	social utility function

Note on superscripts and subscripts: A, B, i: different people; N: Nash Equilibrium; z: endowments.

CHAPTER

15

CAPITALISM

INNOVATION AND INEQUALITY

the proprietors of . . . establishments and their operatives do not stand on an equality, . . . their interests are, to a certain extent, conflicting. The former naturally desire to obtain as much labor as possible from their workers, while the latter are often induced by the fear of discharge to conform to the regulations which [are] detrimental to their health or strength. In other words, the proprietors lay down the rules and the laborers are practically constrained to obey them.

US Supreme Court,
Holden v. Hardy (1898)[1]

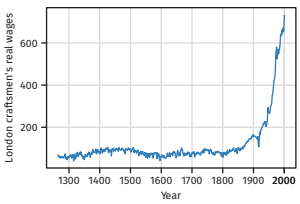
DOING ECONOMICS

This chapter will enable you to:

- Understand why in many countries "the capitalist revolution" brought rapid increases in material living standards.
- Explain how, by placing decision-making power in the hands of wealthy and hence not very risk-averse owners and managers, the capitalist firm can support risk-taking that promotes innovation.
- Explain that measures to reduce inequality such as tax-financed social insurance and publicly provided education and health services *reduce* risk exposure and can also promote innovation.
- Understand Lorenz curves and the Gini coefficient as measures of inequality.
- Use a model of the whole economy to explain how decreased competition among firms may increase inequality and how government support for research and education (raising productivity) can have the opposite effect.
- Analyze the workings of a dual economy with both a capitalist economy sector and an informal sector.
- Explain the exercise of power by employers over workers and lenders over borrowers as well as sociological and psychological aspects of relationships between principals and agents.
- Understand the advantages and limitations of a worker-owned and democratically managed cooperative as an alternative to the capitalist firm.

Figure 15.1 London Craftsmen's Real Wages from 1264 to 2001. Shown is an index whose value is set to 100 for the year 1850. Not having changed much over the six centuries prior to 1850, wages rose to almost seven times higher in the 150 years after 1850.

Source: Allen (2001).



✓ **FACT CHECK** Look carefully at the Figures 15.1 and 15.2. In the UK GDP per capita began its sustained growth about a century before the wages of London craftsmen's started to rise. Workers did not get much of a share in the growth of income until their bargaining power increased in the mid-nineteenth century.

! REMINDER In Chapter 14 we discussed the centrally planned economy of, for example, the Soviet Union, that for most of the twentieth century represented an economic system that was an alternative to capitalism. Today, with the exception of North Korea, Cuba, and possibly China and Vietnam, most economies in the world are capitalist by our definition, for example, all of the countries listed in Figure 15.2 and later in the chapter in Figure 15.11.

15.1 INTRODUCTION: CAPITALISM AND HISTORY'S HOCKEY STICK

In the fourteenth century, Ibn Battuta (who you met earlier in Chapter 4), one of the leading geographers and explorers of his age, traveled widely in Asia, Africa, the Middle East, Russia, and Spain. In 1347, he visited the land we now call Bangladesh. "This is a . . . country . . . abounding in rice," he wrote. He described traveling along its waterways passing "between villages and orchards, just as if we were going through a bazaar." Six and a half centuries later, one-third of the people of Bangladesh were undernourished, and the country was among the world's poorest.

About the time of Ibn Battuta's visit to Bangladesh, Europe was reeling under the impact of the bubonic plague, which took the lives of one-quarter or more of the residents in many cities. Manual workers in London, probably among the better off anywhere on the continent, consumed fewer than 2,000 calories per day (a physically active man needs close to 3,000 calories per day to maintain his health).[2]

The shortage of labor following the plague boosted real wages through the middle of the next century. But over the next four centuries real wages of laborers did not rise in any European city for which records exist: in most cities wages fell by substantial amounts and in northern Italy wages fell to half their earlier level.

But, beginning in the nineteenth century, real wages rose dramatically first in England and later by even greater amounts in other European cities. Figure 15.1 illustrates the increase, showing the real wages of London craftsmen from 1264 to 2001 including the recent sevenfold increase. The astounding increase in wages in London was replicated in similar increases in average living standards (measured by percapita income) in Great Britain, followed by even more dramatic increases in Japan and Italy. Some of the evidence can be seen in what we call the hockey stick of history, shown in Figure 15.2.

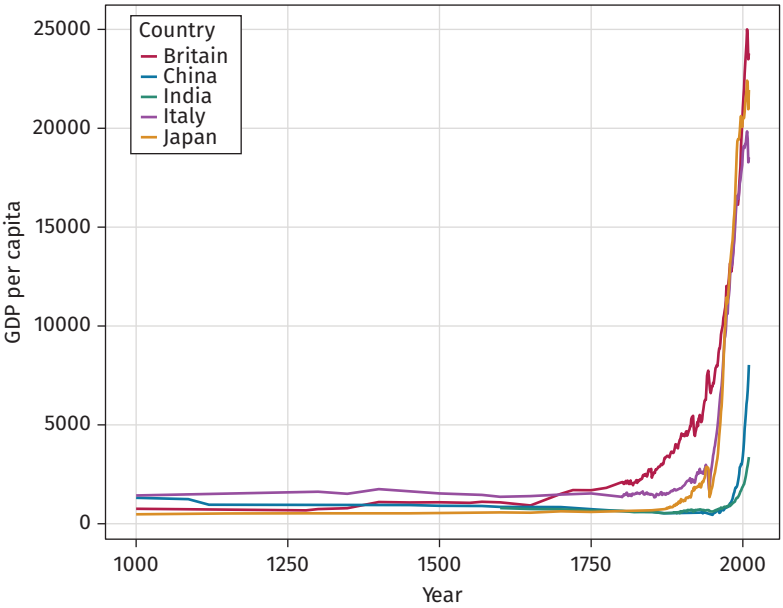
As Figure 15.2 shows, at the time of Ibn Batutta's visit to Bangladesh (1346) the world was flat, economically speaking: countries and regions did not differ much in their average living standards. The main inequalities were between the rich and the poor within a country: landlords and farmers, masters and enslaved people, men and women.

Starting around the middle of the eighteenth century this pattern began changing as vast differences in average incomes began to develop between the rich and poor countries and regions.

How did this happen?

Figure 15.2 History's hockey stick: GDP per capita in Britain, China, India, Italy, and Japan from the year 1000 to 2004. Over most of the last thousand years, based on the available data in the figure, people living in what is now Italy were the richest in the world measured by per-capita gross domestic product. Go to the different version of the "history's hockey stick" figure at tinyco.re/27937150 where you can see that prior to the late twentieth century the incomes of Indian and Chinese people were on average falling.

Source: CORE, The Economy, based on data from the Maddison Project and Our World in Data.[3]



REMINDER As in Chapter 14, in this chapter we represent the economy as an entire system, with many markets. But the system we describe here—capitalism—bears little resemblance to the perfectly competitive general equilibrium model. To understand capitalism we will draw upon what you have learned about risk (Chapter 13) and the process of competition (Chapter 9) and about how markets connecting principals and agents especially the labor market and the credit market (Chapters 11 and 12), as these are keys to understanding how different capitalism is from the model of perfect competition.

CHECKPOINT 15.1 The differing hockey sticks of various countries and global inequality

a. Use the "hockey stick" figure to explain the meaning of "the world was flat" for most of the last millennium.

b. What does the figure tell you about inequality among the peoples of the world in the years 1750, 1850, and 1950?

c. Are there any clues in the figure about why global inequality has decreased in the past half century (it has)?

! REMINDER Institutions are the laws, social norms, and "rules of the game" that govern how people behave and interact in social interactions.

› EXAMPLE In this video (tinyurl.com/y2tprqrg), Suresh Naidu, an economic historian, explains how population growth, technological development, and political events interacted to produce the real wage hockey stick (from the CORE project. www.core-econ.org).

✓ FACT CHECK In a 2018 YouGov survey in the US 44 percent expressed a "very" or "somewhat favorable view of capitalism." Among those polled who were between 18 and 29 years of age "socialism" was viewed more favorably (35 percent) than "capitalism" (30 percent). We do not know, of course, what the survey respondents meant by either of these words.[4]

15.2 CAPITALISM'S SUCCESS: INNOVATION AND ECONOMIC GROWTH

The answer, very briefly, is **capitalism**. In this chapter we introduce capitalism as an economic system and focus on two attributes that are frequently associated with this set of rules of the game:

- growth in living standards supported by *innovation*; and
- *inequalities* in both material living standards and economic decision-making powers.

Capitalism is an economic system in which most production takes place in privately owned **firms** that employ labor in return for wages or salaries to produce goods and services to be sold on markets to make a profit for the owners of the firm's capital goods.

As a historical and contemporary entity, capitalist economic systems have included work done by government officials, unpaid work in the home, and the work of enslaved people. Capitalism contrasts with government ownership of capital goods in a centrally planned economy, where private firms and markets are relatively unimportant. Another contrast: in a slave economy, most of the work is done by people who are *not* hired for wages but, instead, like the land on which they work, are the property of another person.

Other forms of economic organization coexist with firms in a capitalist economic system, but they are not firms: family or individual production (they do not hire others); nonprofit organizations (they do not seek to make profit or sell their output on a market); cooperatives (labor is not hired, work is done by members); government bodies (they do not seek profit; capital goods are not privately owned)

Three characteristics of a capitalist firm help to explain why capitalism as a system has been so innovative. They concern:

- *Control rights*: Who makes *decisions* about how the firm is run, what it produces, what technologies are used, and so on?

CAPITALISM Capitalism is an economic system in which most production takes place in privately owned firms that employ labor in return for wages or salaries to produce goods and services to be sold on markets to make a profit for the owners of the firm's capital goods.

FIRM A business organization in which private owners of capital goods hire and direct labor to produce goods and services for sale on markets in order to make a profit is called a firm (or sometimes a capitalist firm).

- *Residual claimancy*: Who *owns* the *net revenues* (accounting profits) from the sale of what is produced after the payments for the inputs have been made?
- *Competition and survival*: The survival of the firm depends on its making profits, which requires producing goods that are in demand at a cost comparable to or lower than competing firms.

In a capitalist firm, the *owners* of the capital goods used in the production process have the control rights and are residual claimants on the revenues of the firm. This is what ownership of a firm means.

The "capitalist revolution": Creative destruction

Capitalism is an economic system of recent origin, having its roots in the urban economies of northern Italy, England, Belgium, and the Netherlands starting around 500 years ago. Capitalism expanded rapidly, first in Europe, later in the places where European migrants located or were colonized by European nations, and eventually to most economies in the world. Capitalism inaugurated a new economic era as different from what preceded it as did the emergence of agriculture and the spread of the new institutions associated with settled agriculture roughly 11,000 years before. One of the most striking outcomes of the "capitalist revolution" was the rapid increase in the productivity of labor, making possible an extraordinary and prolonged increase in people's material living standards, especially in those countries where workers' bargaining power was augmented by the expansion of workers' political rights. This happened in Europe and not in Bangladesh.

This accomplishment of productivity and improved well-being is not controversial even among the most severe critics of capitalism—Marx and Engels stressed it in their 1848 *Communist Manifesto*. But adopting a capitalist economic system was neither sufficient nor necessary for rapid economic growth.

- *Not sufficient*: Not all capitalist economies have prospered, for example, many Latin American economies over the twentieth century experienced very modest economic growth.
- *Not necessary*: Some other economic systems have also fostered rapid economic growth. For example, the Soviet Union under centralized economic planning, from the Great Depression until the 1970s, or Vietnam and China under a mixture of markets and centralized economic planning since the 1980s.

What capitalism accomplished, and what accounts for much of its productive success, is that capitalism provided conditions under which some individuals would innovate on a grand scale. When people *innovate* they introduce new technologies, new products, and new ways of organizing

Figure 15.3 Joseph Schumpeter (1883–1950). Schumpeter was Finance Minister of his native Austria before becoming a professor of economics at Harvard University. Like Karl Marx, he described capitalism as an economic system that was constantly in motion that would eventually succumb to what he called "the march into socialism." Unlike Marx, he was no friend of socialism: his most famous book *Capitalism, Socialism and Democracy* celebrated the successes of capitalism.[5]

Photo: Granger Historical Picture Archive/Alamy Stock Photo.

EXAMPLE In this video (tinyurl.com/y39gremt), Lynne Kiesling, a historian of economic thought, discusses Joseph Schumpeter (from the CORE project. www.core-econ.org).

HISTORY Joseph Schumpeter is considered the father of the field of innovation economics. In 1950 he wrote: "in capitalist reality, as distinguished from its textbook picture, it is not [price] competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization . . . competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits . . . but at their foundations and at their very lives."[6]

production and marketing. Capitalism created *incentives* for people to innovate by giving people who had sufficient resources:

- a reasonable expectation of reaping substantial *profits* if they successfully innovated;
- while bearing the *costs* if they failed.

This process was based on substantial inequalities in wealth combined with credit and other financial markets that allowed a single individual or a small group of people to amass substantial resources under unified control and take risks on a grand scale. Labor markets allowed these material resources to be put to use to employ vast numbers of workers, so the owners of the capital goods of a firm could reap the rewards of technological innovation and economies of scale.

The key to the success of capitalism in raising living standards in the countries shown in the figure is that it brought about a permanent technological revolution, a never-ending process whereby new products, new technologies, and new forms of organization would be introduced and widely diffused, resulting in a long-run increase in amount of output produced in an hour of labor. The increase in the productivity of labor permitted but did not guarantee that the benefits of the permanent technological revolution would be widely shared. You can see from Figure 15.1 that wages did not begin to rise in London until at least half a century after the epoch-making innovations of the mid- to late eighteenth century such as Watt's steam engine and Hargreaves's spinning jenny.

For the people directing these business projects, their own personal wealth and their ability to borrow funds made the risks of innovation tolerable. For the first time in history, surviving in the competition among members of the economic elite depended on one's success in introducing unprecedented ways of organizing production and sales, new technologies, and novel products. Joseph Schumpeter referred to the process of innovation as "creative destruction." The creative part was the new products, new employment opportunities, and increased productivity associated with the innovation. The destruction part was the fate of the losers in this process: those whose jobs are lost, or businesses bankrupted.

The success of these arrangements of work and innovation hinged critically on the relative security of possession associated with the **rule of law** (including private property rights), accomplished in large part

RULE OF LAW Under the rule of law all people—including those who make the laws, police, heads of state, and other government officials—are subject to the law. In game theoretic terms rule of law means that irrespective of the personal identity of the players the rules of the game govern the interaction for all players, including rules governing how the rules of the game can change.

by the increasingly powerful nation-states that grew in conjunction with capitalist economic institutions.

The rule of law was important: contracts had to be enforced. But capitalism's success did not hinge on contracts being complete. Quite the contrary, capitalism fostered the rapid diffusion of new techniques through a competitive process where firms that imitated the innovators captured much of the increased economic rents generated by innovators.

This was possible because patents, copyrights, and trademark law provided little protection for the intellectual property rights of the innovators. As a result, as innovators they could not effectively monopolize their innovation or prevent imitators for adopting their innovations. Weak, unenforceable, or even totally absent intellectual property rights were essential to the process of diffusion of new goods and ways of producing things.

In the countries where it became the main economic system, capitalism expanded the scope of both labor and credit markets, both of which (as you know from Chapters 11 and 12) are characterized by Pareto-inefficient Nash equilibria. The secret of the capitalist revolution was not that it avoided market failures and allowed a Pareto-efficient allocation of resources. Instead, capitalism promoted innovations and investments in capital goods that radically increased the productivity of labor from one year to the next.

To depict the connection between capitalism and history's hockey stick using the feasible sets of Chapter 6 you would not show a movement along a production possibility curve (feasible frontier) to some point at which the $mrs = mrt$ rule obtained. You would show a process of specialization, exploitation of economies of scale, and most of all the rapid shifting outwards of the feasible frontier.

CHECKPOINT 15.2 **Risk-taking and capitalism** What did Schumpeter mean by "creative destruction"?

HISTORY Petra Moser, an economic historian, studied the engineering and other innovations on display at Crystal Palace Exhibition known as the Great Exhibition or the Great Exhibition of the Works of Industry of All Nations in London in 1851. She found that countries with no or weak intellectual property rights were at least as innovative as those with stronger patent laws and other intellectual property rights, and concluded "that when patent rights have been too broad or too strong, they have actually discouraged innovation."[7]

EXAMPLE In this video (tinyurl.com/y6fkvx79), Petra Moser discusses copyright protection for nineteenth century Italian operas (from the CORE project. www.core-econ.org).

15.3 CAPITALIST FIRMS AS INNOVATORS: EMPLOYMENT AS INSURANCE

Capitalism promoted innovation by concentrating economic decision-making power in the hands of wealthy employers—capitalists—who had both the risk tolerance and the resources necessary to undertake the risky investments which contribute to the process of innovation.

The resources to undertake investments—building factories, developing new products—were available because the owners of firms were personally wealthy. And additionally, as you would expect from Chapter 12, as a result of their wealth, they could borrow substantial sums at moderate rates of interest. The investments that implement an innovation are risky. The

REMINDER From Chapter 13, we know that the slopes of the indifference curves—with risk on the horizontal axis and expected return on the vertical—are a measure of risk aversion: steeper means greater risk aversion.

Figure 15.4 Reminder: The indifference curves and risk choice of a risk-averse actor. The green risk-return schedule is the feasible frontier of the feasible set of risk and expected income. The indifference curves are upward-sloping because the individual is risk averse. The choice of Δ_a maximizes the individual's utility. Declining absolute risk aversion means that people with greater wealth are less risk averse.

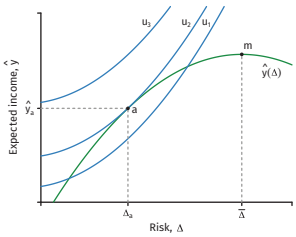
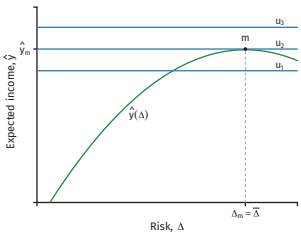


Figure 15.5 Reminder: the indifference curves and risk choice of a wealthy risk-neutral actor. The decision maker's indifference curves are horizontal because she is risk neutral. As a result she chooses Δ_m with the maximum expected income feasible.



reason is that time and money have to be committed to some novel project before knowing what the outcome will be. In Chapter 13 we modeled this as two possible outcomes—the income of the decision maker—that could occur as a result of a "good state" happening, or a "bad state" occurring (with equal probability).

The expected income is just the average of the two equally likely outcomes. (In the margin "Reminder" notes we recall the basic terms we introduced for dealing with risky decisions.) The level of risk is the difference in realized income between the good and the bad state.

In Figure 15.4 we show the risk-averse choice of a decision maker, and in Figure 15.5 the contrasting case of a risk-neutral decision maker, whom we term "wealthy" because of diminishing absolute risk aversion (the tendency or risk aversion to fall as wealth increases). Most people are more like the risk-averse person in the first figure than the risk-neutral person in the second.

The firm in a capitalist economy allows the collaboration among a large number of people, most of whom—the workers and lower level managers—are risk averse (due to lack of wealth) and some of whom—owners and top managers—are less risk averse or even close to risk neutral. The firm's total revenue fluctuates between good and bad states, and the rules of the game defining the capitalist firm mean that the resulting risk is experienced differently:

- workers and low-level managers are paid a wage or salary that is independent of the state that occurs, so as long as they retain their jobs they are exposed to modest levels of risk;
- owners and top managers are the residual claimants on the firm's revenues, so they experience the differences in firm's revenues between a good and a bad state.

There is also a division of labor: workers and lower-level managers produce the goods or services based on the decisions made by owners and top managers. Among these decisions are choices of how much risk to take, represented by selecting a level of Δ in the margin figures.

There are two important consequences of this setup. First, workers are willing to work for a firm engaged in risky investments because they are to some extent protected—by their fixed wages or salaries—from the differences in revenues realized in good and bad states. Without this protection wages would have to be substantially higher to motivate workers to provide effort to the production process. Owners are effectively providing insurance in return for paying lower wages than would be possible were workers to bear the same risks as the owners.

Second, the risky decisions are made by people who are not very risk averse, thus promoting the substantial risk-taking that is required for

innovation to flourish. This is how the capitalist firm in competition with other similar firms serves as an innovation machine.

This way of promoting risk-taking and innovation is based on two kinds of inequality:

- Wealth inequality between owners (and top managers) and workers (and low-level managers) is the basis for the limited risk aversion that guides the firm's decisions.
- Inequalities in decision-making powers, whereby it is the relatively risk-neutral top mangers and owners that make the risky decisions.

CHECKPOINT 15.3 Risk-taking and capitalism

a. Explain how Figures 15.4 and 15.5 illustrate declining absolute risk aversion. (If you are not sure, return to Chapter 13.)

b. In what ways does a language instructor or electrician working for a wage or salary at a large firm bear *less* risk than if they were in business *on their own*?

15.4 CAPITALISM AND INEQUALITY

Unlike the information on per-capita income (in Figure 15.2) consistent data on inequality of income going back a thousand years do not exist. But information on the fraction of total wealth held by the richest 1 percent (in Figure 15.6) give us a picture of inequality spanning most of the period in which capitalism has been the dominant economic system in the countries in question.

Prior to World War I in the countries shown the wealthiest 1 percent owned somewhere between one-third and two-thirds of all the wealth. Since then the top wealth shares have fallen to about one-fifth.

We also have data on the share of income of the top 1 percent since the beginning of the twentieth century. In all countries shown in Figure 15.7, the richest 1 percent received a smaller share of income around 1980 than in 1910, in some cases dramatically so. Most of the fall in the income share of the very rich took place in the four decades following the Great Depression in the 1930s, often called the golden age of capitalism.

As you can see comparing the two panels of Figure 15.7, in the non-European countries (panel (a)) there was a substantial increase in inequality after 1980, while in the European countries (panel (b)) this was less the case.

In Figures 15.7 and 15.6 we have used the income and wealth shares of the richest 1 percent because information on the rest of the distribution of income–the share of the very poor, for example, or of the middle third of the population–is not available until recent years. Using more recent data we can have a more complete picture of economic inequality.

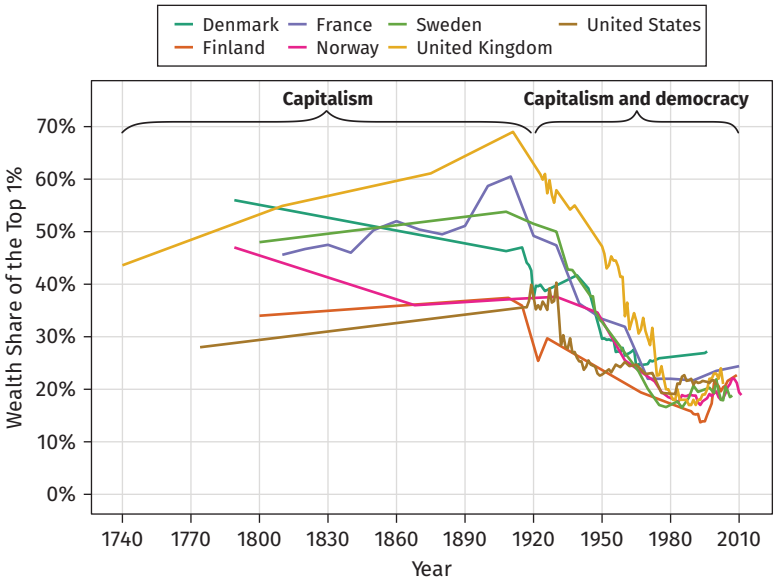
! REMINDER In section 13.9 we showed how a wealthy individual who is initially not exposed to risk could agree to take on some of the risk exposure of a less wealthy and highly risk-exposed person. The capitalist firm, among other things, represents a similar kind of exchange.

! REMINDER Wealth is the value of the land, buildings, machinery, intellectual property (e.g. patents), and other assets that provide income for their owner. Income is a flow—some amount per a given time period—while wealth is a stock, without a time dimension.

✓ FACT CHECK With a single exception (New Zealand) no country had universal suffrage (voting for all men and women) prior to the twentieth century. Many countries abolished property ownership as a condition for the right to vote around the time of World War I also extending the vote to women at around the same time or a bit later. For an account of the advance (and occasional retreat) of democracy, see: tinyurl.com/yxb7qsgx

Figure 15.6 The wealth shares of the top 1 percent of the wealth distribution. After a long period of rising wealth inequality in most of the countries shown, the half-century ending in 1970 witnessed significant declines in wealth inequality, measured as shown.

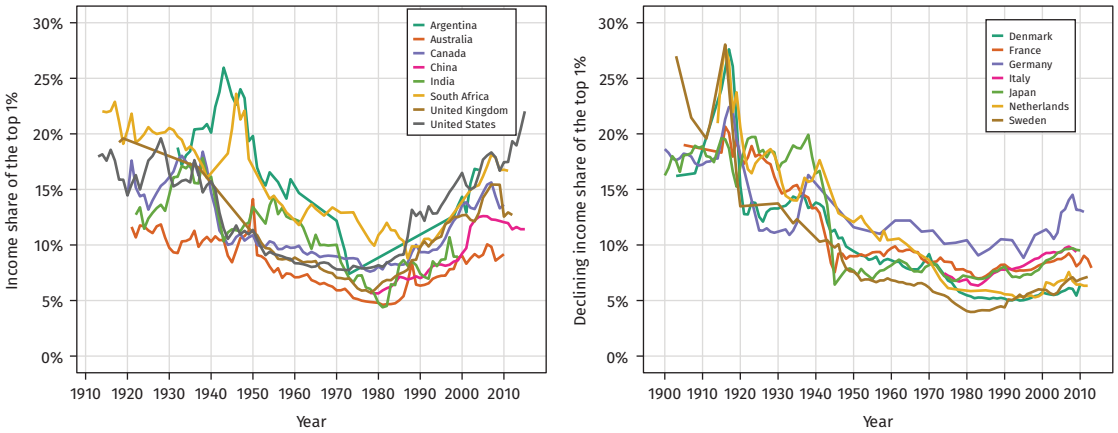
Source: Roine and Waldenström (2015).



HISTORY The use of the word "capitalism" has had its booms and busts over the course of history; and it is often, but far from always associated—in newspaper articles, for example—with the word "democracy." Cognitive scientist Simon DeDeo tracks what he calls "the marriage (and divorce) of capitalism and democracy" using computer science methods applied to centuries of newspaper text. See the video at cmu-lib.github.io/dhlg/project-videos/dedeo/

Figure 15.7 Income share of top 1 percent. Shown is the market income (prior to taxes and government transfers) of the richest 1 percent as a fraction of total market income. As was the case for wealth shown in Figure 15.6 inequality in market income fell in every country over most most of the twentieth century, and then after 1980 rose substantially in some countries (panel (a)) and modestly in others (the European nations shown in panel(b)).

Source: Alvaredo et al. (2017).



(a) Countries with marked increase in inequality post 1980

(b) Countries with little increase in inequality post 1980

But what does it mean to say, for example, that economically speaking Germany is more equal than the US or that since 1980 economic inequality has increased in India and China?

15.5 APPLICATION: MEASURING INEQUALITY—THE GINI COEFFICIENT AND THE LORENZ CURVE

Economic inequality is most commonly measured by differences in wealth or income either among members of a population, for example, a nation, or between distinct groups of members of a population, for example, between men and women, or people of different religions or ethnic identity.

Two widely used ways of measuring the extent of inequality are:

- the **Gini coefficient** that provides a single number measuring how much disparity there is among the members of the population; and
- the **Lorenz curve** that represents the entire distribution of income or wealth in the population.

The Gini coefficient

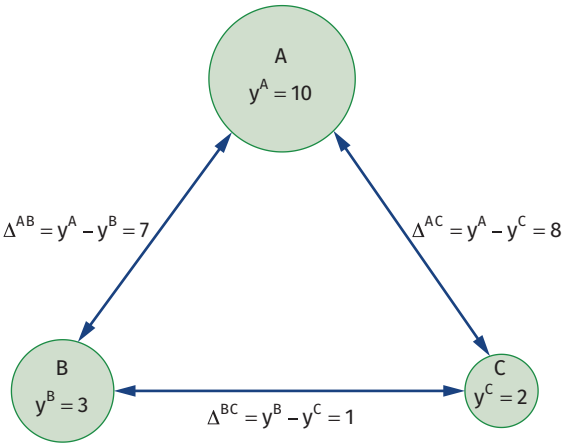
To understand the Gini coefficient think about a group of people (or families) and their income. Then ask about each of the possible pairwise comparisons among them: How different are the incomes of these two individuals, relative to the mean income. This way of seeing the Gini coefficient is shown for a three-person population in Figure 15.8. We represent the population as a network. The circles (called the "nodes" of the network) are individuals or families–Ali (A), Brown (B), and Cohen (C)–and the size of the circle is proportional to the amount of income they have. One of them, Mr. Ali, might be the employer whose income is profits (after paying taxes) made by hiring a worker from the Brown family whose income (also after taxes) is the wages that Mr. Ali pays her, and a member of the Cohen family is the unemployed person receiving some kind of government assistance (financed from the taxes the others pay).

! REMINDER Market income is income before the payment of taxes or the receipt of transfers from the government; it includes earnings (wages and salaries from employment) as well as income from self-employment and from the ownership of assets (interest, rents, or dividends).

LORENZ CURVE The Lorenz curve summarizes the distribution of income or some other measure across a population, mapping the cumulative (poorest to richest) population shares and corresponding cumulative income shares.

GINI COEFFICIENT This measure of inequality (using income as an illustration) is the average difference in income between every pair of individuals in a population relative to mean income, multiplied by one-half. The Gini coefficient is usually calculated as the area between the Lorenz curve and the perfect equality line, divided by the total area under the perfect equality line. See Figure 15.10.

Figure 15.8 Inequality measured as pairwise differences between people. There are three households: the Ali, Brown, and Cohen families. Each household is represented by a circle: the larger the circle, the more income (y) the household has ($y^A = 10$, $y^B = 3$ and $y^C = 2$ units of income). Each of the double-headed arrows indicates a unique pair of households: A and B, A and C, and C and B. The numbers on the arrows show the income difference between the indicated households (Δ^{ij}). In the figure the number of households is 3, the number of pairs is also 3, the total differences between pairs is 16 (8 + 7 + 1), therefore the average difference is $16/3 = 5.33$. The total income is 15 (10 + 3 + 2) and therefore mean income is $15/3 = 5$. The relative mean difference is the average difference divided by the mean income, and one half of that gives us the Gini coefficient: $(16/3)(3/15)(1/2) = 0.53$.



M-CHECK Imagine two people, A and B, dividing a pie. Their slices, expressed as a *fraction* of the pie, necessarily sum to 1: $s^A + s^B = 1$. Let's say that A gets a slice equal to or larger than B such that $s^A \geq s^B$. The *difference* between the two slices is $s^A - s^B$, which because $s^B = 1 - s^A$ we can rewrite as $s^A - s^B = s^A - (1 - s^A) = 2s^A - 1$. This is the *mean difference* (there is only this single difference); the average slice size has to be $\frac{1}{2}$ so Equation 15.1 rearranged becomes $s^A = \frac{G+1}{2}$ or $s^B = \frac{1-G}{2}$.

The important information about the network is on the arrows (called the "edges") between the circles: the labels on the arrows indicate the difference in income between the two individuals connected by the arrows. The Gini coefficient is the sum of the income differences among all of the pairs of households divided by the average income ($\bar{y}$) and then multiplied by one-half (see M-Note 15.1):

$$G = \frac{\text{sum of differences}}{\text{number of pairs}} \frac{1}{\bar{y}} \frac{1}{2}$$

$$= \text{"relative mean difference" x } \frac{1}{2} \tag{15.1}$$

The differences among the three individuals in the figure give you some idea of the degree of inequality represented by a Gini coefficient of 0.53. This is just a bit larger than the Gini coefficient for incomes before government taxes and transfers in the US just prior to the COVID-19 pandemic. From Equation 15.1 you can also see that the relative mean difference is

$2 \times G$; so a Gini coefficient of 0.53 (as in Figure 15.8) means that the average pairwise difference between people in the population is 1.06 times the mean income.

To get a better idea of what the Gini coefficient means imagine you are dividing a pie with just one other person. We show in the M-check that:

$$\text{The smaller slice, as a fraction of the pie:} \quad s^B = \frac{1-G}{2} \tag{15.2}$$

You can check that when one person gets the entire pie, $G = 1$ and when it is evenly spit, $G = 0$.

EXAMPLE In this video (tinyurl.com/y5kogheo), Thomas Piketty, Professor of Economics at the Paris School of Economics, explains how he 'tries to be useful' by collecting long-run data on the distribution of wealth (from the CORE project. www.core-econ.org).

M-NOTE 15.1 Inequality as differences between people

Figure 15.8 represents a population of just three people. But the Gini coefficient, like the Lorenz curve, is used to measure inequality in populations of millions. To see how this is done, we take the following steps.

- Let Δ^{ij} be the absolute difference in income between family i and family j, meaning the income of the richer family minus the income of the poorer family (called a pairwise difference).
- We then define the sum of the differences between all pairs as

$$\Delta = \sum_{i \neq j} \Delta^{ij}$$

  where the summation is over all of the unique (i, j) pairs in the population excluding the cases where i is equal to j (this would be the "difference between a family's income and its own income!).
- If there are n members of the population then the total number of pairs is $\frac{n(n-1)}{2}$, shown as the three edges among the n = 3 families in the figure.
- Then the average difference is Δ divided by the number of pairs or $\frac{\Delta}{(n^2-n)/2}$.

If we then let $\bar{y}$ be the average income, then we have the following measure of the Gini coefficient:

$$\text{Gini coefficient} \equiv G = \left(\frac{\Delta}{(n^2-n)/2}\right)\left(\frac{1}{\bar{y}}\right)\frac{1}{2} \tag{15.3}$$

This means the Gini coefficient is the mean difference among all pairs (the first term: total differences divided by total number of pairs) relative to (divided by) the mean value of y (the "relative mean difference") times one-half.

CHECKPOINT 15.4 Understanding the Gini coefficient

A pie will be divided between two people (so the average size of a slice is $\frac{1}{2}$). Use Equation 15.1 to confirm that if the Gini coefficient of pie slices between the two is 0.53 then the smaller of the two slices will be 0.235 of the whole pie.

Figure 15.9 The smaller of two slices of the pie: What does the Gini coefficient tell us? The M-Check explains the calculation of the slices. In these images B obtains what is "left over" of the pie (s^B is the remaining amount of the pie). A had already taken $s^A = 1 - s^B$. In empirical estimates, even in relatively small groups, we almost never observe the $G = 1$ and $G = 0$ slices. But in 2019 the Gini coefficient for wealth in Germany was about 0.78 and in the United States was about 0.85. Among the high-income countries Gini coefficients for years of schooling are less than 0.21.[8]

Source: Design by www.twistopen.in.

The Lorenz curve

Unlike the Gini coefficient, which summarizes inequality among a group of people with a single number, the Lorenz curve gives us a picture of the disparity of income across the whole population. The **Lorenz curve** shows the entire population lined up along the horizontal axis from the poorest to the richest. At any point on the horizontal axis, the height of the curve indicates the fraction of total income received by the fraction of the population given by that point on the horizontal axis.

Two Lorenz curves (in green) for the Netherlands are shown in Figure 15.10. In the figure the diagonal blue 45-degree line is how the Lorenz curve would look hypothetically if everyone had the same income: for example, it shows that 10 percent of the population receive 10 percent of the total income, 50 percent of the population receive 50 percent of the income, and so on. This is called the *perfect equality line*.

For any Lorenz curve we can calculate the Gini coefficient as the area between the Lorenz curve and the perfect equality line divided by the area under the perfect equality curve or from the figure:

$$\text{Gini} = \frac{A}{A + B} \tag{15.4}$$

The Lorenz curve shows how far a real distribution of income departs from this line of perfect equality. In Figure 15.10 on the left we have the distribution of market income (that is, before payment of taxes or receipt of government transfers).

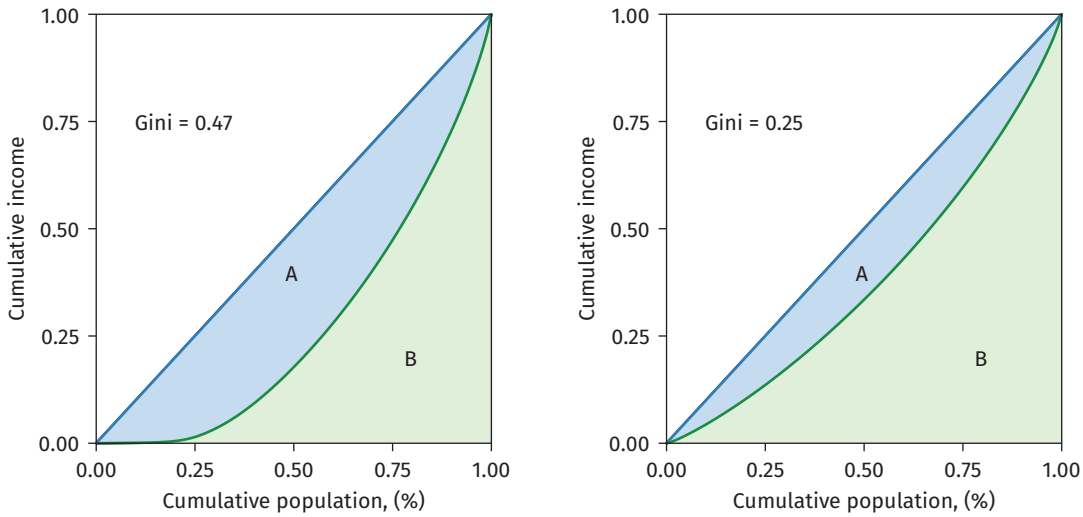
The market income Lorenz curve indicates that the *poorest* 25 percent of the population (0.25 on the horizontal axis) receives about 1.5 percent of total income (0.015 on the vertical axis), and the lower-earning half of the population (0.50) has less than 20 percent of income.

In Figure 15.10 (b) we show the Lorenz curve for *disposable income*, which better captures living standards. Disposable income is the maximum a household can spend ("dispose of") without borrowing or selling something they own (like their car or house), after paying tax and receiving transfers (such as unemployment insurance and pensions) from the government. Notice from the figure that in the Netherlands, almost one-fifth of the households have a near-zero market income, but most nonetheless have a substantial level of disposable income. The Lorenz curve for disposable income is much closer to the perfect equality line than is the Lorenz curve for market income, meaning that the system of taxes and transfers

LORENZ CURVE The Lorenz curve summarizes the distribution of income or some other measure across a population, mapping the cumulative (poorest to richest) population shares and corresponding cumulative income shares.

Figure 15.10 Distribution of market and disposable income in the Netherlands (2010). The Gini coefficient of disposable income is less than that of market income, which means taxes and transfers help reduce income inequality in the Netherlands.

Source: LIS. Cross National Data Center. Calculations were made for the CORE Project by Stefan Thewissen (University of Oxford) in April 2015. Household market (labour and capital) income and disposable income are made comparable through PPP comparisons and are top- and bottom-coded.



(a) Lorenz curve and Gini coefficient of market income

(b) Lorenz curve and Gini coefficient of disposable income

in the Netherlands reduces income inequality. To see how progressive government policy is in the Netherlands, the figure shows that the bottom 20 percent receive only one-half of one percent of market income but over 10 percent of disposable income.

The Netherlands is not exceptional in its progressive policies. Figure 15.11 shows the Gini coefficient of market income and of disposable income for a large sample of countries. Countries differ substantially in the extent to which taxes and government transfers reduce inequality in disposable income. Compare South Korea and Taiwan, on the one hand, where market incomes are the most equally distributed in the sample of countries but taxes and transfers are close to distributionally neutral (the Gini coefficient for disposable income is close to that for market income), with, on the other hand, Sweden and Germany where market incomes are much more unequally distributed but the effect of government is highly progressive. Market income is as unequally distributed in Germany as it is in the US, but inequality of disposable income is one-third greater in the US than in Germany.

M-CHECK The representation of the Gini coefficient using the Lorenz curve and Equation 15.4 does not work for very small populations while the network representation in Figure 15.8 using Equation 15.3 in M-Note 15.1 can be used on populations of any size.

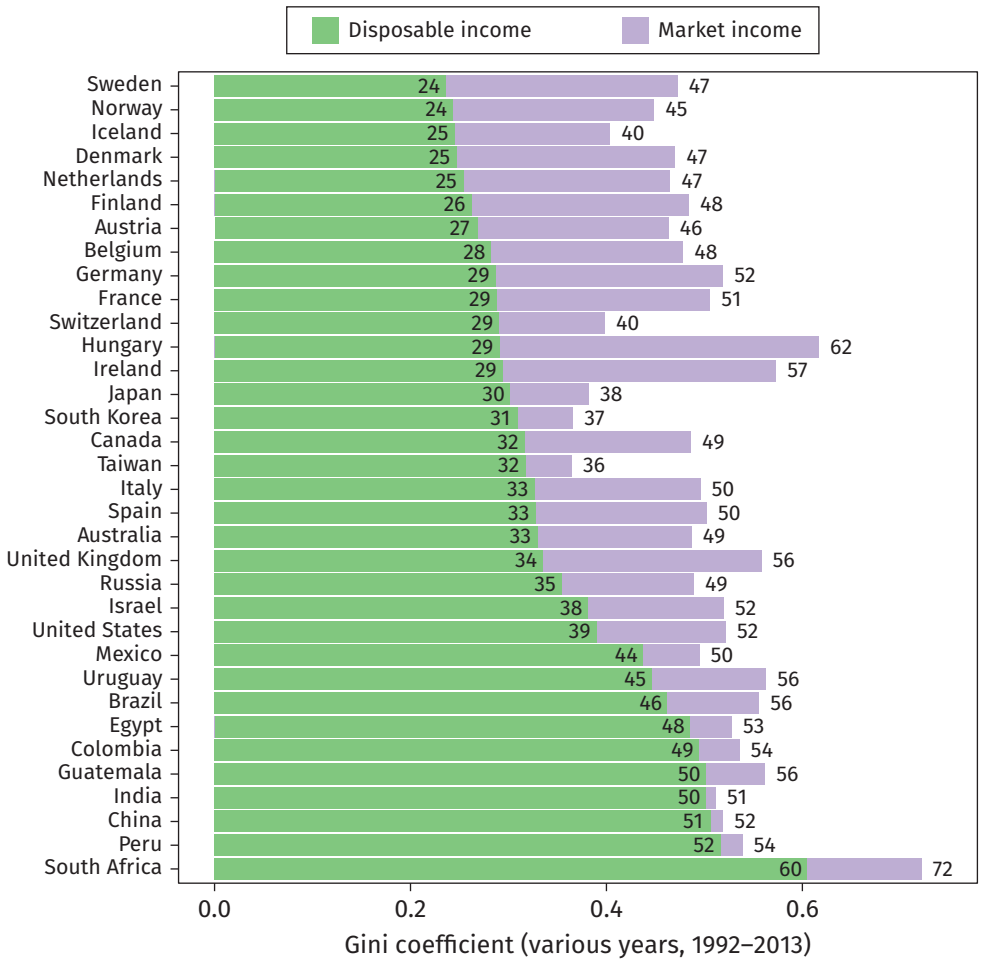
✓ **FACT CHECK "Lost Einsteins"** A study of inventors in the US and how the environment in which one grows up, as well as one's race, family income, and gender affect the likelihood of becoming a major inventor concluded that "there are many 'lost Einsteins'—people who would have had highly impactful inventions had they been exposed to innovation in childhood—especially among women, minorities, and children from low-income families."[9]

15.6 INNOVATION AND EQUALITY

In our discussion of the capitalist firm as a source of innovations we explained that wealth inequality can promote the process of innovation by concentrating decision-making powers in the hands of the owners and managers of firms. These people are sufficiently wealthy that they are *not* very risk averse, and are therefore willing to take a chance on introducing a new product, technology, or form of organization.

Figure 15.11 Gini coefficients for disposable and market incomes. The level of income inequality experienced by people is given by the length of the green bars: that is, the Gini coefficient for disposable income. The right-hand end of the purple bar is the Gini coefficient of incomes before taxes and transfers. The numbers at the end of the bars are the Gini coefficient that the bar represents, times 100.

Source: CORE, The Economy.



On this basis, you might expect that economies like Germany, where workers have a substantial consultative voice in the management of large firms, should lag behind other countries in terms of technology and innovation. And Sweden, with one of the least unequal distributions of disposable income should also rank poorly as an innovator. The most innovative among the major high-income countries would be the US, the UK, Russia, and other countries among the most *unequal* in disposable income, as you can see from Figure 15.11.

But Figure 15.12 shows that that is not the case. The top innovators according to the Bloomberg Innovation Index do *not* include the US, UK, and Russia. With the exception of Israel, all the most innovation countries—including Germany, Sweden, and South Korea—have more equal distributions of disposable income.

There is other evidence that countries with relatively equal distributions of disposable income are also highly innovative:

- Another commonly used measure of innovation is the number of "triadic" patents (those filed for the same innovation in the US, the European Community, and Japan) per head of population. On this measure Sweden and Finland outrank the US with Denmark not far behind.
- All three of these Nordic countries with modest inequality in disposable income outrank the US on research and development (R&D) by businesses as a percentage of GDP, researchers per 1,000 workers, and venture capital as a percentage of GDP.[10]
- Innovation is essential to the growth in labor productivity (output per hour) which in turn contributes to the rate of growth of total output (GDP) per capita. By this measure the more equal countries have a slight advantage, as can be seen in Figure 15.14 (a). The same is true if we focus on middle-income economies (panel (b)) where labor productivity may grow rapidly through a process of technological borrowing from world leaders.

These data show that there are a good number of countries with modest levels of inequality in disposable income that are highly innovative. Partly as a result, these economies have experienced rapid growth in living standards made possible by impressive improvements in labor productivity.

Commonly mentioned explanations of this include:

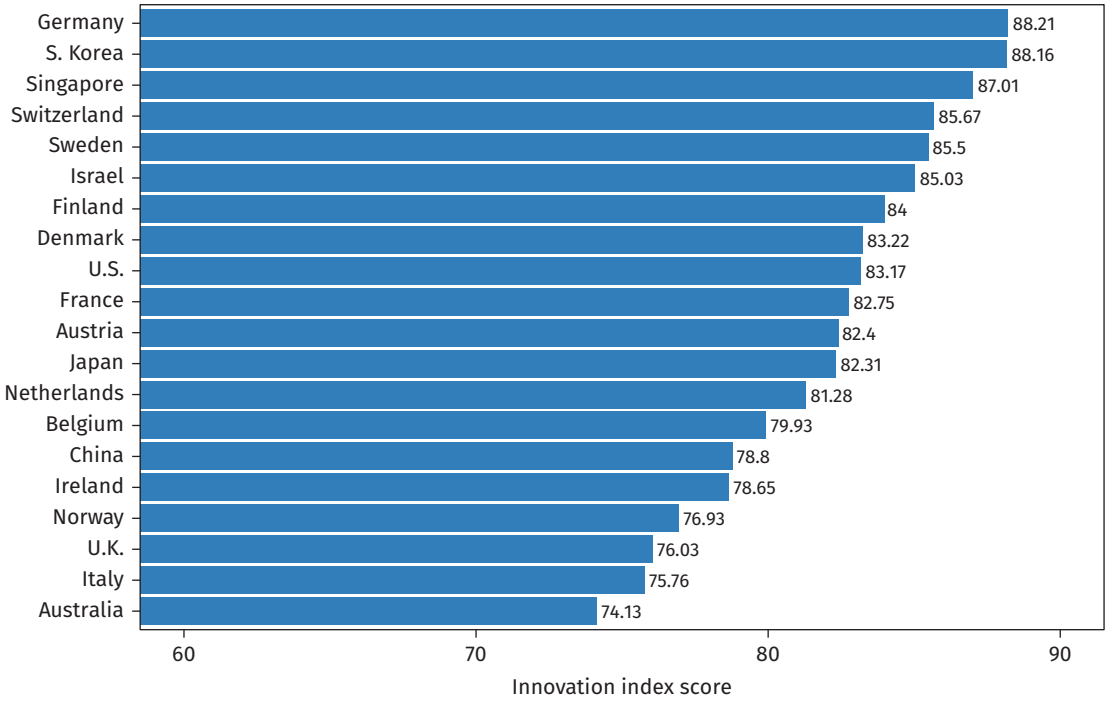
- high-quality education for virtually all citizens (as in Finland and South Korea);
- substantial governmental support for basic research and communications infrastructure; and
- high-wage policies that force low-productivity businesses to close, along with retraining and reemployment for displaced workers, as in Sweden.

HISTORY Most people see policies of income taxation to finance transfers to the general population as a redistribution of income between the well-off and the less well-off. But in 1944, when the welfare state was in its infancy, Richard Musgrave and Evsy Domar adopted a different way of looking at the process, similar to the way we modeled taxes and transfers in Chapter 13. They showed that by redistributing income from the lucky to the unlucky, progressive taxation and transfer policies would reduce risk exposure and therefore could promote greater risk-taking.[11]

› EXAMPLE In this video (tinyurl.com/y4jsx4kd) watch Michigan State University economist Lisa Cook explain how prior to World War I, a rich stream of contributions by African-American inventors was brought to an end by a wave of anti-black violence (from the CORE project. www.core-econ.org).

Figure 15.12 Twenty "most innovative countries": Bloomberg Innovation Index 2020. The index is based on seven measures of innovation including R&D spending, researchers as a fraction of population, science and engineering graduates as a percentage of total graduates, density of high-tech firms (e.g. biotech, aero space, software), and patenting activity. The US was number 1 when the list was first compiled, in 2013. Bloomberg is a New York-based media, software, and data company. The top 20 are not as different one from another as they appear in the figure because the horizontal axis is truncated at 60 rather than extending to 0. In the sample of 60 countries, the lowest score is Macao with a score of 46.

Source: Bloomberg. https://www.bloomberg.com/news/articles/2020-01-18/germany-breaks-korea-s-six-year-streak-as-most-innovative-nation.



Returning to our model of decision-making about the level of risk (Chapter 13) the innovativeness and economic growth of less economically unequal countries like Germany, Sweden, and Finland should not have surprised us. Risk-taking, even by quite risk-averse people, can be promoted by providing opportunities for insurance. For example you saw in Chapter 13 that the elimination of tuition for higher education combined with a progressive "graduates income tax" would reduce the risk of investing in one's own higher education (one only pays the tax after graduating). Insurance, we also showed there could also take the form

of a *linear tax* and *lump-sum transfer*. Figure 15.13 reminds you of the basic logic of why the insurance provided by these public policies supports higher levels of risk-taking. Without insurance the person adopted a limited level of risk and had, as a result, a modest level of expected income (point **a**). With insurance on the terms given by the orange insurance line, the person took a greater level of risk with a higher level of expected income (point **c**) but as a result of the insurance *experienced* both a lower level of risk exposure and a higher level of expected income (even after paying the insurance premium).

CHECKPOINT 15.5 Equality and innovation What do you think are some reasons why the most innovative countries in the world (according to Bloomberg) include many that are more equal than most in their distribution of disposable income?

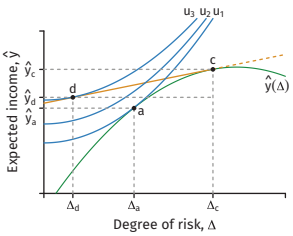
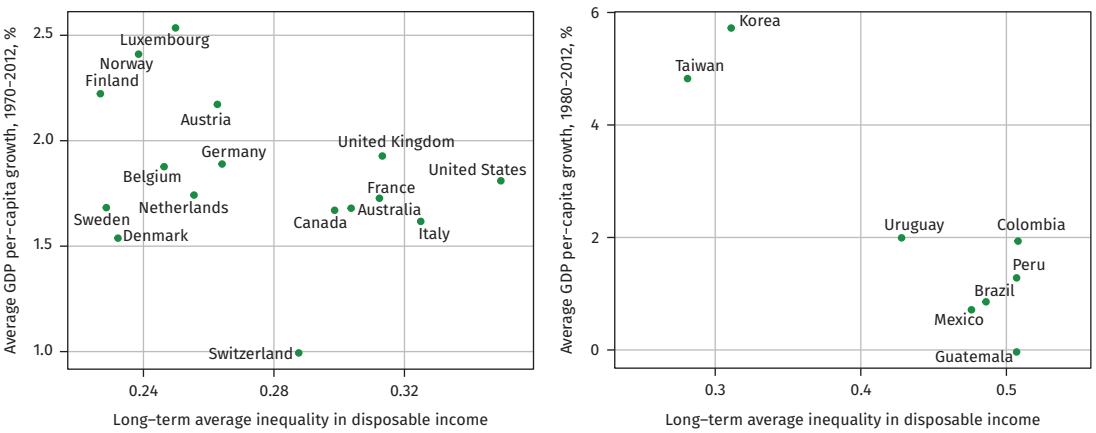


Figure 15.13 Reminder: how insurance reduces risk exposure and promotes risk-taking (from Figure 13.14). Chapter 13 also showed how a linear tax and redistribution (tax revenues divided equally among citizens) provides a similar kind of insurance.

Figure 15.14 Inequality and GDP per-capita growth for high- and middle-income countries. Notice that in order to show the very rapid growth in per-capita income in S. Korea and Taiwan, the vertical scale in panel (b) is more than twice that in panel (a). A better measure would be labor productivity—output per hour—but we lack comparable hours of work data for all of the countries. For countries that experienced a substantial fall in work hours—Netherlands and Germany for example, see Figure 7.10—the figure significantly understates the growth of labor productivity. The high growth rates achieved by Taiwan and South Korea were in part made possible because they could borrow new technologies originating in the richer nations. The Latin American countries shown did not manage to benefit from these "catch up" opportunities.



(a) Growth and inequality among higher income countries

(b) Growth and inequality among middle and low income countries

! REMINDER The labor discipline model (the Ford model in Chapter 11) explains how wages, work effort, and the number of workers hired are determined in the interaction of a profit-maximizing employer and a team of utility-maximizing workers. The model of limited competition among firms (in Chapter 9) explains how the degree of competition among firms determines the price markup over costs and the level of economic profit in the economy. The whole economy model in Chapter 11 puts these two components together in a single model.

15.7 THE MICROECONOMICS OF INEQUALITY AND THE MACROECONOMY

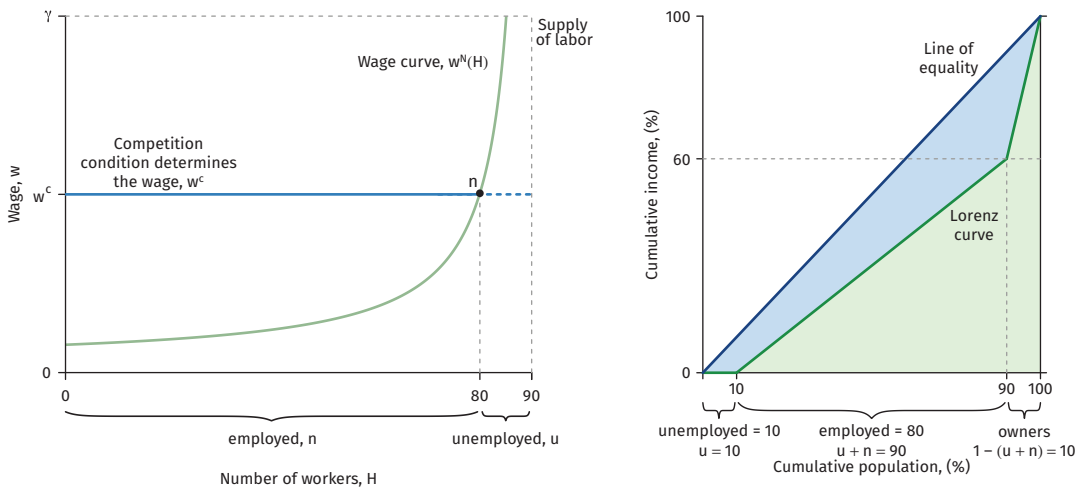
We can provide a framework for understanding the differing degree of inequality in the world's economies by using the model of the whole economy presented in section 11.11, shown in the left panel of Figure 15.15. To do this we translate the main variables of that model—wages as a share of total output, the level of employment, unemployment, and the number of employers—to a Lorenz curve, shown in the right panel of Figure 15.15.

Here we look at inequality in market incomes, that is before taxes and transfers, and we ask: What attributes of an economy make market incomes more or less unequal?

The whole economy model and the Lorenz curve

As an illustration, think about an economy in which there are no self-employed people and nobody works for the government. Also there are no

Figure 15.15 The whole economy model and the Lorenz curve. In panel (a) we show the wage curve and the wage determined by the competition condition for an economy in which there are 90 workers of whom 80 are employed, receiving 60 percent of total income (that is, the wage share), and 10 are unemployed, receiving no income. There are also 10 employers who are owners of 10 firms each of which employs 8 of the workers ($8 \times 10 = 80$). Together, these 100 people form an economy and have access to income from work that can be depicted using a Lorenz curve (panel (b)) with its corresponding Gini coefficient. The Gini coefficient for the model economy shown is 0.36.



(a) The whole economy model: wage curve and competition condition Curve

(b) The Lorenz curve associated with the whole economy model

taxes or government expenditures, so the only income is either wages or profits. As a result, everyone in the economy is included in the three groups of people in the labor discipline model and we assume that they have the following incomes:

- *the unemployed*: they receive nothing (remember, this is market income, that is before government transfers);
- *the employed workers*: workers who receive some share of the value of the goods they produce, σ_w, called the **wage share**; and
- *employers*: who receive the complementary share of the value of goods produced, that is one minus the wage share or $1 - \sigma_w$ called the **accounting profit share**.

The wage share, σ_w, is:

$$\sigma_w = \frac{\text{hourly wage}}{\text{value of output produced by a worker in an hour}} = \frac{w}{\gamma} \tag{15.5}$$

The Lorenz curve for income in this economy is depicted in Figure 15.15. The Lorenz curve for this economy is made up of three line segments with the beginning point (at the lower left of the figure) having coordinates of (0, 0) and the endpoint (at the upper right) having the coordinates (100 percent of the people, 100 percent of the output). The first line segment is a portion of the horizontal axis because the poorest segment of the population (the unemployed) have no income at all.

The first kink in the curve occurs when we have counted all the unemployed people, so everyone else has some income. The second kink in the Lorenz curve is the interior point, whose coordinates are (percent of total number of the population (90 percent), the wage share (60 percent)). The curve between the interior point and the upper-right corner is steeper than the other two segments of the curve because the employers receive more income than workers, so adding a given population fraction along

WAGE SHARE The wage share is the fraction of total income that is received in the form of labor earnings (wages plus salaries).

ACCOUNTING PROFIT SHARE The accounting profit share is the fraction of total income that is received as accounting profits by the owners of the capital goods used in production; it is often decomposed into the capital share and the economic profit share.

ECONOMIC PROFIT SHARE The economic profit share is the fraction of total income received by the owners of the capital goods used in production in excess of the opportunity cost of capital, or: economic profit share = accounting profit share minus capital share.

the horizontal axis accounts for a larger increase in the share of income accounted for.

M-Note 15.3 provides us with an equation that translates data on the share of population unemployed, working, and employers and their shares of total income into the resulting Gini coefficient.

The "kinked" curve in the figure is a simplification designed to show how the Cournot model of competition among firms and the labor market with its three groups of actors allow us to better understand what influences the level of inequality. Lorenz curves based on actual data are smooth like the one shown earlier in this chapter for the Netherlands. This is because we have recognized just a single kind of heterogeneity among the population, the three types of individuals from our principal-agent model of the labor market: owners, employed workers, and unemployed workers. We assumed that within each of these three groups, there are no differences. But we need to keep the following in mind:

- Workers are *not* identical, they are *diverse* or "*heterogeneous*." They have more or less valuable skills, more or less marketable education or credentials, and they differ in race, gender, where they live, and many other ways that affect their pay.
- Employers too are heterogeneous; they differ in the extent of their wealth and in whether assets they own produce goods in growing or declining demand, whether they are an innovation leader or follower, their management skills, and other things affecting the rate of profit they earn.
- Even if workers and employers were identical, they would end up with differing wages and rates of return simply by the luck of the market (see section 14.11).

In M-Note 15.3 we derive a relationship between inequality and the population and income shares determined by the labor market and the goods market:

$$G = u + n - (1-u)\sigma_w - (1+n)\sigma_B \tag{15.6}$$

Equation 15.6 says that the Gini coefficient (G) depends on the proportion of the population that is unemployed (u), the proportion of the population that is employed (n), and the income shares of the unemployed(σ_B) and the employed (that is, the wage share, σ_w).

Equation 15.6 therefore allows us to study how changes in the structure of the economy will affect the degree of inequality as measured by the Gini coefficient. Using this equation, for example, we can see how changing some dimension of the economy while holding other dimensions constant will alter the degree of inequality. We do this in M-Note 15.4, and find:

- An *increase* in the wage share, (σ_w), and the associated *decline* in the profit share will lead to a **decrease** in the degree of inequality as measured by the Gini coefficient.
- An *increase* in the percent unemployed, (u), along with a decline in employment, will lead to an **increase** in the Gini coefficient.
- An *increase* in the percent employed, n, along with a reduction in the number of employers, will lead to an **increase** in the Gini coefficient.
- An *increase* in the percent of income going to the unemployed (σ_B) will lead to a **decrease** in the degree of inequality as measured by the Gini coefficient.

M-NOTE 15.2 The wage share and the accounting profit share

For simplicity, we assume that unemployment benefits, B, are zero as in Figures 15.15 and 15.19. Our variables are $\{\rho, \gamma, b\}$, where ρ is the opportunity cost of capital, γ is output per hour (which equals the inverse of the labor time required to produce one unit), and b is barriers to entry. Recall from Equation 11.21 and M-Note 11.8 that the wage consistent with the competition condition is:

$$\text{Competition condition} \qquad w^c = \frac{(1-b)\gamma}{(1+\rho)}$$

From Equation 15.5 we know that the wage share, σ_w, consistent with the competition condition is therefore:

$$\text{Wage share} \qquad \sigma_w = \frac{w^c}{\gamma} = \frac{1-b}{1+\rho}$$

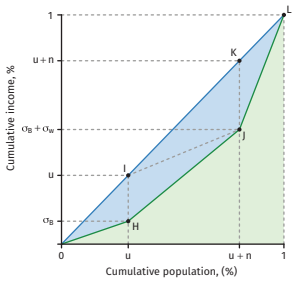
The corresponding accounting profit share is:

$$\begin{aligned}\text{Accounting profit share} \qquad \sigma_A &= 1-\sigma_w \\ &= 1-\frac{1-b}{1+\rho} \\ &= \frac{1+\rho}{1+\rho}-\frac{1-b}{1+\rho} \\ &= \frac{\rho+b}{1+\rho}\end{aligned}$$

Table 15.1 Data for the Lorenz curve in Figure 15.16. Share means "fraction of." There are only the three groups shown in the population as recipients of income, so the employers' share is 1 minus the shares of the other two groups.

	Unemployed	Employed	Employers
Population share	u	n	$1-u-n$
Income share	σ_B	σ_w	$1-\sigma_B-\sigma_w$

Figure 15.16 Lorenz curve given the data in Table 15.1 and used in M-Note 15.4. Notice that here (unlike in Figure 15.15) the unemployed receive some share of total income.



M-NOTE 15.3 An equation for the Gini coefficient as a function of population and income shares

Given the data in Table 15.1, we can draw the Lorenz curve as in Figure 15.16. To calculate the Gini coefficient from these data, notice that the blue-shaded area of the polygon 0HJL (where 0 corresponds to the origin, 0) between the (green) Lorenz curve and the (blue) perfect equality line is made up of four triangles.

We derive an expression for the Gini coefficient by summing the area of these four triangles and then dividing by the area of the large triangle under the perfect equality line.

The vertical distance between H and I and between J and K, respectively are:

$$|HI| = u - \sigma_B$$
$$|JK| = u + n - (\sigma_B + \sigma_w)$$

You can see that the area of the triangle on the lower left

$$S(\triangle_{0HI}) = \frac{1}{2}(u^2 - u\sigma_B) = \frac{1}{2}|HI|u$$

By analogous reasoning about the other triangles, the area of the polygon 0HJL is:

$$\begin{aligned} S &= S(\triangle_{0HI}) + S(\triangle_{HIJ}) + S(\triangle_{IJK}) + S(\triangle_{JKL}) \\ &= \frac{1}{2}|HI|u + \frac{1}{2}|HI|n + \frac{1}{2}|JK|n + \frac{1}{2}|JK|(1-u-n) \\ &= \frac{1}{2}(|HI|(u+n) + |JK|(1-u)) \end{aligned}$$

Therefore, dividing this area by the total area under the perfect equality line (which is $\frac{1}{2}$) the Gini coefficient is:

$$\begin{aligned} G = 2S &= |HI|(u+n) + |JK|(1-u) \\ &= (u - \sigma_B)(u+n) + (u+n) - (\sigma_B + \sigma_w)(1-u) \\ &= u(u+n) - \sigma_B(u+n) + (u+n)(1-u) - (\sigma_B + \sigma_w)(1-u) \\ &= u + n - (1-u)\sigma_w - (1+n)\sigma_B \end{aligned}$$

M-NOTE 15.4 The Gini coefficient and its determinants

Using Equation 15.6, the Gini coefficient is:

$$G = u + n - (1-u)\sigma_w - (1+n)\sigma_B$$

We will examine the effect on G of increases in each of four different variables, holding constant the values of the other variables:

- u: percent of population unemployed, fewer employers
- n: percent of population employed, fewer employers
- σ_B: share of income to unemployed
- σ_w: share of income to employed

continued

We examine the effect of each variable on G by looking at their respective partial derivatives:

$$u: \quad \frac{\partial G}{\partial u} = 1 + \sigma_w > 0$$

$$n: \quad \frac{\partial G}{\partial n} = 1 - \sigma_B > 0$$

$$\sigma_B: \quad \frac{\partial G}{\partial \sigma_B} = -(1+n) < 0$$

$$\sigma_w: \quad \frac{\partial G}{\partial \sigma_w} = -(1-u) < 0$$

Hence, increases in u and n increase inequality, while increases in σ_B and σ_w reduce it.

CHECKPOINT 15.6 **Determinants of inequality** Using either Equation 15.6 and M-Note 15.4, or by redrawing Figure 15.15, show that the bulleted statements above (about "changing some dimension of the economy") are true.

15.8 MARKET POWER AND THE DISTRIBUTION OF INCOME

An important trend in many of the high-income economies of the world in the four decades following 1980 has been the increase in market income inequality. This might have occurred, as you saw in the bullets immediately above, because of an increase in unemployment. But in a number of countries with rising inequality—the US and UK, for example—unemployment did not rise until the COVID-19 pandemic of 2020-2021: in both countries unemployment was substantially lower in 2019 than in 1980.[13] What could account for these developments?

Evidence on declining competition and rising inequality

Here are some clues.

- Figure 15.17 shows that in the US, along with the increase in market inequality (measured by the Gini coefficient), the price markup over costs tripled over the 35 years since 1980, consistent with a reduction in the degree of competition in product markets.
- In Figure 15.18 we show, also for the US, economic profits as a share of total income, increasing threefold from the 1980s until 35 years later.

We see in M-Note 15.5 that in our model of competition the share of economic profits in total income (σ_E) is directly related to our measure of barriers to competition (b) which is the probability that a firm attempting to enter an industry will fail. We show that $\sigma_E = b$ so the evidence in Figure 15.18 suggests that barriers to entry (that is, b) could have tripled over this period, a very substantial reduction in the degree of competition in US markets.

✓ **FACT CHECK** The whole-economy model captures some long-run dynamics of a capitalist economy. But adjustment to a new equilibrium may take decades as the following example shows. The rapid increase in imports of manufactured goods from China in the 1990s destroyed many local economies in the US, for example those producing furniture and other products competing with Chinese imports. A study of these local labor markets concluded: "trade-induced manufacturing declines . . . are not, over the course of a decade, largely offset by sectoral reallocation or labor mobility . . . employment-to-population rates fall at least one-for-one with the decline in manufacturing employment, and generally by slightly more."[12]

Figure 15.17 The markup ratio and the Gini coefficient in the United States. Gini coefficient data from 1979 to 2015 are for market income (income before taxes and government transfers) and for 1967 to 1978 are for money income (income after government cash transfers but before taxes) adjusted upwards so as to be comparable to market income inequality. The mark-up ratio, that is $(p-c)/c$, is for all firms in the US weighted by firm size. The fact that the mark-up ratio and the Gini coefficient tend to move together over time is consistent with the theory of inequality conveyed by our whole economy model, but we cannot conclude that changes in the mark-up ratio are causing the changes in the Gini coefficient without taking account of other possible influences on the two series.

Source: De Loecker et al. (2020), US Census Bureau (2020), and Congressional Budget Office (2020).

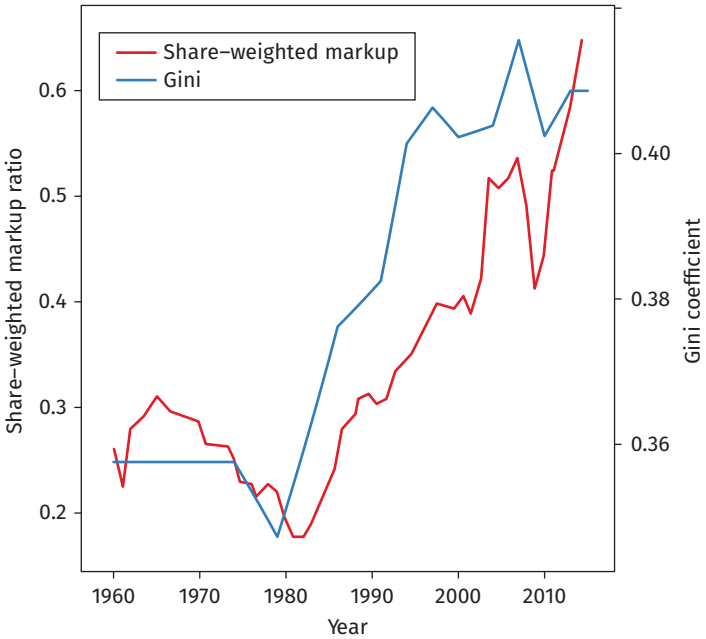
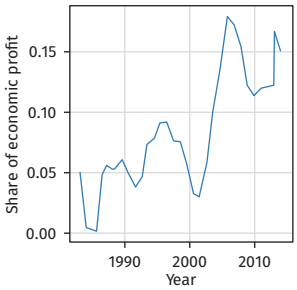


Figure 15.18 Reminder: the share of economic profits in total income, US over the period 1984–2014 (from Chapter 9). The other shares are wages and the difference between economic profits and accounting profits, namely the profits corresponding to the opportunity cost of the capital goods used in production.

Source: Barkai (2020).

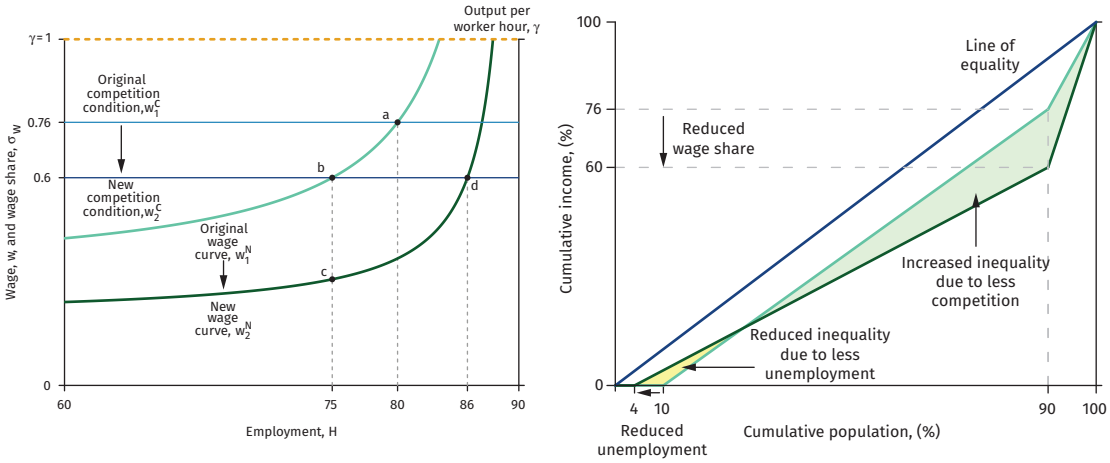


Putting the model to work

Motivated by these data and results, we can use the whole economy model and the Lorenz curve in Figure 15.15 to suggest a possible explanation based on two processes occurring over these years. Figure 15.19 presents our hypotheses about the causes of the increase in inequality.

- *Less competition:* A decline in competition–increase barriers to entry–raised the equilibrium markup of prices over costs and shifted the competition condition downward, reducing the wage share.
- *Less bargaining power for workers*: Fewer impediments to workers being fired and a reduction in government transfers to those out of work resulted in a downward shift of the wage curve.

Figure 15.19 Effect of shifts in the wage curve and the competition condition on the Lorenz curve. Panel (a), presenting the whole economy model, shows the effect of changes in the degree of competition and in the bargaining power of employers and workers on the number of people employed (remember there are 10 employers and 90 workers, some of whom will be unemployed). Panel (b) shows the Lorenz curve: its horizontal axis is the fraction of the total population (not numbers of workers). Note that because the Lorenz curve represents market (not disposable) income, the unemployed receive no income. The initial state of the economy is point **a** in panel (a), with the corresponding green Lorenz curve in panel (b). The combined effect of the downward shift in the competition condition and in the wage curve is to move the economy to point **d** in panel (a), with just 4 workers unemployed (or 4 percent of the population in the panel (b)). The new dark green Lorenz curve associated with point **d** shows the effect on inequality. The yellow triangle is the effect of reduced unemployment, lowering inequality. The green-shaded area is the increase in inequality due to the reduction in competition (resulting in a reduced wage share of total output and an increased accounting profit share). To make this figure we have set worker productivity, $\gamma = 1$, so the wage share $\frac{w^c}{\gamma} = w^c$. Also, because $\gamma = 1$ output is 80 before the change at point **a** and output is 86 at point **d**.



(a) Shifts in the competition condition and the wage curve

(b) Changes in the Lorenz curve

We offer this explanation as an illustration of how the model works and can be applied to real developments in the economy. While it is consistent with the available data (e.g. the decrease in measures of competition) there are many competing explanations which we do not exclude.

If occurring in isolation, the shift in the competition condition would result in a movement from the original equilibrium at point **a** to a new equilibrium at point **b** with a lower wage and lower level of employment. But if the wage curve also shifted down, then point **b** would no longer be an equilibrium. To see why, consider point **c**, which indicates that the wage

! REMINDER The competition condition determines the level of the real wage that is consistent with the degree of competition in the product market, given the productivity of labor and the opportunity cost of capital.

Table 15.2 Characteristics of the start and end points in Figure 15.19. Recall that w^c is the wage determined by the competition condition, and u and n are, respectively, the fractions of the population that is unemployed and employed. The equation for the Gini coefficient is Equation 15.6 in the text, with σ_B, the share of income going to the unemployed, set equal to zero (as we have assumed).

Point	w^c	Total income	Wage share	u	n	Gini coefficient
Point ***a***	0.76	80	0.76 = 61/80	0.10	0.80	0.10 + 0.80 − (1 − 0.10)(0.76) = 0.216
Point ***d***	0.6	86	0.6 = 52/86	0.04	0.86	0.04 + 0.86 − (1 − 0.04)(0.6) = 0.32

required to motivate workers to work is lower than the wage determined by the competition condition. Firms would therefore lower their wage, resulting in the rate of profit now exceeding that which is low enough to deter firm entry in the product market. So firms would enter, pushing up employment. This would continue until the economy reached the new Nash equilibrium at point **d**.

Thus a *simultaneous* shift in the wage curve and the competition condition results in a transition from point **b** to point **d** with *lower* unemployment at the wage given by the new competition condition. The consequences of these changes for inequality are evident in the Lorenz curve shown in panel (b):

- A small *reduction* in unemployment from 10 to 4 percent of the population; and
- A *reduction* in the wage share from 0.76 to 0.60.

The result of these changes taken together is to increase the Gini coefficient from 0.216 to 0.32.

M-NOTE 15.5 Entry barriers and economic profit as a share of income

Recall from M-Note 15.2 that the accounting profit share is the complement of the wage share:

$$\text{Accounting profit share} \quad \sigma_A = 1 - \sigma_w = \frac{\rho + b}{1 + \rho}$$

A portion of accounting profits is the opportunity cost of the amount of capital used. In the model used here labor employed prior to the sale of its products is the only capital cost, so the opportunity cost of capital used to produce one unit of output is wages per unit of output paid in advance of production times the opportunity cost of capital. Expressing the hours of labor required to produce one unit of output as a_l this is: $\rho w^c a_l$. Because a_l is just $\frac{1}{\gamma}$ we have the following expression for the **capital share** or σ_E:

continued

CAPITAL SHARE The capital share of total income is the fraction of output accounted for by the opportunity cost of the capital goods used in the production of the output.

$$\frac{\text{Opportunity cost of an hour of labor}}{\text{Output of an hour of labor}} = \frac{\rho w^c}{\gamma} = \sigma_E$$

The economic profit share is the accounting profit share minus the capital share:

$$\text{Economic profit share} \quad \sigma_E = \sigma_A - \sigma_\rho$$

$$\text{substituting in the values derived above} \quad = \frac{\rho + b}{1+\rho} - \frac{\rho w^c}{\gamma}$$

$$\text{and rearranging} \quad \sigma_E = \frac{\rho + b}{1+\rho} - \frac{\rho}{\gamma}\frac{(1-b)\gamma}{(1+\rho)}$$

$$= \frac{\rho + b - \rho + \rho b}{1+\rho}$$

$$= \frac{b(1+\rho)}{1+\rho}$$

$$\text{Economic profit share} \quad \sigma_E = b$$

where b is the extent of barriers to entry.

M-NOTE 15.6 **Disposable income shares in the whole-economy model**

The unemployment benefit (B) paid to the unemployed (u) is paid out of a tax on the employed (n). The pretax income of the employed is $\frac{w^c}{\gamma}$. The n employed people therefore pay a tax equal to a share of unemployment benefits paid to each unemployed person of $\frac{u}{n}$ multiplied by the benefits each unemployed person receives, B, such that the tax share paid by the employed, $\tau = \frac{u}{n}\mathrm{B}$, equals the share paid to the unemployed.

We therefore have the following shares:

$$\text{Pretax wage share} \quad \sigma_{w^c} = \frac{w^c}{\gamma}$$

$$\text{Unemployed share} \quad \sigma_B = \frac{\tau}{\gamma}\frac{\left(\frac{u}{n}\right)\mathrm{B}}{\gamma}$$

$$\text{Wage share (after tax)} \quad \sigma_{w^N} = \frac{w^c - \tau}{\gamma}$$

$$= \frac{w^c}{\gamma} - \frac{\left(\frac{u}{n}\right)\mathrm{B}}{\gamma}$$

$$= \frac{w^c - \left(\frac{u}{n}\right)\mathrm{B}}{\gamma}$$

$$\text{Labor share} = \text{Wage share (after tax)} + \text{Unemployed share}$$

$$\sigma_{w^c} = \sigma_{w^N} + \sigma_B$$

$$= \frac{w^c - \left(\frac{u}{n}\right)\mathrm{B}}{\gamma} + \frac{\left(\frac{u}{n}\right)\mathrm{B}}{\gamma}$$

$$\text{Accounting profit share} \quad \sigma_A = \sigma_E + \sigma_\rho = b + \frac{w^c \rho}{\gamma}$$

$$\text{Economic profit share} \quad \sigma_E = b$$

continued

Sum of shares $$1 = \frac{\overbrace{w^c - \left(\frac{u}{n}\right)B}^{\sigma_{w^N}}}{\gamma} + \frac{\overbrace{\left(\frac{u}{n}\right)B}^{\sigma_B}}{\gamma} + \overbrace{b}^{\sigma_E} + \frac{\overbrace{w^c\rho}^{\sigma_\rho}}{\gamma}$$

Multiply both sides by γ $$\gamma = w^c - \left(\frac{u}{n}\right)B + \left(\frac{u}{n}\right)B + b\gamma + w^c\rho \qquad (15.7)$$

Equation 15.7 shows that the product of an hour's of labor is divided into four parts attributable to the employed, unemployed, economic profits, and opportunity cost of capital.

Wages and labor productivity

! REMINDER Remember the following notation. Real wage determined by the competition condition: w^c; productivity of labor (per hour): γ; barriers to competition: b, opportunity cost of capital: ρ Wage (per hour) determined by the no-shirking condition: w^N; disutility of working at the no-shirking level: $\underline{u}$; unemployment benefit: B; probability of being terminated if shirking: t; probability of not finding a job if terminated: $j = 1 - H$.

Our explanation is supported not only in the data on the markup and economic profit share, but in evidence on wages and productivity. Figure 15.20 shows that from the end of World War II to the end of the 1970s, labor productivity and real wages in manufacturing in the US grew approximately in tandem (real wages actually grew slightly faster than productivity over much of that period).

But since the 1980s, in the US the real wages of workers have been roughly constant, while productivity has much more than doubled. This is consistent with an *increase* in barriers to entry and, as a result, a *decline* in the degree of competition in goods markets. The results predicted by our models are:

- a decrease in the wage share; and
- an increase in the share of economic profits.

These changes follow because, as we have shown in M-Notes 15.2 and 15.5, the share of economic profits equal to b the degree of barriers to entry, and the wage share is equal to $(1-b)/(1+\rho)$.

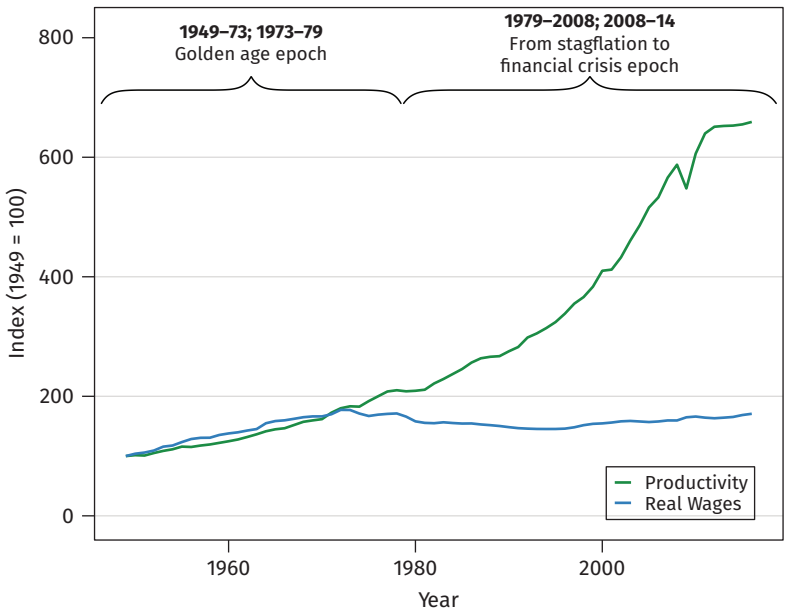
CHECKPOINT 15.7 Explaining the surge in US inequality Explain how a shift downward in the wage curve along with a shift downward in the competition condition could result in a small increase in employment and a large increase in inequality. (Figure 15.19 will be helpful here.)

15.9 PUBLIC POLICY TO RAISE WAGES AND REDUCE UNEMPLOYMENT AND INEQUALITY

We can use the model to study how public policy could alter the wage rate, the profit rate, and level of employment and reduce inequality in a way that is sustainable in the long run. Sustainability here means that the desired policy outcome—higher real wages, or less unemployment—is consistent with maintaining the workers' motivation to work hard and well and the

Figure 15.20 Real wages and productivity of manufacturing workers over time in the US 1949 to 2016. Productivity is an index measuring real output per hour of all persons. Real wages and productivity increased in tandem during the period 1949 to 1979. Afterward, wages remained stagnant while productivity kept increasing.

Source: Bowles et al. (2017).



owners' motivation to invest in the firms' capital goods sufficiently and to hire additional workers even at a higher real wage.

Sustainable policy therefore requires that the Nash equilibrium be moved, which can only happen if one or both of the wage curve and the competition condition are shifted.

Recall that the Nash equilibrium of the model of the whole economy is determined by the competition condition and the wage curve:

$$\text{Competition condition} \qquad w^{c} = \frac{(1-b)\gamma}{(1+\rho)} \tag{15.8}$$

The wage curve is based on the no-shirking condition.

$$\text{No-shirking condition} \qquad w^{N} = B + \underline{u} + \frac{1-t}{tj}\underline{u} \tag{15.9}$$

Letting the probability of remaining jobless if terminated, j, be the fraction of the labor supply that is unemployed, $1-H$, gives us the wage curve:

$$\text{The wage curve} \quad w^{N}(H) = B + \underline{u} + \frac{1-t}{t(1-H)}\underline{u} \tag{15.10}$$

Changing the Nash equilibrium requires changing one or more of the variables in one of these two equations. We will consider three ways that the Nash equilibrium might be changed by public policies, that is, by altering barriers to entry (b), the opportunity cost of capital (ρ), or labor productivity (γ).

Competition policy to reduce barriers to entry (b)

The competition condition immediately above includes the term b, which is a measure of barriers to entry in an industry and therefore reflects how uncompetitive an industry is. This can be a target of economic policy.

As we saw in Chapter 9, the Anti-Trust Division of the Department of Justice in the United States and Competition Commissions in other countries pursue policies to reduce barriers to entry in an industry by means of legislation or legal action. If successful, the effect, as Equation 15.8 shows, is to raise the real wage that is consistent with the competition condition.

Here is how that works. As you know from Chapter 9 lower barriers to entry means that more firms will enter the economy, and the effect will be to reduce the price that will maximize profits for the firms' owners. The resulting reduction in the profit-maximizing markup ratio over cost means that the share of profits in total income falls, and the share of wages increases. This is why the competition condition would shift upward.

At a higher real wage, the level of employment that is consistent with sustaining the incentive to work (shown by the wage curve, that is Equation 15.10) increases. The resulting decrease in the level of unemployment along with the shift in the **labor share** of income then has the effect of reducing the level of inequality in the economy, as shown by the Gini coefficient, as shown in Table 15.3.

Table 15.3 Policy effects on equilibrium. Each line shows a policy, the *parameter* that the policy targets, which curve it shifts and the direction of the shift, and the effect on the real wage, unemployment, and the Gini coefficient.

Policy target	Effect	Shift	Wage	Unemployn	Gini
Competition policy	Reduce b	w^c up	up	down	down
Opportunity cost of capital	Reduce ρ	w^c up	up	down	down
Research, education	Raise γ	w^c up	up	down	down
Disutility of effort	Reduce $\underline{u}$	w^N down	None	down	down
Barriers to job termination	Increase t	w^N down	None	down	down

LABOR SHARE The term labor share—distinct from wage share—refers to unemployment benefits plus wages as a fraction of total income.

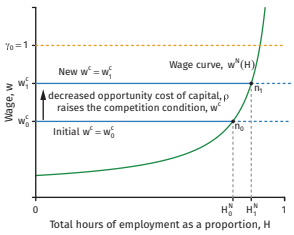
Monetary policy and other interventions to reduce the opportunity cost of capital (ρ)

Equation 15.8 also allows us to see the effect of a central bank's monetary policy. If the central bank (for example, the Federal Reserve Bank in the US) wishes to promote higher levels of investment, it may reduce the borrowing costs of commercial banks and hence lower the interest rates at which other firms and individuals can borrow. This lowers the opportunity cost of capital ρ.

The decrease in ρ means that the wage that is consistent with the competition condition is now higher. This shifts *up* the competition condition. Just as was the case with the upward shift in the competition condition resulting from a more competitive economy, this will allow a *higher* level of employment and higher wages as more firms enter or as firms expand their employment.

The resulting *decrease* in the Gini coefficient occurs because the level of employment has increased and the wage has also increased.

Figure 15.21 Decreased opportunity cost of capital raises the competition condition, w^c.



Research, education, and training to raise labor productivity (γ)

Consider a government policy that supports basic research resulting in improved technologies or additional education and training, all of which will raise γ, the output produced by a worker per hour of work at the required effort level.

Equation 15.8 shows that an increase in γ will shift the competition condition upward, meaning an increase in the real wage consistent with firms neither leaving nor entering markets. If wages do *not* rise, then additional firms will enter the industry or existing firms will increase employment thereby *raising* the real wage necessary to motivate workers to provide effort. The result will be a new Nash equilibrium with *higher* wages and employment. As a result, the Gini coefficient will decrease. (We show in Figure 15.24 (b) the effects of an increase in γ.)

CHECKPOINT 15.8 Putting the models to work Redraw the two figures—the whole-economy model and the associated Lorenz curve comparing the status quo and the effect of the following:

a An increase in labor productivity; and

b An improvement in working conditions that reduces the disutility of effort.

15.10 APPLICATION: TRADE UNIONS, INEQUALITY, AND ECONOMIC PERFORMANCE

In many industries, the contract under which workers work is not between individual workers and the employer, but instead between a

Figure 15.22 Beatrice Webb (1858–1943) was an English economist who invented the term "collective bargaining." She and her husband Sidney were leaders of the Fabian Society, advocating democratic socialism. She was one of four co-founders of the London School of Economics.

Photo: Granger Historical Picture Archive/Alamy Stock Photo.

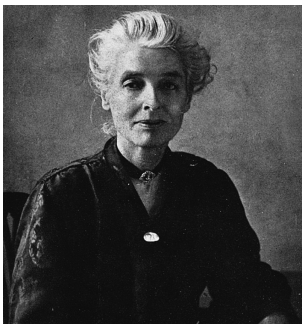
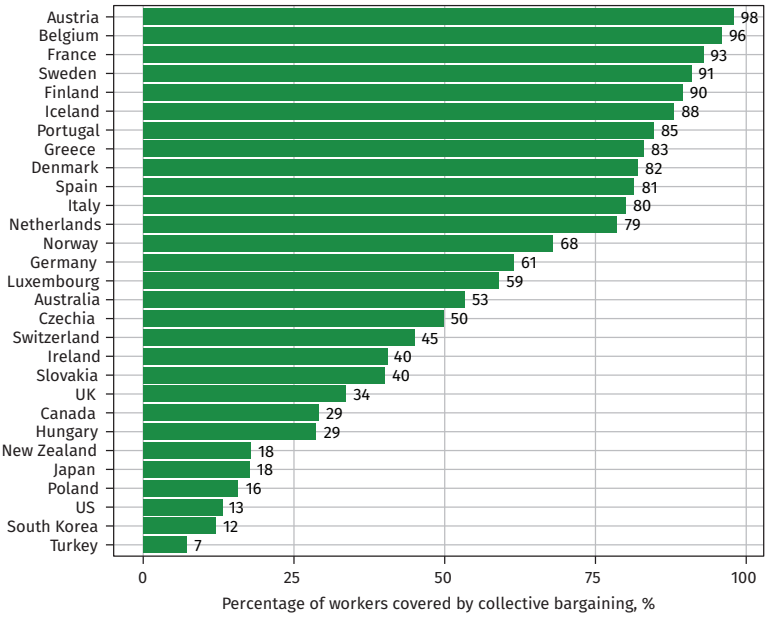


Figure 15.23 The proportion of workers covered by collective wage agreements in different countries. The countries range from Turkey, South Korea, and the US at the low end to France, Belgium, and Austria at the upper end.

Source: OECD Labor Force Statistics.[14]



labor union representing the workers and the employer. In this case the firm is not a wage setter as we described in Chapter 11, but instead engages with the union in what is called collective bargaining to determine the wage rate and other conditions of work that are specified in the contract. Figure 15.23 shows that countries differ markedly in the fraction of all employed workers that are working under a collectively bargained contract.

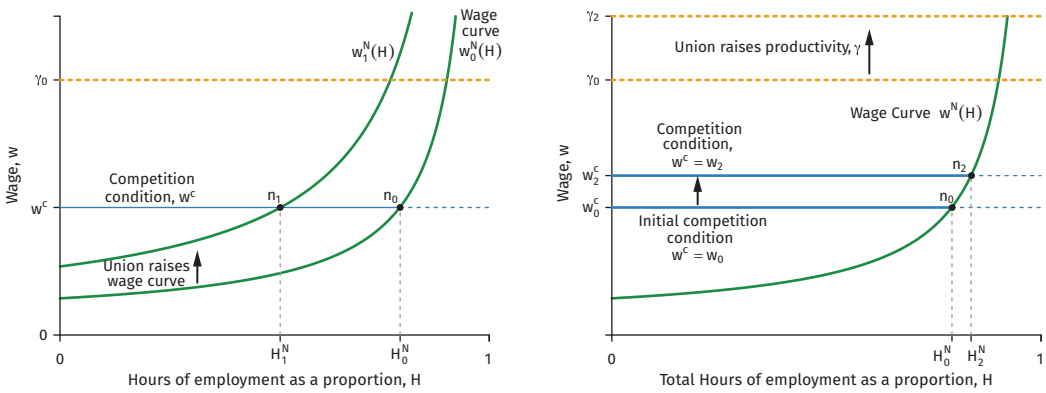
If the economy is not at the Nash equilibrium, **n**, in Figure 15.15, and the wage is below the competition condition line, a trade union may bargain with employers to raise the wage. But Figure 15.15 shows that if trade unions succeed in bargaining a wage higher then w^C, then in the long run firms will leave the industry or the economy and not be replaced by new entrants. As a result the bargained wage $w > w^C$ will not be sustainable.

Limits to what trade unions can do

In the long run, the effect of labor unions on the labor market occurs because of effects on the Nash equilibrium, that is, by shifting either the

LABOR OR TRADE UNION A labor/trade union is an organization of workers who together bargain with one or more employers about wages and working conditions, a process known as collective bargaining.

Figure 15.24 Effects of labor unions on employment and wages. In panel (a), the union decreases t from 1 (the shirker is sure to be dismissed) to $t_1 < 1$. From Equation 15.9, you can see that, as a result, the wage curve shifts *upward* and there is a new Nash equilibrium, $\mathbf{n_1}$ with the same wage and lower employment. In panel (b), when a union increases the productivity of workers (γ) the competition condition shifts upward because now a profit rate equal to the opportunity cost of capital is achieved with a higher real wage. The result is a new Nash equilibrium, $\mathbf{n_2}$, with higher wages (w_2) and higher total employment (H_2^N) than the previous Nash equilibrium $\mathbf{n_0}$.



(a) Negative effect of unions

(b) Positive effect of unions

wage curve or the competition condition. For example, based on Equation 15.9, the union might think that if it were made *more difficult* for the firm to dismiss shirking workers, the employers would be forced to pay *higher* wages to motivate workers to actually work. Reducing the termination probability, t, the equation shows, will raise the "no shirking wage."

This is a case where studying the equilibrium of the whole market, not simply the one-on-one interaction of employer and worker, is required. Equation 15.10 shows that making termination more difficult (lowering t) will *raise* the wage curve for all levels of employment. But Figure 15.24 (a) shows that this does *not* have the effect the trade union intended. On the contrary, in the model, making it harder for employers to fire workers would shift the Nash equilibrium to the left, reducing employment and leaving the wage unaffected.

! REMINDER The equation for the no-shirking wage (the wage curve) is: $w^N = B + \underline{u} + \frac{1-t}{t(1-H)}\underline{u}$. If t decreases, then the denominator in the third term is *smaller* and the numerator is *larger*; therefore the wage curve shifts *up*.

Bargaining for a share of the gains to cooperation

But there are trade union strategies that can *reduce* inequality. Recall (from Chapter 11) that the Nash equilibrium of the labor discipline (Ford) model is not Pareto efficient. Figure 15.25 shows the yellow Pareto-improving lens of possible outcomes with higher wages and greater worker effort that are preferred by both workers and owners. A trade union can sometimes bargain with an owner to allow such mutual gains to be made.

For example, unions may be able to bargain with employers to improve working conditions by providing:

- *material amenities* such as a safer workplace, air-conditioning, and flex-time scheduling, or
- *social amenities* such as a voice in company decision-making and a respectful and fair-minded approach to their workers.

Either will reduce the disutility of effort: working hard for a kind and respectful boss is a lot less unpleasant than working for an indifferent and insulting one.

The result would be a more cooperative workplace environment and workers' willingness to work harder. This is called the **union voice effect**. If this reduced the disutility of working at the "no-shirking effort level" ($\underline{u}(\underline{e})$) then the result, as Equation 15.10 shows, would be a shift down the wage curve.

Why would *lowering* the wage curve be in the interest of workers? Here is another case in which looking at the equilibrium of the whole market or economy is essential. From Figure 15.24 you can see that increasing voice (by decreasing the disutility of work) would *not* affect the wage but it would *increase* the level of employment. As $\underline{u}$ decreases, the numerator of the no-shirking wage decreases, therefore the wage curve shifts down, and the level of employment increases.

An alternative—also favorable—outcome of policies that reduced the disutility of effort would be that the employer decided to increase the level of effort expected of workers *without* lowering the wage. If workers work harder in an hour, then the output produced in an hour γ—the productivity of an hour of (non-shirking) labor—will increase. The result is a shift upward in the competition condition, raising w^c and increasing employment.

Trade unions can also shift up the competition condition by providing a means of communication and conflict resolution between owners and workers about work rules and other practical aspects of the production process. The effect is to increase the productivity of the workers' effort by reducing wasted effort, and this raises the wage that can be offered consistent with the competition condition. This outcome is illustrated in Figure 15.24. In this case, as the figure shows, the effect on the union's negotiation has been to change the Nash equilibrium such that it raises both the level of employment and the wage rate.

The conclusion is that trade unions can affect the functioning of the labor market both positively and negatively. The positive effects of the

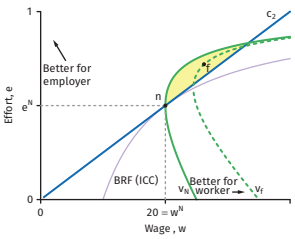


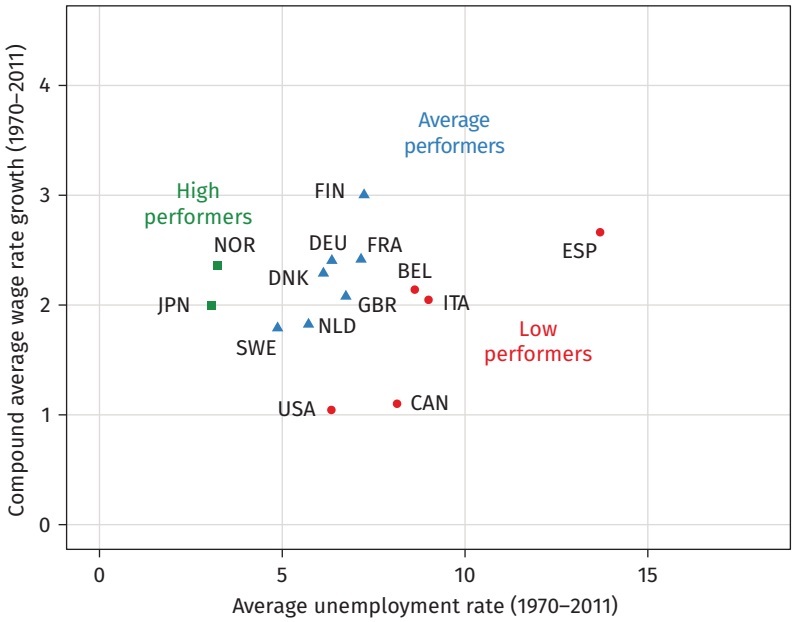
Figure 15.25 A Pareto-provement over the Nash equilibrium in the labor discipline model. Cooperation between employers and workers—possibly facilitated by a labor union—are a change in the rules of the game that may support a Nash equilibrium in the Pareto-improving lens.

UNION VOICE EFFECT The union voice effect occurs when a trade union, by providing a 'voice' to otherwise unheard workers, improves their treatment by employers and their job satisfaction (in our model, decreasing the disutility of work), with, as a result, greater work effort provided by workers, and an increase in output per worker hour.

Figure 15.26 Real wage growth and unemployment in high-income economies.

ESP refers to Spain, GBR to Great Britain, DEU to Germany.

Source: CORE's, *The Economy* (tinyurl.com/y6jf6vpz). Based on data from the US Bureau of Labor Statistics and OECD Labor Statistics.



union voice effect and bargaining with employers to raise productivity are part of the explanation for the evidence in Figure 15.26. By the standard of rapid wage growth and low unemployment over the long run, countries in which a majority of workers work under collective bargaining agreements (Norway, Germany, Denmark, Finland, and Sweden) have outperformed countries with weak trade unions (the US, Canada, and the UK). The figure also shows, however, that some countries where labor union negotiated contracts cover a small fraction of the labor force—Japan is an illustration—perform well compared to countries with stronger unions—such as Spain, Italy, and Belgium.

CHECKPOINT 15.9 Policymaking with labor markets

a. It's not too difficult to explain why the economic conditions experienced by the citizens of Norway and Japan as depicted in Figure 15.26 were preferable to those experience by Americans and Canadians. But how would you describe the differences in the conditions experienced by Finns and Swedes, or Americans and Spaniards?

b. If you were to draw your own personal indifference curves—your evaluation of the countries' record with respect to unemployment and real wage growth—what would they look like?

15.11 THE RULES OF THE GAME AND THE DISTRIBUTION OF RENTS

We can bring together the effects of trade unions, monopoly, and monopsony on the distribution of rents in a single model.

A monopolistically competitive monopsony

To do this we determine the fallback options and incomes received by the owners and workers in a single firm shown in Figure 15.27 (a), that:

- faces a downward-sloping demand curve so that it is a *monopolistic competitor* in the market for its output; and
- hires a significant fraction of the local supply of labor and so is a *monopsonistic employer* in the market for labor, facing a rising average cost of labor hours.

To study this case we return to the model of monopsony hiring introduced in sections 11.12, 11.13, and 11.14, as shown in Figure 15.27. We assume that:

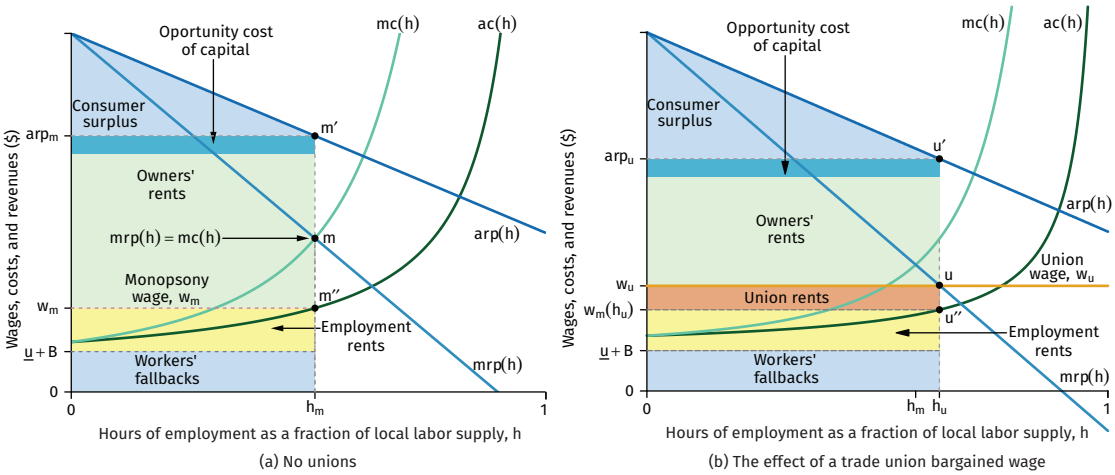
- the only input to the production of the firm's output is labor effort;
- the cost per hour of effort at the level required by the owners of the firm is given by the average cost of labor $ac(h)$ shown in Figure 15.27;
- when workers provide effort at the level required by their employer they produce an amount of output equal to γ;
- the firm sells the product of an hour of employment at a price p so the average revenue product of hours of employment, arp, is $p\gamma$; and
- that is shown as the curve $arp(h)$ in Figure 15.27, which declines because the more hours of employment that the firm hires the more it produces and the lower must be the price at which the increased output can be sold.

Faced with the marginal cost of hours of labor ($mc(h)$) and marginal revenue product of hours of labor ($mrp(h)$) the profit-maximizing firm will hire h^m hours so as to equate $mc(h) = mrp(h)$ as shown at point **m** in the figure. If the firm employs h_m, then it must pay $ac(h_m)$ to motivate the workers to provide the no-shirking level of effort. Given the output produced by hiring h hours, the maximum price that the firm can charge for the product of an hour's work is arp_m. So the Nash equilibrium is that the firm hires h_m hours and pays a wage of w_m per hour and sells the product of an hour's work at the price arp_m.

Owners and workers' fallback options and rents

Because we are interested in the distribution of rents received by the owners and the workers we need to know their fallback options. The fallback option for the workers is to receive an unemployment benefit of B and to avoid the disutility of providing the level of "no-shirking" effort

Figure 15.27 Market power, incomplete contracts, and the distribution of rents among workers, employers, and consumers. The limited competition is illustrated in two ways. First, it means that the price at which the firm can sell its product $p_m(h)$ declines the more it sells. This is why the average revenue product curve $arp_m(h) = \gamma p_m(h)$ is downward-sloping, or what is the same thing, why the $mrp(h)$ is below the $arp(h)$. Second, in panel (a), limited competition in the labor market means that the wage it must pay to motivate workers increases the more workers it hires. So, the average cost of labor hours increases the more the firm hires, and the marginal cost of hours exceeds the average cost. In panel (b), the trade union negotiated wage exceeds the average cost of hours given by the no-shirking condition (except for very high levels of hiring). So, the average and marginal cost are both equal to the union wage.



required by the employer $\underline{u}$. So the total opportunity cost to the workers hired for h^m hours is $h^m(B + \underline{u})$. This is shown in Figure 15.27 (a) as the light-blue-shaded rectangle at the bottom.

The fallback option for the owners of the firm is to invest in some other project the amount that it spent hiring labor at the beginning of production (selling the product is possible only after the time that production takes). If, as before, ρ is the expected rate of profit on the alternative investment (that is, ρ is the opportunity cost of capital), then the opportunity cost of hiring one hour is ρw and the opportunity cost of hiring h_m hours is $\rho w h_m$. This is shown in Figure 15.27 (a) as the dark-blue-shaded rectangle above the owner's rents (the owner's rents must be *net of* the opportunity cost of capital).

The three types of rents shown in the left panel of Figure 15.27 are:

- *Employment rents* received by workers, that is, the excess of the wages rate paid by the monopsony firm over the workers' fallback summed over the hours hired.

REMINDER You studied consumer surplus in Chapters 7 through 9. In those cases the price of a unit of the good sold was on the vertical axis and the quantity number of units of the good sold was on the horizontal axis. Here the hours of employment is on the horizontal axis, and the value of the average revenue product of employment on vertical axis is the amount the buyer pays for the amount of the output produced in an hour, that is, price per unit of the good p times number of units of the good produced in an hour y.

- *Owners' market power rents*, that is the excess of accounting profits over the opportunity cost of capital due to the limited competition in both the labor market and the output market.
- *Consumer surplus*, that is, the excess of the maximum willingness to pay (the height of the average revenue product of labor) for the product of an hour's employment and what they actually do pay.

A trade union's labor market power rent

In panel (b) we introduce a trade union that has enough bargaining power so that the firm's owners have agreed to a wage (w_u) that, at the level of employment of the firm in the absence of the trade union (h_m), is *above* the minimum they need to pay in order to motivate workers to provide the no-shirking level of effort. In Figure 15.27 (b) you can identify the following effects of the trade union's bargain:

- At the current level of hiring (h_m), the firm no longer has monopsony power: the union negotiated wage it must pay is now independent of the level of employment.
- As a result, the firm's owners will maximize their profits by increasing the level of employment to h_u, the point at which the marginal revenue product of labor is equal to the wage.
- The total wages received by the firm's workers is now $w_u h_u$ which is now composed of–in addition to the workers' fallbacks and employment rents–a new rent, the trade union's labor market power rent the size of which is the area of the orange-shaded rectangle in the figure.
- The owner's rents are now less because the higher wage reduces accounting profits and also increases the total opportunity cost of the capital devoted to paying the wages in advance of the sale of the product ($\rho w_u h_u$).
- Because more is now produced we know from the inverse demand curve that the price at which it can be sold is lower, so what consumers pay for the product of an hour of labor arp_m is now lower.
- Because consumers now buy more at a lower price the rents constituted by the consumer surplus triangle have now increased.

Figure 15.27 provides a summary of how differing institutional environments–represented by monopsony, monopolistic competition, and trade union bargaining–affect how different distributions of income are implemented. The impact of differing rules of the game extend beyond the distribution of income to include differences in the ways that people experience economic interactions depending on whether they are an owner, a manager, or a worker.

CHECKPOINT 15.10 The rules of the game and the distribution of rents

Redraw Figure 15.27 for the cases where:

a. The firm is a monopsonist in the labor market but faces unlimited competition in the product market so that the inverse demand curve is horizontal.

b. The firm is a monopolistic competitor but is not a monopsonist in the labor market, so the "no-shirking wage" that is the average cost of an hour of labor does not depend on the hours of labor that it employs.

c. The unemployment benefit B is eliminated.

15.12 CAPITALISM AS A SOCIAL SYSTEM: DISPARITIES IN WEALTH AND POWER

Our model of the capitalist firm provides a look into the politics, the sociology, and the psychology of the workplace.

The economy: Markets *plus* organizations

In 1951 a paper by Herbert Simon pioneered the study of exchanges with incomplete contracting.[15] Forty years later, he imagined a mythical visitor from Mars approaching earth in a spaceship:

> equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas … market transactions show as red lines connecting the firms forming a network in the spaces between them.[16]

What would the Martian see, mused Simon?

> No matter whether our visitor approached the United States or … urban China, or the European Community, the greater part of the space below it would be within the green areas, for almost all of the inhabitants would be workers, hence inside the firm boundaries. Organizations would be the dominant feature of the landscape.

The moral of the story, for Simon, is about the proper subject matter of economics:

> A message sent back home [by the Martian], describing the scene would speak of large green areas interconnected by red lines. It would not speak of a network of red lines connecting green spots.

Economics, Simon insisted, should be at least as much about the *structure* of organizations—including the exercise of power—as it is about *voluntary* exchanges on markets.

HISTORY Herbert Simon (1916–2001) was a Nobel Laureate in economics though his undergraduate and PhD degrees were in political science. He was a pioneer in fields as diverse as artificial intelligence and organizational theory and is best known for stressing people's limited cognitive capacities and incomplete information when making decisions, what he termed "bounded rationality." He favored replacing taxes on wages and salaries by a tax on the value of land.

"Power" in economics

The idea that the political structure of firms should play a central role in the analysis of the economy is a recent development in economics. But in other fields the point seems obvious. The US Supreme Court stated it quite clearly in the head quote to this chapter in 1898.

HISTORY A decade after Samuelson's remark, John Kenneth Galbraith chided his fellow economists for not having asked "why power is associated with some factors [of production] and not with others?"[18]

But many economists have considered the exercise of power by employers over workers to be illusory. In 1957, Paul Samuelson wrote "Remember that in a perfectly competitive market, it really does not matter who hires whom; so have labor hire capital."[17]

An important mid-twentieth-century microeconomics text by Armen Alchian and William Allen may have surprised some students with the following:

> Calling the employer the boss is a custom derived from the fact that the "boss" specifies the particular task. One could have called the worker the boss because he orders the employer to pay him a specific sum if he wants services performed. But words are words.[19]

HISTORY The seventeenth-century philosopher Thomas Hobbes wrote, "Riches joyned with liberality [generosity] is Power; because it procureth friends, and servants." In Hobbes's day the term "servant" referred to any worker.[20]

We doubt very much that the authors would disagree with the Supreme Court's assessment as an empirical account. Like Samuelson they were describing the logic of a *model*, not an *empirical* aspect of the economy.

In the approach modeled here, Alchian and Allen's example would look quite different. The employer would simply refuse any pay demanded by the worker unless it happened to be $w^N(H)$, the least wage the employer could offer consistent with the worker providing effort at the employer's chosen non-shirking level. There are identical workers ready to take the place of the overly demanding worker.

Economists recognize that domination of some people by others was an essential characteristic in many past economic systems, slavery on plantations in the US before the Civil War or feudalism in medieval Europe, for example. These were both economic systems in which one party—the slave owner, the feudal lord—could threaten dire consequences to any of "their" enslaved people or serfs who did not obey their commands.

! REMINDER For B to have power over A, it is sufficient that, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B. Where B has this capacity because B is on the short side of a non-clearing market, we say that B has short-side power.

But in a capitalist economy exchanges are voluntary, not coerced at gunpoint (or swordpoint), and parties to any exchange are free to walk away. That is why the participation constraint must be satisfied for any exchange to take place: as we saw in Chapter 5, this is because participation in an exchange is has to be motivated by the prospect of doing better as a result.

Even in a voluntary exchange among private parties, however, power can be exercised. We have seen that where contracts are incomplete, one of the actors, the principal, acting as a first-mover offers terms that induce the agent to do something in the principal's interest that could not be secured by enforcing the terms of a contract.

The principal does this not by threatening physical harm to the agent but by committing to terminate the interaction if the agent does not do what the principal asks. The harm threatened is economic: if the relationship ends, then the agent loses the enforcement rent that she received as part of the interaction.

To explain how power can be exercised even when either party to an exchange can walk away we will use an economic model of the firm that we owe to an unlikely pair. The first is Karl Marx the nineteenth-century socialist revolutionary; the second is twentieth-century University of Chicago economist Ronald Coase whose work is often used to advocate a lesser role for the government in addressing economic problems.

HISTORY Ronald Coase won the Nobel Prize for his contributions to the economics of institutions. Upon his death (at the age of 102) *Forbes* magazine called him "the greatest of the many great University of Chicago economists."[21] Our explanation of market competition (in Chapter 9), firms and labor markets (Chapter 11), and bargaining (Chapter 14) owes much to Coase's work.

Marx was the first to stress the fact that the employment contract stipulated time on the job, but it did not cover such things as the *amount* or *quality* of work done. Rather, the employment contract specified the *hours* during which the worker agreed to submit to the *authority* of the employer. Under the employment contract, the employer does not purchase the worker's work, he rents the worker's time.

› EXAMPLE For example, should a car manufacturer also manufacture tires for the car (produce them in-house) or should they purchase the tires that another manufacturer produces?

According to Marx the worker's supply of effort to the production process is not secured by contract but is instead an "extraction" that "only by misuse could . . . have been called any kind of exchange at all." Anticipating the logic of Henry Ford's "five dollar day" as well as late twentieth-century developments in economic theory, Marx pointed out that an increase in the wage might *reduce* the cost of labor per unit of output.

Like Marx, Coase (1910-2013) stressed the central role of authority in the firm's contractual relations: "note the character of the contract into which a factor enters that is employed within a firm. . . . [T]he factor . . . for certain remuneration agrees to obey the directions of the entrepreneur." Indeed, Coase defined the firm by its *political* structure:

> If a workman moves from department Y to department X, he does not go because of a change in prices but because he is ordered to do so . . . the distinguishing mark of the firm is the suppression of the price mechanism.[22]

Coase sought to understand why firms exist at all, and what determines the extent of what he called these "islands of conscious power in this ocean of unconscious cooperation."

The size of the firm is determined by the decisions of its owners and managers when confronted with the question: should we purchase this input from another supplier or should we make it in-house. The more the firm produces in-house, the larger will the firm be, for a given level of final sales of the product. The reason why we have large firms according to Coase, is that for many inputs *the suppression of the price mechanism* within the firm in favor of a centralized system of control makes in-house production more cost-effective than acquiring the same input on the market.

EXAMPLE The "suppression of the price mechanism"—the expression that Coase used to describe the firm—sounds like something that was done in the Soviet Union when the economy was centrally planned by government officials, a so-called *command* economy *without* market-determined prices. But Coase pointed out that the private owners of firms in a capitalist economy also suppress the price mechanism by not having prices *within* the firm determine which worker does what task. In a firm, workers do what managers or supervisors tell them to, they are not guided by prices.

Given the benefits of specialization and economies of scale, economic activity is necessarily social rather than individual: as a result economics is about organizations as much as it is about individuals making exchanges on markets. What we have learned from Coase, Simon, and Marx is that the types of institutional arrangements governing production and exchange–including relationships between employers and workers in firms–reflect the fact that the conflicts of interest among the participants are governed by incomplete contracts.

The combined effect of incomplete contracts and conflicts of interest is that the determination of the allocation resulting from an exchange depends on who exercises what kind of powers in the transaction. Power in the firm is generally exercised by those who hold what are called the residual rights of control, meaning the right to determine what is not specified contractually. These are the owners of the firm or their delegated managers.

Control over assets and power over people

What we have explained is why control rights over assets confers power over people. Samuelson's claim asserts the contrary–it does not matter who hires whom.

Samuelson is right about "a perfectly competitive market" if we add the proviso "with complete contracts." In the perfectly competitive general equilibrium model to which he was referring the labor contract is assumed to be complete, so the notion of "hiring" simply means "buying." "What does it mean," Oliver Hart asked, "to put someone 'in charge' of an action or decision if all actions can be specified in a contract?"[23]

This basic point also explains why, in Marx's terms, contractual transactions on competitive markets appear to be a free exchange among equals ("a very Eden of the innate rights of man" is the expression he used), while in the workplace the two parties to the employment contract take on a different appearance: the employer is boss and the worker is "his laborer."[24]

REMINDER Recall that in a market with incomplete contracts one side of the market will be the *short side* because it is the side on which the desired number of contracts is lowest, whereas the other side of the market is the *long side*. For example, in employment markets, employers desire fewer contracts than workers, and so employers have short-side power over the long-side workers who they employ.

Samuelson's and Marx's picture of the economy as a social system could not be more different. For Samuelson, the economy is a level playing field politically speaking, no actor has any power or authority over any other. For Marx the economy is also a political system in which power is exercised by those with wealth.

Expressed in modern economic terms, we update Marx to say that those on the short side of a market–employers in the labor market, lenders in the credit market for example–exercise short-side power over those on the long side of the market with whom they transact–workers and borrowers.

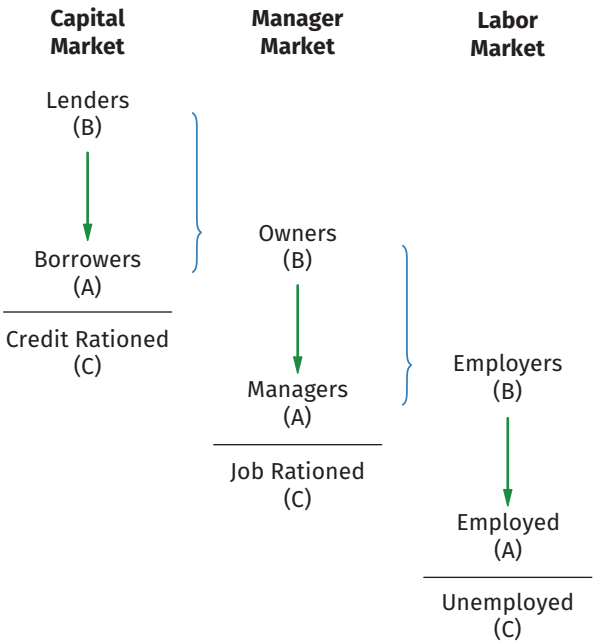
The political dimension of the economy is depicted in Figure 15.28 as a downward cascade of short-side power beginning with wealthy lenders who exercise power over borrowers wealthy enough to secure a transaction. The wealthy and the successful borrowers, then exercise power

over managers (those who secured employment), who in turn, along with owners, exercise short-side power over workers. Here we introduce the market for managers as a distinct kind of contingent renewal contract (following the structure of the contract between employers and workers in Chapter 11).

The owners of a firm are the principals and the managers are the agents. The principal would like the agent to skillfully manage the owner's assets to maximize the owner's wealth. But the manager's action cannot be subject to a complete contract. The other necessary ingredient of a principal-agent relationship—conflict of interest—is also present, because the manager has interests other than to maximize the owner's wealth. These interests include the manager's own leisure, frequent first-class air travel and luxurious accommodation, and self-promotion activities that will improve his fallback position (that is, his prospects for employment in a different firm).

HISTORY The term "other people's money" is from Adam Smith. He foresaw the principal-agent problem faced by owners who would like the managers ("directors") of their assets to maximize the owners' wealth.[25]

Figure 15.28 The incomplete contracts model of the economic and political structure of a capitalist economy. The Bs are short-side principals exercising short-side power over the As (the long-side agents with whom they transact). The Cs are quantity-constrained long-siders: the unemployed, job-rationed, or credit market excluded. The green arrows show the direction in which power is exercised—that is, by principals over agents. The blue brackets indicate that the As and Bs in the market on the left become the Bs in the market on the right.



✓ **FACT CHECK** This is often the model followed by franchising: managers of a franchise pay the owners of a brand a fixed amount for use of the brand, but are residual claimants on the profits of their franchise.

› **EXAMPLE** The powers wielded by principals (owners) over agents (workers) go far beyond Coase's example—directing a person to work in department Y rather than department X. These power include requiring the worker during the COVID-19 pandemic to come to work, exposing the worker and her family to infection.

To reduce the conflict of interest, owners typically compensate managers not only with a salary but also with payments or stock options that will increase with the value of the firm.[26] But unless the manager owns a very substantial fraction of the entire firm, the conflict of interest between principal and agent will remain an important aspect of the owner-manager relationship. The only way to eliminate the conflict of interest would be to make the manager the sole residual claimant on the value of the firm's assets, while paying to the owners some contractually fixed amount independent of the firm's value.

In Figure 15.28, the short-siders (B) exercise short-side power over the long-siders with whom they transact (A), while the excluded long-siders (C) are quantity-constrained.

The vertical dimension (indicated by the vertical arrows) is between those who have short-side power (the principals) and those over whom this power is exercised (the agents), that is, between the lenders and borrowers, the owners and the managers, and the managers and the workers. Principals (Bs) and agents (As) differ in wealth; possessing wealth is a reason why one has he option of being a lender, an owner, or an employer.

But among the long-siders, the As and the Cs—those who are able to make a transaction and those who are excluded—may be identical. This is the case in the labor market model, for example, in which workers—both employed and unemployed—are identical except that some of them have jobs and others do not.

We have come to three conclusions:

- *Those with wealth are more likely to be able to become principals* in principal–agent relationships.
- As actors on the short side of markets that do not clear, *principals exercise power over agents and they benefit from the first-mover advantage that their position confers.*
- *Wealth and power are concentrated in the same hands*, even in an environment of competitive markets in which the powerful cannot secure participation in exchanges by coercion.

The sociology and psychology of short-side power

The model of power developed here provides a reason to doubt the old adage quoted in Chapter 12: "The wealthy are different from everybody else; they have more money." Wealth does indeed determine the position of one's budget constraint and wealth commands more goods and services. Substantial wealth gives a person a large feasible set—for say consumption, free time, and other valued things. Having a wider range of choice because of an enlarged feasible set, we can say that wealthy people have more freedom.

But those wealthy enough to engage in their own projects or to borrow large amounts at the going rate of interest enjoy more than superior purchasing power. They may command people as well as goods. Their access to capital allows them, but not others, to become employers of managers and workers, and as such to occupy positions of short-side power in non-clearing markets.

Those without wealth tend to be constrained not only by a more limited feasible set of consumption choices, but also by the fact that as workers or borrowers they are subject to the exercise of short-side power by others.

These disparities in power show up in many noneconomic realms of our lives. Working at a particular kind of job is a relationship that persists over many years even decades and people change as a result of their experiences at work. The workplace is a cultural environment in which workers' and employers' preferences and beliefs evolve. Workplaces are no different in this respect from schools or neighborhoods; they are environments in which we spend a lot of our waking hours, and, as we will see, how we interact there influences how we develop as people and how we raise our children.

An empirical example will suggest the importance of these effects. Over a period of three decades the social psychologist Melvin Kohn and his collaborators studied the relationship between a position in the authority structure of your workplace—giving as opposed to taking orders—and the your valuation of self-direction and independence in your children, as well as your own intellectual flexibility, and personal self-directedness. They concluded that "the experience of occupational self-direction has a profound effect on people's values, orientation, and cognitive functioning."[27] They found, for example, that those who routinely take orders on the job place a large value on obedience in raising their children, while those who give order place a higher value on independence.

His collaborative study of Japan, the US, and Poland (when it was still under Communist Party rule) yielded cross-culturally consistent findings: people who exercise self-direction on the job also value self-direction more in other realms of their life (including child rearing and leisure activities) and are less likely to exhibit fatalism, distrust, and self-deprecation. Kohn and his co authors reason that "social structure affects individual psychological functioning mainly by affecting the conditions of people's own lives." Kohn concludes that:

> The simple explanation that accounts for virtually all that is known about the effects of job on personality ... is that the processes are direct: learning from the job and extending those lessons to off-the-job realities.[28]

As the personality dimensions mentioned by Kohn are part of individuals' preferences explaining how they raise their children, what kind of leisure

activities they engage in and the like, this is strong evidence for the effects of workplace organization on our values. This is an example of endogenous preferences: our preferences being altered by our economic or other experiences.

CHECKPOINT 15.11 **Mayors and managers** How is the power that employers exercise over workers similar to or different from the exercise of power by governments (e.g., the mayor of a city) over citizens?

15.13 APPLICATION: A WORKER-OWNED COOPERATIVE

To some, the power of employers over workers has seemed inconsistent with both democratic principles and the dignity of the workers. The nineteenth-century philosopher, economist, and author of *On Liberty*, John Stuart Mill, wrote: "To work at the bidding and for the profit of another, without any interest in the work . . . is not, even when wages are high, a satisfactory state to human beings of educated intelligence."[29] He went on to predict in his 1848 *Principles of Political Economy*, perhaps the first textbook in economics, that the "relation of masters and work-people will be gradually superseded by [an] association of labourers among themselves."

Here is an example of how that might look. In 1921, a group of loggers, carpenters, and mechanics in Olympia, Washington in the US formed the Olympia Veneer plywood cooperative. In return for an investment of $1,000, a cooperative member gained the right to work in the plywood plant and to share equally in any profit. Members wishing to leave had to sell their shares, and prospective members, if approved by the membership, were required to purchase shares, which by 1923 were selling for $2,550.[30]

A change in the rules of the game

This was a generation before Samuelson, but these workers certainly would not have agreed that "it really does not matter who hires whom." For them it very much mattered, and they set out to do exactly what Samuelson would later whimsically suggest: "have labor hire capital."

Olympia Veneer was not a capitalist firm: members owned the buildings and equipment with which they made the plywood. They were their own employers.

The conventional and cooperative plywood firms exemplify differing assignments of the relevant rights. In worker-owned cooperatives, both residual claimancy and control is assigned to the member-owners who supply labor. This contrasts with conventional firms in which the suppliers of capital and labor are distinct individuals, and residual claimancy and control is assigned to the capital suppliers.[31]

In 1939, 250 workers in nearby Anacortes invested $2,000 each in a second cooperative plywood mill. Strong wartime demand for plywood boosted the value of their shares to $28,000 in 1951, and members were paying themselves at rates double the union wage in nearby conventionally organized (capitalist) plywood mills.

Stimulated by the success of Olympia Veneer and Anacortes, between 1949 and 1956 twenty-one more co-ops entered the plywood industry in the states of Washington and Oregon, nine of them by buying out existing conventional firms. Banks offered loans to prospective cooperative members looking for a way to buy the shares necessary for membership. The borrower's home—if they owned it—could be used as collateral.

This is exactly what the kind of transfer of ownership rights that the Coasean bargaining approach would predict. If the cooperative organization is a more effective way to organize plywood production, then the assets of a plywood factory—the buildings, machinery, and trademark—would be worth more to a team of cooperative workers than to the erstwhile capitalist owners of the firm. So the workers—if they could find a way to borrow the necessary funds—would purchase the assets, and convert the firm to a cooperative.

At mid-century, about half of the plywood firms were co-ops, the rest being conventional firms. In some of the conventional firms the workers were members of labor unions, and in others, not. Though the co-ops and conventional firms used virtually identical machinery, the co-ops specialized in the more labor-intensive "sanded" plywood because, as one analyst of the co-ops commented, sanded plywood "puts a premium on worker effort."[32]

The structure of the typical plywood co-op was both egalitarian and democratic. With few exceptions, worker-owners received equal pay, and jobs were often rotated. Management was elected by the body of worker-members. Some nonmembers were hired under conventional wage contracts, their numbers making up an average of one-quarter of the total workforce. High levels of productivity were maintained through a strong work ethic among members, enforced by peer pressure and mutual monitoring. The resulting saving in supervision costs was substantial: when one conventional firm converted to a co-op, the number of supervisors was reduced to one-quarter of its previous level.

✓ **FACT CHECK** James Andreoni and Laura Gee explored how mutual monitoring works in an experimental Public Goods Game with students in the US where the team members are able to appoint one of the team members as a 'hired gun' with the ability to punish the lowest-performing team member. Subjects in the experiment were willing to pay for a "gun for hire" over 70 percent of the time and the mechanism increased average payoffs between 15 percent to 40 percent. The institution of the hired gun, like a manager in a cooperative, ensured that the experimental subjects cooperated in the Public Goods Game.[33]

Why cooperative production succeeded (while it lasted)

The ownership shares could be purchased only by a person wishing to work in the factory, and most of those jobseekers were not wealthy enough to put up a lot of money for a share. This limited the demand for shares and therefore lowered the share prices.

As a result joining a co-op was a good investment: an individual who purchased a share and worked in a co-op for a number of years had a

much higher long-term average income than an individual who put the value of a share in a Portland savings bank and worked at union wages in a conventional firm.

The coexistence of cooperatives and conventional firms producing the same goods using virtually identical technologies over a period of three-quarters of a century provides a remarkable opportunity for us to compare and contrast the success and failures of different institutional structures. Conventional firms and cooperatives alike were able to attract both labor and capital over this period.

But the firms differed markedly in a number of ways.

- *Productivity*: The **total factor productivity** of the co-ops—a measure of the productivity of labor and capital goods combined—was substantially higher than of conventional firms (from 6 percent higher to 45 percent higher, depending on the method of estimation).
- *Equality and security*: Cooperatives also adjusted to insufficient product demand in a very distinctive way: rather than laying off members, they reduced hours and pay of all workers, thereby spreading the impact of negative shocks among the membership so as to avoid any member bearing the cost of joblessness.

In this particular case, contrary to Samuelson, it mattered very much "who hired whom."

Reasons why the cooperative firms were more productive than their conventional competitors include the superior work effort and reduced cost of monitoring of the co-op workers. This would occur because each co-op member shares in the income that they and their fellow workers produce, so as a result:

- worker-owners have a greater incentive to work hard and well; and
- to assist monitoring other workers; and also
- they may experience less disutility of effort because they are working under a system of discipline that they have devised and agreed to, not one imposed by an outsider.

Eventually, some co-ops either transformed themselves into de facto conventional firms, or sold out to conventional firms. For example, by mid-century the remaining handful of member-owners of Olympia Veneer were employing 1,000 workers on conventional wage contracts, remaining a cooperative in name only. In 1954, they sold their shares to the US Plywood Corporation. In the sale, 23 early members realized a return averaging

TOTAL FACTOR PRODUCTIVITY Output divided by a weighted sum of the inputs (the weights being each input's relative contribution to producing output).

$652,000 (in 1954 dollars) on their average initial investment (of $1,415 (in 1954 dollars). The entire industry moved from the Northwest to the Southeast in the 1980s and 1990s; none of the firms in the new location were cooperatives. (We return to the closely related problem of team production in the next chapter.)

A cooperative economy

The democratic management of the cooperatives by worker-owners provided both a higher level of income and an opportunity for workers to participate in decision-making about their work. We could hypothetically represent an economy made up of cooperatives just as we have in Figure 15.15 (the whole-economy model) with two changes:

- The increased productivity of the cooperative workers—an increase in γ—would (as you can see from Equation 15.8) result in a shift upward in the competition condition, that is, higher w^c. The new Nash equilibrium results in higher wages and less unemployment
- The greater satisfaction with the more democratic rather than top down decision-making process could reduce the disutility of labor ($\underline{u}$), which would (as you can see from Equation 15.9) shift downward the wage curve. The result would be a further increase in employment at the Nash equilibrium, and no further change in the wage.

The result of greater employment at higher wages would be a larger pie, with less unequal slices.

With these benefits to cooperative members and improvements in the performance of the whole economy the spread of cooperative production would seem to be not simply a good public policy. It would occur spontaneously through the actions taken by groups of workers, as it did in the plywood coops in the Pacific Northwest of the US a century ago.

CHECKPOINT 15.12 **Pros and cons of cooperatives** What would be the advantages and drawbacks of being a member of a worker-owned cooperative compared to being an employee of a conventional firm? Assume that both people have limited wealth, so the co-op member owns shares in the co-op and has borrowed money from a bank to buy her home (with the home as collateral) while the conventional employee purchased his home without borrowing.

! REMINDER In Chapter 12, we saw that a borrower who has more wealth to invest as *equity* in a project or to post as collateral will be able to borrow at a lower interest rate (lower cost of capital) than a borrower with less wealth.

15.14 RISK AND REDISTRIBUTION

There are two reasons why John Stuart Mill's prediction that worker-owned cooperatives would become the dominant form of economic organization did not occur: both arising from the fact that most workers do *not* own substantial amounts of wealth:

- *Risk aversion*: Not having substantial wealth, workers tend to be risk averse. They may view employment for a given wage as a kind of insurance, and prefer to be employed as wage earners rather than residual claimants on a risky income stream based on the firm's revenues.
- *Credit constraints*: Workers are not able to borrow substantial sums of money at interest rates as favorable as those available to wealthier borrowers, or potentially are not able to borrow at all, so it is difficult for them to become owners of the firm's assets.

We now will see why a combination of risk aversion and limited access to credit explains why few workers create or join worker-owned cooperatives.

Would a wealth-poor person want to hold a risky asset?

Would a worker currently employed on a fixed wage contract—meaning, receiving w^c with certainty—prefer instead to be the owner of the capital goods with which she worked as would be the case if she were to become a member of a worker-owned cooperative? If the worker were considering forming a cooperative she would do this with a team of other workers. But for simplicity we here consider the case where there are no economies of scale, so she could just borrow the funds, purchase or rent the capital goods, and go into business on her own.

- The worker would now become the *residual claimant* on the income (y) resulting from the project.
- She would also have *control rights* about how the project was conducted, that is, she would choose the risk level.
- She would bear all of the *risk* of the project.
- As owner of the capital goods (k) she would have to consider the opportunity cost of the value of the capital goods that she uses.

Because the former worker has limited wealth, the opportunity cost of capital for her ($\bar{\rho}$) is more than it would be for a wealthy owner, namely ρ. The next best use of the funds she invests in the asset could be, for example, to purchase a car or home. If instead the funds are used to buy the asset, she will have to borrow at a high interest rate to purchase the car or home. So the opportunity cost of investing k in the asset, meaning having to borrow money for the car or house ($\bar{\rho}k$), would be more than if she were wealthy and could borrow at a lower rates of interest.

For simplicity we assume that the amount of work she does on the project is the same as a worker or as an owner-operator. Because she is an owner she is now residual claimant on the revenues of her project and she therefore owns the income resulting from her work.

As an owner-operator the following would be true:

- her expected income is therefore the expected profit from the project minus the opportunity cost of capital or $\hat{y} = \hat{y}(\Delta) - \bar{\rho}k$; and
- her utility would be $u(\hat{y}, \Delta) = u((\hat{y}(\Delta) - \bar{\rho}k), \Delta)$.

We want to know how her utility as a worker—namely $u(w^c, 0)$—compares to her utility as an owner-operator who selects risk level Δ, that is $u((\hat{y}(\Delta) - \bar{\rho}k), \Delta)$.

In Figure 15.29 we show indifference curves for two possible owner-operators, who we shall call Ana (A) and Beata (B). Because her indifference curves are steep, we know that Ana is very risk-averse while Beata is only modestly risk-averse (her indifference curves are flatter).

For each of the two we ask if she would prefer bearing the risk associated with owning the assets, as she would were she to join or form a worker-owned cooperative, or instead work for a fixed wage:

1. from the tangency of her indifference curves and the risk-return schedule, we determine which level of risk she would take under each circumstance; and then
2. we ask is she better off with the resulting expected income and risk level than she would be with a certain wage equal to w^c?

Given the same risk-return schedule $\hat{y}(\Delta)$, Ana and Beata will choose, respectively, points **a** and **b**, with risk levels Δ_a and Δ_b and corresponding expected incomes $\hat{y}_a$ and $\hat{y}_b$. Risk-averse Ana will choose less risk and as a result have a lower expected income than will less risk-averse Beata ($\hat{y}_a < \hat{y}_b$).

To see how these workers would evaluate the prospect of being a wage worker rather than an owner-operator, compare the certainty equivalent of points **a** and **b** (chosen by the two owner-operators) with the utility of the certain wage that each would receive as a worker.

We see that risk-averse Ana would prefer to be a wage worker: the certain wage (w^c) as a worker is higher than the certainty equivalent of the best she could do as owner-operator, namely choosing point **a** with certainty equivalent $\overline{w}^A$. You can see this in Figure 15.29 (a), which shows that Ana's indifference curve going through w^c, where she works for a wage, provides her a higher utility (on indifference curve u_2^A) than if she were an owner-operator using capital goods at point **a** (on indifference curve u_1^A).

Beata, though, is on a higher indifference curve (u_2^B) at point **b** as an owner-operator than she would be as a wage worker receiving w^c (on indifference curve u_1^B). Less risk-averse Beata would therefore prefer to be the owner of the capital goods. So, if she were initially a wage worker and could borrow funds to purchase the capital goods at rate $\bar{\rho}$, she would do so.

But if most workers lack wealth and therefore are more like Ana, few would be willing to become owner-operators like those who risked losing

their homes when they used the value of their house as collateral in order to get a loan to purchase shares in Olympia Vaneer, the first plywood co-op.

Comparing Ana and Beata, and their situations there are two changes that would give Ana good reason to become an owner-operator:

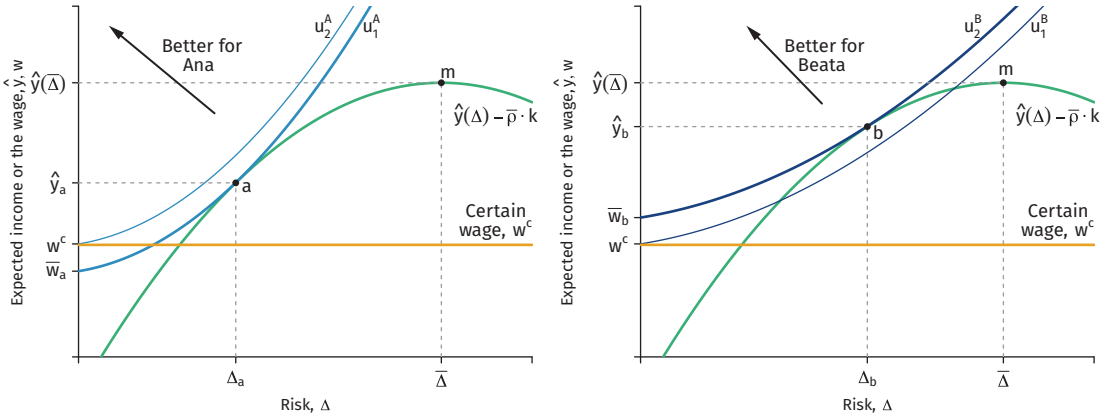
- a reduction in the rate of interest at which she could borrow, $\bar{\rho}$; or
- a reduction in her risk aversion.

Risk, redistribution, and innovation

Both could be accomplished by a redistribution of wealth. Suppose the ownership of the capital goods were simply transferred to a prospective owner-operator like Ana, so that she would be much richer than before: Would this wealth redistribution make it attractive to retain ownership rather than selling?

The fact that she now owns the capital goods does not mean that the opportunity cost of capital is irrelevant to her (she could, for example, sell the capital good, and pay off even more of the debt she has incurred). Let

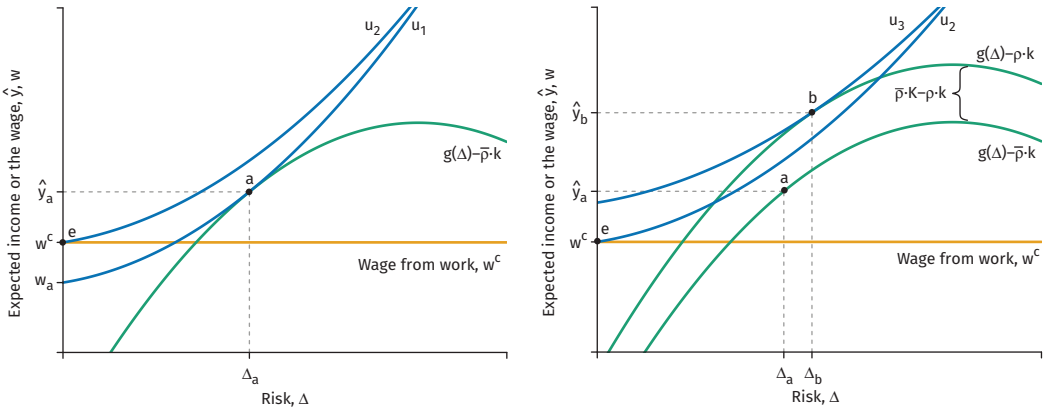
Figure 15.29 The risk and income choices of two owner-operators. Ana is very risk averse and is considering her choice of point **a** on indifference curve u_1^A as an owner operator and a certain wage w^c on indifference curve u_2^A with $u_2^A > u_1^A$. Her utility is higher at the wage than at the combination of expected income and risk $(\hat{y}^A, \Delta^A)$, so she would prefer to be employed—the less risky option—rather than to be an owner-operator. Beata is less risk averse and is comparing the best she could do as an owner operator that is point **b** on indifference curve u_2^B and a certain wage w^c on indifference curve u_1^B with $u_2^B > u_1^B$. Her utility is higher when she is an owner-operator with expected income and risk $(\hat{y}^B, \Delta^B)$ than at the certain wage, so she would prefer to hold the risky asset and be an owner-operator to being a wage worker.



(a) Ana chooses to work for a wage

(b) Beata chooses to be self-employed

Figure 15.30 Effects of redistribution of wealth to one owner-operator. Panel (a) repeats panel (a) of Figure 15.29. In panel (b), redistribution of wealth makes ownership with risk exposure preferable to employment at the wage determined by the competition condition, w^c. This is due to two effects. First is that her increased wealth *reduces* the opportunity cost of capital from $\overline{\rho}$ to ρ and shifts up the green risk-return schedule as shown. Second, because Ana the owner-operator is richer she is *less risk-averse* (flatter blue indifference curve). The shift in the risk-return schedule and the change in the slope of the indifference curves combine to make ownership more attractive than wage employment, and the certainty equivalent of her chosen risk level, and the resulting expected income, exceeds the wage.



(a) Ana's choice before wealth redistribution

(b) Ana's choice with greater wealth

us assume that she were simply given ownership of the capital goods she would be wealthy enough to borrow at the same rate that was available to the rich person, namely ρ. This is her opportunity cost of capital.

In Figure 15.30 (a), the status quo before any redistribution, is shown as point **e**. As in Figure 15.29 (a), Ana would choose wage employment receiving a wage w^c on indifference curve u_2 rather than acting as an owner-operator at point **a** with risk and expected income $(\Delta_a, \hat{y}_a)$ on u_1.

After the redistribution, as shown in Figure 15.30 (b), two things happen:

- *The risk-return schedule shifts up*: With increased wealth she has a lower opportunity cost of capital and therefore the risk-return schedule shifts upward.
- *Her risk aversion decreases*: People with greater wealth experience lower risk aversion (their indifference curves are flatter, more like Beata's).

After the redistribution, Ana therefore chooses point **b**, meaning a risk level of Δ_b with expected income $\hat{y}_b > \hat{y}_a$. At this risk level and corresponding expected income at point **b** she is better off (on indifference curve u_3) than being a wage worker where she received w^c at point **e** (on indifference curve u_2).

A transfer of wealth (by a government for example) that made Ana wealthier would therefore result in her preferring being an owner-operator rather than a wage worker. She would have no incentive to sell her capital good and become a wage worker.

The hypothetical redistribution of assets is a vehicle for exploring the interaction of credit constraints, risk aversion, and ownership. It is *not* a policy design. Design of actual policies of asset distribution would need to address the policy's administrative aspects as well as general equilibrium and long-term dynamic effects not considered here. For example, whether the once-poor would adopt savings and investment strategies which would preserve, enhance, or consume their assets would need to be considered.

! REMINDER In section 12.2 we introduced evidence that an unexpected increase in wealth—an inheritance or the increase in the value of one's home, for example—led people to take up self-employment rather than working at a fixed wage.

Also, we have focused on the relationship between a single owner and a single worker. But, due to economies of scale, an economically viable cooperative will typically employ a large number of workers. They would face the challenge of motivating team members to work hard even though the revenues of the co-op would be shared among all members.

CHECKPOINT 15.13 Redistribution and risk Explain how a redistribution of wealth (raising the wealth of workers) could make it attractive for a worker to own the capital goods with which she works, possibly as a member of a cooperative instead of being employed for wages, which she preferred when she was less wealthy.

15.15 APPLICATION: THE DUAL ECONOMY AND HISTORY'S HOCKEY STICKS

The name we have used for "the whole-economy model" is not really correct. What the model *includes* is the part of the economy that is made up of private firms–their owners, customers, and workers–and the unemployed. The model leaves out those working for governments or nonprofit organizations, independent producers who are neither employers nor workers, and the unpaid work done in families.

The informal sector in a dual economy

In many lower income economies a substantial fraction of all the work done is in what is called the **informal sector**, meaning work done independently of an employer (whether a private firm or a government), and without hiring workers. Examples of those working in the informal economy are farming families, small shopkeepers, and other small businesses relying on unpaid

INFORMAL SECTOR The informal sector is comprised of the economic activities of family farmers, small shopkeepers, and others who work independently of an employer—whether a private firm or a government—and without hiring workers.

family members for work but not regularly hiring workers. We provided examples of informal sector work in India in section 6.9.

As these examples suggest the informal economy is a prominent characteristic of less developed economies where farming is a major form of production. But informal sectors exist in all economies. In high-income economies, for example, much of what is termed gig work is part of the informal economy: driving for Uber or Lyft, those finding work through platforms such as Task Rabbit, and others paid by the task completed rather than by the hour. Gig work, however, is an insignificant fraction of the total amount of work done (in the US not more than 2 percent of the economically active population).

By contrast, the informal sector in India makes up something like a half of the entire economy.[34] (Remember, by the conventional definition, this excludes housework and care work done at home.) India is not unusual among the lower-income countries in what is called its **dual economy**, that is, an economy in which employment done outside the home and not for government is of two kinds: work in the informal sector and employment for wages and salaries in what is called the capitalist sector.

The institutions of the informal economy share with the rest of the private economy two aspects of the rules of the game:

- *markets*: people buy and sell goods and services on markets; and
- *private property*: what they buy and sell, and the tools, land and other capital goods they use in production are privately owned.

But the informal economy differs from the capitalist and governmental sectors of the economy in an important way:

- *Self-employment.* Those working in the informal sector are neither employers nor workers and they are not compensated by wages or salaries.

In the informal sector, incomes take forms other than wages and salaries. Included are revenues above costs of production from producing crops, profits associated with buying goods at wholesale prices and selling them at a profit, or pay for some service provided.

In India and other low-income economies, the informal sector is for the most part poor due to the relatively low productivity of the technologies,

✓ **FACT CHECK** In the higher-income countries governments employ between roughly 15 and 30 percent of the workforce. From Figure 3.1 you can also see that the in the US, time spent on work done in the home—"Housework" and "Care work"—is a substantial fraction of the time spent in paid work even for men. More that one-third of the working time of women—work at home plus paid work—is at home. Caring for one's own family members and other family-related work at home is not considered to be part of the informal economy. When these kinds of work in the home are added to the hours worked in the informal economy and work for governments we see that the total hours worked that is not in the capitalist portion of a modern high income economy is on a par with the hours worked in the capitalist economy. In most lower income countries it is much greater.

DUAL ECONOMY The dual economy is one in which work, except for housework and government employment, is of two kinds: work compensated by incomes other than wages or salaries in the informal sector and employment for wages and salaries in what is termed the capitalist sector.

limited use of capital goods, and inability to exploit economies of scale that are typical of the informal sector.

Modeling the dual economy

We can repurpose and extend the whole-economy model to include an informal sector. The capitalist sector of the economy is as before: the owners of private firms employ workers to produce goods that will be sold with the intention of making profit. So the rules of the game make it capitalist; but it is no longer the entire economy.

The new element in the model is the informal sector itself. People work independently (or in family groups) in the informal sector and they receive an income which we will call γ^I which is less than wages in the capitalist sector. Now think of a farming family typical of the informal sector, with one or two members working for wages in the capitalist sector. A worker in the capitalist sector who loses their job returns to working with their family in the informal sector receiving, like the others, the average income.

In this setup there is no unemployment. Every person is working in either the capitalist or the informal sector. And receiving the average income in the informal sector is the fallback option for employed workers in the capitalist sector.

The size of the capitalist sector as before, is described by the wage curve and the competition condition. These determine how many people (or hours of employment) are engaged in the capitalist economy. But the wage curve is no longer based on the fallback option of unemployment along with an unemployment benefit from the government. Instead the height of the wage curve is determined by the average productivity of work in the informal sector, γ^I. This is because work in the informal sector is the fallback option of those employed in the capitalist sector.

Figure 15.32 panel (a) introduces the new model. As before, intersection of the competition condition w^c and the wage curve $w(H)$ determines the level of employment in the capitalist sector. The remainder of the economically active population works in the informal sector. Notice that we have calibrated the new model so that those employed in the informal sector are about half of the total. This contrasts with the initial single-sector model in which the number who do not find employment is typically less than 15 percent.

The distance between the green wage curve and the dashed γ^I line is a measure of the loss suffered by a capitalist sector worker who loses her job and returns to the informal sector. And because we assume that the person works equally hard in the two sectors (so the disutility of labor is the same) this income difference is the employment rent (per period) enjoyed by the capitalist sector worker.

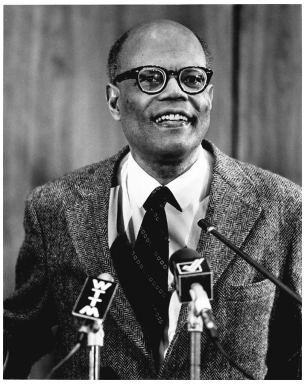


Figure 15.31 Arthur Lewis (1915–1991) won a Nobel Prize and knighthood from the British queen for his work in economics, but economics was not his field of choice. Born in the Caribbean island (and then British colony) of Saint Lucia, Lewis finished secondary school at age 14. As a young boy, required to remain at home for a couple of months due to illness, he learned so much studying by himself and being taught by his father that when he returned to school he was transferred from grade 4 to grade 6. "I wanted to be an engineer," he wrote, "but this seemed pointless since neither the government nor the white firms would employ a black engineer." At the London School of Economics he was a student of Fredrich Hayek.

Photo: Keystone Press/Alamy Stock Photo.

Behind the hockey sticks: The capitalist revolution in a dual economy

The hockey sticks in Figure 15.2 were propelled by a combination of the following:

- *technical change in the capitalist sector*, raising the level of productivity of labor in the capitalist sector of the economy; and
- *shrinkage of the informal sector* where (for the most part) labor productivity was lower, so a larger fraction of people were working at relatively higher productivity jobs.

This process is often described as the movement from farming into industrial production. But it can also be seen as a shift from working under one set of rules of the game—self-employment in the informal sector—to a different set of institutions—capitalism in the capitalist sector.

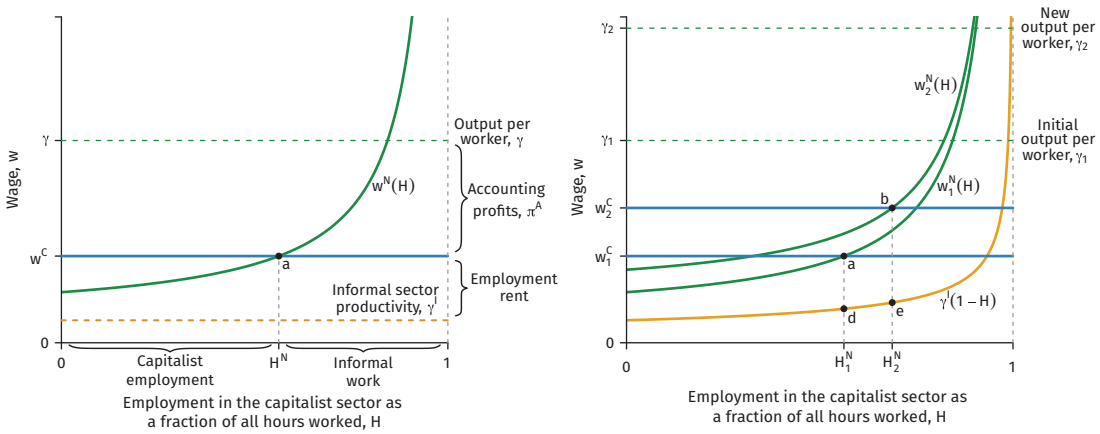
To see how this works turn to Figure 15.32 (b), where we illustrate the process of economic growth in a dual economy. We now take account of the fact that as the capitalist sector *grows* and the informal sector *shrinks*, then the average income in the informal sector may *increase*. Why would it increase? Because, for example, each person farming the land now on average has more land to farm, or a larger pool of customers to whom they can sell their services. This is shown by the upward-rising orange line.

To analyze the process of growth in the dual economy remember that we have explained why the rules of the game in the capitalist economy will promote innovation. Beginning in Figure 15.32 (a) with the economy at point **a**, here is how the dual economy may be transformed:

- *Innovation*. An advance in technology or the organization of production raises labor productivity in the capitalist sector from γ_1 to γ_2 (in panel b).
- *Increase in economic profits* in the capitalist sector. As long as the wage is unchanged the productivity increase raises profits.
- *Formal sector expands, informal sector shrinks*. The increase in economic profits leads new firms to enter the capitalist economy and existing firms to expand, drawing in additional labor from the informal sector.
- *Average productivity in the informal sector increases*. For example, the amount of land per farmer is now greater. The resulting increase in γ^I is shown by the movement from point **d** to point **e**.
- *Upward shift in the wage curve*. The effect of greater productivity and higher incomes in the informal sector is to raise the fallback option of the employed workers (in the capitalist sector). This means that to motivate capitalist sector workers to provide effort on the job, the wage offered must be higher.

HISTORY Arthur Lewis's 1954 paper introducing a dual economy with a modern "capitalist" sector and a "subsistence" sector is the inspiration for our model presented here. In that paper Lewis sought to return economic thinking to its roots, writing in the opening paragraph: "The classics, from Smith to Marx enquired how production grows through time [and] determined simultaneously income distribution and income growth, with the relative prices of commodities as a minor by-product."[35]

Figure 15.32 The dual economy. Panel (a) is a "snapshot" of the dual economy with employment in the capitalist economy determined by the intersection of the wage curve and the competition condition at point **a**. The height of the wage curve at point **a** is determined in part by the level of productivity in the informal sector when employment in the capitalist sector is H^N so that $1-H^N$ are working informally. Panel (b) shows the effect of an increase in the productivity of labor in the capitalist economy, expanding employment and raising the average productivity in the informal sector as fewer people are working there. The effect is the upward shift in the wage curve from $w_1^N(H)$ to $w_2^N(H)$ and the dual economy moving from point **a** to point **b**.



(a) Labor in the capitalist and informal sectors of a dual economy.

(b) Technical change and economic growth in a dual economy

- *Increase in the wage rate consistent with the competition condition.* The increase in productivity of workers in the capitalist sector (the change that initiated this process) means that the wage rate consistent with firms neither entering nor leaving the economy is now higher.

To explain the process we have listed the above changes as steps in a process. But other than the first step being innovation, the remaining steps do not take place in any order, they happen at once, possibly at varying speeds.

Where does the process end? In Figure 15.32 (b) we show the new Nash equilibrium of the dual economy at point **b**.

- *Informal workers*: Those who were initially in the informal sector have higher incomes for one of two reasons: they either moved to the capitalist sector where their wages exceed the average income in the informal sector, or they remained in the informal sector where average incomes have risen.

- *Workers in the capitalist sector*: They now have higher wages, sharing the increase in their productivity with their employers as before (assuming that there was no change in the degree of competition among firms).
- *Employers*: Profits are higher because the output of the capitalist sector is higher and the share of profits is unchanged.

The increase in profits provides savings (owners of firms consume just a small fraction of their income) that allow investment in further expansion of the capitalist sector.

CHECKPOINT 15.14 Stagnation and inequality in a dual economy Innovation in the capitalist sector could have effects quite different from the scenario sketched above. Reconstruct the above set of steps with a somewhat different narrative given by the alternatives below (consider them singly, not jointly).

a. *Change in the degree of competition:* the productivity increase was made possible by a new technology with strong economies of scale, so that the process of competition became a winner-take-all game and barriers to entry rose.

b. *A capitalist sector technology that competes with informal sector goods.* The new technology in the capitalist sector requires very little labor to produce large quantities of a good that was initially a major source of income in the informal sector.

Now return to Figures 15.1 and 15.2, and recall that from the second that real output per capita in Great Britain begins to take off before 1750 while real wages in London do not start climbing for another century. Could your responses to the alternative scenarios above provide a hypothesis about why this might have occurred?

15.16 CONCLUSION

With few exceptions (Cuba, North Korea; possibly China, Vietnam) and with many variants (US, Germany, Russia, South Korea) capitalism is the economic system of the world today. Knowing how the institutions of capitalism work–promoting innovation, sustaining substantial economic inequalities, transforming our biosphere, and its other consequences–is an essential starting point for modifying the rules of the game so that the economy better serves the needs and interests of all. In our final chapter we will use the understanding you have gained to explain how well-designed policies can promote this objective. We will also explore why policymaking and other governmental interventions are subject to limitation.

MAKING CONNECTIONS

Gains from exchange and conflicts over their distribution: The institutions of the capitalist economy both facilitate the exploitation of mutual gains based on innovation, specialization, and exchange and influence how these gains are shared among the people making up the economy.

Institutions of capitalism: Models of behavior under risk and models of the labor, credit, and other markets with incomplete contracts provide frameworks for understanding both the dynamism of capitalism and its characteristic forms of economic inequality.

Economic models and public policy: The models of risky decisions, labor discipline, and the whole-economy model provide ways of systematically studying the effects of policy interventions designed to raise income while reducing inequality and insecurity.

Risk and risk aversion: The wealthy and not-very-risk-averse owners of the capitalist firm make it an effective risk-taking "innovation machine." By reducing risk exposure, insurance in the form of taxes and transfers as well as more conventional forms of insurance can promote risk-taking among the less wealthy.

Economics as an empirical science: The Lorenz curve and the Gini coefficient provide measures of inequality allowing comparisons across countries and over time. Both highly unequal countries and countries with limited economic inequality are among the world's most innovative.

Efficiency: The success of capitalism in raising living standards (the hockey stick) is explained by the way that this economic system (when combined with the rule of law) promotes innovation in the long run, not by its success in implementing Pareto-efficient outcomes in any given period.

Incomplete contracts, social preferences, and power: The incomplete contracts that give rise to the principal–agent relationships characteristic of the capitalist economy—especially the employment of managers and workers—mean that both social norms and the exercise of power are important in determining economic outcomes, providing a political and social dimension to the economic system.

Comparison among institutions: A worker-owned and democratically managed cooperative is an alternative form of production providing advantages in motivating difficult to monitor work activities. But risk aversion and lack of access to credit due to the limited wealth of most workers make it unlikely that co-ops will proliferate in the absence of a significant equalization of wealth and new forms of insurance.

IMPORTANT IDEAS

cooperative	ownership	economies of scale
economic profits	accounting profits	insurance
coordination failure	specialization	credit-rationed
Gini coefficient	history's hockey stick	endogenous preferences
barriers to entry	Lorenz curve	democracy
monitoring	economic profit share	risk-sharing
risk-return schedule	risk	schedule
risk averse	expected income	inequality
wealth redistribution	risk neutral	short-side/long-side
short-side power	innovation	sanction
competition condition	residual claimancy	labor union
union voice effect	wage curve	unemployment benefit
wage share	collective bargaining	no-shirking condition
labor share	creative destruction	peer monitoring
capital share	incomplete contracts	dual economy
informal economy	control rights	residual claimancy

MATHEMATICAL NOTATION

Notation	Interpretation
K	value of total capital invested in a project
a_l	labor hours required to produce one unit $= \frac{1}{\gamma}$
γ	productivity of labor (output per hour) $= \frac{1}{a_l}$
x	units of output of a project
y	income
$\bar{y}$	average income
Δ	difference in income between good and bad states (risk)
$\bar{\Delta}$	amount of risk that maximizes the expected income
ρ	opportunity cost of capital
w	hourly real wage
w^c	hourly real wage consistent with the competition condition
w^N	hourly real wage consistent with the (no-shirking) wage curve
n	number of people and proportion of the population that is employed
G	Gini coefficient
b	probability of failure of an entering firm (barriers to entry)
p	price of a good
H	level of employment in the economy as a whole.
σ_w	wage share
σ_B	income share of the unemployed
u	number of people or proportion of the population that is unemployed
B	unemployment benefit
t	probability of being fired if shirking
j	probability that a "terminated" worker will not find a job
$\underline{u}$	disutility of effort

Note on superscripts and subscripts: E: economic; c: competition condition; A,B,C,D: different people. A "hat" (such as $\hat{y}$) means "expected."

CHAPTER

16

PUBLIC POLICY AND MECHANISM DESIGN

Mechanism-design theory aims to give the invisible hand a helping hand.

The Economist, explaining the Nobel Prizes in economics (2007)

DOING ECONOMICS

This chapter will enable you to:

- Explain how mechanism design can provide new rules of the game that will support Nash equilibria that are improvements over the status quo.
- See this process as an inversion of standard economic practice, reverse engineering the set of institutions that will achieve some desired social objective rather than predicting outcomes based on given rules of the game.
- Explain how in the case of public goods provision a mechanism designer accomplishes this task by internalizing uncompensated external effects and by other means.
- Understand the limits of mechanism design and why economics and policy makers cannot escape from the second-best world of imperfect solutions to societal problems.
- See that mechanism design is a modern variant of the approach of the eighteenth-century philosopher-economists who proposed institutions for societal coordination that would result in socially desirable outcomes.
- Write a blog post or an editorial for your university or local newspaper or video yourself doing a TED Talk using what you have learned to explain the pros or cons of some economic policy idea about which you are passionate.

16.1 INTRODUCTION: SEAT BELT SURPRISES

Figure 16.1 A mechanism designed by a watch-maker.

Photo: Josh Redd on Unsplash.

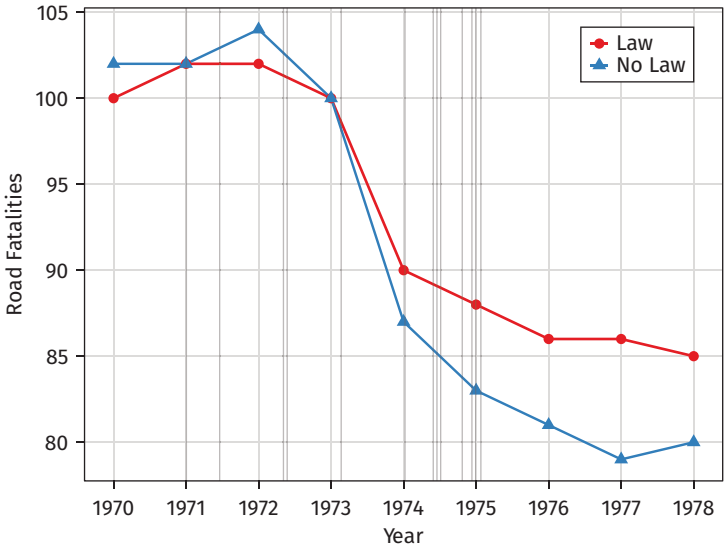
Seat belts in cars are now standard equipment. But they are a recent addition to the safety features of an automobile. The Australian state of Victoria was the first to introduce a mandatory seat belt law, in 1971. Over the next two decades, spurred by claims that seat belts would save thousands of lives per year ("10,000 to 20,000" in the US according to advocates), over 80 jurisdictions implemented similar laws, applying to the vast majority of automobiles in the world.

The claimed life-saving effects of seat belts were based on simulated crashes in which–to take one example–"for belted occupants the deaths were reduced by 77 percent in full frontal crashes and 91 percent in rollovers." In the UK, the Royal Society for the Prevention of Accidents summarized the evidence: "no other single practical piece of legislation could achieve such dramatic savings of lives and serious injuries." The truly global spread of mandatory seat belt laws seemed like a case of evidence-based public policy at its best. But was it?

In Figure 16.2, we look at the evidence on road fatalities in 17 countries during the 1970s, 13 of which passed mandatory seat belt legislation and

EXAMPLE In Germany buses with seat belts were allowed a top speed of 100km/h (60 m/h) while those without belts were restricted to 80km/h. Curiously the Royal Society for the Prevention of Accidents, whose advocacy of the seat belt laws in the UK was quoted above, seemed to endorse this not as a bug but as a feature of the safety devices that "allowed [buses] to travel faster . . . thus allowing drivers to cover more miles in the hours they are allowed."[2]

Figure 16.2 Seat belts and road accident deaths. Shown are indices of the number of road traffic deaths in 17 countries, with a value of 100 for 1973 in the case with seat belt laws. The average of the indices for the 13 countries that passed seat belt laws are shown by the red line, the vertical lines representing the dates at which each of these countries introduced (and enforced) a seat belt law. The blue line represents the average indices for the four countries which did not impose mandatory use of seat belts.[1]



the other four of which did not. At the time the legislation was passed, together these 17 countries accounted for 80 percent of the world's cars. In almost all of the countries road deaths fell, in part due to the substantial increase in the price of gas, and the reduction in both legally permitted and actual speeds of driving. But the drop in fatalities was much greater in the countries that had *not* passed the seat belt laws.

How could this have occurred?

The answer, it seems, is that people drive faster when they are wearing seat belts. The result was that accidents are more frequent, and while the occupants of cars are more likely to survive an accident if one occurs, the greater frequency of accidents results in greater fatalities including among pedestrians and non-occupants. In some cases the compensating effect of the seat belts–driving faster–was even enacted into law.

An experiment provided evidence consistent with the increased driving speed explanation. In the Netherlands, before belt use was mandatory, a group of people who had never used seat belts participated in an experiment. The drivers were randomly selected to either wear seat belts or not. They were told that the purpose was to judge the comfort of seat belts, but in fact the experimenter measured the speeds at which they covered a specific 105 km course. Drivers wearing seat belts drove faster.

! REMINDER A "mechanism" in ordinary language is a way of getting something done; in economics it refers to a set of rules of the game—possibly designed deliberately—that implements some allocation as an outcome of an economic interaction.

16.2 MECHANISM DESIGN: THE CLASSICAL INSTITUTIONAL CHALLENGE, 2.0

We do not conclude that seat belt laws should be abolished: mandatory seat belts and enforced moderate speed limits surely reduce road fatalities. But the surprising results of the introduction of seat belt laws teach an important lesson: effective policy design should take account of the diverse *unintended effects* of the policy (in this case, including greater speed), not simply the intended effects (in this case, fewer fatalities among car occupants in a crash).

Adam Smith wrote (quoted in the introduction of Part I of this book) that the policymaker:

> *The man . . . enamored of his own ideal plan of government, . . . seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board . . . but . . . in the great chess-board of human society, each single piece [on the chess board of society] has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it.*

Today the field of mechanism design is carrying the tradition initiated by Smith and the other founders of economics: addressing the institutional challenge of developing rules of the game under which a free people can best coordinate their activities. The unexpected consequences of seat belt

› EXAMPLE A mechanism for the fair division of a cake: Suppose two people are going to share a small cake. What would be a good rule for making sure that the cake is divided fairly? A mechanism that is widely used is to let one person cut the cake and the other person choose which piece to take. Because the person cutting the cake will get the second piece, that is, the piece not chosen by the other, she will have an incentive to try to make both pieces be equally desirable.

! REMINDER Recall that if one party can set both the price and the quantity transacted, then this person—with take-it-or-leave-it power—is constrained by the other person's participation constraint, so the resulting allocation must be Pareto efficient.

Figure 16.3 Eric Maskin (1950–) went into the field of mechanism design because it was "an irresistible combination: … the precision, rigor, and sometimes the beauty of pure mathematics [that] … also addressed problems of real social importance."[3] His research on intellectual property rights suggests that "society and even inventors themselves may be better off without [patent] protection" in the software industry.[4] Along with another Nobel Prize winner, Amartya Sen, he has proposed ranked-choice voting and changes in the rules for determining winners in elections so as to better reflect the intensity and diversity of citizen's political preferences.

Photo: NICHOLAS KAMM/AFP via Getty Images.

laws and Adam Smith's warning illustrate some challenges in designing public policy. They also provide guidelines for effective policy interventions, the subject of this final chapter. We can group government activities by their intended purposes:

- *Providing the economic framework* that governs how people and economic organizations interact, including the judicial system that interprets and enforces contracts and property rights, policies regulating how firms and other economic entities compete and the conditions under which people are employed, and reducing uncertainty about the level of aggregate demand and the value of the currency.
- *Addressing market failures* including through the provision of public goods (including schooling and basic research), competition policies, and internalizing the harmful external effects of economic activity on the biosphere.
- *Addressing unfairness* in the distribution of income, wealth, or some other valued aspect of our living standards that arises due to the working of the economic framework adopted, including through taxes and transfers, direct government provision of some services (schooling, fire and police protection) and setting prices (rent control, minimum wages).

Thinking back to the allocation of goods between Ayanda and Biko in Chapter 4, the above "providing the economic framework heading" would include determining why Ayanda—not Biko—was the first mover, why Biko could not simply take Ayanda's goods by force, and whether, when Ayanda set a high price for her good, Biko had any other ways of acquiring the good (from a competing supplier, for example).

In that example, "addressing market failure" could take the form of making it easier for Ayanda to impose both a price and a quantity on Biko, giving her "take-it-or-leave-it power" rather than just price-setting power. If, as a result, Ayanda's advantages were deemed to be unfair, then "addressing unfairness" could take the form of the government imposing a maximum price that Ayanda could charge, or possibly redistributing some of Ayanda's endowment to Biko.

Reverse-engineering good outcomes

Mechanism design is a branch of economics that seeks to design policies, decision-making protocols (like majority rule), laws, property rights, and other so-called mechanisms that will implement outcomes that are judged to be desirable.

The primary focus of the field has been to provide novel rules of the game that will result in Pareto-efficient Nash equilibria in cases where market exchange or bargaining among private parties fail to accomplish this result. In previous chapters we have followed the common practice in economics, namely, start with the rules of the game for some economic interaction

and a description of people's objectives and knowledge, and then figure out what allocation will result. In other words, take the following steps:

- Start with a set of rules of the game or some other description of the institutional setting in which an interaction will take place, for example, the principal-agent game describing an employer and worker interacting about the wage and level of effort performed by the worker.
- Describe people's objectives (by their utility functions), constraints (by the set of feasible actions or strategies open to them), and beliefs.
- Use concepts like best response and Nash equilibrium to determine the outcome that will result from the interaction.

Mechanism design reverses these steps, starting with a desired outcome and then working back to see what rules of the game would bring that outcome about, given people's preferences and beliefs. So the steps above are taken up in reverse order:

- Start by defining a desired outcome, for example an allocation of goods among members of a population, or their use of some environmental resource that is an improvement over the status quo.
- Then devise a set of rules of the game that will lead people pursuing their own private objectives to implement the desired outcome as a Nash equilibrium.

The logic of mechanism design is simple: if the status quo is a market failure, or some other coordination failure, there must be some other allocation that is Pareto superior to the status quo. Or there may be some other feasible allocation that is preferable on grounds of fairness.

The task of mechanism design is then to reverse-engineer that desired allocation. This requires finding a set of rules of the game under which the desired allocation will be a Nash equilibrium, and therefore could be implemented by introducing the mechanism (the new rules of the game) she has discovered. The creation of a new superior equilibrium by a change in the rules of the game is called implementation by Nash equilibrium.

The policymaker who practices implementation by Nash equilibrium is respecting Adam Smith's dictum that the government cannot simply order people how to act (like moving chess pieces around on the chess board). Instead the government can alter economic outcomes by changing the circumstances under which people themselves decide what to do.

You have studied another case of implementation by Nash equilibrium in Chapter 5 where the policymaker designed a tax that would deter overharvesting fish from a lake. As in the Prisoners' Dilemma game, the fundamental problem was that actors did not take account of the effect of their decisions on others. In the case of fishing, this so-called external effect

› EXAMPLE King Solomon and Buddha as mechanism designers: King Solomon in the Old Testament (the Hebrew Bible) was asked to determine which of two women should be awarded a young boy, whom both claimed was their child. He asked for a sword and said he would award each woman half of the infant. The first woman who spoke agreed with the proposal. But the second, in tears, objected and told Solomon to stop, and to give the child to the first woman. Solomon then knew that the first woman, who agreed to dividing the child, could not possibly be the true mother, and awarded the child to the second woman. In an almost identical story from India, Buddha makes a similar ruling (the two women would have a tug of war over the child). Solomon and Buddha are wise mechanism designers in these stories, implementing a good outcome (the child goes to the true mother) by devising a way to get the necessary information to do this (who is the true mother).

EXAMPLE
Sixteenth-century mechanism designers: Under the Mughal rulers of much of the Indian sub-continent between the sixteenth and nineteenth centuries tax authorities sought to avoid taxing crops as this would *reduce* the farmers' incentives to work hard caring for his crops. But taxing based on the value of the land (rather than the crops grown) was infeasible because the crops *actually* grown depended on rainfall and other things that farmers did not control. So a tax based on the value of the land, to be paid in a drought year would have imposed serious hardship on farmers. An ingenious solution adopted by the Mughal state was to level taxes based on the crops raised on neighboring fields, not the farmers' own field; these would reflect rainfall and other growing conditions, but the tax would not reduce the farmers' incentive to work, as the tax he paid would be unrelated to the size of his own crop.[5]

occurred because each person fishing more meant that the other caught fewer fish. The policy that addressed this problem was the obligation to pay a tax equal to the costs that each of their fishing time imposed on the other.

This is called "internalizing the external effects" of a person's actions, and it is the key idea of mechanism design. This is accomplished *not* by changing people's preferences, but instead by devising rules so that each person pays for the social costs imposed on others or is rewarded for the benefits to others resulting from his actions. Internalization therefore accounts both for the *private* costs and benefit and for the *external* costs and benefits.

Mechanism design has come into prominence in recent decades for two reasons.

- First, economists now recognize that the simple institutions that twentieth-century economists used to illustrate Adam Smith's invisible hand idea—perfect competition and well-defined private property rights in anything that matters—represent ideals that are hardly ever realized in existing economies. The invisible hand as *The Economist* pointed out, needs "a helping hand."[6] Where markets fail, mechanism design provides this helping hand in suggesting more complex institutions including auctioning a limited number of permits for carbon emissions to address the challenge of climate change and designing kidney exchanges to match organ donors and those needing a kidney replacement.[7]
- Second, with the growth of government's role in the economy the unintended consequences of policy interventions—often ignored by policy advocates—have become increasingly evident. A common error in these policy designs is to ignore a fundamental precept of mechanism design. Policymakers must make sure that the desired outcome of some policy will be a Nash equilibrium once the policy is introduced. If this is not the case, the intended effects of the policy will be undone by the actions of private actors.

CHECKPOINT 16.1 **Adam Smith: mechanism designer** Explain how King Solomon (in the margin note above) or Adam Smith (in the quote above) fits the description of the mechanism designer.

16.3 OPTIMAL CONTRACTS: INTERNALIZING EXTERNAL EFFECTS OF PUBLIC GOODS

Designing a policy—that is, a mechanism—to address the problem of private under-provision of public goods illustrates the method. Examples of public goods include the knowledge generated by basic research, weather reports, or other information broadcast on an open access platform, and public safety. Remember: if a good is public, then what any *one* gets *everyone* gets.

Private provision of public goods: A coordination failure

As we showed in Chapter 5, in the absence of a subsidy or other public policy, private economic actors will typically *not* provide public goods at all, or will provide them in insufficient quantity. Under-provision means that there is some greater level of provision of the public good—for example with each person contributing more—such that all citizens would be better off.

The reason that public goods are under-provided is that they are costly to provide, but the benefits that result from any person's contribution to the public good are shared by everyone:

- There are private costs and private benefits that a decision maker will take account of in choosing to contribute to a public good.
- But there are also external benefits, so the social (total) benefits exceed the private benefits.
- And unless contributions are subsidized, the person is *not compensated* for the external benefits that her contributions create.

Here is a specific example of a public good and how it benefits a particular citizen, call her Bridget, who is one of n identical citizens. The amount of the public good available to Bridget and every other citizen is equal to a positive constant (γ) multiplied by the sum of all contributions or:

$$\text{Total public good provided} = \text{A} = \gamma \cdot (a^1 + a^2 + a^{\text{B}} + \ldots + a^n) \tag{16.1}$$

where a^{B} is Bridget's contribution and each other a is the contribution of the $n-1$ other citizens.

We can distinguish among these three essential concepts about the benefits of the public good.

- *Marginal private benefit* of contributing: γ. From Equation 16.1 we can see that if Bridget contributes (a^{B}), then, compared to the case in which she does not contribute, the amount of the public good that she enjoys will increase by an amount γ times the increase in her contribution. This is her *marginal private benefit*.
- *Total marginal external benefits* of contributing, $\gamma(n-1)$: From the fact that consumption of a public good is non-rival, we know that each person who is *not* Bridget will (like Bridget) receive a benefit γa^{B} from Bridget contributing the amount a^{B}. As there are $n-1$ other people, the marginal external benefit that Bridget confers by contributing is the marginal external benefit multiplied by the number of people who receive it, that is, $\gamma(n-1)$. This the *marginal external benefit* of her contribution.
- *Marginal social benefits of contributing* are the marginal private benefits plus the marginal external benefits $\gamma + \gamma(n-1) = n\gamma$.

✓ **FACT CHECK** San Francisco's rent control law passed in 1995 is estimated to have reduced citywide housing supply by 15 percent and raised average rents by 5 percent. This is an unintended consequence, similar to the fact that the introduction of seat belts led some people to drive faster. We illustrated this case in Figure 8.22 in Chapter 8.[8]

! REMINDER You know from Chapters 2 and 5 that a public good is one that is both *non-excludable* and *non-rival*, meaning that it is impossible or very costly to exclude anyone from access to a public good, and the consumption of the good by one citizen does not reduce the amount available to others. Return to Table 2.3 if you need to refresh your memory about what is distinctive about public goods.

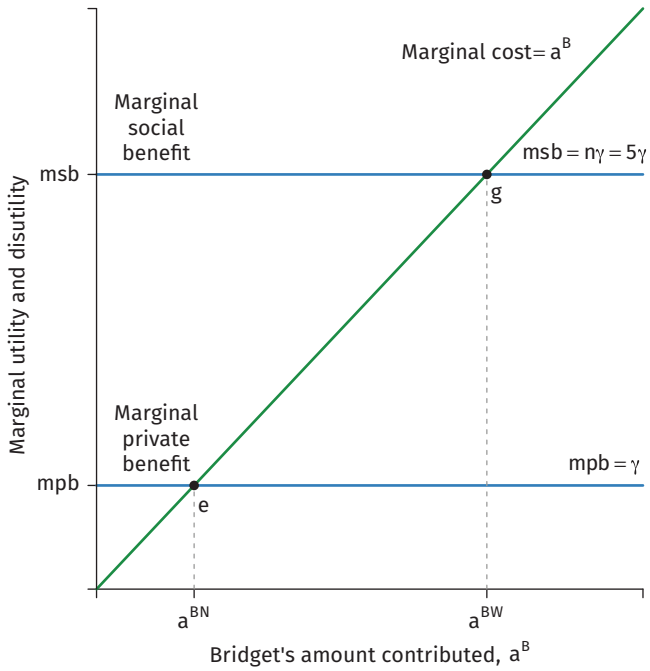
! REMINDER In Chapters 5 and 14 we highlighted how marginal social costs include both the marginal private costs and the external costs. The private costs were the costs to the individual person, fishermen, or firm, the external costs were the costs imposed on others that the person did not pay for. The marginal social costs were the sum of these two costs. Marginal private benefits, external benefits, and marginal social benefits can be thought of in a similar way.

The final essential concept is the marginal cost of the effort, time, or money that contributing requires. In M-Note 16.1 we provide an equation showing how the marginal disutility to Bridget of contributing increases as the size of the contribution increases. So, if she is contributing nothing, then contributing a little bit does not impose much disutility on her. If she is already contributing a lot, then contributing more is very costly. We have chosen a utility function such that the marginal cost (disutility) of contributing is a^B itself. So we have:

- *Marginal private cost of contributing*, a^B.

To see why a coordination failure results we will study Bridget, and her contribution a^B, as illustrated in Figure 16.4. The marginal private benefit

Figure 16.4 The marginal benefits and marginal costs of contributing to a public good. Bridget will contribute up to such a point as her marginal costs, measured by a^B, equal her marginal private benefits ($mpb = \gamma$). Therefore, her Nash equilibrium contribution is a^{BN}. If there are $n-1$ other citizens, then the socially optimal outcome occurs where Bridget's marginal private cost a^{BW} equals the marginal social benefit $msb = n\gamma$, which includes the benefit that Bridget confers on the four other citizens as a consequence of the positive external effect of her contributing to the public good. The Nash equilibrium contribution is lower than the socially optimal contribution. In the figure, $n = 5$.



from her contributing is the lower horizontal line. The private marginal cost is the upward-sloping green line.

She considers contributing nothing. But then the marginal private benefit should she contribute some small amount would exceed the cost. Similarly, were she to contribute a large amount the marginal costs would far exceed the marginal private benefits. So she will contribute up to the point that the marginal private costs equal the marginal private benefits (this is shown in M-Note 16.1):

$$\text{Marginal cost } = a^{\mathrm{B}} = \gamma = \text{ Marginal private benefit} \tag{16.2}$$

All of the other $n-1$ members of the population, being identical to Bridget and facing the same incentives, also contribute the same amount, which we will now call simply a^{N}, the amount contributed by each in the Nash equilibrium without any subsidy. So the total amount contributed is na^{N}, and the total amount of the public good provided to each citizen is $na^{\mathrm{N}}\gamma$. And because it is a public good, this amount is enjoyed by all n citizens. So the total utility experienced by the whole population is $n(na^{\mathrm{N}}\gamma)$.

M-CHECK $a^{\mathrm{i}} = a^{\mathrm{B}} = \gamma$ is the Nash equilibrium of the game, and it is also the dominant strategy equilibrium. Each citizen is doing the best they can and this is independent of what the other citizens are doing. This is because there is a dominant strategy—to contribute γ—that is not affected by the actions of others. You can confirm this by noticing that the condition for the choice of the level of contribution (Equation 16.2) does not involve any terms about the contributions of others.

M-NOTE 16.1 **Public goods: a coordination failure**

From Equation 16.1, the amount of the public good available to Bridget and every other citizen is equal to $\gamma \cdot (a^1 + a^2 + \ldots + a^{\mathrm{B}} + \ldots + a^n)$. Bridget incurs a cost (in disutility) of $(a^{\mathrm{B}})^2/2$ when she contributes an amount a^{B} to the public good.

So Bridget's utility is:

$$\text{B's utility function} \qquad u^{\mathrm{B}} = \gamma(a^1 + a^2 + \ldots a^{\mathrm{B}} + \ldots + a^n) - \frac{(a^{\mathrm{B}})^2}{2} \tag{16.3}$$

To find how much Bridget will contribute we differentiate Equation 16.3 with respect to a^{B}, and set the result equal to zero (finding the first-order condition that defines the maximum private benefit minus private cost):

$$\text{Using the notation } u^{\mathrm{B}}_{a^{\mathrm{B}}} \equiv \frac{\partial u^{\mathrm{B}}}{\partial a^{\mathrm{B}}}$$

From the first-order condition we have:

$$u^{\mathrm{B}}_{a^{\mathrm{B}}} = \gamma - a^{\mathrm{B}} = 0$$
$$\therefore\ \gamma = a^{\mathrm{B}}$$
$$\text{Marginal private benefit} \quad = \quad \text{Marginal private cost}$$

Therefore, Bridget's Nash equilibrium level of contribution is $a^{\mathrm{BN}} = \gamma$ (We add the superscript N for Nash equilibrium as in earlier chapters).

To see why Bridget and other citizens contributing this amount is a coordination failure think about the effect that Bridget's contributing just a little more would have on Bridget's utility and the utility of the $n-1$ other citizens. We know from the first-order condition used to determine her level of contribution that:

$$\text{Effect on her own utility} \qquad u^{\mathrm{B}}_{a^{\mathrm{B}}} = 0$$
$$\text{but the total effect on the } n-1 \text{ others' utility is:} \ = \gamma(n-1)$$

continued

! REMINDER Remember the Auctioneer who regulates how markets work in the general equilibrium model in chapter 14? We also called him imaginary because there is no such entity. The Auctioneer is a device for helping us think clearly about how markets might ideally work. We call the Mechanism Designer imaginary to stress that she is not intended to represent how governments or public policymakers really work, but instead to convey in simple terms some aspects of how they ideally might intervene to address coordination problems. Don't confuse the Mechanism Designer with the equally hypothetical Impartial Spectator, who provides an "above-the-fray" viewpoint on what a fair allocation might be. The Mechanism Designer has the more down-to-earth job of designing rules of the game to implement good outcomes.

So Bridget's utility would be virtually unaffected if she contributed a small amount more; but her contribution and the increase in the public good that results would add an amount γ to the utility of each of the $n-1$ other citizens. If everyone could agree to contribute a little more, everyone would be better off. Therefore the Nash equilibrium is not Pareto efficient, and the result is a coordination failure.

16.4 ENTER, THE MECHANISM DESIGNER: AN OPTIMAL SUBSIDY

We now introduce an imaginary actor: the Mechanism Designer, charged by the citizenry with the task of devising a change in the rules of the game that will address the under-provision of public goods. The Mechanism Designer, we will assume, is committed to treating each citizen equally and regards the utility of each as comparable and equally worthy of being maximized. So the Mechanism Designer will attempt to maximize the sum of the utilities of the members of the population.

The Mechanism Designer's task is:

- *Objective—The desired outcome:* Determine the level of the public good that, if implemented, would maximize the *sum* of the utilities of the citizens. This is the *socially optimal level* of public goods provision.
- *Implementation by a mechanism that makes the desired outcome a Nash equilibrium:* Devise a policy that will alter the citizens' utility functions—adding a subsidy in this case—so that the citizens acting privately will have a sufficient incentive to contribute the desired amount.

Notice that, as we said above, the Mechanism Designer *starts* with the desired level of contributions and then works backward to discover a mechanism—new rules of the game—that make the desired outcome a Nash equilibrium.

To determine the socially optimal level of public goods provision, we simplify the Mechanism Designer's task by assuming that citizens are identical so the result she implements will be the same for everyone. Choosing a given level that every citizen will contribute to the public good in order to maximize the utility of any particular citizen—say, Bridget—is therefore the same thing as choosing a level for *every* citizen to contribute that will maximize the total utility of *all* citizens. So we can simplify by just considering the level of contribution that maximizes the utility of one citizen, Bridget.

The optimal total amount of the public good provided can be determined by following these three rules (return to Figure 16.4). Consider a particular level of Bridget's (and everyone else's) contribution and the level of marginal social benefits and marginal private costs when contributing this amount. There are three possible cases:

- $n\gamma > a^B$: if the marginal social benefits are *greater* than the marginal private costs of contributing, *increase* the contributions (the effect on increasing the benefits will exceed the effect on increasing costs);
- $n\gamma < a^B$: if the marginal social benefit of contributing is *less* than the marginal cost, then *reduce* the contribution level;
- $n\gamma = a^B$: if the marginal social benefit is equal to the marginal cost, then do *not* change anything: this is the socially optimal level of contributions.

M-Note 16.2 explains this result further.

M-NOTE 16.2 The optimal contribution to the public good

To maximize the sum of the utilities the Mechanism Designer would choose a level that every citizen would contribute, $a^1 = \cdots = a^n = a$ and then design a mechanism to make these contributions a best response by each of the citizens, each maximizing their own utility independently.

Because citizens are identical we take Bridget as a representative citizen and determine what level of contribution would maximize her utility u^B assuming that the $n-1$ other citizens did the same so the total contributed will be na.

Therefore, the Mechanism Designer's problem to find the desired outcome is to vary a to maximize each person's utility:

$$\max_a u^B = \gamma(na) - \frac{(a)^2}{2} \tag{16.4}$$

Differentiating Equation 16.4 with respect to a and setting the result equal to zero, we find that the socially optimal contribution $a = a^I$ is:

$$\frac{\partial u^B}{\partial a} = \gamma n - a = 0 \Longrightarrow a = a^I = n\gamma \tag{16.5}$$

Thus we see that the socially optimal level is that each citizen contributes $n\gamma$, not just γ which is what the individual maximizing her own utility does in the *absence* of the introduction of the mechanism. We have used the superscript "I" because we used I for the "impartial spectator" previously and here the Mechanism Designer is pursuing a similar goal to that which the Impartial Spectator did: the maximization of social welfare.

! REMINDER You have seen this idea of internalizing external effects in action right from Chapter 1 where you learned how public policy could convert a Prisoners' Dilemma game, with its undesirable mutual defect outcome, to an assurance game with a Pareto-superior Nash equilibrium. The 'mechanism' was a liability law that required the player to pay for the external costs that she imposed on another person, in this case the costs imposed on a cooperator by defecting on them.

CHECKPOINT 16.2 Population size and optimal public goods provision

Explain the economic reasons why the socially optimal level of the public goods provision is larger if the population is larger.

! REMINDER In Chapter 13 we introduced costs of taxation and redistribution: some of the taxes collected "leak out" and do not get spent on their targeted purposes. Here for simplicity we drop this realistic feature: there are no leaks in the bucket.

A subsidy to internalize the external benefits

Now that the Mechanism Designer has found Bridget's—and every citizen's—socially optimal level of contribution to the public good, she will determine what mechanism could implement their levels of contribution as a Nash equilibrium. The Mechanism Designer needs to diagnose the coordination failure and to propose a remedy:

! REMINDER The subsidy proposed here is like the tax on fishing time to prevent overharvesting fish stocks in the lake (studied in Chapter 5), but instead of internalizing the external effect by making those who impose costs on others *pay* those costs, it *rewards* those who confer benefits on others.

- *Diagnosis*: The under-provision of the public good by private actors occurs because contributing is costly to the contributor and produces external benefits for which the contributor is *not* compensated.
- *Solution*: Find a way to internalize these external effects by *compensating* each citizen for the benefits that their contribution confers on *others*.

We saw in M-Note 16.1 that the external benefits conferred by Bridget contributing a^B are $a^B\gamma(n-1)$. Therefore, to compensate for the otherwise *uncompensated* external benefits, the subsidy per unit of contribution should be $\gamma(n-1)$. M-Note 16.3 shows how the Mechanism Designer determines this amount.

To determine the socially optimal subsidy the Mechanism Designer does *not* take account of the subsidy each person received as a contribution to each citizen's utility. The reason is that for every subsidy received, some citizen has paid an equivalent tax (to provide the government revenues for the subsidy). The subsidy is a *transfer* among citizens, it has no effect on the total utility of the citizenry. Each citizen will pay the given tax however much they decide to contribute in response to the subsidy incentive. In equilibrium they will all contribute the same, so the subsidy received will equal the tax paid.

The subsidy addresses private under-provision of the public good by shifting the Nash equilibrium to a Pareto-efficient outcome. The public good example, however, is a particularly simple case: the actions taken by each citizen do not depend on what the other citizens did. Most problems of mechanism design are complicated by the fact that how one person responds to the policy depends on what everyone else is doing, including how they respond to the policy, as was the case of overfishing the lake. We therefore now turn to how more complex and realistic social interactions affect the results of public policies.

M-NOTE 16.3 **The Mechanism Designer's socially optimal subsidy**

To find the optimal subsidy we modify Bridget's utility function to include the subsidy of ω (the Greek letter "omega") for each unit of her contribution a^B (the last term on the right):

$$\text{Subsidized utility} \quad u^B = \gamma \cdot (a^1 + a^2 + \ldots + a^B + \ldots + a^n) - \frac{(a^B)^2}{2} + \omega a^B \quad (16.6)$$

Therefore her modified utility function is: Utility = (Private benefits from the public good) − (private costs of contributing a^B) + (subsidy for contributing a^B).

We can now find Bridget's marginal utility with a subsidy, again by differentiating her utility function with respect to her amount contributed and setting the result equal to zero. This gives us the first-order condition for Bridget's contribution level when the subsidy is introduced:

continued

$$u_{a^B} \equiv \frac{\partial u^B}{\partial a^B} = \gamma - a^B + \omega = 0 \tag{16.7}$$

$$\Longrightarrow \gamma - a^B + \omega = 0$$

$$\therefore \gamma + \omega = a^B \tag{16.8}$$

$$\text{Private and subsidized marginal benefit} = \text{Marginal cost}$$

Equation 16.8 is Bridget's best-response function, giving the amount she will contribute for any value of the subsidy that the Mechanism Designer might choose. The Mechanism Designer had already calculated the optimal contribution (Equation 16.5): $a^N = \gamma n$. Therefore, to find the level of subsidy that will implement the socially optimal contribution, we equate Bridget's best-response function to the socially optimal level of contribution:

$$a^N = \gamma + \omega = \gamma n = a^I \tag{16.9}$$

The Mechanism Designer solved Equation 16.9 for ω to find the amount of subsidy that would implement the social optimum:

$$\omega = \gamma n - \gamma$$

$$\omega = \gamma(n-1)$$

The subsidy that implements the optimal level of public goods contributions exactly compensates each contributor for the external benefits that their contributions generate, that is γ for each of the other citizens, of which there are $n-1$.

CHECKPOINT 16.3 Internalizing external benefits

a. Explain in your own words why the subsidy $\omega = \gamma(n-1)$ internalizes the external benefits of each citizen's contribution.

b. If the population grew to $2n$, double its current size, explain why the optimal subsidy would also increase.

c. Would the optimal subsidy double with the larger population?

16.5 THE SOCIAL MULTIPLIER EFFECTS OF PUBLIC POLICIES

To see how the Mechanism Designer works in this more realistic environment, consider the case of what are called sin taxes, that is, taxes intended to raise the cost of what are widely considered to be bad habits, like excessive drinking of alcoholic beverages or smoking. The key fact here is that drinking and smoking are social activities, so the enjoyment of the activity depends in part on smoking or drinking with others.

The fact that consumption is a social activity was illustrated in Chapter 7 by a model and some evidence about how a person might engage in "conspicuous consumption" as a social signal of status, earning power, or respectability. We return to the social nature of consumption here–smoking with friends is more enjoyable than alone–and see how this

changes the effects of policies designed to discourage this unhealthy form of consumption.

We will suppose, for example that smokers currently smoke a pack (20 cigarettes) a day and the Mechanism Designer has established a target of reducing this to just one cigarette a day. She could of course recommend persuasion campaigns to reduce people's desire to smoke, or ban smoking in public places, making the social enjoyment of smoking less. But let's ask: If taxing cigarettes is the only policy under consideration, how large a tax would be required to get smokers to limit themselves to one cigarette a day?

The tax changes social outcomes by altering the environment in which people decide what to do, that is, how much to smoke, now that it's more expensive and partly as a result, others are smoking less. Our approach will be the same as in the case of public goods: the Mechanism Designer will find a way to implement an outcome by replacing the status quo Nash equilibrium (20 cigarettes a day) by a new Nash equilibrium (one cigarette a day).

We show how taxes can affect behavior through two channels:

- a *direct* channel through raising prices and as a result, the marginal costs of the action, smoking; and
- an *indirect* channel through changing other people's behavior (which affects the marginal *benefits* of the action).

The Mechanism Designer will pay attention to both of these effects.

✓ **FACT CHECK** Econometric estimates suggest that in other forms of "social consumption" the social multiplier may be large. A study of "heavy drinking" by Russian men, for example, estimated that the effect of a 50 percent permanent increase in the price of vodka would reduce heavy drinking by about 30 percent. The social multiplier—that is the indirect effect of reduced peers' drinking on one's own alcohol consumption—accounts for one-third of this effect.[9]

The social multiplier

Suppose we want to determine the effect of a cigarette tax on the amount of smoking that people do. We know that the amount of smoking will be reduced by both a higher price of cigarettes and by other people smoking less.

The total effect of a price increase on a person's smoking will therefore be greater than if the price increase were experienced (hypothetically) only by a single person. The reason is the additional indirect effect from the tax directly reducing how much other people smoke, and this indirectly reducing the pleasure of smoking. The total effect of the tax is the sum of the direct (cigarettes cost more) effect and the indirect (smoking is less pleasurable because fewer are smoking) effect. The indirect effect is due to what is called the **social multiplier**. As a quantitative magnitude, the social multiplier is the difference between the direct and total effects.

SOCIAL MULTIPLIER When there are indirect effects of a policy through its effects on other people's behavior the presence of a social multiplier means that the total effect of the policy will differ from the direct effect (hypothetically holding constant the behavior of others). As a quantitative magnitude, the social multiplier is the difference between the direct and total effects.

Figure 16.5 The social multiplier of a tax on cigarettes. The figure shows the first-round effects of a tax on cigarettes smoking both directly (by making smoking more expensive) and via the social multiplier (by making smoking less enjoyable, because there will be fewer people to smoke with). The negative signs show a negative relationship between taxes and smoking: as taxes increase, smoking decreases; as taxes decrease, smoking increases. The positive sign shows that there is a positive relationship between others' smoking and your own smoking: so if other people smoke more, you smoke more, if other people smoke less, you smoke less. What is shown is just the first-round effect: in the next and later rounds, *your* smoking is now less than before, so now others will find smoking less enjoyable and reduce *their* level of smoking, and so on.

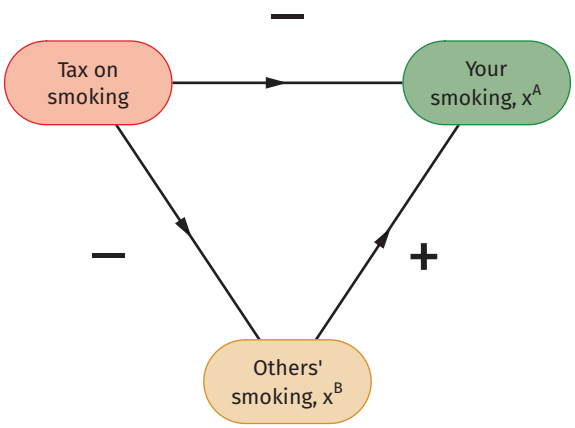


Figure 16.5 illustrates the:

- *direct effect*: the top horizontal arrow; and
- *indirect effect*: the two arrows to and from the others' smoking.

It does not show the *total effect* including the follow-on rounds of effects, your smoking then affects others' smoking, which in turn affects yours, and so on.

How could the Mechanism Designer calculate the impact of a tax on the level of smoking?

A model of smoking as a social activity

Let us assume, for simplicity, that there are two smokers (Ana and Burak). Each smoker has an income, y^A and y^B. The price of each cigarette is p. Ana and Burak smoke x^A and x^B cigarettes daily respectively. As a social activity, the utility of smoking depends *positively* on the amount of smoking of the other person. Both smokers choose between some amount of smoking or using their money on any other good.

Figure 16.6 shows how Ana will choose a level of smoking. The downward-sloping solid curve is Ana's marginal benefit of smoking cigarettes (mb^A).

Her marginal benefit curve is downward-sloping because the 20th cigarette smoked in a day is less enjoyable to her than the first, or the 15th. To draw this figure, we also assume a given level of smoking for her friend, Burak, x_a^B.

The marginal cost of smoking is the price of cigarettes, p, each cigarette having the same price and therefore the same marginal cost. Before the tax is imposed marginal cost is $mc_0 = p_0$.

Ana will smoke more as long as the marginal benefit of smoking exceeds the marginal cost of smoking. She maximizes her utility where her marginal benefits equal her marginal costs, as shown by point **a**. At point **a** Ana smokes x_a^A cigarettes. M-Note 16.4 explains her choice using calculus and a particular utility function.

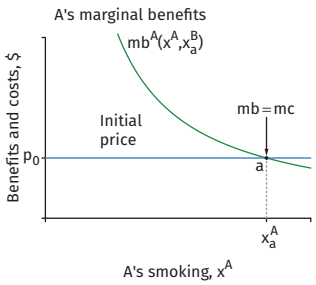


Figure 16.6 Initial choice of how much to smoke given another smoker's level of smoking. The figure is a graphic representation of Equation 16.10 for a given level of smoking by Ana's friend, Burak, x_a^B. The downward-sloping mb curve is the left-hand side of the equation and the horizontal p line the right-hand side.

M-NOTE 16.4 A smoker's utility-maximizing choice of smoking

For simplicity, we assume that for A the utility gained by the money spent on other goods equal to the income remaining after paying for cigarettes $(y - px^A)$ and similarly for B. Then a utility function for Ana that expresses the idea that smoking is a social activity is:

$$\text{Ana's utility} \quad u^A(x^A, x^B) = (\alpha + \beta x^B)\ln x^A + y^A - px^A$$

The parameter $\alpha > 0$ measures how much the smoker enjoys smoking alone (that is when $x_B = 0$) while $\beta \geq 0$ measures how much more pleasurable smoking is if other people smoke more. The fact that A's utility depends on the natural logarithm of her smoking $(\ln x^A)$ means that her marginal utility of smoking diminishes the more she smokes. (Remember that $\frac{d\ln X}{dX} = \frac{1}{X}$, so the marginal utility of smoking falls the more smoking she does.)

To maximize her utility Ana will pick x^A, the only variable that she can select. To calculate the x^A that maximizes Ana's utility, we differentiate her utility function and set the result equal to zero:

$$\text{First-order condition} \quad \frac{\partial u^A}{\partial x^A} = \frac{\alpha + \beta x^B}{x^A} - p = 0$$

$$\text{Marginal utility} = \text{Marginal benefit} - \text{Marginal cost} = 0$$

$$\therefore \text{ Marginal benefit} = \text{Marginal cost}$$

$$\frac{\alpha + \beta x^B}{x^A} = p \tag{16.10}$$

$$\text{Rearranging to isolate } x^A(x^B) \quad x^A = \frac{\alpha + \beta x^B}{p} \tag{16.11}$$

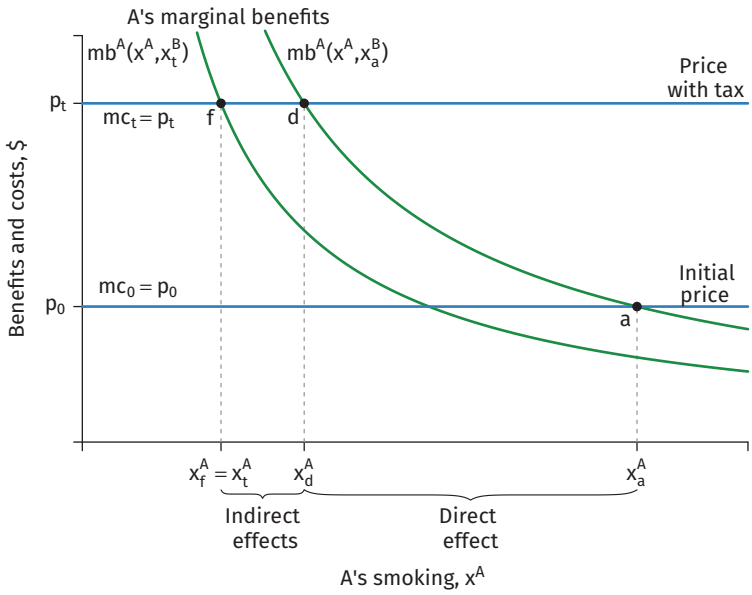
Equation 16.11 is Ana's best-response function. Her smoking depends negatively on the price of cigarettes (because p is in the denominator of the expression), and positively on the smoking of her friend Burak (because x^B, the smoking by Burak, is in the numerator with a positive sign).

Burak's best-response function is the mirror image of Ana's, as was the case for the fishermen in Chapter 5 and for the firms in Chapter 9.

A tax affects the marginal benefits and costs of smoking

Suppose the smoking tax raised the price of cigarettes, as shown in Figure 16.7, increasing p_0 to p_t. There would be a series of effects on Ana's level of smoking:

Figure 16.7 The social multiplier of a tax on cigarettes. The figure shows how both a price increase and others' smoking less will reduce smoking. The first-round and subsequent indirect effects are shown as 'indirect effects' in the figure. Before the tax, she chooses a level of smoking that equates her marginal benefit of smoking to the marginal cost of smoking (shown by point **a**). Point **d** shows how much she would cut back in the hypothetical case that others' level of smoking did not change. Point **f** gives her level of smoking when both direct and indirect effects of the tax are taken into account.



- A *first-round direct effect*: Because smoking is more expensive ($p_t > p_0$), she would do less ($x_d^A < x_a^A$).
- A *first-round indirect effect*: Because in response to the tax, others would smoke less too, she would enjoy smoking less, and so would reduce her smoking.
- *Subsequent indirect effects*: Now, because Ana has cut back, others would enjoy smoking even less and reduce their smoking even more, in response to which she also would reduce her level of smoking even further, and so on.

Figure 16.7 shows both of the direct and the indirect effects of the tax.

The *direct* effect is shown by the shift up of the price line from p_0 to p_t, with a movement from the initial utility-maximizing point **a** to point **d**, the intersection of the marginal cost (the price) and her original marginal benefit curve. This movement assumes that the tax does *not yet* change what Burak does and simply shows the direct effect of increasing the marginal cost of smoking on Ana's level of smoking, decreasing her smoking from x_a^A to x_d^A.

The second and subsequent *indirect* effects are shown by the shift *downward* in her marginal benefit of smoking curve: for any level of smoking that she does, the marginal benefit is now less because Burak is smoking less. The sum of the two effects results in Ana reducing her smoking from x_a^A to x_f^A

The Nash equilibrium level of smoking

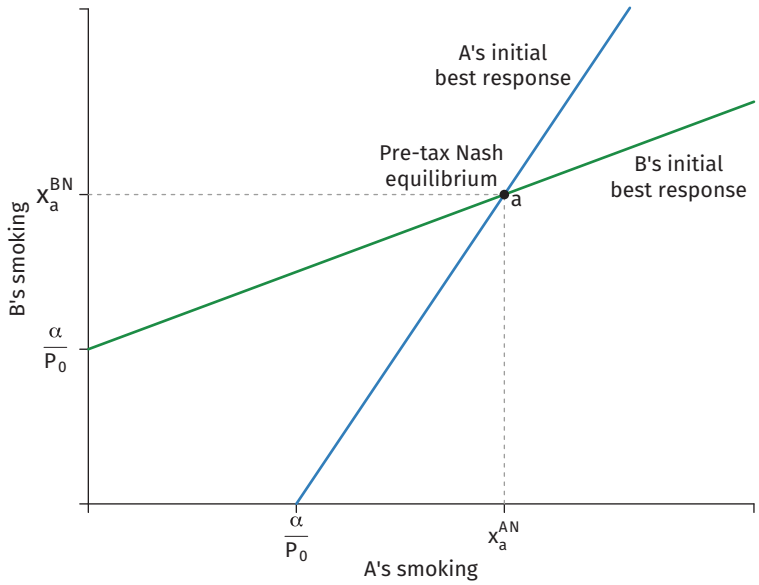
Figure 16.7 shows how Ana changes her level of smoking when the price increases. But it does *not* show how Burak responds to the tax and to the fact that Ana now smokes less.

To find out, we show in Figure 16.8 how each smoker responds to the level of smoking of the other, assuming the price of cigarettes is *fixed*. The steeper of the two upward-sloping lines is Ana's best-response function, showing the levels of smoking she will do in response to Burak's level of smoking. We derived Ana's best-response function in M-Note 16.4. Ana's best-response function slopes upward because, holding the price of cigarettes constant, the more Burak smokes, the more Ana will smoke. The x-axis intercept of her best-response function shows how much she would smoke if Burak did *not* smoke at all.

M-CHECK We call Ana's best-response function "steeper" and Burak's "flatter." But the two smokers are identical, so their best-response functions are just mirror images of each other. They would exactly coincide if we plotted Burak's on a graph in which his smoking amount rather than Ana's was the x-axis variable.

The flatter of the upward-sloping lines is Burak's best-response function, showing how he responds to each possible level of Ana's smoking. The two

Figure 16.8 The Nash equilibrium level of smoking before a cigarette tax is imposed. Ana's best-response function is the steeper upward-sloping blue line. Burak's best-response function is the flatter upward-sloping green line. The Nash equilibrium occurs at the intersection of the two best-response functions at point **a**. A's best-response function is Equation 16.11 and B's is symmetrical to A's.



best-response functions give us all the information we need to determine the Nash equilibrium level of smoking (x^{AN}, x^{BN}). Remember, the Nash equilibrium is a mutual best response so it must be a point on both of the smokers' best-response functions. There is only one Nash equilibrium which occurs at the intersection of the two best-response functions.

Figure 16.9 shows how the two smokers might respond to each other's smoking were they not at the Nash equilibrium. You can see that there is a plausible set of steps (shown by the arrows) that might lead them to a mutual best response.

M-NOTE 16.5 The Nash equilibrium of an interaction between smokers

Since it is a mutual best response, the Nash equilibrium of the interaction between Ana and Burak is the intersection of the two best-response functions in Figure 16.8. To find the Nash equilibrium we could substitute Burak's best response function $x^B(x^A)$ in Ana's best-response function, and solve that for x^{AN}.

But there is a shortcut: the two are identical so at the Nash equilibrium, $x^A = x^B$. We can therefore, substitute $x^B = x^A$ into Ana's best-response function (Equation 16.11) to find the Nash equilibrium levels of smoking x^{AN} and x^{BN}.

$$\begin{aligned} x^A &= \frac{\alpha + \beta x^A}{p} \\ \left(1 - \frac{\beta}{p}\right) x^A &= \frac{\alpha}{p} \\ \left(\frac{p-\beta}{p}\right) x^A &= \frac{\alpha}{p} \\ x^{AN} &= \frac{\alpha}{p-\beta} = x^{BN} \end{aligned} \tag{16.12}$$

For the Nash equilibrium to exist the best-response functions must intersect, requiring that $\frac{\beta}{p} < 1$. This is true because of the slope of the best-response functions is $\frac{\beta}{p}$ for Ana and Burak. If these slopes are greater than one, the best-response functions do not intersect. Equation 16.12 shows that the Nash equilibrium level of smoking (if it exists) will be positive because $\alpha > 0$ and $\frac{\beta}{p} < 1$ (or what is the same thing) $\beta < p$, so the denominator of Equation 16.12 must be positive.

CHECKPOINT 16.4 A different dynamic Think about another setup in which smoking alone is *unpleasant* so $\alpha < 0$ but the positive effect of the other smoking (β) is large so that $p < \beta$.

a. Redraw Figure 16.9, carefully labeling the intercepts.

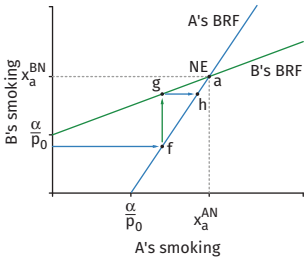
b. What is the slope of Ana's best-response function?

c. Imagine a situation in which neither of them smoked. Would either of them start smoking?

d. What are the Nash equilibria of this interaction?

e. Let $\alpha = 0$ and $p = \beta$. Is there a Nash equilibrium? (Hint: there could be more than one.)

Figure 16.9 Getting to the Nash equilibrium. To see how the two smokers might arrive at the Nash equilibrium imagine that for whatever reason in the first instance Ana was not smoking at all and Burak was smoking the amount indicated by the intercept of the blue arrow and the y-axis. Ana's best response to Burak's smoking would be to smoke the amount indicated by point **f** on her best-response function. But if Ana were smoking that amount, then Burak's best response would be to increase his smoking to point **g** on his best-response function. And so on until they arrived at point **a**, the Nash equilibrium, a mutual best response, so their level of smoking would then be stationary (unchanging).



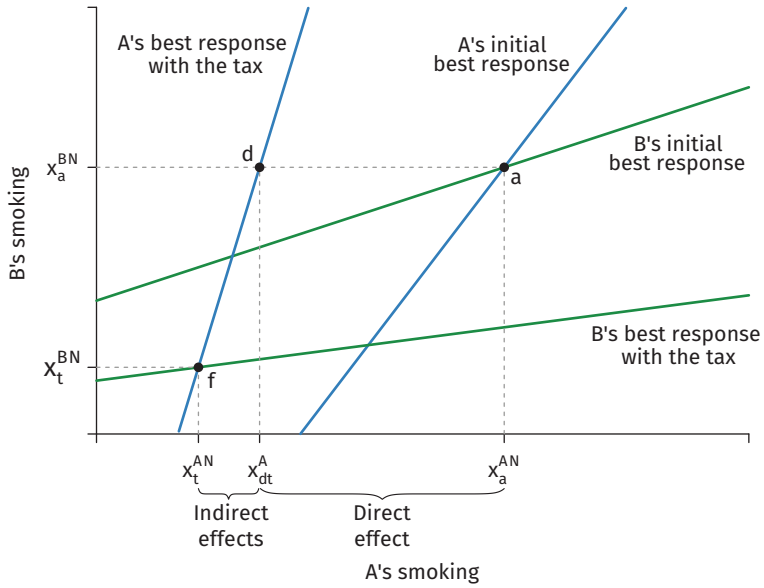
16.6 MECHANISM DESIGN WITH SOCIAL INTERACTIONS: A CIGARETTE TAX

Figure 16.10 shows how the increase in price caused by the tax changes the best-response functions of the two, and how as a result the Nash equilibrium level of smoking is reduced. There are two effects of the price increase on Ana's best-response function

- it reduces the amount of smoking she would do if Burak did not smoke at all, reducing the x-intercept; and
- it reduces the effect of Burak's smoking on her own smoking, making the best-response function steeper. This means that the tax *reduces* the effect of Burak's smoking on her smoking for any level at which Burak smokes.

The effect of the tax-induced price increase on Burak's level of smoking is similar. In Figure 16.8 it is shown as a shift downward of his best-

Figure 16.10 The Nash equilibrium level of smoking before (point a) and after (point f) a cigarette tax is imposed. The effect of the tax is shown in the shift to the left of Ana's best-response function (the blue line) and the shift down of Burak's best-response function (the green line). The direct (first-round) effect is shown by the horizontal distance between points **a** and **d** (the points have the same meaning as in the previous figure) which shows the effect of the higher cost of cigarettes on Ana's smoking if there was not social multiplier (i.e. if for some reason Burak's smoking was unaffected). But we know that Burak will smoke less, which reduces Ana's utility from smoking. The movement from **d** to **f** captures these indirect effects of the tax on A's smoking decision.



response function and a flattening of the slope. The new Nash equilibrium is the intersection of the two *new* best-response functions at point **f** in Figure 16.8.

The total change in A's smoking x_a^{AN} to x_t^{AN} is made up of two effects:

- the *direct effect*: shown as a reduction from x_a^{AN} to x_{dt}^{A}, and
- the *indirect effect*: shown as the reduction from x_{dt}^{A} to x_t^{AN}.

First, we can see that the total effect of the tax is given by comparing the initial Nash equilibrium (point **a**) with the Nash equilibrium after the tax (point **f**). How do we identify the direct and indirect effects of the tax? We separate the effects as follows.

How much would Ana smoke *with the tax* if Burak had smoked the same level as he did at the initial Nash equilibrium? If Burak smoked x_a^{BN} while Ana's best-response function had pivoted to the left, we can identify that she would have smoked x_{dt}^{A} at point **d**. Comparing points **a** and **d** we can find the difference $x_a^{AN} - x_{dt}^{A}$, which is the *direct effect* of the tax to reduce Ana's smoking.

But the tax also affects Burak: his best-response function pivots downward. We can identify the *indirect* effect as the difference between points **d** and **f** which includes Burak's reduced smoking on Ana's smoking, with the difference $x_{dt}^{A} - x_t^{A}$ being the indirect effect of Burak's smoking less on Ana's smoking, the effect of this on Burak's smoking, the effect of his further reduction in smoking on Ana, and so on.

M-NOTE 16.6 Direct, indirect, and total effects of a cigarette tax

The direct and indirect effects of the cigarette tax shown in Figure 16.10 can be identified mathematically using the *total derivative*.

To understand the effects of the tax on smoking, we would need to know how the tax affects the price of cigarettes, which will depend on the degree of competition and the nature of the cigarette firms' cost functions. To simplify, then, we will assume that we know the change on smoking that results from a change of the price ($\frac{dx^N}{dp}$).

Using Equation 16.11, Ana's best-response function, we calculate dx^A, the change in smoking of Ana given a change in the price dp:

$$\text{Change in A's smoking} = \text{Change via price} + \text{Change via B's smoking}$$

$$dx^A = \frac{\partial x^A}{\partial p} dp + \frac{\partial x^A}{\partial x^B} dx^B$$

Divide through by dp

$$\frac{dx^A}{dp} = \frac{\partial x^A}{\partial p} + \frac{\partial x^A}{\partial x^B} \frac{dx^B}{dp} \tag{16.13}$$

$$\Delta \text{ smoking for } \Delta\text{Price} = \text{Direct effect} + \text{Indirect effect}$$

Choosing a tax to implement a target

Equipped with this information on the effect of taxes on smoking taking account of both direct and indirect effects, the Mechanism Designer is almost ready to recommend a tax level. Recall that there is a target level for smoking in the population—a reduction of smokers' cigarette use one per

day—then to determine the tax that she should introduce, the Mechanism Designer needs two further pieces of information.

- What will be the effect of the tax on the price of cigarettes?
- Given the price effect of the tax, and the target level of smoking—one cigarette per day—what is the tax she should recommend?

She already has the answer to the first question, and as a result answering the second is now straightforward. The target level of smoking must be a Nash equilibrium, so, setting $x^N = 1$ in Equation 16.12 solve for the price. Using a numerical example, we show in M-Note 16.7 how this would be done.

Our analysis of subsidies to internalize the external benefits of contributing to a public good, or taxes to discourage smoking conveys the idea that policymaking is a simple matter. We have shown how improved—even Pareto-efficient—allocations can be implemented by changing the rules of the game so that the desired outcome is a Nash equilibrium.

We have focused on the promise of mechanism design as an aspiration. We now turn to some reasons why these models are difficult to apply in practice, beginning with the aptly named "theory of the second best."

M-NOTE 16.7 The Mechanism Designer chooses a cigarette tax

We know from Equation 16.12 that the Nash equilibrium level of smoking is:

$$x^N = \frac{\alpha}{p - \beta}$$

Suppose we have $\alpha = 1$, $\beta = \frac{9}{20}$, and p is \$0.50 per cigarette. Then, without a tax, we have:

$$x^N = \frac{1}{0.5 - \frac{9}{20}} = 20$$

a Nash Equilibrium level of smoking of 20 cigarettes (one pack) per day.

The Mechanism Designer has been assigned the task of reducing smoking to one cigarette per day. Using the superscript ω to differentiate the post tax from the status quo level, implementing this one-smoke-a-day objective would require setting the tax so that the new price p^ω satisfies:

$$x^{N\omega} = 1 = \frac{\alpha}{p - \beta}$$

$$\text{Rearranging} \qquad \alpha + \beta = p \tag{16.14}$$

To implement the target $x^{N\omega}$ of one cigarette per day, we insert the values of α and β into Equation 16.14 and see that the Mechanism Designer has to impose a tax such that:

$$p = \alpha + \beta = 1 + \frac{9}{20} = p^\omega = \$1.45$$

continued

Just to check, we can see at this tax level, our new Nash equilibrium, as intended would be:

$$x^{N\omega} = \frac{\alpha}{p^{\omega} - \beta} = \frac{1}{\left(1 + \frac{9}{20}\right) - \frac{9}{20}} = 1$$

CHECKPOINT 16.5 **A tax on "conspicuous consumption"** Imagine a situation like that the one in section 7.6 in which some item of luxury consumption—designer clothing, for example—is visible by others (that is, conspicuous) and people preferences are such that the more others practice conspicuous consumption, the more pressure people feel to do the same.

a. Make a new version of Figure 16.5 showing how you could adapt the smoking model with conspicuous consumption in place of smoking.

b. Explain why there would be an indirect effect of one person consuming more of the conspicuous luxury good.

c. Use a modified version of Figure 16.10 to explain the effect of a tax on luxury consumption.

16.7 THE THEORY OF THE SECOND-BEST AND PUBLIC POLICY

Many public policies seek to reduce the extent of uncompensated external effects that occur when contracts are incomplete, and thereby to mitigate market failures and other coordination failures. Examples that you have studied include taxing producers who overexploit our natural surroundings and consumers whose eating or smoking habits lead to illnesses the treatment of which imposes costs on others.

Public policies also seek to address market failures due to limited competition in markets. Recall that the owners of monopolistically competitive firms as well as duopolies and monopsonies studied in Chapter 9 as well as the monopsonies studied in Chapter 11 will maximize profits by limiting their outputs so as to sustain a price above marginal cost. The result is that there are would-be buyers willing to pay *more* than the marginal cost of the good, who *do not* buy the good.

! REMINDER A second-best outcome is one that is Pareto inferior to another ("first best"). In our example here, it may be that the first-best outcome while technically feasible cannot be implemented—perhaps due to political opposition—so no policy at all is a second-best outcome. Any outcome that is Pareto inferior to the first-best outcome is called second best.

Second best: The idea

Government policy might wish to address coordination failures arising from both limited competition and uncompensated external effects. But, what is called the theory of the second best states that unless policies can address both problems completely (totally eliminating uncompensated external effects and ensuring the prices will equal marginal costs), the government's intervention may make the situation worse. The "first-best" option is to address both problems completely, but if this is not possible,

! REMINDER Recall that from Chapter 8 you know that if $p > mc$ then there will be a deadweight loss and from Chapter 14 that marginal cost pricing is also the basis of the "invisible hand" reasoning formalized in the first welfare theorem (see Equation 14.4).

then a "second-best" option—not addressing either—may be preferable to addressing the problems singly.

This "all-or-nothing" conclusion is surprising. But its basic logic is simple. Think about the two examples above: those overexploiting nature and imposing external costs on others are producing too much, while firms facing limited competition are producing too little. If that is the case, then a firm that is over-exploiting nature (producing too much) and *also* facing limited competition (producing too little) may be producing the right amount.

Let's continue with the above example: a firm emitting pollution when it produces a good, so that its production has negative external effects on the environment. The marginal social costs of production (the total costs to society which include the external costs) exceed the marginal private costs (to the firm), the firm produces more than the Pareto-efficient level of output.

But imagine that the firm is a major petroleum company. Now take account of the fact that the firm faces limited competition, so it is *restricting* output, and setting a price in excess of its marginal private cost. This could correct the problem of the firm's negative environmental effects.

Correcting the producing-too-little (limited competition) problem without correcting the producing-too-much (negative external effects) problem may be worse than doing nothing. To see how this could be the case return to Equation 14.6 in Chapter 14 repeated below, which expresses why the combination of *perfect competition* and *complete contracts* yields a Pareto efficient equilibrium. Remember these are the two conditions under which a market equilibrium is Pareto-efficient according to the first welfare theorem.

! REMINDER c_x^P and c_y^P are the private marginal costs of goods x and y, and c_x^S and c_y^S are the social marginal costs, meaning the costs taking account of the negative external effects imposed (e.g. by carbon emissions) on others.

$$mrs(x,y) = \overbrace{\frac{u_x}{u_y} = \frac{p_x}{p_y} = \underbrace{\frac{c_x^p}{c_y^p} = \frac{c_x^s}{c_y^s}}_{\text{Complete contracts}}}^{\text{Perfect competition}} = msrt(x,y) \tag{16.15}$$

Let good y be the 'innocent' good: it is sold in markets with unlimited competition and its production and use imposes no external costs (or benefits) on others. The problem is that good x violates both of the conditions: it is produced by a monopoly whose production process imposes external carbon emission costs on others. This means both that:

- because it is a monopoly: $p_x > c_x^P$; price exceeds private marginal cost and
- because it is a polluter: $c_x^P < c_x^S$ private marginal costs are less that social marginal costs.

So taking account of these two sources of market failure Equation 16.15 becomes:

$$mrs(x,y) = \frac{u_x}{u_y} = \frac{p_x}{p_y} > \overbrace{\underbrace{\frac{c_x^p}{c_y^p} < \frac{c_x^s}{c_y^s}}_{\text{Incomplete contracts}}}^{\text{Limited competition}} = msrt(x,y) \tag{16.16}$$

From Equation 16.16 you can see that limited competition makes the relative price of the x-good too high, while the incomplete contract (not taking account of the environmental costs) make the price of the x-good too low. The offsetting effects of limited competition and incomplete contracts could result in the price ratio approximating the ratio of the two marginal social costs.

It is therefore possible that it would be preferable to leave both of the inequalities in Equation 16.16 in place—the "no-policy" option—rather than adopting policies that would replace just one of them by an equality. Replacing just one of the inequalities in Equation 16.16 by an equality could mean either:

- *competition policies*: the price ratio would now be equal to the ratio of private marginal costs (replacing the first inequality by an equality); or
- *environmental policy*: this would alter private costs so that the ratio of private to social marginal cost would now be equal (replacing the second inequality by an equality).

To explore these possibilities we return to ideas from Chapters 7 and 8: especially consumer surplus, economic profit, and deadweight loss.

Policies to deal with a monopolistic polluter

We can see how a single policy might be worse than no policy at all in Figure 16.11, which illustrates the case of a firm—the producer of the x-good—facing a downward-sloping demand curve (due to limited competition) with marginal private costs lower than marginal social costs (due to the external environmental costs imposed on others because they are not covered by a complete contract). We assume a constant average (and therefore marginal) cost to simplify the figures. In the absence of any policy, the firm will equate marginal revenue to marginal private cost, at point **b** producing the amount X^{MP} (it is a monopoly, so the firm's output x is the same as the industry output, X), and selling at the price p^{MP}.

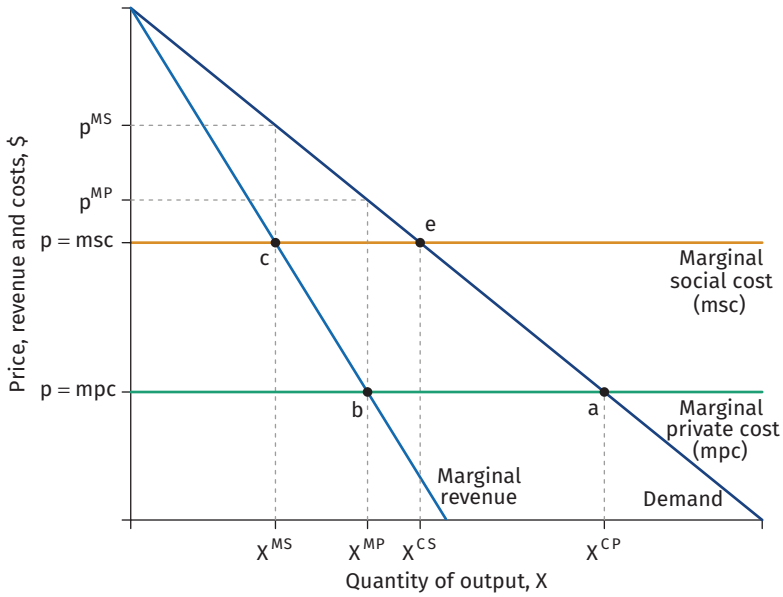
Ideally a government would introduce both:

- A *competition policy, reducing barriers to entry*: ensuring that the firm faces sufficient competition so that it would produce up to the point where the price is equal to the marginal private cost, eliminating deadweight losses.
- *An environmental policy, green taxes*: imposing costs on the firm equal to the negative external effects of its emissions, so that the firm's marginal

HISTORY Here is the gist of what has come to be called the *general theorem of the second best* advanced in 1956 by Richard Lipsey (1928–) and Kelvin Lancaster (1924–1999): if there is a single violation of the relevant efficiency conditions (that is, one of the equalities in Equation 16.15) then fulfilling the remaining marginal conditions may result in an allocation that is Pareto inferior to an allocation implementable by more extensive violations of the efficiency conditions. In our example, the environmental coordination failure may offset the monopoly coordination failure.[10]

! REMINDER In Chapter 9, we spoke about competition policy as a way of *increasing* the number of firms in a market that had high concentration (few firms competing). In the US, this competition policy has typically been called antitrust policy because monopolies were called trusts. Currently, many governments—South Africa, the UK, the Philippines, Singapore, India, the EU—have a "competition commission" that administers competition policy.

Figure 16.11 The theory of the second best. The lowercase letters indicate the outcomes under no policy (**b**, for bad), environmental policy (green tax) only (**c**), competition policy only (**a**), and both policies (**e**, for efficient.)



private costs would now include the environmental costs (the marginal external costs) imposed on others and its private costs would now be the higher marginal social costs (*msc*).

These policies *together* would correct the market failure. The firm would produce at point **e** instead of point **b**. This is *the first-best outcome*.

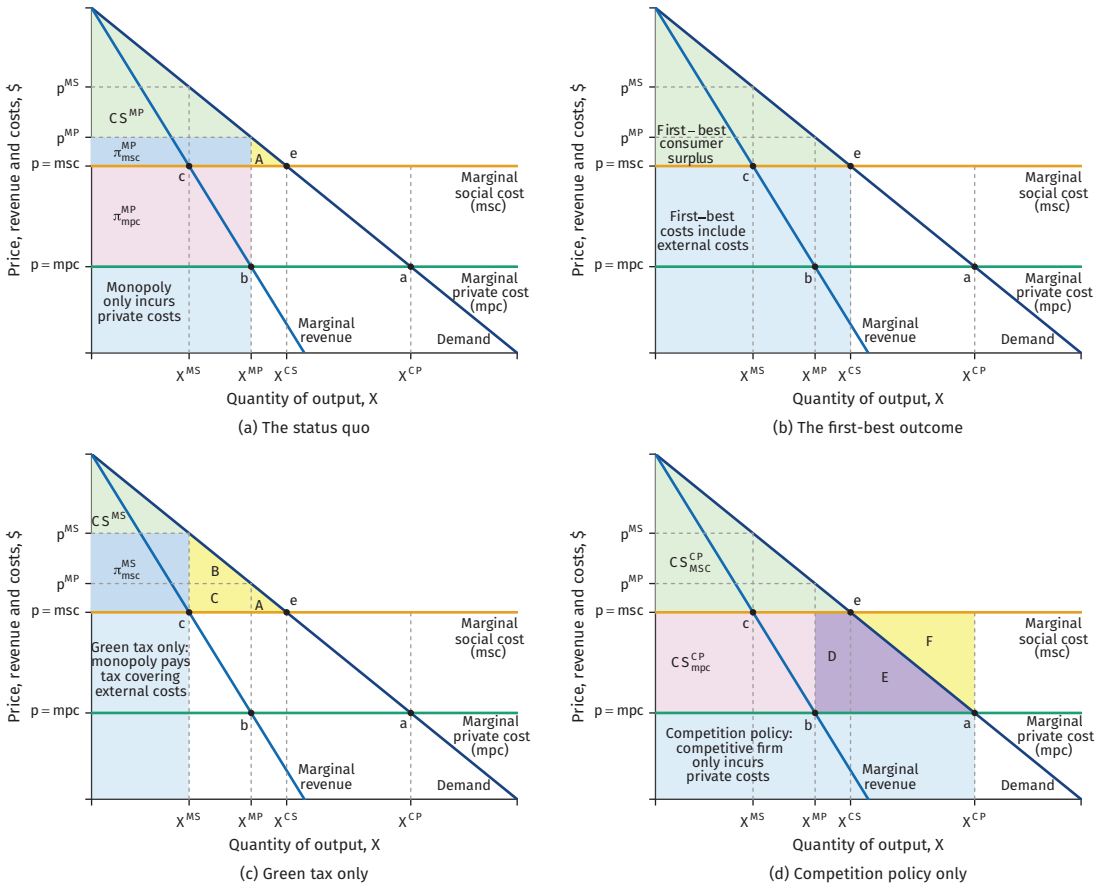
But it may be *impossible* to enact both policies. Reasons could include the limited information or administrative and legal capacities of the government, opposition from the owners of firms whose profits are elevated by limited competition or depend on carbon-intensive technologies. In the absence of an environmental policy, a competition policy alone would induce the firm to produce at point **a** where its marginal private costs equals the price on the demand curve. The firm would choose the competitive output level (x^{CP})—producing and emitting more. This could *reduce* welfare rather than *enhance* welfare relative to point **b**.

And an environmental policy in the *absence* of a competition policy would impose the marginal social costs on the firm, leading the firm to *restrict* production even more and produce at point **c** rather than point **b**, expanding the preexisting deadweight losses.

The welfare outcomes of the different policies

Figure 16.12 helps us to diagnose these differences. Panel (a) is the status quo, while panel (b) shows the case of the first-best outcome where price equals marginal social cost ($p^{CS} = msc$) and the output is X^{CS}. Panels

Figure 16.12 The theory of the second best: comparison of profit, costs, consumer surplus, and deadweight losses. The theory of the second best illustrated by a comparison of competitive markets and monopoly when the marginal social costs exceed the marginal private costs because of the negative external effects imposed on others. The areas in yellow are welfare losses relative to the first-best outcome.



(c) and (d) respectively show the results if just the environmental policies or the competition policies are introduced.

Comparing (b) with Figures 16.12 (a)–(d), we can identify the following areas:

- *Area* A in panel (a): is deadweight loss (foregone consumer surplus) resulting from limited competition in the status quo. The firm produces *less* ($X^{MP} < X^{CS}$) than it would if competition were unlimited *and* its private marginal costs included the external costs it imposes on others. At its chosen level of output X^{MP} there are consumers whose willingness to pay exceed the social marginal cost and who therefore would have bought the

good at $p = msc$, but could not because the firm was monopolistic and charges $p^{MP} > p^{CS} = msc$.

- *Areas* B *and* C in panel (c): are the additional deadweight loss (foregone consumer surplus and economic profit) from enacting *green taxes only*. The monopolistic firm incurring costs *msc*, selling at price p^{MS}, and producing at X^{MS} loses economic profits that the firm had in the status quo case shown by area C. The reduction in output also deprives consumers of surpluses shown by the area B. The total deadweight loss in this case is A + B + C (welfare losses compared to the first-best option), an amount B + C greater than the status quo deadweight losses.
- *Areas D, E, and F* in panel (d): the sum of these areas is the total external costs if only the competition policy is introduced, that is, producing $X^{CP} > X^{CS}$. But areas D and E are *also* consumer surplus gained by producing $X^{CP} > X^{CS}$. Their areas are based on the difference in the willingness and the marginal private cost, and therefore consumer surplus when the firm produces X^{CP}. So D + E are increased consumer surpluses exactly offset by external costs imposed on others.
- *Area F* in panel (d): is deadweight loss from the implementation of *competition policy only*, that is, the increase in the total external costs (D + E + F) that is *not* offset by gains in consumer surplus. Because F is larger than A, the competition policy alone is *worse* in welfare terms than the status quo.

! REMINDER The deadweight loss measured by area F is different from the deadweight losses studied in Chapter 8, which resulted from firms producing too *little*. Here the firm produces *too much*, and the deadweight loss is the sum of the difference between the marginal social costs and the (lesser) willingness to pay of the buyers.

The four possible outcomes—no policy, both policies, and one but not the other policy—and their effects are summarized in Table 16.1. The table and Figure 16.12 on which it is based illustrate a case in which both of the single policy outcomes—the competition policy without the environmental policy and the environmental policy without the competition policy—are worse in welfare terms than the original outcome (X^{MP}, p^{MP}).

Table 16.1 Policymaking in a second-best world. The rows are the four policy options discussed in the text. The entries in the welfare loss column refer to the yellow shaded areas in figure 16.12. If introduced alone, green taxes and competition policy both result in an increase in the size of the deadweight losses over the "no-policy" option.

Policy	First-order condition	Point in fig	Price	Output	Welfare loss	Reason for loss
No policy	mr = mpc	**b**	p^{MP}	X^{MP}	A	Too little output
Green taxes only	mr = msc	**c**	P^{MS}	X^{MS}	A + B + C > A	Even less output
Competition policy only	p = mpc	**a**	P^{CP}	X^{CP}	F > A	Too much output
Green taxes + competition	p = msc	**e**	P^{CS}	X^{CS}	None	

Imposing either competition policy or green taxes in the *absence* of the other is worse than doing nothing: area A—the status quo deadweight welfare loss—is smaller than either F—the loss from adopting the competition policy alone—and of course smaller than A + B + C, the loss from adopting the environmental policy alone.

To illustrate the idea of the second best, we have chosen a case in which doing nothing is better than adopting either of the two policies singly. But this need not be true in other cases (see the next Checkpoint). Moreover, the theorem of the second best does not question the idea that public policy ought to address market failures arising from lack of competition and uncompensated external effects. But, it does show that treating market failures in isolation rather than as a general problem can be counterproductive. It also underlines the fact that the best the policymaker can do may not be to entirely eliminate the market failure but instead to limit its extent.

CHECKPOINT 16.6 Second-best policies

a. Explain why the points **b**, **c**, **a**, and **e** in Figure 16.12 indicate the results of the four policies shown in Table 16.1.

b. It is not always true that doing nothing is better than adopting either policy in isolation. Can you draw the curves in Figure 16.12 so that adopting the environmental policy but not the competition policy will be better than doing nothing? Can you do the same figure for the case in which adopting a competition policy is an improvement over the do-nothing policy even if it is not accompanied by the environmental policy?

16.8 THE PERFECT COMPETITOR AS AN IMPEDIMENT TO EFFICIENT EXCHANGE

The theory of the second-best addressed cases in which mechanism design provided clear advice as to the policies that could implement Pareto-efficient Nash equilibria. But there are cases in which the Mechanism Designer's search for such policies is bound to fail. There may be no mechanism that will implement a Nash equilibrium that is Pareto efficient.

In Chapter 9, we introduced the perfect competitor—an economic actor who never misses an opportunity for private gain. We provided conditions under which an economy made up of perfect competitors would implement Pareto-efficient outcomes.

We showed, for example, that the owner of a firm facing limited competition (and hence a downward-sloping demand curve) like the firm studied in the previous section would *not* restrict output to sustain higher prices if he were a perfect competitor. Instead he would practice price discrimination. This would allow him to expand production up to the point where price = marginal cost. If perfect price discrimination is possible even a monopoly

would therefore implement the conditions that we associate with perfect competition. We also used a model of bilateral exchange with rent-seeking perfect competitors to understand how buyers and sellers could reach a Pareto-efficient outcome in Chapter 14.

We gave the perfect competitor extraordinary powers, the ability of the monopolist, for example, to find out the willingness to pay of each potential buyer of her product, and to impose that price on each. But one power we did not give her: the power to *lie*.

Lying for profit: Why mutually beneficial trades may not occur

And so we confront a problem: Can the perfect competitor implement efficient outcomes if she (and everyone else) can lie?

To see why the answer may be "no," think about the following situation. When traders meet they have no incentive to report how much they value the good to be exchanged. That is, they have no incentive to report what we call their *true valuation* of the good. The reason is that a trader may be able to profit by *misreporting* or withholding information about how much she values the good.

This behavior is quite common in bargaining situations. Think about a prospective buyer of your home. He is not going to say: "Well I'd prefer a lower price, but in fact I'd be willing to pay as much as $200,000 for it." And you, the owner, are not going to offer the information "I'd like the highest price possible, but I'd be willing to part with the house for as little as $100,000 if that is the best I can do."

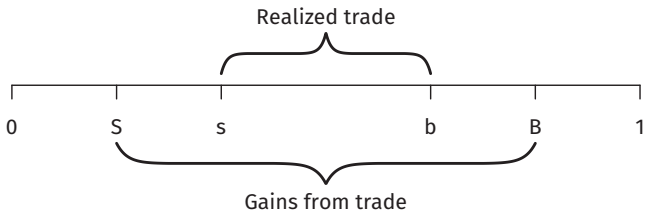
If they are perfect competitors they will certainly recognize withholding or misrepresenting information about an exchange as a rent-seeking opportunity, and falsely report their valuations of the goods. The result, we will show, is that some exchanges that could have benefited both buyer and seller will not happen.

The reason is that the traders' stated valuations will influence the prices at which the traders exchange. As a result, some mutually beneficial exchanges will not occur. The problem is quite general, but it is best illustrated by an institution called a **double auction** involving goods for which the usual impediments to bargaining (such as incomplete contracts) are absent.

A particular good now in the possession of the seller, her house for example, is worth S to a seller who may sell this good to a buyer. The good is worth B to the buyer. These valuations are private information, so the other

DOUBLE AUCTION In a double auction buyers and sellers simultaneously submit to an auctioneer "bids" and "asks" that are the prices at which they are willing to buy and sell, respectively. An auctioneer then chooses a price that clears the market.

Figure 16.13 Bargaining over the gains from trade. For the pair of traders shown in the figure, a mutually beneficial trade is technically possible because the seller's true valuation S is lower than the buyer's true valuation B. Therefore, for any price between those points the exchange is mutually beneficial. If the two announce the values $s < b$, then they will trade at some price between those two numbers. If, however, s had been greater than b, no trade would take place, even though it would be Pareto improving to execute the trade.



trader does not know her trading partner's valuations. The double auction proceeds as follows:

- *Pairing buyers and sellers*: A large number of buyers and sellers are paired for a simultaneous one-shot interaction.
- *Seller*: In each pair, the seller announces the minimum price at which she is willing to sell, s.
- *Buyer*: At the same time, the buyer announce the maximum price at which he is willing to buy, b.
- *Exchange*: An exchange occurs if the buyer's *offer* price is greater than the seller's *sell* price, that is, $b > s$.
- *Price*: The price determined by the two bargaining will be between s and b: for example the bargaining rule could be that the price is midway between the two, so $p = (b + s)/2$.

Figure 16.13 depicts the valuations of the players and the potential for gains from trade.

You can see that in this exchange $B \geq S$ is the participation constraint: if the good is not worth more to the buyer than it is to the seller, then there is no way that the exchange could occur voluntarily. If $B \geq S$ and the exchange does not happen, then this result is not Pareto efficient because then there would exist some change—the sale that did not occur—that would make both traders better off.

The incentive compatibility constraint in this exchange is $b \geq s$ because this inequality restricts the actual exchanges to cases where buying and selling prices that the two will voluntarily choose to announce will allow a trade to occur.

HISTORY Roger Myerson was awarded a Nobel Prize in Economics for demonstrating in 1983 (with his coauthor Mark Satterthwaite) what is called the "Myerson–Satterthwaite theorem," which shows that some mutually beneficial trades will not take place in the double auction because buyers and sellers misrepresent their true values.[11]

The reason why opportunities for beneficial trade will be missed is as follows:

- *Buyers have an incentive to lie*: if a trade takes place, then the price will be more favorable to the buyer, the lower is her announced valuation, b.
- *Sellers have an incentive to lie*: if a trade takes place, then the price will be more advantageous to the buyer, the higher is his announced valuation s.

M-CHECK Both the buyers and sellers know that lying would reduce the *probability* of a transaction. They will account for this tradeoff in their maximization process. In other words, each also has an incentive not to overstate or understate too much. As a result, they would lie to maximize their *expected* gains. This is a choice with risk as we have modeled in Chapters 2 and 13.

The seller therefore has an incentive to *overstate* her valuation of the good and the buyer has an incentive to *understate* his valuation of the good. By misrepresenting their valuations, if a trade nonetheless occurs, both benefit from the resulting increase in their share of the gains from trade that they will get. But this comes at the cost of reducing the probability of a mutually beneficial transaction. As a result, when the two traders meet it may happen that $b < s$, so that trade does not occur, even though $B > S$, so the buyer valued the good more than the seller, and a mutually beneficial trade *could* have occurred were they not deceitful. This inefficient result arises because the announced valuations of the buyers and sellers influence *both*:

- the price at which the good will transact—if a transaction is concluded, and
- whether a transaction will take place at all.

If the price at which the good is exchanged were determined without regard to the announced valuations, then all mutually beneficial transactions would take place.

Conflict over the distribution of the pie may mean a smaller pie

The failure of our rent-seeking perfect competitors to exploit all mutually beneficial exchange opportunities is another example of the general problem that the conflict over how the pie will be divided up often results in a smaller pie. Other examples you have already studied include:

- Rejections of positive but very unequal offers in the Ultimatum Game.
- Pareto-inefficient allocations when one actor has price-setting power (but not take-it-or-leave-it power).
- Conflict over on-the-job effort and resulting unemployment and diversion of resources to monitoring.
- Conflict of interest over repayment of loans and conduct of projects financed by credit leading to exclusion from the credit market of would-be borrowers seeking to finance good projects but lack sufficient wealth to post collateral.

The conclusion is that even *without* the uncompensated external effects and limits to competition that result in market failures, it may be effectively

impossible to achieve the Pareto-efficient outcome envisioned in either a (Coasean) rent-seeking model of perfect competitors or a perfectly competitive model (a Walrasian model) of competitive general equilibrium such as those modeled in Chapter 14. Instead, we live in a second-best world.

CHECKPOINT 16.7 Market failures in the double auction Explain why the following statements are true:

a. If the bids and asks are submitted in sealed envelopes and the price is chosen by an assessor (a person skilled at determining the value of things) before opening the envelopes, following which the buyers and sellers could then decide whether to make the exchange at the selected price, then they would have no incentive to misrepresent their true values.

b. If the potential buyer truly values the good much more than the seller, then even though they will misrepresent their true values, it is likely that they will execute a trade (this means that the trades that are most beneficial are likely to be take place).

Markets would work better if *Homo economicus* had a conscience

But even in this second-best world, we might wonder whether a Mechanism Designer or even a randomly selected citizen might try to encourage people to change their behavior.

For example, in the double auction we described, if one of the many buyers or sellers were asked to choose a strategy for announcing the buying or selling price that everyone (including her) would have to follow, what rule would she choose? She would choose telling the truth.

The reason is that if everyone were to follow the tell-the-truth rule then she would do better on average because she would then be in an economy in which all mutually beneficial exchanges actually happen. So in this case $b = B$ and $s = S$ and every exchange that allowed mutual benefits would then occur, leading to a Pareto-efficient outcome.

But telling the truth is not a best response to others telling the truth in the double auction, as we have seen. So unless everyone acquired a conscience that simply *banned* misrepresenting their values on ethical grounds, the tell-the-truth rule would break down. Nobody would follow it. Economists and philosophers have long sought to devise rules of the game that would provide incentives that would motivate people to truthfully reveal ones preferences, called **truth-telling mechanisms**. But not only has

HISTORY In the early 1970s, philosopher Allan Gibbard and economist Mark Satterthwaite independently proved a theorem demonstrating the impossibility of mechanisms that would make honest revelation of preferences a dominant strategy.[12]

TRUTH-TELLING MECHANISMS Rules of the game that would make it a best response for a self-interested and amoral person to reveal their true preferences (including their true value of a good that they may buy or sell).

the search been unsuccessful: it has shown that truth-telling mechanisms cannot exist.

It is clear that if all parties can agree to restrict the pursuit of their own individual advantage—forgoing the advantage of misrepresenting one's true value—all may benefit. Another example that you encountered in Chapter 1 would be in the Prisoners' Dilemma game, to just outlaw defection.

But there are (and should be) limits to what the government can accomplish in this respect. In many countries, for example, it is illegal to overstate the value of one's property when making an application for a loan. But neither banks nor government typically have the information or enforcement capacities required to make truth-telling a best response. Many environmental policies illustrate the "outlaw defection" option. This works in cases where defection takes a specific and observable form whose banning can be easily enforced—banning the use of lead in vehicle fuel—but not in others.

✓ **FACT CHECK** Volkswagen successfully falsified emissions tests on many of its vehicles for a period of years.[13]

16.9 MECHANISM DESIGN IN AN IMPERFECT WORLD

The impossibility of efficient exchange among perfect competitors in the double auction above illustrates a coordination problem that—by its nature—cannot be solved: it was not obvious at the outset, but on investigation we can see that the failure was built into the game itself.

In other cases there exists a mechanism that could implement the desired outcome, but there are practical obstacles to doing so. **Team production** provides an illustration of why the optimal contracts a Mechanism Designer would introduce may not be possible.

! REMINDER Return to Chapter 10 to remind yourself of the aspects of contracts that make them *incomplete*. Remember that the difficulty of defining and measuring outcomes is an important source of contractual incompleteness.

Team production

Most production in a modern economy involves large numbers of people contributing to a job in ways that make it difficult to assign particular responsibility for each task to any one worker. This is called team production.

The team might be a group of professionals sharing a practice (common among doctors and lawyers) or employees of a restaurant who all contribute to putting the meal on the table. In recognition of the team nature of production many firms base a portion of compensation to employees on some measure of team output. A major airline, for example, paid the ground staff who manage departures and arrivals of planes a team bonus based on the fraction of on-time departures.

! REMINDER You have already studied another example of team production: the worker-owners of the cooperative plywood firms in Chapter 15.

TEAM PRODUCTION A form of production involving two or more people in which the contribution of each person to the output cannot be readily determined, either because it cannot be defined or because it cannot be measured.

For concreteness think about a group of n software engineers working together to write code for new applications. Their work is similar to the production process described in Chapter 11:

- *Disutility of effort*: The engineers devote effort (which they consider to be onerous, that is, a bad (it generates disutility)) to producing an output.
- *Effort produces a good:* They sell the output and with the resulting sales revenue they purchase goods that they consume contributing to their utility.

And:

- *Total team output is verifiable*: Information on the total output the team produces is observable and verifiable.
- *Individual effort is unverifiable*: Effort levels of individual engineers and the quality of their contributions to the team's output are not verifiable.

The team confronts a problem: how to design a compensation system that will motivate the members to work hard and well even when their effort is unverifiable. The problem they face is different from the public goods problem explored in section 16.3, for two reasons.

- *A key input is not verifiable*: In the public goods example we assumed that the Mechanism Designer could *observe* the levels of contribution to the public good and subsidize each member's contribution. This could motivate the desired level of total effort contributed by fines or subsidies based on the amount worked. But because information on the level of effort of each team member is not verifiable, team members *cannot* enforce this kind of contract—either a work level or subsidies and fines—on themselves.
- *The output of the team is rival*: When the team sells its output, the revenues it receives will be *divided* among the members somehow. Each dollar more that a team member gets is one dollar *less* another team member gets. The team's output is *not* a public good.

CHECKPOINT 16.8 Team production in your experience Describe a team production process in which you or someone you know have engaged. Be clear about how this process fits the definition of team production.

EXAMPLE Think about some team sport, football (soccer) for example. Each player contributes something to the team's success in winning games. But how much? Suppose you were asked to determine the contribution of each player to this objective. If you could observe enough games and were free to experiment with the membership of the team you might look at the differences in some measure of success of the team (points scored in excess of opponents' points, perhaps) when a given player was part of the team and when she was not. Thinking about even this simple procedure makes it clear how difficult it is to determine the contribution of an individual member of a team.

How team production works

Each team member contributes some amount of effort to the team's production process, e^j, with $j = 1 \ldots n$ for the n members of the team. The team produces an output, x that is just the sum of all of the contributions of the team members multiplied by a positive constant γ.

$$\text{Production:} \qquad x = \gamma \cdot (e^1 + e^2 + \ldots + e^n) \tag{16.17}$$

From the production function you can see that γ is the average productivity of effort and, because average productivity is a constant, it is also the marginal productivity of effort. To simplify we assume the team has no costs other than paying its members an amount y^j, member j's income. Then the utility of team member j is:

$$\text{Utility:} \qquad u^j = u(y^j, e^j) \tag{16.18}$$

Which says that an individual team member's utility is derived from income (a good, $u_y > 0$) and the amount and quality of effort contributed to the project (a bad, $u_e < 0$.).

M-NOTE 16.8 The marginal utility and disutility of contributing effort

We will use an explicit utility function:

$$u^j = y^j - \frac{1}{2}(e^j)^2 \tag{16.19}$$

from which we can see that:

- the marginal utility of income, denoted u_y, is equal to 1; and
- the marginal utility of effort, denoted u_e, is the derivative of i's utility with respect to effort, equal to $-e^j$ or what is the same thing, the marginal *dis*utility of effort, $-u_e$, is e^j itself.

Similar to the disutility of contributing to the public good in section 16.3, the second bullet says if you are *not* contributing much effort to the teams output, the disutility of working harder is not very great; but if you are providing a lot of effort, the marginal disutility to you is substantial.

The socially optimal effort by a team member

Suppose the members of the team ask the Mechanism Designer to advise them on the compensation system they should adopt, determining the income of each member. They would like the compensation system to provide each member incentives to contribute to the team an amount of effort that will maximize the total utility of its members.

Just as in the case of the public good that you studied earlier in this chapter, the designer would proceed in two steps:

- First, she would determine what is the amount of work effort that, if implemented by each team member, would maximize the *total* welfare (sum of individual utilities) of the team.
- Second, she would find (if one exists) a set of incentives (a "mechanism") that would motivate the team members—each maximizing their own utility—to implement the socially optimal amount of effort.

M-Note 16.9 shows how the Mechanism Designer would determine the socially optimal level of work effort by each team member for the case in

which there are just two members, A and B, and the marginal private cost (disutility) to each of providing effort is the level of effort itself.

The level of effort that the Mechanism Designer determines to maximize the sum of their utilities is that:

$$\text{Marginal disutility of effort} = e^A = e^B = \gamma = \text{Marginal productivity of effort}$$

This is the same thing as requiring that the marginal private cost of effort for each of the two team members is equal to the marginal social benefit (including the external benefits that each one working confers on the other).

M-NOTE 16.9 The optimal effort of a team member

Suppose $n = 2$, so the team is composed of just A and B. Their utilities are:

$$u^A = y^A - \frac{1}{2}(e^A)^2$$
$$u^B = y^B - \frac{1}{2}(e^B)^2$$

The Mechanism Designer selects e^A and e^B to maximize the sum of their two utilities:

$$\begin{aligned} W &= u^A + u^B \\ &= y^A - \frac{1}{2}(e^A)^2 + y^B - \frac{1}{2}(e^B)^2 \\ &= x - \frac{1}{2}\left(e^A\right)^2 - \frac{1}{2}\left(e^B\right)^2 \end{aligned} \tag{16.20}$$

We get the last equation above by noting that the sum of the incomes of the two team members is the total output of the team or $x = y^A + y^B$. Remember that x is γ multiplied by the sum of the players' efforts. Therefore, we can substitute $x = \gamma \cdot (e^A + e^B)$ into Equation 16.20 and find:

$$W = \gamma \cdot (e^A + e^B) - \frac{1}{2}(e^A)^2 - \frac{1}{2}(e^B)^2 \tag{16.21}$$

To find the values of e^A and e^B that maximize the sum of the utilities of the team members, we differentiate W with respect to e^A and e^B and set the result equal to zero:

$$\frac{\partial W}{\partial e^A} = \gamma - e^A = 0$$
$$\frac{\partial W}{\partial e^B} = \gamma - e^B = 0$$

So, rearranging the above equations we find that $e^A = e^B = \gamma$. This requires selecting e^A and e^B such that the marginal *cost* (disutility) of providing effort is equal to the marginal social *benefit* of providing effort (which is γ the marginal productivity of effort).

The $\frac{1}{n}$ problem

Having figured out the optimal level of effort—each member should provide a level of effort equal to γ—is a start. But as we have already seen, the team members cannot simply agree to all work at that level. The reason as your already know is that information about people's effort levels is not verifiable and hence cannot be used in any enforceable agreement or contract.

They first try the idea that the income of each team member is the total output of the team divided by the number of team members. So, according to this $1/n$ rule member i receives:

$$y^i = \frac{x}{n} \tag{16.22}$$

$$= \frac{\gamma \cdot (e^1 + e^2 + \ldots + e^i + \ldots + e^n)}{n} \tag{16.23}$$

Notice that if he works harder not only does he receive *more* income (because his effort contributes to x): so does *everyone else* on the team. In fact they each get *the same* share, namely, $1/n$ of the contribution of his greater effort to increased team output. So we have the following results:

- The *marginal private benefit* of working harder is γ/n ...
- the *marginal external benefit* of working harder (the increased income enjoyed by $n-1$ other members) is $(n-1)\gamma/n$ ...
- and the *marginal social benefit* is the sum of the two, or just γ itself.

How hard will the team member work under the x/n compensation system? Here the following rules give the answer:

- if the marginal private benefit of working harder (γ/n) is *greater* than the marginal cost (e), she will work harder; and
- if the marginal private benefit of working harder is *less* than the marginal cost, she will work less hard; so
- she will work at the level of effort that equates the marginal private benefits and costs of working harder. She will set $e = \gamma/n$.

Each other team member will choose the same level so in the Nash equilibrium of this game, we will have $e^N = \gamma/n$. We clarify the result mathematically in M-Note 16.10.

The Nash equilibrium level of effort is one n^{th} the socially optimal level of effort, which is γ. The fact that she is not compensated for the external benefits that her work confers on other team members means that she (and the other team members) will not put in as much effort as they would if they could *agree* on how much each would work. Every worker could be better off if each worked harder.

The team production problem is another coordination failure. The external effect of working harder and how it results in effort being too *low* is similar to the external effect of the citizen's contribution to the public good and how it means the public good will be under-provided.

M-NOTE 16.10 The team member chooses an effort level

Each of the two team members maximize their utility function, which for A is:

$$\text{Maximize} \quad u^A = y^i - \frac{1}{2}(e^A)^2$$
$$= (e^A + e^B)\frac{\gamma}{2} - \frac{1}{2}e^{A^2}$$

To find the level of effort that maximizes utility we differentiate this function and set the result equal to zero:

$$\frac{\partial u^A}{\partial e^A} = \frac{\gamma}{2} - e^A = 0$$

The same reasoning is true for team member B. Therefore we have:

$$e^A = e^B = \frac{\gamma}{2}$$

Or the marginal private cost is equal to the marginal private benefit. This is the Nash equilibrium level of effort. For n team members, the general the Nash equilibrium level of effort will be:

$$e^N = \frac{\gamma}{n}.$$

CHECKPOINT 16.9 Team size and the extent of the coordination failure

a. Explain in words why there is no coordination failure if the team has only one member.

b. Explain why the extent of the coordination failure—the difference between the socially optimal amount and the amount someone will contribute under the x/n compensation system—is greater, the more team members there are, n.

16.10 WHEN OPTIMAL CONTRACTS FAIL: THE CASE OF TEAM PRODUCTION

We have shown how paying team members a fraction of the output proportional to the total output (one n^{th}) results in a Pareto-inefficient outcome. We therefore want to ask: Is there a payment system that can motivate individual team members to implement the socially optimal level of effort? Surprisingly, there is: pay each member the *entire* value of the output of the team, minus a constant sum.

So now each team member would be paid:

$$\text{Payment} = \text{Total output} - \text{Constant}$$
$$y^i = \gamma \cdot (e^1 + e^2 + \ldots + e^n) - k \tag{16.24}$$

Subtracting a constant sum (k) in Equation 16.24 is necessary because otherwise the team would be required to pay out n times the team's total revenue. Though we do not pursue the question, the Mechanism Designer would set k so that, at a minimum, the members' expected incomes would be sufficient to satisfy the team members participation constraints given the amount of effort they expend.

This seemingly bizarre mechanism ensures the following:

- *Compensation*: Any contribution by a member to the output of the team will be exactly compensated.
- *Incentives*: The compensation gives each team member *the same* incentives as an isolated person who owns the *entire* fruits of his labor (what we called an "owner-operator" in Chapters 11 and 12).

As in the case of the optimal subsidy for the contribution to the public good, the compensation system proposed by the Mechanism Designer succeeds in internalizing the positive external effects of each member's work effort. It does this because each member now is treating the external benefits that their effort confers on others as if it was their own income—*because it is*!

M-NOTE 16.11 An optimal contract

Remember, x is the sum of all members' effort levels multiplied by γ. Therefore, with the new mechanism each person is paid $x - k$ rather than x/n. That is, in our two-person team illustration, each gets $y = \gamma \cdot (e^A + e^B) - k$ rather than $y = \gamma \cdot (e^A + e^B)/2$ (their compensation under the "$1/n$ rule").

So we can rewrite the utility function (see Equation 16.19) of team member A as follows, substituting in $y = \gamma \cdot (e^A + e^B) - k$:

$$u^A = \gamma \cdot (e^A + e^B) - k - \frac{1}{2}{e^A}^2 \tag{16.25}$$

To find the amount that will maximize the utility of member A under this compensation system we differentiate the utility function with respect to effort and set the result equal to zero:

$$\frac{\partial u^A}{\partial e^A} = \gamma - e^A = 0$$

After rearranging, we find that $e^A = \gamma$. The contract is optimal: each member maximizing *their own* utility implements the allocation that maximizes the *sum* of their utilities, as the Mechanism Designer intended.

CHECKPOINT 16.10 **Team production and public goods** Go back to Figure 16.4 which contrasts the socially optimal level of contribution to the public good with that which the citizen will provide in the absence of a subsidy.

a. Draw a similar figure with team members' effort on the horizontal axis and dollars measuring marginal private and social benefits, and private costs of effort on the vertical axis.

b. Draw in lines for the marginal cost of effort, the marginal private benefit of effort (when team members receive x/n as their income), and marginal social benefits.

c. Use the figure to explain why paying the each team member x/n results in a coordination failure and why paying them $x-k$ eliminates the coordination failure.

d. The social benefit of the public good per person ($na^{N}\gamma$) increases as the size of the population increases (with each citizen contributing an unchanged amount, a^{N}) but this is not the case in the team production case where the total output per person is just $e^{N}\gamma$ and does not depend on the size of the team. Why does the Public Goods Game differ from the team production game in this way?

Risk and credit constraints: Mechanism design in an imperfect world

How clever! You might think. But you have probably never heard of such a compensation system in practice. To see why we do not see this kind of contract, we can introduce some real-world risk and credit market constraints on borrowing by the team members.

Suppose the team's production depends not only on the sum of the team members' efforts but also on chance events affecting production but not controlled by team members. Introducing risk changes the problem in unexpected ways. It means that in any period of time—say, week or month—the actual output of the firm—called its realized team income—may be either higher or lower than the average income of the firm over a longer period, say a year—called its expected income.

! REMINDER Shirking is the behavior of *not* exerting effort or not exerting *enough* effort. Someone who shirks is called a shirker.

Call these positive or negative chance events 'shocks' and suppose, realistically, that like each member's effort, the shocks are not observable (or at least not *verifiable*) so the members cannot determine whether the team's unexpectedly low output in some year comes from bad luck or from workers who shirk.

A positive shock may also raise the level of total output significantly. So, for teams of any significant size, each member's realized income in any period could be many times larger than the workers' next-best alternative. Remember the realized income is total output minus some large constant

M-CHECK Here is an illustration. The team has 100 members and suppose under the optimal contract they would each provide one unit of effort. Remember the ith member receives $y^i = \gamma \cdot (e^1 + \ldots + e^{100}) - k$, and γ can take two values: 1.05 or 0.95, occurring with equal probability, so expected value of the team's output is 100. Then, supposing that $k = 99$, a worker's average pay over many years is one unit. But, in a year with a negative shock ($\gamma = 0.95$), the team's output would be reduced to 95. Each team member would therefore not be paid because their income would be negative ($95 - 99 = -4$). They would each instead owe 4 units to the team. 4 times their expected income!

sufficient to balance the budget and consistent with satisfying members' participation constraints.

For example, if a positive shock increases γ by 5 percent and there are 100 members of the team, then the income of the worker in a period with a positive shock could be many times larger than his average income during other periods because he receives $y^i = (1.05)\gamma \cdot (e^A + \ldots + e^N) - k$ in the period with the positive shock. But the potentially large *realized* income is not the problem for the vulnerability of this mechanism.

A *negative* shock could mean that the worker's realized income could also be much *less* than what he would have received in another job where he receives a wage rather than having income that depends on other people's effort and positive or negative shocks. With a negative shock, his income could, in fact, be a large negative number, meaning that the member would have to *pay the team* a substantial amount of money rather than the team paying the member.

The problem arises because, for this mechanism to work, the pay of each member has to be tied to the *entire team's realized output*. But, both negative and positive shocks to total output would realistically be much greater than any person's average compensation. A contract under which, in some periods, a team member would not be paid and instead would be required to pay the team a substantial multiple of her expected salary is not likely to attract many workers even if they were only modestly risk averse. No contract of this type would be voluntarily accepted by the team members.

If the team members could borrow an unlimited amount of money in the bad periods they might be okay with such a contract. But we know from Chapter 12 that this is impossible given that people of modest wealth are either credit market-excluded or limited in how much they can borrow. So, implementing the optimal contract for the team members just displaces the contractual challenge to the analogous problem in the credit market: the incompleteness of contracts.

This is a reminder that economists, policymakers, and citizens concerned about economic policy all work in a *second-best* world, one in which ideal models with Pareto-efficient Nash equilibria are important in teaching basic ideas, but often differ importantly from the best that the policymaker can do. The impossibility of implementing an optimal contract in the case of team production is an example. Another is the fact the incentives on which the Mechanism Designer relies can sometimes be *counterproductive*.

CHECKPOINT 16.11 **How team projects are (optimally) graded** Consider a team project in which five students conduct some research and write up a single paper. The students all receive the same grade on the paper depending on the teacher's assessment of its quality. Explain how this system of 'compensation' is similar to or differs from the optimal compensation for the team production members described above.

16.11 THE LIMITS OF INCENTIVES: CROWDING OUT AND CROWDING IN

As you have seen from this chapter, mechanism design is all about incentives. And we have seen that well-designed incentives can motivate citizens to contribute to a public good, to smoke less, and (in previous chapters) to eat healthier foods and act in more environmentally sustainable ways. But we have also seen that there are *limits* to what incentives can accomplish:

- In the double auction set up in section 16.8 there are no incentives that can make telling the truth a dominant strategy. As a result of buyers and sellers misrepresenting their true values, some potential mutually beneficial exchanges will not occur.
- In the team production case in section 16.9 there does exist a set of incentives that would motivate all members to contribute a Pareto-efficient level of effort to the teams' production. But the optimal contract that would introduce these incentives would expose team members to extraordinary risks, and risk-averse team members would never accept the mechanism.
- And recall from Chapter 2 that imposing a fine on parents arriving late to pick up their kids at Haifa daycare centers led to a doubling of lateness, not a reduction; the introduction of incentives apparently *crowded out* social preferences.

! REMINDER Motivational crowding out occurs when monetary or other material incentives or attempts to control someone diminish that person's other-regarding or ethical preferences.

› EXAMPLE In honoring Myerson (along with Eric Maskin and Leo Hurwicz) the Prize Committee of the Royal Swedish Academy of Sciences in 2007 summarized their findings: "no incentive compatible mechanism which satisfies the participation constraint can produce Pareto-efficient outcomes… In a large class of models Pareto efficiency is incompatible with voluntary participation, even if there are no public goods."[14]

Crowding out

The most plausible explanation of crowding out in the Haifa case is that the fine changed the way the parents thought about lateness. Before the fine, parents may have considered arriving on time to be something that is the "right thing to do" out of kindness to the staff of the daycare center. By contrast, the fine may have suggested that lateness is simply something that is OK to "buy" as long as you pay the price. In other words the message of the fine was "picking up your kids on time is like shopping, not like how you treat family or neighbors." The incentive provided a frame for the decision suggesting appropriate behavior. The result is what psychologists call **moral disengagement.**

There are additional reasons why crowding out occurs. Incentives have a purpose, and because the purpose is often evident, the target of the incentives may infer information about the person who designed the incentive,

MORAL DISENGAGEMENT A process by which, in some particular situations, people come to feel that ethical considerations need not be applied to their own actions or others' actions.

about his beliefs concerning her (the target), and about the nature of the task to be done. This is what the social psychologist Mark Lepper and his coauthors meant when he wrote that incentives may affect preferences: because they indicate "the presumed motives of the person administering the reward."[15]

By implementing an incentive, an employer reveals information about his intentions (own-payoff-maximizing versus fair-minded, for example) as well as beliefs about the target (hardworking or not, for example) and the targeted behavior (how unpleasant or tiring it is, for example). This information may then affect the target's motivation to undertake the task.

The Boston fire commissioner's threat to cut the pay of firemen accumulating more than 15 sick days (also from Chapter 2) conveyed the information that he did not trust that the firemen were doing their very best to come to work. For the firemen, the new situation—working for a hostile boss—seems to have altered their motivation. In other words, the threatened reduction in pay conveyed some bad news: "the commissioner does not trust you." You will recall that the firemen responded with a pronounced spike in sick days.

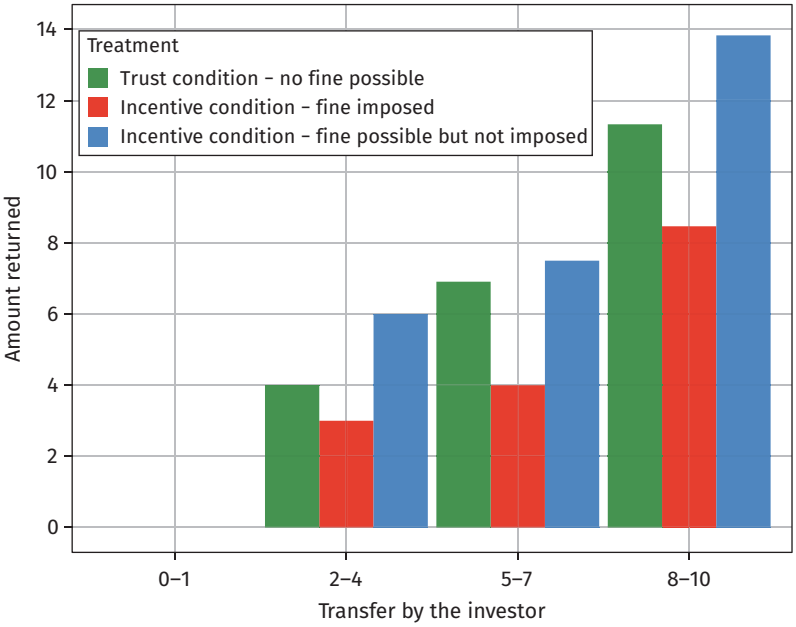
This "bad news" effect commonly occurs in relationships between a principal, who designs incentives (a wage rate, a schedule of penalties for late delivery of a promised service, a system of monitoring, and so forth), and an agent, who is being induced to behave more in the principal's interest than the agent otherwise would. To do this, the principal must know (or guess) how the agent will respond to each of the possible incentives he could deploy, for example the employee's best (effort) response to each wage the employer may offer. The agent knows this, of course, and hence can ordinarily figure out what the principal was thinking when he chose one particular incentive over other possible ways of affecting the agent's behavior.

Here is an example of how this sometimes does not work out well in practice, based on an experiment by economists Ernst Fehr and Bettina Rockenbach. German students in the role of 'investor,' the principal, were given the opportunity to transfer some amount to the agent, called the 'trustee,' also their fellow students. The experimenter then tripled this amount. The trustee, knowing the investor's choice, could then return some (or all or none) of this tripled amount.

When the investor transferred money to the trustee, he also specified a desired level of back transfer. In addition, the experimenters implemented an incentive treatment: in some of the experimental sessions, the investor had the option of declaring that he would impose a fine if the trustee's back transfer were less than the desired amount.

In this "fine treatment," the investor had a further option to decline to impose the fine, and this choice (forgoing the opportunity to fine a nonperforming trustee) was known to the trustee and taken before the

Figure 16.14 Reciprocity in the Trust Game. The height of each bar shows how much the second mover (the "trustee") returned given how much the first mover (the "investor") transferred to the second mover. The figure shows that in the standard Trust condition (no fine possible) the second mover reciprocated the first mover's trust and/or generosity: larger initial transfers were associated with larger back transfers. It also shows that threatening the use of a fine reduced reciprocity and that foregoing the use of the fine when it was available increased reciprocity.



trustee's decision about the amount to back-transfer. There was also the standard "trust" condition, in which no such incentives were available to the investor. Figure 16.14 summarizes the results.[16]

In the trust condition, trustees reciprocated generous initial transfers by investors with greater back transfers. But stating the intent to fine a noncompliant trustee actually *reduced* return transfers for given levels of the investors transfer. The use of the fine appears to have diminished the trustees' feelings of reciprocity toward the investor. Even more interesting is that renouncing use of the fine when it was available increased back transfers (given the amount transferred by the investor).

Only one-third of the investors renounced the fine when it was available; their payoffs were 50 percent greater than those of investors who used the fines. The bad-news interpretation suggested by the authors of the experiment is that both in the trust condition and when the investor renounced the fine, a large initial transfer signaled that the investor trusted the trustee. The threat of the fine, however, conveyed a different message

! REMINDER We referred to the ideas of control aversion in Chapter 11 when we explained the role of monitoring software and how workers might reciprocate distrust by employers with less effort or that workers would dislike being controlled at all (regardless of reciprocity motives).

✓ FACT CHECK In a famous experiment in social psychology, children who did paintings expecting a monetary reward subsequently chose to paint (as opposed to some other kind of activity) half as often as did children in a control group in which there had been no experience of a reward. This is an example of a monetary reward crowding out intrinsic preferences (the joy of painting).[17]

✓ FACT CHECK During the COVID-19 pandemic of 2020–2021, Germans surveyed about their response to policies to combat the spread of the virus reported greater support for vaccination and a cell-phone-tracing app if these are "strongly recommended" rather than legally enforced. Economist and psychologist Katrin Schmelz interpreted her results as evidence of control aversion. Those who had been raised in East Germany under Communist Party rule were less control averse than other Germans of similar age.[18]

like the Boston fire commissioner's—one of distrust—and diminished the trustees' reciprocity.

There are lessons here for the design of institutions and organizations. Crowding out as a result of the bad-news effect may be prevalent in principal-agent settings but can be averted where the principal has a means of signaling fairness or trust in the agent. The Trust Game experiment even gives us a glimpse of how incentives could crowd *in* social preferences: the nonuse of an available fine resulted in greater reciprocity by the trustees than occurred when fines were not in the picture.

A third reason for crowding out is that people value their own autonomy and may feel that incentives are designed to limit their freedom of action, provoking a negative response. This is called "control aversion" which, analogous to risk aversion, is a preference for self-determination and a negative valuation of any attempt to control the person.

To summarize, the problem of crowding out may arise:

- when the presence of the incentive frames the problem as one in which self-interested motives are acceptable or even called for ("moral disengagement"); or …
- when the information that an incentive conveys is off-putting about the person imposing the incentive ("bad news"); or
- when the incentive appears to be an attempt to control the person who responds by acting contrary to the incentive in order to affirm their self-determination ("control aversion").

The economics and the psychology of "getting" and "being"

A combination of game theory and social psychology may help us better understand how incentives work and sometimes why they do not, that is, why crowding out occurs.

Recall from Chapter 2 that in a sequential Prisoners' Dilemma, the second mover most often mimics the first mover, reciprocating cooperation or defection depending on what the first mover did. Why do second movers reciprocate cooperation even though their money payoffs would be larger if they defected? It appears that they place a positive value on the payoffs that the cooperative and trusting first mover will receive (as a result of their reciprocating). Or they may think that in situations like this, reciprocating is the right thing to do.

When second movers defect on defecting first movers, they are taking the action that maximizes payoffs under the circumstances, but the motives for defection go beyond seeking monetary gain. The same people, as we have just seen, would have forgone payoffs that could have been gained by defecting, in order to cooperate with a cooperator.

But cooperating with a defector has a different meaning, identifying the second mover as a "loser," someone easily taken advantage of. Part of the

motivation behind the "mimic the first move" pattern that we observe is what the second mover wants to say about herself: "I am the kind of person who rewards those who cooperate and stands up to defectors who would exploit the cooperation of others."

This kind of motivation goes way beyond how we play games. When people engage in trade, produce goods and services, save and invest, vote and advocate policies, they are attempting not only to get things, but also to be someone, both in their own eyes and in the eyes of others. "Getting" and "being" (or "becoming") are different types of motives. The term **acquisitive motives** refers to the positive value we place on "getting" (acquiring) material things, like monetary payoffs in games. **Constitutive motives** are our desires to "be" or "become" a particular kind of person (to constitute or construct ourselves in a particular kind of way).

Sometimes constitutive and acquisitive motives are closely aligned. For example, the second mover in the sequential Prisoners' Dilemma game who defects on a defecting first mover is both making a statement about who she is (not someone you can take advantage of) and maximizing her payoffs. But, acquisitive and constitutive motives may sometimes clash.

We know from experiments (in Chapter 2) and from observing ourselves and others that being good sometimes is more important to us than doing well in monetary terms. Responding to an incentive in the manner intended (that is, as a money payoff maximizer) may make the responder a victim, so the incentive may not work. This was the case for the use of fines in the Trust Game experiment just described. But not always. A self-interested response to an incentive may constitute the actor as a good citizen or an intelligent shopper, indicating that constitutive and acquisitive motives were closely aligned. The same reasoning, we will see, suggests how we can make incentives and social preferences synergistic.

How acquisitive ends interact with constitutive motives may explain why incentives sometimes work exactly as economists predict on the basis of unmitigated self-interest—and sometimes don't.

Figure 16.15 Rachel Kranton (1962–) as a new PhD wrote Nobel Laureate George Akerlof to say that one of his recent papers was missing an essential concept. The two economists went on to pioneer a new field—identity economics—that places the desire to be a particular kind of person—that is to construct and maintain one's identity—at the center of many of our decisions. Kranton has also explored the economics and neuroscience of how group identities affect the workings of our brain and decision-making, and the political psychology of group identities and biases.[19]

Photo © Duke University.

Crowding in

The good news about crowding out is that there is something to crowd out, namely people's other-regarding preferences and desire to do the right thing. The problem of crowding out would not arise if we were really like *Homo economicus*—self-regarding and amoral—so that there would be nothing to crowd out.

ACQUISITIVE MOTIVES These are the positive value we place on "getting" (acquiring) things, for example, monetary payoffs in games.

CONSTITUTIVE MOTIVES These are our desires to "be" or "become" a particular kind of person (to constitute or construct ourselves in a particular kind of way).

In the Trust Game described above, the 'investor's' threatened imposition of a fine if the 'trustee' did not return enough money to the investor reduced rather than increased the amount the trustee sent to the investor, consistent with the "bad news" interpretation of crowding out. But on closer scrutiny the incentive itself—the threatened fine—seems not to have been the problem.

When we looked closely at the raw data from the experiment to see who among the trustees responded negatively to the incentive, it appears that crowding out was almost exclusively a reaction, not to the incentive per se, but to the apparent greed of the investor.[20] Crowding out occurred when the back transfer demanded of the trustee by the investor along with the threat of the fine would have given most of the sum of the rents from the game (total payoffs for the two) to the investor. There was no backlash against the fines threatened by investors who asked for back transfers that allocated both the investor and the trustee substantial shares of the total rent.

The key difference was the message sent by the fine. Where the stipulated back transfer would have captured most of the rent for the investor, the fine conveyed greed. Where it would have split the total rent more equally, the fine conveyed a commitment to fairness, and perhaps the investor's desire not to be exploited by the trustee. The use of the fine to enforce a seemingly unfair demand provided an acquisitive motive to comply, but to the trustee, it also may have transformed the meaning of complying with the investor's stipulated back transfer. Going along with the investor's demands no longer made the trustee a cooperative and ethical person, as it would have had the investor's demands been modest, but instead possibly a person easily manipulated, or a victim, or a loser.

It therefore appears that it was the conflictual relationship between the investor and the trustee that the threatened fine created, not the threatened fine itself, that was the source of crowding out. That hypothesis is reinforced by evidence from another experimental game: the diametrically opposite reaction to fines in a Public Goods with Punishment experiment that you saw in Chapter 2. The imposition of fines by peers who have to pay to levy them, when they had nothing to gain personally from doing so, appears to have crowded *in* social preferences.

Why is the fine counterproductive when imposed by an acquisitive and overreaching investor in the Trust Game, but, by contrast, so effective when imposed by peers in the Public Goods Game? A plausible explanation is that when punished by a peer who had nothing to gain in payoffs by doing so, players saw the fine as a signal of public-spirited social disapproval by fellow group members. If this were the case, targeted free riders would feel shame, which they would redress by contributing more. If so, the incentive (the prospect of peer-imposed fines) has crowded in social preferences.

We have seen that both experimental evidence and observations outside the laboratory show that incentives can affect people's preferences (e.g. moral disengagement) and beliefs (e.g. "bad news" about the principal). This complicates the Mechanism Designer's task because implementing a desirable allocation now requires that it be a mutual best response once the mechanism is introduced taking account of the manner in which the mechanism itself may change the preferences and beliefs of the actors.

CHECKPOINT 16.12 Crowding out: reasons and responses

a. Can you think of examples from your own or others' experiences that illustrate the three reasons for crowding out (above)?

b. In these examples is there some way that the incentive could have been explained or framed that would have been more likely to produce the intended response?

16.12 BEYOND "MARKET VERSUS GOVERNMENT": EXPANDING THE SPACE FOR POLICIES AND INSTITUTIONS

A theme of this chapter has been the promise of mechanism design. We have also introduced cases in which policies and other interventions to address coordination problems will have limited success—however wisely guided by the Mechanism Designer.

Government versus market

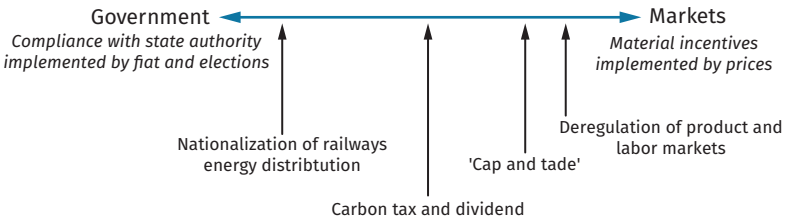
To do this we have contrasted two scenarios. In the first, under some status quo rules of the game self-regarding people interact in noncooperative ways—unable to coordinate their actions in any way. As a result, too much smoking, not enough of the public good, or some other undesirable outcome are the Nash equilibrium. In the second scenario, the Mechanism Designer acting as a government policymaker, proposes a new set of rules of the game in which self-regarding citizens still playing noncooperatively will reach an improved outcome. These two scenarios represent poles of a continuum which are often referred to as "market" (the first scenario) and "government" (the second). In this framework the classical institutional challenge raised by the eighteenth-century founders of economics—how shall we organize our society?—is often posed as a question about the extent to which markets or governments should determine economic outcomes.

HISTORY In 1986, economist Robert Sugden described the Mechanism Designer and policymaker in somewhat unflattering terms: "like the U.S. Cavalry in a good Western [film], the government stands ready to rush to the rescue whenever the market 'fails' and the economist's job is to advise it on when and how to do so. Private individuals, in contrast, are credited with little or no ability to solve collective problems among themselves."[21]

The question motivating these "market versus government" debates, then is where policy should be located along the market-state continuum shown as the horizontal line in Figure 16.16. On the market side of the line we have the policy of removal of government regulations concerning functioning of markets (called "deregulation"). Closer to the government pole we have

Figure 16.16 The government versus markets poles and policy continuum.
Positions on the line are policies or rules of the game that resemble markets (to the right) or governments (to the left).

Source: Bowles and Carlin (2020b).



the government ownership ("nationalization") of critical infrastructure. The carbon tax and dividend proposal that you studied in Chapter 7 is midway between the poles as it is a major intervention into the pricing of carbon but it does not include direct government regulation of the carbon emissions of particular firms.

At the poles of the continuum we list the incentives and mechanisms implementing outcomes, to illustrate the distinctiveness of the poles. For example, what a democratic government implements is determined by elections and the laws and regulations (fiats) that the government adopts. Motives essential to this pole include obedience to government authority. Similarly, markets implement allocations by means of prices and the material incentives that they convey.

EXAMPLE In this video (tinyurl.com/y57ulqm9), Michael Sandel investigates some moral limits of markets in his TED Talk "Why we shouldn't trust markets with our civic life" (from the CORE project. www.core-econ.org).

A new job for the Mechanism Designer

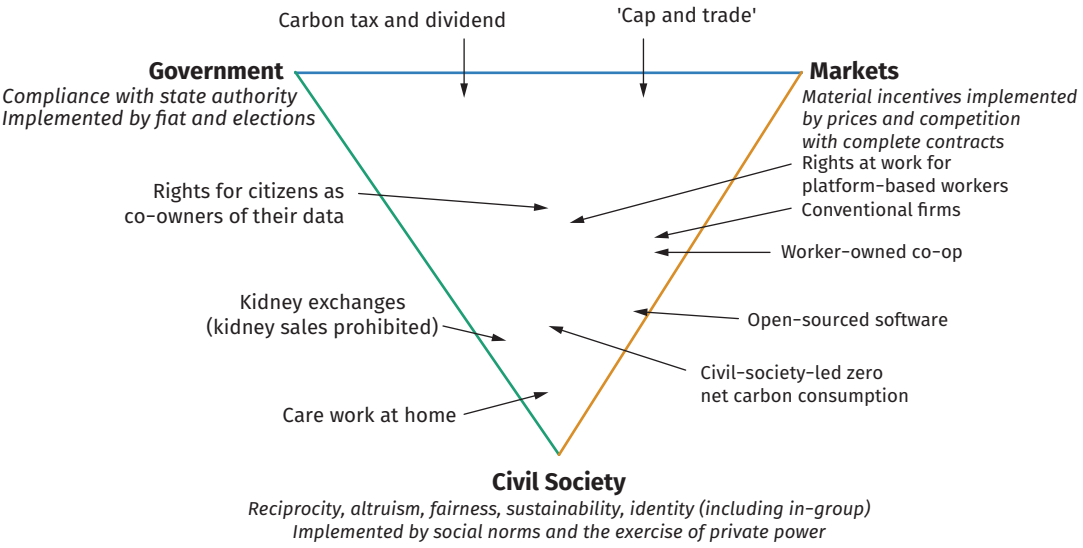
In previous chapters we have described ways of addressing coordination problems—both motives and mechanisms—that do not really lie on the blue government vs. markets continuum.

- In Chapter 5 we showed how repeated interactions among people who are overfishing their fish stocks might provide the basis for their coordinating to fish less, entirely in the absence of government policies.
- In Chapter 11 we represented the firm as a kind of hierarchical organization that while influenced by markets and governments is an entirely different kind of institution.
- Also in Chapter 11 we showed how a worker-owned and managed firm could be an alternative organization of production, similar to the conventional firm in that it is neither a government nor a market.
- In Chapter 14 we illustrated the process of bargaining between private parties and how this could resolve some common coordination failures associated with uncompensated external effects (noisy machinery, late-night music).

EXAMPLE In this video (tinyurl.com/y6s33jo5), Alvin Roth explains how the kidney exchange and other matching markets work (from the CORE project. www.core-econ.org).

Figure 16.17 Government, market, and civil society: an expanded space for policies and institutions. At each of the poles are shown the kinds of motivations and mechanisms implementing outcomes that are characteristic of the poles.

Source: Bowles and Carlin (2020b).



We also showed, in Chapter 2, that in addition to the motives we associate with governments and markets in Figure 16.16—obedience to governments and material self-interest—there are other motives that better describe many aspects of how we interact in face-to-face relationships with neighbors, friends, family, and coworkers. Included are altruism, fairness, commitments to uphold ethical or religious norms, and the desire to establish an identity (including sometimes hostility toward those with different identities). We therefore expand the "government versus markets" framing of our choices in Figure 16.17 by adding a third dimension, called civil society. It might also be called community. Critical to the functioning of institutions that are close to the civil society pole are:

- social preferences such as reciprocity, altruism, and identity; and
- the private (nongovernmental) exercise of power of the type that occurs in principal–agent relationships as we have shown in Chapters 10, 11, and 15.

The two poles—market and government—describe a line, but with the added third pole, the figure represents a triangle enclosing a space. Any point in the space defined by the triangle represents some combination of the "pure cases" depicted at the poles. Just as was the case with the line, in the triangle the closer a point is to one of the poles, the more that policy or institution resembles the rules of the game at the pole.

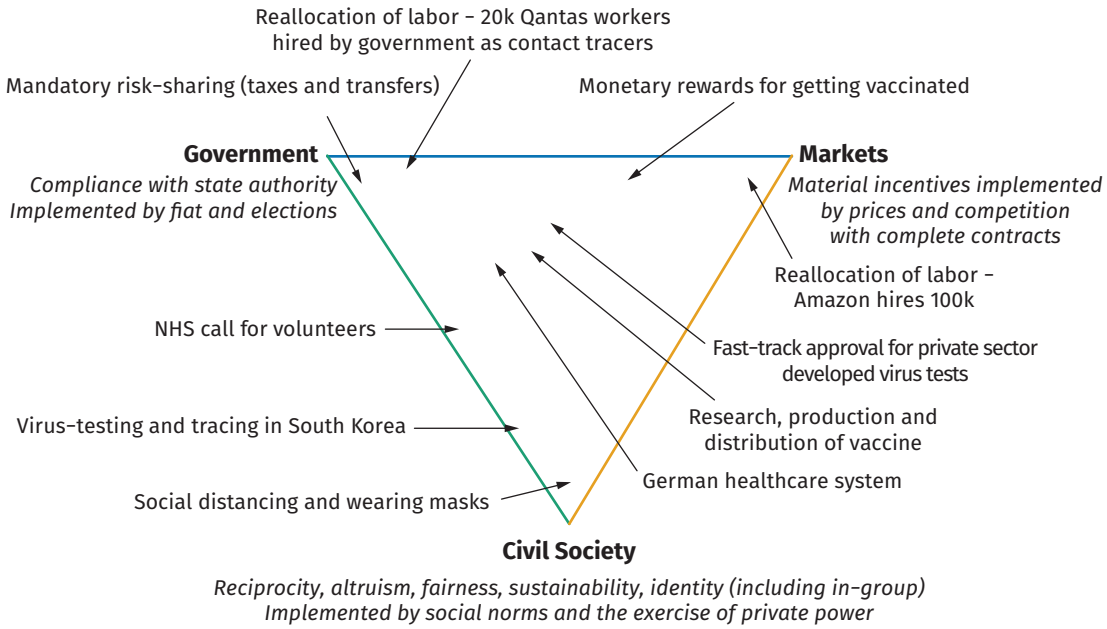
In the interior of the triangle are firms, both worker-owned and conventional firms. So, too, are social movements to promote new lifestyles such as zero net carbon consumption. "Care work at home"—raising one's children, assisting elderly relatives—is also there, as it is neither mandated nor regulated by governments nor does it involve the buying and selling of services.

The kidney exchange shown is not government-organized, but it relies on a government ban on kidney sales. To improve the possibility of finding a match, donors can provide a kidney and establish a claim on a matched kidney from the pool of kidneys for a loved one.

Figure 16.18 illustrates a range of policies adopted during the 2020-2021 COVID-19 pandemic, showing the various combinations of aspects of government, market, and civil society. Social distancing, for example is not something that either markets or a government could effectively

Figure 16.18 Illustrating the new policy space by policies adopted during the COVID-19 pandemic of 2020–2021. Social distancing is an example of policies close to the civil society pole: it could not effectively be required by law, or incentivized by paying people to respect distancing protocols. Where people's values and opportunities to share their concerns about violations of the practice were sufficient social distancing was very effective. The UK National Health Service (NHS) requested 250,000 volunteers to help distribute food and medicine. Recruitment was temporarily shut down four days later after 750,000 people had applied.

Source: Bowles and Carlin (2020a).



implement, but where it had support from large numbers of the members of a community, it became a norm that was widely adhered to.

In this new space of institutions and policies, the Mechanism Designer has an expanded job description. Her new job specifications go beyond taxes, subsidies, legal remedies and prohibitions that governments may propose to alter the feasible sets and benefits and costs of market activities. She will have to consider the formal or informal rules of the game governing the institutions of civil society—whether they be families, conventional firms, neighborhoods, or cooperatives—and how these interact with governments and markets.

CHECKPOINT 16.13 Government, market, and civil society In Figure 16.17 locate in the triangle some organizations of which you are a part, for example the university at which you are studying or a workplace in which you or someone you know has been employed.

16.13 RULES FOR REDISTRIBUTION: ADDRESSING COORDINATION FAILURES

Some of our illustrations of both the promise and the limitations of mechanism design have been of homogeneous populations: identical citizens providing public goods and identical smokers. In other cases we have distinguished between economic actors—buyers and sellers in the double auction, the owners of the monopolistically competitive firm in the model illustrating the theory of the second best, interacting with its buyers and those affected by the firm's carbon footprint.

But we have not yet presented the Mechanism Designer with the challenges arising from an essential dimension of differences among people, that is, disparities in wealth and power. The Mechanism Designer may be tasked with addressing these inequalities, because they are seen to be either unfair or inconsistent with democratic values.

As always, the Mechanism Designer has the objective of avoiding Pareto-inefficient outcomes. Can the Mechanism Designer provide guidance—what we call rules of redistribution—for policies to advance the objectives of both fairness and Pareto efficiency?

Some of the cases we have encountered would suggest a negative answer.

- *Concentrated wealth and power can support a Pareto-efficient Nash equilibrium.* In Chapter 5 private ownership of the common property resource (the lake) by a single individual resulted in a Pareto-efficient but maximally unequal Nash equilibrium. The same result followed giving one of the fishermen take-it-or-leave-it bargaining power.
- *Policies to distribute income to the less well-off may reduce average income.* In Chapter 8 we showed that San Francisco's rent control policy, while

! REMINDER Inequality and innovation: Taking a more dynamic view of efficiency as our standard, we showed in Chapter 15 that the structure of the capitalist firm—with control rights held by wealthy and hence closer to risk-neutral decision makers could promote innovation. But we also showed that some of the most innovative countries in the world today have among the most equal distributions of disposable income.

! REMINDER In Chapter 3 we showed that if citizens have identical cardinal preferences that are comparable among them, then an Impartial Spectator that wished to maximize the sum of their utilities would implement an equal distribution of wealth.

! REMINDER The residual claimant owns whatever remains (the residual) after all fixed claims (in this case the wage paid by the owner) are settled.

› EXAMPLE In this video (tinyurl.com/yy8ggzwt), James Heckman describes why investing in the early years of disadvantaged children's lives is both fair and efficient (from the CORE project. www.core-econ.org).

conferring substantial benefits on lower-income renters reduced the stock of rental housing; giving the renters who benefited a larger slice or a smaller pie.

The cost of inequality when contracts are incomplete

But other models we have studied show that it is possible to increase the size of the slice going to the less well-off while *at the same time* enlarging the pie. To see how, return to a central theme of Parts III and IV of this book. This theme is that–whether in the credit, labor, or other markets–Pareto-inefficient incentive structures are built in to the principal-agent relationships in economies based on highly unequal wealth distributions.

The source of this incentive inefficiency is that the agents performing actions that cannot be specified in a complete contract–providing high quality of a good produced, hard and careful work effort, prudence in the conduct of a project, for example–are *not* the residual claimants on the consequences of their actions. They *do not own* the results of hard work done (in the Ford model in Chapter 11), or a high-quality product delivered to a principal (in the Benetton model in Chapter 10). Incentives to perform the action well are compromised as a result.

An example from Chapters 11 and 12 will illustrate why this is the case and why wealth inequality makes addressing the incentive problem difficult. Think about an owner of a "machine" (like the one in Chapter 12) who hires a single worker to operate it. The worker, we assume, has no wealth. The worker has little reason to supply a high level of effort, since she is paid a given wage; working harder will simply increase the machine owner's profits.

Without costly monitoring of the worker, therefore, little work will get done. But monitoring the worker to ensure hard work uses up resources that could have otherwise been productively employed. Moreover, if her work effort is at all complex or difficult to measure, even intensive monitoring will not ensure a job well done.

To avoid this particular incentive problem the owner might instead introduce a rental contract. The worker would rent the machine from the owner for a fixed sum and become the residual claimant on the entire income stream generated by her labor and the machine. As residual claimant she would not lack incentives to work hard and well: she would own the results.

But this solution to the effort-incentive difficulty simply displaces the conflict of interest and the problem of incentives from getting the worker to work to how the worker cares for the machine. The reason is that the worker would then be residual claimant on the income that she produces with the machine, but not on the value of *the machine itself*. So, she would have little incentive to maintain the machine, running it too fast, for example, and risking its destruction.

A solution to the problem that would certainly work, is that the worker should purchase the machine from its current owner, and thereby become the residual claimant on both the income resulting from her working with the machine, but also on the value of the machine itself. But since the worker has no wealth, she can neither purchase it using her own funds nor borrow the necessary funds. She *cannot* be the owner of the machine.

The unavoidable problem here is that behaviors critical to high levels of productivity—hard work, maintenance of productive equipment, risk-taking, the production and use of knowledge, and the like—are difficult to monitor and hence cannot be fully specified in any contract enforceable at low cost. As a result, key economic actors, workers, people providing goods of variable quality, for example, cannot capture the effects of their actions as they would if, for instance, they were residual claimants.

To become residual claimants they would no longer be agents acting at the direction of a principal, but instead owners, acting as independent economic actors. But they cannot become owners because they do not have enough wealth. The consequence of these incentive problems is that a highly concentrated ownership of capital goods is often inefficient.

There is a second reason—apart from the associated incentive problems—why high levels of wealth inequality make it difficult to address the coordination failures arising in principal-agent relationships. The reason is that the kinds of preferences that are likely to prevail in highly unequal settings heighten the conflicts of interest and worsen the barriers to cooperation.

To see why this is important, remember that coordination failures arise because people do not take account of the effect of their actions on others, and are engaged in a noncooperative game, a game in which they are *unable* to agree on a common course of action. If people are intrinsically motivated to do a good job, to tell the truth, and to care about and trust one another, then they can more easily deal with the coordination failures that result from contractual incompleteness. These preferences could facilitate their internalizing the external effects of their actions on others, and provide the basis for common agreements that together would go a long way toward mitigating the incompleteness of the contract. But these preferences are difficult to sustain between principals and agents of vastly differing wealth levels. An example is the fact—in the Trust Game experiment reported above—that feelings of reciprocity, which could form the basis for less conflictual and more cooperative relationships, were crowded out by "investors" claiming what appeared to be an unfair share of the mutual gains made possible in the experiment.

Summing up: recall that in section 14.15 we showed that the Coase theorem and the second welfare theorem support the conclusion that a Pareto-efficient Nash equilibrium can be implemented irrespective of the distribution of wealth. This gives the Mechanism Designer—and economists—a reason to believe that they can implement Pareto-efficient outcomes without addressing the problem of inequality.

HISTORY In 1971 Kenneth Arrow wrote: "It is useful for individuals to have some trust in each other's word. In the absence of trust it would be very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be forgone . . . norms of social behavior, including ethical and moral codes [may be] . . . reactions of society to compensate for market failures."[22]

! REMINDER In section 15.14 we showed that a redistribution of wealth could support a Nash equilibrium with ownership of capital goods by the workers who operate them—either as owner operators or as a member of a cooperative. The wealth distribution reduced both the risk aversion and the opportunity cost of capital for the once wealth-poor worker, promoting greater risk-taking and as a result higher expected income.

But we have seen in the previous chapters that the nature and distribution of property rights critically affect the performance of the economy including along the dimension of Pareto efficiency. Where hard work, innovation, maintenance of an asset, and other behaviors essential to productivity cannot be specified in costlessly enforceable contracts, some distributions of property rights are more efficient than others. In particular it seems likely that there exist distributions of wealth that are both more egalitarian and more effective in avoiding coordination failures than the concentrated asset-holding observed in most capitalist economies.

The cost of inequality: Enforcing the rules of the game

Broadening our focus from principal-agent relationships to look at all of the social interactions making up a society, there is a further reason why inequality is costly. The cost is the substantial resources that highly unequal societies devote to sustaining the rules of the game and containing conflicts between the society's members.

In the US, for example, a large fraction of the economy's productive potential—its labor and capital goods—is devoted to enforcing the rules of the game from which the inequalities flow. The people who enforce these rules of the game are collectively called "guard labor"—soldiers, police officers, prison guards.[23]

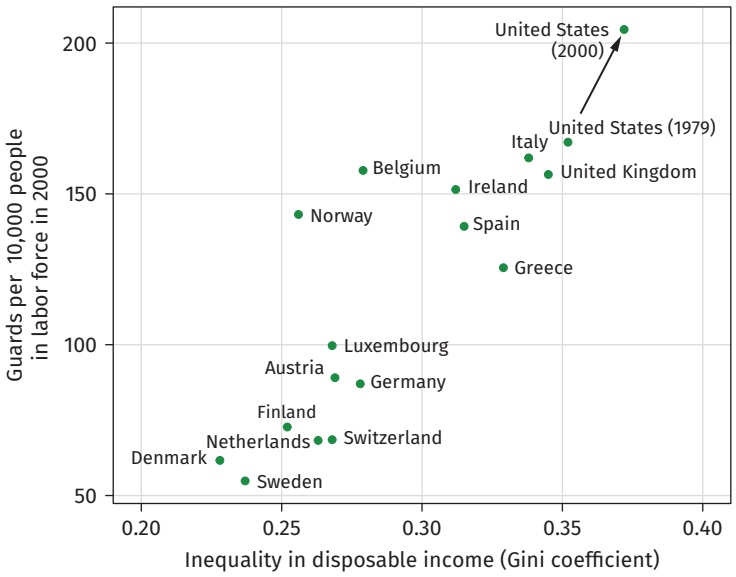
The private sector also incurs costs in enforcing inequality, in such forms as high levels of expenditure on work supervision and security personnel. One could add the labor devoted to producing the weapons and other equipment used by these private and public security personnel. Indeed, one might count levels of unemployment above what is termed frictional unemployment itself as one of the enforcement costs of inequality, because the threat of losing a job contributes to employers' labor discipline strategies. In less conflictual conditions, unemployed workers might be allocated to productive activities.

In Figure 16.19 we present data on the number of people employed literally as guards (not the broader "guard labor" category above) and the level of inequality of disposable income in the set of countries for which comparable data are available. There is an unmistakable positive association of inequality and the level of guard labor.

CHECKPOINT 16.14 **The costs and benefits of inequality** You are assigned to debate pro or con on the following proposition: Micro-economic theory shows that high levels of inequality are necessary to provide incentives for a highly productive economy. Take one side or the other and list the main points that you will raise. (The question is about the theory; it is not about the facts).

Figure 16.19 Employment of guards and inequality in disposable income. The measure of inequality is the Gini coefficient of household disposable income based on estimates in the early 2000s. Guards are technically termed "protective service employees." The earlier data for the US is shown to check if the movement in these two measures over time in a single country resembled the positive statistical association of the cross section of countries at the same time. We can see (the arrow) that it does.

Source: Bowles and Jayadev (2014).



✓ **FACT CHECK** The positive correlation shown in Figure 16.19 is not sufficient to establish that inequality is the *cause* of the heightened level of employment of guards. For this we would need data at differing points of time to see if the employment of guards rises at times when (or soon after) inequality has risen. The only country on which these data are available is the US where it is the case that the recent *increase* in employment of guards coincided with an *increase* in inequality (the two US points in the upper right of the figure for 1979 and 2000).

16.14 WHY GOVERNMENTS ALSO FAIL: A CAVEAT

This chapter has demonstrated how the concepts and models you have learned can be used to design rules of the game, whether for private organizations or as new public policies. But before concluding, we need to walk back our too-good-to-be-true representation of the Mechanism Designer. The imaginary figure we introduced skillfully deployed the tools of economics to address market failures and unfairness in the economy. But you may have wondered: If the clever mechanisms that she designed are available, then why are market failures so common and why is unfairness such a feature of real economies?

The short answer (which is all that we can provide here) is that the Mechanism Designer is not the government. A conceptual device for understanding ideal policies is not going to be an accurate representation of what governments actually do, any more than the Auctioneer describes how markets work, or the "perfect competitor" gives an accurate picture of how firms really compete.

HISTORY In a 1896 essay the Swedish radical reformer and founder of neoclassical economics, Knut Wicksell (1851–1926) challenged the common assumptions that policymakers are entirely public spirited "with no thought other than to promote the common weal… members of the representative body are… precisely as interested in the general welfare as are their constituents, neither more nor less." A century later the libertarian James Buchanan in his Nobel Prize lecture underlined this theme: "Wicksell's message was clear, …Economists should cease proffering policy advice as if they were employed by a benevolent despot."[24]

We have stressed what governments can do. Well-designed government economic policies introduced over the past century–things like social insurance and the stabilization of aggregate demand in the macroeconomy–have contributed substantially to the quality of people's lives. And economists can be proud of our contribution to these advances.

But there remain significant instances of market failure–global climate change, for example–and also unfairness–disparities between the well-off and the deprived, group inequalities based on accidents of birth. This is not a limitation of mechanism design but a government failure.

To understand the difference keep in mind:

- The Mechanism Designer is tasked with maximizing the utilities of citizens or at least avoiding outcomes that are Pareto inefficient when evaluated using the citizen's utility functions. Policymakers in governments have their own private objectives, and there is no reason to expect them be any more (or less) other-regarding than the citizenry from which they were selected.
- The mechanisms in place to ensure that government policies are designed to promote the interests of citizens (rather the policymakers themselves)–in a democracy, for example, majority rule in fair elections–work imperfectly. This is true in part for intrinsic unavoidable reasons, just as the failure of the double auction market stemmed from an intrinsic problem: the impossibility of ensuring truth-telling. But the limits of democratic mechanisms also arise in part because governing elites often succeed in subverting rules of the game intended to make them accountable to the citizenry.
- A successful mechanism implements its outcome by the use of incentives in ways that are consistent with the citizens' participation constraint (as well as their incentive compatibility constraints). The Mechanism Designer cannot propose policies that literally force citizens to contribute to the public good, or to stop smoking. By contrast, the distinctive characteristic of the government is that it can legitimately attempt to coerce people to do or not do particular actions: laws are enforced by the threat of incarceration, that is forced confinement.

HISTORY In 1919, Max Weber one the founders of sociology, defined the state (meaning, government) as the only entity that has a legitimate monopoly on the use of force within some given geographical territory. This is still the standard definition used in the social sciences.[25]

This fact that democratically elected governments can legitimately attempt to coerce is essential to the government doing its job well–requiring the payment of taxes or the universal purchase of health insurance, or, when required, service in the military, for example. But it also means that the scope for harm done by a government is substantial.

Fortunately we also know, and more important, most political leaders know, that compliance with government directives–whether it be honestly paying taxes or social distancing during a pandemic–depends critically on the citizenry trusting the government and endorsing its objectives.

CHECKPOINT 16.15 Government: part of the solution/part of the problem "Any government with sufficient powers to address problems of coordination failures and unfairness will also be powerful enough to pursue the interests of a political elite against the interests of the vast majority." Comment on this idea, and suggest ways of addressing the problem.

16.15 CONCLUSION

We began this book with the classical institutional challenge posed by Adam Smith, David Hume, and the other great philosopher-economists of the eighteenth century: How should society be organized?

We used modern economic language to update an important strand of the challenge they posed: "How can social interactions be structured so to avoid Pareto-inefficient Nash equilibria resulting from people's free choice of their own actions?" Achieving a just distribution of the economy's burdens and bounties was a second strand of the challenge they raised that we have also taken up.

Mechanism design carries on this tradition in economics (though substantially narrowed in its scope and made more precise with the aid of mathematics), seeking new rules of the game to better coordinate how we interact.

The eighteenth- and nineteenth-century outpouring of new ideas in the field of studies we now call economics was a response to a rapidly changing environment. New technologies introduced by the Industrial Revolution, new types of social interaction associated with the emergence of urban life and factory production, and new values of individual autonomy stemming from the Renaissance and the first halting steps toward political democracy.

Today we face no less epoch-making changes. The COVID-19 pandemic of 2020–2021 underscores the fact that humanity is increasingly interconnected in ways that go far beyond the conventional focus of economics: buying and selling goods on markets. We now face the challenges of climate change and economic injustice, and the vast opportunities afforded by our ever-expanding knowledge. These—and many others—are challenges that economics must now, as it did then, seek to understand, and then to address for the betterment of all.

MAKING CONNECTIONS

Nash equilibrium: is a key concept for policymaking because policies work by altering the environment in which people make their decisions, resulting in a Nash equilibrium that differs from the status quo.

Internalizing external effects: so as to prevent market failures and other coordination failures is a key idea in mechanism design.

Limited information: is both the fundamental reason for coordination failures (because it makes complete contracts impossible) and also a challenge facing the Mechanism Designer (or policymaker).

Fairness and Pareto efficiency: are widely shared values that people would like to see public policy advance, though people differ in the weights they would place on the two values. Some policies can advance the two together (government-provided or tax-financed subsidies for public goods or insurance to "democratize risk-taking") while in some cases the values may conflict, imposing a trade-off (rent control).

Models and abstract thinking: in economics these are illustrated by the imaginary figure of the public-spirited and well-informed Mechanism Designer; but the Mechanism Designer is not the government.

The classical institutional challenge: originated in the seventeenth and eighteenth centuries; but it is still with us as we seek to find a way that free people can coordinate their activities so as to achieve desired societal outcomes including addressing the challenges of climate change and economic injustice.

IMPORTANT IDEAS

Nash equilibrium	mechanism design	public goods
private benefits and costs	social benefits and costs	theory of the second best
social multiplier	optimal subsidy	economic profit
deadweight loss	consumer surplus	1/n rule
double auction	team production	moral disengagement
optimal contract	crowding out/in	civil society
control aversion	trust game	cost of inequality
community	guard labor	
government failure	internalizing external effects	

MATHEMATICAL NOTATION

Notation	Interpretation
$u(\ )$	utility function
γ	productivity (of individual contributions to the public or a team member's effort)
n	number of people
a	contribution to a public good
ω	subsidy to public good
y	income
p	price
x	number of cigarettes smoked, output of the team
α, β	parameters of utility functions
X	output of the monopolistically competitive firm
m	social multiplier
e	level of effort
W	Mechanism Designer's welfare function
k	constant sum subtracted to the income of team members
B	maximum willingness to pay of a double auction buyer
b	double auction buyer's offer price
S	double auction seller's minimum price at which they are willing to sell
s	double auction seller's minimum announced price at which they are willing to sell

Note on superscripts and subscripts: A, B: different people; N: Nash equilibrium; M: monopoly; C: competition; P: private benefits and costs; S: social benefits and costs.

GLOSSARY

Absolute advantage A person or a nation has an absolute advantage in the production of a particular good if, given a set of available inputs, they can produce more of it than another person or country. See also *comparative advantage*. 303

Accounting profit Accounting profit is the difference between sales revenue and the direct cost of the inputs used to produce output, excluding the opportunity cost of the funds tied up in financing long-lived assets such as buildings, intellectual property, and equipment. See also *economic profit*. 416

Accounting profit share The accounting profit share is the fraction of total income that is received as accounting profits by the owners of the capital goods used in production; it is often decomposed into the capital share and the economic profit share. 883

Acquisitive motives These are the positive value we place on "getting" (acquiring) things, for example, monetary payoffs in games. 973

Adverse selection In a principal–agent relationship, adverse selection is a problem confronting a principal who lacks information on some relevant attribute of an agent such that the terms the principal offers may induce the agents with whom the principal could have most advantageously interacted to drop out. Also referred to as the 'hidden attributes' problem. See also *moral hazard*. 556

Aggregate demand The sum of expenditures on goods and services produced in a country, including demand from the rest of the world. 719

Allocation In a game, an allocation is a particular distribution of goods or other things of value to the players. 164

Assurance Game An Assurance Game is a two-person, symmetric, strategic interaction with two strict Nash equilibria, one of which is Pareto superior to the other. 33

Asymmetric information Information is asymmetric if something that is relevant to the parties in an economic interaction is known by one actor and is not known by another. 29

Average cost Average cost is the cost per unit of output produced. 418

Average product The average product of an input is total output divided by the total input. 291

Average revenue Average revenue is the revenue per unit of output, which is the price. 430

Backward induction Backward induction is a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen by other players in subsequent steps in response to her choice. 73

Bargaining power The extent of a person's advantage in securing a larger share of the economic rents made possible by an interaction. 180

Bargaining set The set of all allocations that are Pareto improvements over the players' fallback (no-bargain) options and the utilities associated with these allocations is termed the bargaining set. 185

Barriers to entry Anything making it difficult for new firms to enter a market, including intellectual property rights, economies of scale in production, and predatory pricing. 439

Beliefs Beliefs are an individual's understandings of the relationship between an action she may take and the outcome of the action. 58

Best response A strategy is a player's best response to the strategies adopted by others if no other strategy available would result in higher payoffs. 19

Bilateral monopoly In a bilateral monopoly transaction there is a single transactor on each side of the market—one potential buyer and one potential seller. 525

Budget constraint An equation that represents all combinations of goods and services that one could purchase that exactly exhaust one's budgetary resources. 138

Budget constraint (compensated) A compensated budget constraint after a price change takes the new prices of goods as given (so it is parallel to the budget constraint after the price change), but gives the person just sufficient

income to purchase a bundle on their original indifference curve, at a new point of tangency. 375

Bundle A bundle is a list of an individual's goods (or bads). 115

Capital share The capital share of total income is the fraction of output accounted for by the opportunity cost of the capital goods used in the production of the output. 891

Capitalism Capitalism is an economic system in which most production takes place in privately owned firms that employ labor in return for wages or salaries to produce goods and services to be sold on markets to make a profit for the owners of the firm's capital goods. 866

Cardinal utility A cardinal utility function assigns a number to each bundle, such that, with a cardinal utility function, $u(x,y) = 10u(x',y')$ means that (x,y) is preferred ten times as much as (x',y'). See also *ordinal utility*. 114

Cartel A group of firms that collude to set output and/or prices in order to raise profits. 507

Centralized economy A centralized economy is one in which the government decides what should be produced, where, by whom, and when, and how the resulting goods should be distributed among the population. See also *decentralized economy*. 796

Certainty equivalent The level of *certain* income that would be valued by the person equally to each of the other combinations on the same indifference curve (involving more risk and more expected income). 748

Ceteris paribus A Latin term that means 'other things equal.' In an economic model it means an analysis that 'holds other things constant'. 84

Club good A club good is non-rival and excludable. See also *common property resource*, *public good*, *private good*. 80

Collateral An asset that a borrower pledges to a lender as a security for a loan. If the borrower is not able to make the loan payments as promised, the lender becomes the owner of the asset. See also *lien*. 669

Common property resource A common property resource is rival and non-excludable. See also *public good*, *private good*, *club good*. 79

Comparative advantage A person or a nation has a comparative advantage in the production of a particular good if the opportunity cost to them of producing it (in terms of foregone output of another good) is lower than it is for another person or a country. See also *absolute advantage*. 303

Comparative statics A method which analyzes the process of change by comparing the status quo Nash equilibrium with a new equilibrium after some change in the underlying data of the problem. 235

Competition condition The competition condition is an equation giving the unique relationship of prices to wages—that is, the real wage rate w^c—such that the number of firms will neither increase nor decrease given the productivity of labor and the opportunity cost of capital. 644

Competition policies Government policy and laws to limit monopoly power and prevent cartels or to otherwise regulate the process of competition. Also known as antitrust policy. 533

Complements in consumption Goods are complements in consumption if an increase in the quantity consumed of one raises the marginal utility of the other. See also *substitutes in consumption*. 375

Complete contract A contract is complete if it (a) covers all of the aspects of the exchange in which anyone affected by the exchange has an interest, and (b) is enforceable (by the courts) at close to zero cost to the parties. 546

Complete preferences Complete preferences specify for any pair of possible outcomes that a person's actions may bring about, A and B, that A is preferred to B, B is preferred to A, or they are equally preferred. 61

Consistent preferences Preferences are consistent if whenever an individual prefers a bundle of goods A to another bundle B, and bundle B to a third bundle, C, they cannot prefer C to A. Consistent preferences are also known as transitive preferences. 62

Constant Returns to Scale In the case of constant returns to scale, increasing all inputs

by some constant increases output proportionally. 291

Constitutive motives These are our desires to "be" or "become" a particular kind of person (to constitute or construct ourselves in a particular kind of way). 973

Constrained optimization A constrained optimization problem is one in which a decision maker chooses the values of one or more variables to achieve an objective which is subject to a constraint that determines the feasible set of actions or outcomes. 108

Consumer surplus The consumer's willingness to pay for a good minus the price at which the consumer bought the good, often summed across all units sold. 391

Contingency A contingency is a state of the world that may or may not happen and that affects the payoff to some action. 63

Convexity assumption The indifference curves for two goods are "bowed inward" toward the origin and the production possibility frontier (the frontier of the feasible production set) is "bowed outward" from the origin. 821

Cooperative A cooperative is a business organization or other association whose members together own the assets of the organization; they share the income resulting from their activities and jointly determine how the organization will be run (possibly through the democratic election of a manager). 276

Cooperative game A strategic interaction in which the players' choice of a strategy is subject to a binding (enforceable) agreement. 14

Coordination failure A coordination failure occurs when the non-cooperative interaction of two or more people results in an outcome that is worse for at least one of those involved and not better for any. 6

Coordination problem A coordination problem is a situation in which people could all be better off (or at least one be better of and none be worse off) if they jointly decide how to act—that is, if they coordinate their actions—than if they act independently. 4

Costless bargaining Bargaining is costless when the parties to the bargain do not incur costs in executing a trade other than the price of the good exchanged. Efficient bargaining is often used interchangeably with the term "costless bargaining." 836

Credible threat A threat is credible if carrying it out is the best response if the target of the threat takes the action that the threat was intended to deter. 186

Credit constraints A person or business is said to be credit constrained if: (a) they are excluded from borrowing entirely, or (b) they face limits on how much they can borrow, or (c) they pay extraordinarily high rates of interest when they do succeed in getting a loan. 671

Credit market In a credit market, lenders (banks, payday lenders, and other financial institutions) provide loans to borrowers (individuals and companies) who commit at some future date to repay the amount borrowed plus an additional percentage of that amount, termed the interest on the loan. 670

Decentralized economy In a decentralized economy, who produces what, when, how, and for whom is determined by the uncoordinated decisions of owners of individual firms, employees, and other private economic actors. See also *centralized economy*. 797

Decreasing risk aversion The tendency of a person to be less risk averse if she has more income (or wealth) than if she has less. 748

Demand curve (individual) A demand curve provides the answer to the hypothetical question: what is the maximum amount of a good that can be sold at each price. The individual demand curve refers to the purchase of a good by a person given the prices of the other goods and the individual's budget. 137

Demand curve (market) A demand curve provides the answer to the hypothetical question: what is the maximum amount of a good that can be sold at each price? The market demand curve is the horizontal summation of the individual demand curves of actual or potential buyers. 383

Demand function A demand function shows how the amount of a good purchased by an individual varies with the prices of all goods and the individual's budget. *See also Demand curve*. 352

Demand-side economies of scale occur when the value of a firm's product is greater to a buyer the more other buyers of the product there are. 534

Diminishing marginal utility A property of some utility functions according to which each additional unit of a given variable results in a smaller increment to total utility than did the previous additional unit. 120

Disagreement Game A game in which there are two or more Pareto-efficient Nash equilibria that are ranked differently by the players so that in the two-person case the players prefer different equilibria is a Disagreement Game (which also goes by other names). 36

Diseconomies of scale When production exhibits diseconomies of scale, multiplying all inputs by some number greater than one increases output by a factor less than that number. See also *economies of scale*. 290

Disposable income Disposable income is the maximum a household can spend ('dispose of') without borrowing, after paying tax and receiving transfers (such as unemployment insurance and pensions) from the government. See also *market income*. 775

Distributional neutrality Distributional neutrality is a characteristic of the perfectly competitive general equilibrium model in which the distribution of wealth following competitive exchange is identical to the distribution of wealth prior to exchange. 819

Distributional outcome How the gains from exchange—the economic rents—are distributed among the people in an exchange. 165

Dominant strategy A strategy is dominant if it yields the highest payoff for a player for any strategy chosen by the other players. Weak dominance refers to the case where there are one or more strategies yielding the same payoff. 21

Dominant strategy equilibrium A dominant strategy equilibrium is a strategy profile in which all players play a dominant strategy. 22

Double auction In a double auction buyers and sellers simultaneously submit to an auctioneer 'bids' and 'asks,' that are the prices at which they are willing to buy and sell, respectively. An auctioneer then chooses a price that clears the market. 956

Dual economy The dual economy is one in which work, except for housework and government employment, is of two kinds: work compensated by incomes other than wages or salaries in the informal sector and employment for wages and salaries in what is termed the capitalist sector. 919

Duopoly When there are just two firms selling the same output, we call the industry a duopoly and we call each firm a duopolist. 479

Durable asset A durable asset is one that remains valuable over a long period of time. 524

Dynamics Refers to the study or process of change. 236

Earnings This term—sometimes called "labor earnings"—refers to income from employment by a firm, government or some other employer, whether in the form of wages or salaries. 661

Economic profit Economic profit is accounting profit minus the opportunity cost of funds tied up in long-lived plant and equipment evaluated at the opportunity cost of capital, ρ. See also *accounting profit*. 415

Economic profit share The economic profits share is the fraction of total income received by the owners of the capital goods used in production in excess of the opportunity cost of capital, or: economic profits share = accounting profits share minus capital share. 883

Economic rent A participant's economic rent is the payoff they receive in excess of what they would get in their fallback position. 28

Economies of agglomeration The advantages that firms may enjoy when they are located close to other firms in the same or related industries. 311

Economies of scale When production exhibits economies of scale, multiplying all inputs by some number greater than one increases output by a factor greater than that number. See also *diseconomies of scale*. 290

Endogenous enforcement of contract When the parties to an exchange—employers and workers, buyers and sellers, borrowers and lenders—themselves adopt strategies to ensure favorable terms of an exchange for aspects of it not covered by a contract, enforcement is endogenous. 548

Endogenous preferences If one's experiences result in durable changes in preferences, then they are termed endogenous. See also *exogenous preferences*. 102

Endowment allocation The ownership of goods at the start of a game (or at the status quo) is called the endowment allocation. 181

Enforcement rent In a principal–agent relationship an enforcement rent is the excess of the value of the transaction to the agent over the agent's fallback. The fear of losing the enforcement rent induces the agent to act in the principal's interest. 568

Equilibration Equilibration is the process of getting to an equilibrium from a nonequilibrium situation. 518

Equilibrium An equilibrium is a situation that is stationary (unchanging) in the absence of a change external to the model. 18

Equity One's own wealth (rather than borrowed funds) invested in a project. There is a second entirely different use of the term, meaning the character of being fair, as in "an equitable division of the pie." 701

Excess demand Excess demand exists when, at the prevailing price, the amount demanded exceeds the amount supplied. 515

Excess supply Excess supply exists when, at the prevailing price, the amount supplied exceeds the amount demanded. 515

Excludable A good is excludable when a potential user may be denied access to the good at a low or zero cost. 78

Excluded borrower A borrower who is unable to obtain credit. 706

Exogenous enforcement of contract Exogenous enforcement of the terms of an exchange is done by courts or another third party—not the parties to an exchange themselves—and is a defining characteristic of a complete contract. 547

Exogenous preferences Preferences are exogenous if they change in response only to influences external to the economy or at least outside of the economic subject matter under study. See also *endogenous preferences*. 102

Expected payoff In a situation of risk, the expected payoff to an action is the sum of payoffs occurring under each contingency multiplied by the probabilities that each contingency occurs. 64

Expected profit In a situation of risk, expected profit is the sum of profits occurring under each contingency multiplied by the probabilities that each contingency occurs. 494

Extensive-form representation of a game An extensive-form representation of a game includes, in addition to the strategies with payoffs associated with each strategy profile, the time dimension—who knows what, when they know it, and the sequence of moves as described by a game tree. See also *normal-form representation of a game*. 73

External effect An external effect occurs when a person's action confers a benefit or imposes a cost on others and this cost or benefit is not taken into account by the individual taking the action. External effects are also called externalities. 9

External validity Results of experiments or other scientific research that can be generalized to circumstances outside (external to) the laboratory or other setting in which the research was produced are said to be externally valid. 97

Factor intensity A production function A is more labor-intensive than production function B if for any given ratio of wages to the price of capital goods, the cost-minimizing choice of inputs will be to hire more labor hours when using A than when using B. 330

Factor of production Any input into a production process is called a factor of production. In the past economists often referred to land, labor, and capital goods as primary factors of production, but this usage is overly narrow given the essential role of other production inputs such as our natural environment beyond "land" and knowledge. 317

Fallback position A player's fallback position (or reservation option) is the payoff they receive in their next best alternative. 28

Feasible frontier The boundary of a feasible set. In the case of two goods, it is the curve made of points that defines the maximum feasible quantity of one good for a given quantity of the other. 129

Feasible set All of the combinations of the actions or outcomes that a decision maker could choose, given the economic, physical, or other constraints. 129

Firm A business organization in which private owners of capital goods hire and direct labor to produce goods and services for sale on markets in order to make a profit is called a firm (or sometimes a capitalist firm). 866

First mover A player who can commit to a strategy in a game before other players have acted is a first mover. 72

First welfare theorem A perfectly competitive equilibrium of an economy with complete contracts is Pareto efficient. 812

Free rider A free rider is a person who benefits from the cooperation or generosity of others, while not reciprocating in a cooperative or generous way, for example, not contributing in a Public Goods Game. 81

Game theory Game theory is the branch of applied mathematics that studies strategic interactions. 11

General equilibrium General equilibrium analysis is a study of two or more markets and their interactions. See also *partial equilibrium*. 799

Giffen good Over some range of prices, purchases of a Giffen good increase if the price rises, and fall if the price falls. See also *law of demand*. 382

Gini coefficient This measure of inequality (using income as an illustration) is the average difference in income between every pair of individuals in a population relative to mean income, multiplied by one-half. The Gini coefficient is usually calculated as the area between the Lorenz curve and the perfect equality line, divided by the total area under the perfect equality line. See also *Lorenz curve*. 873

Group inequality Economic differences between sets of people distinguished by some common attribute—men and women and people of different nations, ethnic or racial groups—are called group inequalities. 39

Guard labor Those employed as police, private security personnel, the armed forces, and others whose job is enforcing and perpetuating the rules of the game. 787

Incentive-compatibility constraint The incentive-compatibility constraint, ICC, describes the limits on the outcomes that a first mover may implement by showing how a second mover will respond to each of the choices that the first mover might make, also known as the second mover's best-response function. 199

Income The largest amount that a household or person can consume over a given period of time without reducing the value of wealth (their stock of assets, minus any outstanding debt). See also *wealth*. 701

Income effect When the price of a good changes, this alters people's real income, expanding or shrinking the feasible set of purchases. The effect of this change in real income (with no change in price) on the goods purchased is the income effect. See also *substitution effect*. 373

Income–offer curve An income-offer curve describes consumption or other choices made by an individual for varying levels of income. 354

Incumbent firms Firms already selling in a market. 439

Indifference When a person is indifferent between two outcomes, they do not prefer one over the other. 61

Indifference curve The points making up an individual's indifference curve are bundles—indicated by (x, y), (x', y'), and so on—among which the person is indifferent, so that $u(x, y) = u(x', y')$ and so on. 115

Industry An industry is a set of firms producing similar products. 473

Inequality aversion A preference for more equal outcomes and a dislike for both disadvantageous inequality that occurs when others have more than the actor and advantageous inequality that occurs when the actor has more than others. 89

Informal sector The informal sector is comprised of the economic activities of family farmers, small shopkeepers, and others who work independently of an employer—whether a private firm or a government—and without hiring workers. 918

Institutions Institutions are the laws, informal rules, and mutual expectations which regulate social interactions among people and between people and the biosphere. 11

Insurance Any costly action one can take that reduces the level of risk to which one is exposed. 686

Interest factor The interest factor, δ, is one plus the rate of interest. 678

Inverse demand function The inverse demand function (curve) answers the hypothetical question: what is the highest price at which a

given amount of some good could be sold? 352

Invisible Hand Game An Invisible Hand Game has a single Nash equilibrium that is Pareto efficient. 32

Iso-social welfare curve Iso-social welfare curves show constant or equal ("iso") levels of welfare for different combinations of utility among those involved. 178

Isocost line A line that represents all combinations of inputs that cost a given total amount. 327

Isoprofit curve An isoprofit curve shows combinations of prices and quantities sold of a good yielding equal profits to the owners of a firm. 425

Isoquant An isoquant gives the combinations of two inputs that are just sufficient to produce a given level of output. 318

Labor (effort) The amount of actual work devoted to production. Labor is measured in units of effort, not in hours. 607

Labor discipline model A model that explains how employers set wages so that workers receive an economic rent (called an employment rent), which provides workers an incentive to work hard and well in order to avoid job termination. 609

Labor or trade union A labor/trade union is an organization of workers who together bargain with one or more employers about wages and working conditions, a process known as collective bargaining. 896

Labor share The term labor share—distinct from wage share—refers to unemployment benefits plus wages as a fraction of total income. 894

Law of demand The law of demand holds that a decrease in the price of a good will result in an increase in the quantity of the good purchased. See also *Giffen good*, an exception to the law of demand. 382

Law of one price The law of one price states that in equilibrium identical goods or services will transact at the same price. 475

Lien A lien is a property right in some good held by a lender to secure the repayment of a debt. Collateral is a form of lien. 727

Linear tax and lump-sum transfer A tax that is proportional to income (a linear tax), the proceeds of which are divided equally and transferred to citizens (a lump sum). 776

Long side of a market The long side of the market is the side—either supply or demand—on which the number of desired transactions is greater, given the price. 515

Lorenz curve The Lorenz curve summarizes the distribution of income or some other measure across a population, mapping the cumulative (poorest to richest) population shares and corresponding cumulative income shares. See also *Gini coefficient*. 873

Loss aversion Loss aversion is present when the loss of some given amount reduces a person's utility by more than a gain of the same amount would have raised their utility. 792

Lottery In game theory a lottery is a set of uncertain outcomes and the probabilities that each will occur. 740

Marginal borrower A marginal borrower is a borrower with just enough wealth to secure a credit contract with a lender. 706

Marginal cost Marginal cost is the effect on total cost of producing a small amount more of output. It is the slope of the total cost function at each point. 418

Marginal cost of labor The marginal cost of labor is the change in total wages paid associated with employing (a small amount) more labor, that is, effort. 615

Marginal product The marginal product of an input is the change in total output associated with a small change in the input. 291

Marginal rate of substitution The marginal rate of substitution is the negative of the slope of an indifference curve. It is also the maximum willingness to pay for a small increase in the amount x expressed as how much of y the person would be willing to give up for this. In a model with y as money, this is called the offer price. 121

Marginal rate of technical substitution The marginal rate of technical substitution is the negative of the slope of the production isoquant and is equal to the ratio of the marginal product of the input on the x-axis to the marginal product of the input on the y-axis. It shows how much more of the y-axis input must be added to compensate for the

withdrawal of one unit of the x-axis input so that output is unchanged. 322

Marginal rate of transformation The marginal rate of transformation is the quantity of some good that must be sacrificed to acquire more of another good. It is equal to the negative of the slope of the feasible frontier (constraint). See *opportunity cost*, 3.2. 130

Marginal revenue Marginal revenue is the change in total revenue associated with a small change in sales. 430

Marginal revenue product of labor The marginal revenue product of labor is the change in total revenue associated with a small change in labor employed. 615

Market clearing A market clears when the amount supplied is equal to the amount demanded. 452

Market income Market income is income before the payment of taxes or the receipt of transfers from the government; it includes earnings (wages and salaries from employment) as well as income from self-employment and from the ownership of assets (interest, rents, or dividends). See also *disposable income*. 776

Markup The markup is the difference between the price at which a good sells and its cost (including the opportunity cost of the capital goods used). 436

Markup ratio The firm's markup ratio is the profit per unit of output (the markup) divided by unit costs. 436

Mechanism A mechanism is a set of rules of the game—possibly designed deliberately—that provide the incentives, constraints, and information that will result in an allocation—often a preferred allocation—being implemented. 823

Merit good A merit good is one that it is thought on moral grounds should be available to all irrespective of their income. 455

Monetary policy Policies implemented by a central bank affecting the rate of interest at which businesses and others can borrow and the amount of borrowing, thereby regulating aggregate demand to moderate the business cycle and regulate inflation, are termed monetary policy. 720

Monopolistic competition Monopolistically competitive firms are the only sellers of the particular good they produce, but they compete with other firms that sell similar products. 424

Monopsony A firm is a monopsony if it is the only buyer (or just one of a small number of buyers) in a particular market for some good or service. 650

Moral disengagement A process by which, in some particular situations, people come to feel that ethical considerations need not be applied to their own actions or others' actions. 969

Moral hazard If there is a conflict of interest between a principal and an agent over the agent taking some action that cannot be ensured by a complete contract, then the principal faces a moral hazard problem. Also referred to as the 'hidden actions' problem. See also *adverse selection*. 552

Motivational crowding out Motivational crowding out occurs when monetary or other material incentives or attempts to control someone diminish that person's other-regarding or ethical preferences or intrinsic motivation. 100

Nash equilibrium A Nash equilibrium is a profile of strategies—one strategy for each player—each of which is a best response to the strategies of the other players. 20

No-shirking wage The no-shirking wage is the wage that is just sufficient to motivate a worker to provide effort at the level specified by their employer. 632

Normal-form representation of a game The description of a game by a matrix of strategies with payoffs associated with each strategy profile is the normal-form (or strategic) representation of a game. See also *extensive-form representation of a game*. 73

Opportunity cost Where x and y are both valued positively, the opportunity cost of x in terms of y is how much y a person must give up to get a unit more of x. 109

Opportunity cost of capital The opportunity cost of capital is the accounting rate of profit that a wealth holder would make on his next best alternative use of the funds used to acquire the capital goods used in production. 414

Ordinal utility Let $a \succ b$ mean "a is preferred to b." An ordinal utility function ranks bundles,

e.g. $(x,y) \succ (x',y') \succ (x'',y'')$, without specifying how much (x,y) is preferred to (x',y') or (x',y') is preferred to (x'',y''). The assignment of numerical utilities representing ordinal preferences is meaningful only to express an ordering: $u(x,y) > u(x',y')$ implies only that the first bundle is preferred to the second but not by how much. See also *cardinal utility*. 114

Other-regarding preferences A person with other-regarding preferences, when evaluating the outcomes of her actions, takes into account their effects on the outcomes experienced by others as well as the outcomes she will experience. 60

Pareto efficiency A Pareto-efficient allocation is an allocation with the property that there is no alternative technically feasible allocation in which at least one person would be better off and nobody would be worse off. If an allocation is Pareto efficient, then there is no alternative allocation that is Pareto superior to it. 23

Pareto-efficient curve The points making up the Pareto-efficient curve represent all of the allocations that are Pareto efficient. 173

Pareto-improving lens The set of allocations that are (at least weakly) Pareto superior to the fallback options of the players is the Pareto-improving lens. 182

Pareto-superior Outcome A is Pareto-superior to outcome B (it Pareto-dominates outcome B) if, in outcome A, at least one player is better off than in outcome B without anyone being worse off. 24

Partial equilibrium Partial equilibrium analysis is the study of a single market. See also *general equilibrium*. 799

Path dependence A process is path dependent if the most likely state of something this period depends on its state in recent periods. 38

Perfect price discrimination Occurs when a firm can make a separate take-it-or-leave-it offer to each individual buyer for each unit to be sold at a price equal to the consumer's maximum willingness to pay. 510

Perfectly competitive equilibrium In a perfectly competitive equilibrium, supply equals demand, and neither buyers nor sellers can benefit by altering their price or quantity. 450

Permit A permit allows a firm or person to engage in an activity: it gives permission. 255

Piece rate Under a piece-rate contract, a worker is paid a fixed amount for each unit ("piece") of the product made. 608

Poverty trap A poverty trap occurs when identical people may experience either an adequate living standard or poverty, depending only on chance events of their histories. Poverty in this case is a result of a person's circumstances, not personal attributes. 38

Power If, by imposing or threatening to impose sanctions on A, B can affect A's actions in ways that further B's interests, while A lacks this capacity with respect to B, then B has power over A. 582

Predatory pricing Predatory pricing occurs when an incumbent firm charges a price lower than its (marginal) costs, seeking to drive its competitors out of business. 439

Preferences Preferences are evaluations of outcomes of one's actions that provide motives for taking one course of action over another. 59

Preferences, beliefs, and constraints approach According to this approach, from the feasible set (which includes all of the actions open to the person given by the economic, physical, or other constraints she faces), a person chooses the action that she believes will bring about the outcome that she values most as given by her preferences. 56

Price discrimination Selling the same product at different prices, for example, charging more to buyers with a greater willingness to pay. 509

Price elasticity of demand The price elasticity of demand is the ratio of the percentage change in quantity demanded to the percentage change in price, $\eta_{Xp} = \frac{\Delta X/X}{\Delta p/p}$. 386

Price-making Price-making is a strategy that an economic actor may follow, altering the price at which they offer to buy or sell, or altering the level of output in ways that change the price at which they can transact. 461

Price-setting power A first mover with price-setting power can commit to a price—or in the case of barter, the ratio at which goods will be exchanged—but not the quantity that will be transacted at that price. 199

Price-taking Price-taking is a strategy that an economic actor may follow, taking as given the prices at which one might buy or sell. 450

Price-offer curve The price-offer curve shows every utility-maximizing consumption bundle for each price of the goods under consideration. 141

Principal–agent relationship A principal–agent relationship (also called an agency problem) arises when two conditions hold: (a) Conflict of interest: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent; and (b) Incomplete contract: the agent's actions or attributes are not known to the principal (or, if known, are not verifiable) and so cannot be subject to enforceable contract. 551

Prisoners' Dilemma A Prisoners' Dilemma is a 2-by-2 social interaction in which there is a unique Nash equilibrium (that is also a dominant strategy equilibrium), but there is another outcome that gives a higher payoff to both players (and a higher total sum of payoffs than any other outcome), so that the Nash equilibrium is not Pareto efficient. 26

Private cost The private cost (marginal or average) is the cost that the decision maker bears as a result of some action that he or she takes. 249

Private good A private good is rival and excludable. See also *common property resource*, *public good*, *club good*. 79

Private property The right and expectation that one can enjoy one's possessions in ways of one's own choosing, exclude others from their use, and dispose of them by gift or sale to others who then become their owners. 164

Probability distribution A probability distribution for n contingent outcomes of a decision is a list of non-negative numbers $\{P_1, P_2, \ldots, P_n\}$ that add up to 1. These probabilities express the decision maker's belief about the likelihood that each of the of n contingent outcomes will occur. 63

Product differentiation Product differentiation is a business practice aimed at making the firm's product appear more distinct from or less similar to substitute products. 424

Production Production is the process by which we transform the resources of the natural world using already produced tools, facilities, and other inputs to meet human needs. 316

Production function A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows. 127

Production possibilities frontier The production possibilities frontier (PPF) for two goods shows the maximum amount of one good that can be produced given the output of the other good. The production possibilities frontier is the boundary of the producer's feasible set and is an alternative name for the feasible frontier when we study production. 294

Progressive policies A system of taxes and transfers or other policies that reduce disposable income inequality is called progressive. See also *regressive policies*. 380

Public good A public good is non-rival and non-excludable. See also *common property resource*, *private good*, *club good*. 79

Quadratic, quasi-linear utility function A quadratic, quasi-linear utility function is quadratic in one variable and linear in another variable. 369

Quantity-constrained An actor is quantity constrained if they are unable to transact the quantity they would like at the going price. 515

Quasi-linear function A quasi-linear function depends linearly on one variable and nonlinearly on another variable. 190

Rate of economic profit Economic profits divided by the value of the capital stock. 416

Rational A rational person has complete and consistent (transitive) preferences and can therefore rank all of the outcomes that their actions may bring about in a consistent fashion. 61

Regressive policies A system of taxes and transfers or other policies that increase disposable income inequality is regressive. See also *progressive policies*. 380

Relative price A relative price is a ratio of one price to another. 349

Rent-seeking Any activity undertaken to gain a rent for the actor is called rent seeking,

including changing a price or quantity in a non-clearing market, creating barriers to entry to reduce competition, introducing an innovation to reduce costs, price discrimination, and lobbying or other political activities aimed at granting the actor some kind of legal or other advantage. 518

Rent control A policy regulating the rent that a landlord can charge, most commonly limiting the size of a rent increase that is permitted. 455

Residual claimant The residual claimant is whoever gets what is left over (the residual) from the revenue (or other benefit) of a project when all of the costs that have been contracted for are paid. 637

Returns The term "returns to risk" (or just returns) is the realized income or expected income resulting from an investment or some other risky choice. 752

Risk The term risk is conventionally used in economics to describe situations in which payoffs depend on contingencies, and the probabilities of each contingency occurring are known. 63

Risk aversion A risk-averse person dislikes uncertainty about outcomes and will choose a certain outcome valued at $x over some lottery whose expected value is greater than $x. 742

Risk-dominant strategy The strategy in a 2×2 game that yields the highest expected payoff when the player attributes equal probability to the two actions of the other player. 69

Risk exposure The difference between the better and worse outcome when the two are equally likely. We also term this the level of risk. 739

Risk neutral A risk-neutral person is indifferent between receiving $x with certainty and playing an uncertain lottery with the same expected value. A risk-neutral person is not risk averse. 743

Rival A good is rival when more people using the good reduces the benefits available to other users. 78

Rule of law Under the rule of law all people—including those who make the laws, police, heads of state, and other government officials—are subject to the law. In game theoretic terms rule of law means that irrespective of the personal identity of the players the rules of the game govern the interaction for all players, including rules governing how the rules of the game can change. 868

Scarce A good is scarce if it is valued and there is an opportunity cost of acquiring more of it. 810

Second welfare theorem Given complete contracts and the convexity assumption about production and preferences, any Pareto-efficient allocation can be implemented by some assignment of the endowments of all parties, followed by a perfectly competitive market exchange process. 822

Self-regarding preferences When choosing an action, a self-regarding actor considers only the effect of her actions on the outcomes experienced by the actor, not outcomes experienced by others. 60

Set A set (in mathematics) is a collection of objects defined either by enumerating the objects, or by a rule for deciding whether any particular object is in the set or not. For example, the set of positive, even integers less than or equal to 10 is, $\{2, 4, 6, 8, 10\}$. 12

Sharecropper A sharecropper is a farmer who cultivates land owned by another person with whom he or she contracts to give a share (often one half) of the crop produced. 55

Shirking When a worker does not work as hard as the employer requires, economists call this "shirking." 630

Short run The term does not refer to a period of time, but instead to what is exogenous, meaning fixed, that may become endogenous in the long run. About a firm's costs, for example, we assume that its stock of capital goods and technology is exogenous (constant) in the short run, but may be varied in the long run. 418

Short side of a market The short side of the market is the side—either supply or demand—on which the number of desired transactions is least, given the price. 516

Social cost The social cost is the private cost that the decision maker bears plus any costs imposed on others as negative external effects. 249

Social multiplier When there are indirect effects of a policy through its effects on other

people's behavior the presence of a social multiplier means that the total effect of the policy will differ from the direct effect (hypothetically holding constant the behavior of others). As a quantitative magnitude, the social multiplier is the difference between the direct and total effects. 940

Social welfare function A social welfare function is a representation of "the common good" based on some weighting of the utilities of the people making up the society. 177

Solution concept A solution concept is a rule for predicting the outcome of a game, that is, how a game will be played. 14

Stable equilibrium An equilibrium is stable if a sufficiently small displacement away from the equilibrium is self-correcting (leading to movement back toward the equilibrium). 405

Strategic complementarity Strategic complementarity exists when (a) a strategy is a strategic complement to itself: the payoff to playing a particular strategy increases as more people adopt that strategy as a result of some form of positive feedbacks, or (b) one strategy and another are strategic complements to each other. In this case, for two activities A and B, the more that A is performed, the greater the benefits of performing B, and the more that B is performed, the greater the benefits of performing A. 34

Strategic interaction An interaction is strategic when participants' outcomes—their profit, standard of living, or some other measure of their well-being—depend on the actions that both they and others choose, and this interdependence is known to the actors. A shorthand expression for the term strategic is: mutual dependence, recognized. 12

Structural unemployment Structural unemployment is the unemployment that results from the fundamental structure of the economy. 650

Substitutes in consumption Goods are substitutes in consumption if an increase in the quantity consumed of one reduces the marginal utility of the other. See also *complements in consumption*. 376

Substitution effect When the price of a good changes, the change in the consumption of the good that is due to the change in relative prices (holding constant the buyer's real income) is the substitution effect. See also *income effect*. 373

Supply curve (individual firm) A supply curve provides the answer to the hypothetical question: What amount will be supplied for each given price? The individual firm's supply curve is the portion of the firm's marginal cost curve that is not less than the average variable cost. 442

Supply curve (market) A supply curve provides the answer to the hypothetical question: What amount will be supplied for each given price? A market supply curve is the horizontal sum of the firms' supply curves. 448

Take-it-or-leave-it power A player with take-it-or-leave-it power (TIOLI power) in a two-person bargaining game can specify the entire terms of the exchange—for example, both the quantity to be exchanged and the price—in an offer, to which the other player responds by accepting or rejecting. 186

Team production A form of production involving two or more people in which the contribution of each person to the output cannot be readily determined, either because it cannot be defined or because it cannot be measured. 960

Technical efficiency A technique of production is technically efficient if there is no other technique with which the same output can be produced with less of at least one input and not more of any input. 318

Technical progress A reduction over time in the quantity inputs required to produce some given quantity of output. 334

Technique of production A technique of production is a particular way—bundle of inputs—of producing some given amount of output. 318

Technology A technology is a description of the relationship between inputs—including work, machinery, and raw materials—and outputs. 223

Tipping point An unstable equilibrium at the boundary between two regions characterized by distinct movements in some variable. 71

Total cost Total cost is the minimum cost of producing an output level at given prices of inputs and the opportunity cost of capital. When we use the term cost without an

adjective ("economic" or "accounting") we mean economic cost. 417

Total factor productivity Output divided by a weighted sum of the inputs (the weights being each input's relative contribution to producing output). 912

Total revenue Total revenue is total sales times price per unit sold. 430

Trade-off A trade-off is a situation in which having more of something desired (a "good") requires having less of some other "good" or more of something that the actor would like to have less of (a "bad"). 119

Tragedy of the commons The tragedy of the commons is a term used to describe a coordination problem in which self-interested individuals acting independently deplete a common property resource, lowering the payoffs of all. 8

Transaction costs Costs that impede the bargaining process, including costs of acquiring information about the good to be traded, and costs of enforcing a contract. 836

Truth-telling mechanisms Rules of the game that would make it a best response for a self-interested and amoral person to reveal their true preferences (including their true value of a good that they may buy or sell). 959

Uncertainty The term uncertainty describes situations where the decision maker decision maker does not know and cannot learn the probabilities of the contingencies affecting their payoffs. 69

Union voice effect The union voice effect occurs when a trade union, by providing a 'voice' to otherwise unheard workers, improves their treatment by employers and their job satisfaction (in our model, decreasing the disutility of work), with, as a result, greater work effort provided by workers, and an increase in output per worker hour. 898

Usury Unreasonably, unethically, or illegally high interest rates. 670

Utility function A utility function is an assignment of a number $u(x,y)$, to every bundle (x,y) representing a person's valuation of that bundle. This means that if given the choice between two bundles (x,y) and (x',y'), the individual will choose the first if $u(x,y) > u(x',y')$. 112

Utility possibilities frontier The utility possibilities frontier is composed of the feasible Pareto-efficient combinations of utilities of the members of a population. 176

Verifiable information Information that can be used in legal proceedings to enforce a contract or other agreement. 29

Voluntary exchange An exchange is voluntary if neither party can coerce the other to participate nor require them to renew the exchange beyond an agreed upon duration. 164

Wage maker Employers are called wage makers because they typically decide on ("make") the wage to offer to particular workers either unilaterally, or through bargaining with at trade union. The term is intended to contrast with what would be a wage-taking firm (like a price-taking firm) that cannot alter the wage or price to its advantage. 651

Wage share The wage share is the fraction of total income that is received in the form of labor earnings (wages plus salaries). See also *labor share*. 883

Wealth Stock of things owned (or the value of that stock) that yields a flow of income or other valued services to the owner. See also *income*. 701

Wealth constraints Any restriction on the kinds of contract one can engage in due to a lack of wealth is called a wealth constraint. 671

Welfare economics A branch of economics that studies the effect of economic policies and institutions on individual and societal well-being ("welfare"). 834

Willingness to pay The maximum amount a person would pay to acquire a unit of a good. 367

Winner-take-all competition is a competitive process that results in a monopoly or near monopoly. 535

Zero-profit condition This condition requires that when barriers to entry are absent—the case of unlimited competition—expected economic profits are zero in a Nash equilibrium. 647

NOTES

Preface

1. See Hall and Jones (1999).

Part 1

1. See Smith (320, 233–234).

Chapter 1

1. Hume (186, 304).
2. See Schultz (1980, 639).
3. A comprehensive study of life in Palanpur is offered in Lanjouw and Stern (1998) with follow-up work twenty years later in Himansu et al. (2018). The visitor was Bowles.
4. See Banerjee and Duflo (2011).
5. See Smith (1982 [1776], 456).
6. See Hardin (1968).
7. The work that constituted Von Neumann's main contribution to economics is Von Neumann and Morgenstern (1944).
8. See Bonnefoy (2020).
9. See Nash (1950, 1951); Nasar (1998).
10. See Jaffe (1972).
11. See Pareto (2014 [1906]).
12. This language is from King County Recorder's Office (1929, 348). See Moyer (2020) for a brief explanation of restrictive covenants.
13. See Bharti (2018).
14. See Schelling (1980, v-vi).
15. See Acheson (1988).
16. See Ostrom (2015 [1990]).
17. See Baland et al. (2006).
18. See Platteau and Seki (2001).

Chapter 2

1. See Young and Burke (2001).
2. See Mill (1848, 149).
3. See Bardhan (1984) and Banerjee et al. (2002).
4. See Herndon and Weik (1889, 439).
5. See *A Treatise on Probability*, Keynes (1921, 52–53).
6. See Fehr and Fischbacher (2002, 14).
7. See Ledyard (1995).
8. See Fehr and Rockenbach (2003) and commentary by Bowles and Gintis (2002).
9. See Mill (1844, 97).
10. See a variety of sources including Güth et al. (1982); Roth et al. (1991); Slonim and Roth (1998); Carpenter et al. (2005); Gneezy et al. (2016).
11. For an explanation of why this is true, see Fehr and Fischbacher (2002).
12. The results with three months wages as the stake from Indonesia are from Cameron (1999). The results from the United States with stakes of are from Hoffman et al. (1996), Carpenter et al. (2005) and Camerer (2011). The results from France are from Munier and Zaharia (2002). The results from India are from Andersen et al. (2011).
13. See Henrich et al. (2010b) and Henrich et al. (2010a).
14. See Henrich et al. (2001).
15. See Cohn et al. (2019).
16. We summarize the work in Andreoni and Miller (2002).
17. See Forsythe et al. (1994, 362) and Engel (2011).
18. See Anderson and Putterman (2006).
19. See Levitt and List (2007, 155) and Falk and Heckman (2009).
20. The results on trust and reciprocity are based on Fehr and List (2004).
21. See Fehr and Leibbrandt (2011).
22. The results from the Bale Oromo in Ethiopia are found in Rustagi et al. (2010).
23. See Belkin (2002) and Greenberger (2003).

24. These results are discussed in two papers, Gneezy and Rustichini (2000a) and Gneezy and Rustichini (2000b).
25. See Fehr and Gächter (2000b, 162) and Camerer (2011).
26. See Loewenstein et al. (1989, 433).
27. See Ariely (2019).
28. See Stigler and (1977, 76) and Becker (1998).

Chapter 3

1. The data for the figure and related empirical work can be found in de V. Cavalcanti et al. (2008).
2. See Friedman (1974).
3. See Leopold (2020 [1949], xxi).
4. See Mill (1998 [1861], 81).
5. See Smith (1976 [1759], 110).
6. See Robbins (1932, 123,124).
7. See Kahneman et al. (1997).
8. See Kahneman and Krueger (2006) and Lucas et al. (2003).
9. See Oswald and Wu (2010).
10. See Krueger and Schkade (2008), Blanchflower and Oswald (2008), Oswald and Wu (2010).
11. See Kahneman et al. (2004).
12. See Friedman (1953).
13. See Keynes (1897) and Friedman (1953).
14. See Bentham (1996 [1789], 11).

Chapter 4

1. See Pareto (1971, 341).
2. See Ibn Battuta (2013 [1929], 151).
3. See Giri (1983, 23).
4. See Edgeworth (1881).

Chapter 5

1. See Tierney (2000).
2. See Thunberg (2019).
3. The quote is Hardin (1968, 1244, 1247). For the alternative perspective see Ostrom et al. (1999).
4. See Ouchi (1980); Taylor (1987); Ostrom (2015 [1990]).
5. See Murphy (2017).
6. See Frey and Bohnet (1995), Frey and Bohnet (1996), Sally (1995), and Kollock (1998).
7. See Cardenas et al. (2004).
8. See Alesina and Ferrara (2000).
9. These two examples are from Bardhan and Dayton-Johnson (2002) and Bardhan and Dayton-Johnson (2007).
10. See García-Barrio and García-Barrios (1990).
11. See Ostrom et al. (1999).

Part 2

1. See Voltaire (1734 1961, 16–18).
2. See Bar-On et al. (2018).

Chapter 6

1. See Smith (1982 [1776], 31).
2. See McDonald and Kotha (2015).
3. See Kapur (2002).
4. See Hausmann and Rodrik (2003).
5. See Banerjee (2006).
6. See, for example, Perese (2010).
7. See Isaacson (2014) and Cortada (2019).
8. See Leontief (1983, 30).
9. See Salomons and Auto (2017).
10. See Acemoglu and Restrepo (2020, 2191–92;2241)
11. See Atkin et al. (2017) and Atkin et al. (2017).

Chapter 7

1. See Smith (1982 [1776], 25).
2. See Veblen (1934 [1899], 71–72).
3. See Perez-Truglia (2020).
4. See Veblen (1934 [1899], 81).
5. See Feldstein et al. (2017).
6. See Fremstad and Paul (2019, 96).
7. See Jensen and Miller (2008).
8. For the BMI data see Abarca-Gómez et al. (2017).
9. See Schumpeter (1954, 842).
10. See Grubs (2017).
11. See Colchero et al. (2016).
12. See Smed (2012).
13. See Jensen and Smed (2013).
14. See Clark (1991, 9).

Chapter 8

1. See Marshall (1890 1920, 290).
2. See Lafond et al. (2020).
3. See Blinder et al. (1998, 218).

4. See Samuelson (1948, 484).
5. See Robinson (1962, 41,45) and Mas-Colell (1989).
6. See Gates (2018).
7. See Marshall (1890 1920).
8. See Diamond et al. (2018, 4).
9. See Samuelson (1948, 457)
10. See Teitelbaum and Axtell (2005, 33) and Axtell and Guerrero (2022).

Chapter 9

1. See Hayek (2014 [1948], 96).
2. See Jobs (2005).
3. See Isaacson (2014).
4. See Tirole (2017).
5. See Attanasio and Pastorino (2018, 7).
6. See Kahneman et al. (1986).
7. See Valentino-DeVries et al. (2012).
8. See Keynes (1923, 80).
9. See Mian and Sufi (2014).
10. See Khan (2016).
11. See Isaacson (2014).
12. See Hayek (2014 [1948], 96).
13. See Schonfeld (2011).
14. See Furman and Orszag (2018).
15. See Mihet and Philippon (2019).
16. See Sherman (1890, 2457).

Part 3

1. See Durkheim (1893, 189,193).

Chapter 10

1. See Coase (1988, 6–8).
2. See Cronon (2009).
3. See Akerlof (1970) and Akerlof (2001).
4. See Rawls (2009 [1971], 137).
5. See Smith (1976 [1759], 74).
6. See Stewart (2012).
7. See Vicini (2012).
8. See Milgrom and Roberts (1990).
9. See Markovits (2020).
10. See Parrott and (2018).
11. See Bernstein (1992, 115).
12. See Smith and Williams (1992).
13. See Brown et al. (2004).
14. See Kollock and O'Brien (1992, 341).
15. See Coase (1992, 717).

Chapter 11

1. See Coase (1937, 387,389).
2. See Raff (1988).
3. See Simon (1951).
4. See Lazear (2000).
5. See Laffont and Matoussi (1995).
6. See Shearer (2004, 514).
7. See Shapiro and Stiglitz (1984), Bowles (1985), and Weisskopf et al. (1983).
8. See US Bureau of Labor Statistics (2020a), U.S. Bureau of Labor Statistics (340).
9. See Jacobson et al. (1993).
10. See Farber (2005).
11. See Winkelmann and Winkelmann (1998).
12. See Grossman and Hart (1986).
13. See Baker and Hubbard (2004).
14. See Burdin et al. (2020).
15. See Falk and Kosfeld (2006), Ziegelmeyer et al. (2012), and Burdin et al. (2018).
16. See Blanchflower and Oswald (1994) and Bowles (1991).
17. See Bowles et al. (2017).
18. See Editorial Board (1987).
19. See Krueger and Ashenfelter (2018).
20. See Card and Krueger (1994).
21. 21% agreed while 5% strongly agreed; see Initiative on Global Markets (IGM), http://www.igmchicago.org/surveys/15-minimum-wage/.
22. See Samuelson (1948, 454).
23. See Coviello et al. (2020).
24. See Azar et al. (2019).
25. See Ruffini (2020, 4).
26. See Starr et al. (2019, 59).
27. See Starr et al. (2019).

Chapter 12

1. Corkery and Silver-Greenberg (2014).
2. Silver-Greenberg (2014).
3. See Vega (1999).
4. See Blanchflower and Oswald (1998). The amount of inheritance, 5,000 pound sterling in 1981, is inflated to 2020 prices (19,600 pound sterling) and converted to dollars using the purchasing power exchange rate: $\frac{19{,}600}{0.716} = \$27{,}374$ OECD data (2020).
5. See Holtz-Eakin et al. (1994).

6. See Black et al. (1996, 64).
7. See McKenzie (2017).
8. See Banerjee and Duflo (2010).
9. See Gross and Souleles (2002).
10. See Gross and Souleles (2002, 181).
11. See Gross and Souleles (2002, 152–53).
12. See Guiso et al. (1996).
13. See https://www.bls.gov/opub/ted/2021/14-6-million-private-sector-jobs-lost-during-second-quarter-2020.htm
14. See US Bureau of Labor Statistics (2016).
15. See Campbell and Mankiw (1989).
16. This section draws upon Ransom and Sutch (2001).
17. See Ransom and Sutch (2001).
18. See Ransom and Sutch (2001, 170).
19. See Hemingway (1961, 421) and Brody (2009).

Chapter 13

1. See Powell (2017).
2. See S&P Dow Jones Indices LLC (2020).
3. See Tooze (2018, 156), Mian and Sufi (2014).
4. See *Economist* (2009).
5. See Tooze (2018, 144).
6. See Liu and Zuo (2019).
7. See Booth and Nolen (2012).
8. Rosenzweig and Wolpin (1993).
9. See McKenzie (2017), and see https://www.npr.org/sections/money/2016/05/20/478883658/episode-702-nigeria-you-win
10. Tooze (2018).
11. See Liu and Zuo (232).
12. See Okun (1975, 91).
13. See Grant et al. (2010).
14. See Iversen and Soskice (2001).
15. See Smith (1976 [1759], 110).
16. See Bowles and Jayadev (2014).
17. See Feynman (1988).

Chapter 14

1. See Smith (1982 [1776], 456).
2. See Nixon et al. (1959).
3. See Carroll (2000 [1865], 92).
4. See Hayek (1975, 442).
5. See Arrow and Hahn (1971, vi–vii).
6. See The Editors of Encyclopaedia Britannica (2019).
7. See Hayek (1945).
8. See Arrow (1951), Debreu (1951), and Arrow and Debreu (1954).
9. See Debreu (1984).
10. See Sen (1985, 11).
11. Banerjee et al. (2002).
12. Grabowski (2002); İşcan (2018).
13. See Roemer (1988, 54).
14. See Arrow (1971, 6).
15. Gauthier (1986, 93,96–97)
16. See Coppage (1915, 17).
17. See Arrow and Hahn (1971, 95).
18. See Foley (1994, 322).
19. See Smale (1976, 212).
20. See Coase (1992, 715).
21. See Smith (2013).
22. See Lewin (2019).
23. See Coase (1960, 31).
24. See Coase (1960, 19).
25. See Coase (1992, 717).
26. See Coase (1960, 15).
27. This setting is inspired by Farrell (1987).
28. See Coase (1992, 717).
29. See Buchanan and Tullock (1962, 47–48).
30. See Farrell (1987).
31. See World Health Organization (2018).
32. See Coase (1992, 714).
33. Barone (1935).
34. Schumpeter (1942, 167, 172).
35. Pareto (2014 [1906], 364).
36. See Trotsky (1933).
37. See Stiglitz (1996, 10).
38. See Arrow and Hahn (1971, 325).

Chapter 15

1. See Holden (1898).
2. See Zelman (2008).
3. See Our World in Data (2018) and Inklaar et al. (2018).
4. See YouGov (2018).
5. See Schumpeter (1942).
6. See Schumpeter (1954, 84).
7. See Moser (2012) and Moser (2016, 241).
8. See Credit Suisse (2020).
9. Bell et al. (2018, 648).
10. See Acemoglu et al. (2017), Maliranta et al. (2012), and Bassanini and Garnero (2013).

11. See Domar and Musgrave (1944).
12. See Autor et al. (2016, 224).
13. US Bureau of Labor Statistics (2020b); Office for National Statistics (2021)
14. See Visser (2015).
15. See Simon (1951).
16. See Simon (1991, 27).
17. See Samuelson (1957, 894).
18. Galbraith (1967, 47).
19. See Alchian and Allen (1969, 320).
20. See Hobbes (1968 [1651], 150).
21. See Smith (2013).
22. See Coase (1937, 387, 389).
23. See Hart (1995, 62).
24. Marx (1967 [1867], 195).
25. See Smith (1982 [1776], 311).
26. Prendergast (1999).
27. See Kohn (1990, 967), Kohn and Schooler (1983), and Kohn (1989).
28. See Kohn (1990, 59).
29. See Mill (2002 [1859], 20–21).
30. See Craig and Pencavel (1992), Craig et al. (1995), and Pencavel (2002).
31. See Dow and Putterman (2000), Dow (2003).
32. See Bellas (1972, 30).
33. See Andreoni and Gee (2012).
34. This calculation is based on data in the 2018–19 Periodic Labour Force Survey (PLFS) Government of India (2002, 53).
35. Lewis (2000).

Chapter 16

1. See Adams (1982, 2827).
2. See Adams (1982, 2824–38).
3. See Maskin (2008).
4. See Bessen and Maskin (2009).
5. See Richards (1995, 2).
6. *Economist* (2007).
7. See Diamond et al. (2018).
8. See Diamond et al. (2018).
9. See Yakovlev (2018).
10. See Lipsey and Lancaster (1956).
11. See Myerson and Satterthwaite (1983).
12. See Gibbard (1973) and Satterthwaite (1975).
13. Gates et al. (2017).
14. See Nobel Prize Committee (2007).
15. See Lepper et al. (1973).
16. See Fehr and Rockenbach (2003).
17. See Lepper et al. (1973).
18. See Schmelz (2021).
19. See Akerlof and Kranton (2010).
20. See Bowles and Polania-Reyes (2012).
21. See Sugden (1986, 3).
22. See Arrow (1971).
23. See Jayadev and Bowles (2006) and Bowles and Jayadev (2007).
24. Wicksell is quoted in Buchanan (1986).
25. Weber (2004).

BIBLIOGRAPHY

Abarca-Gómez, Leandra, Ziad A. Abdeen, Zargar Abdul Hamid, et al. (2017). "Worldwide Trends in Body-Mass Index, Underweight, Overweight, and Obesity from 1975 to 2016: a Pooled Analysis of 2416 Population-Based Measurement Studies in 128.9 Million Children, Adolescents, and Adults." *The Lancet* 390(10113): 2627–2642.

Acemoglu, Daron, and Pascual Restrepo (2020). "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128(6): 2188–2244.

Acemoglu, Daron, James A. Robinson, and Thierry Verdier (2017). "Asymmetric Growth and Institutions in an Interdependent World." *Journal of Political Economy* 125(5): 1245–1305.

Acheson, James M. (1988). *The Lobster Gangs of Maine*. Lebanon, NH: University Press of New England.

Adams, John G. U. (1982). "The Efficacy of Seat Belt Legislation." *SAE Transactions* 91(3): 2824–2838.

Akerlof, George A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3): 488–500.

Akerlof, George A. (2001). "Writing the 'The Market for Lemons': A Personal and Interpretative Essay." *NobelPrize.org*. Accessed March 29, 2021. https://www.nobelprize.org/prizes/economic-sciences/2001/akerlof/article/

Akerlof, George A., and Rachel E. Kranton (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton, NJ: Princeton University Press.

Akino, Masakatsu, and Yujiro Hayami (1975). "Efficiency and Equity in Public Research: Rice Breeding in Japan's Economic Development." *American Journal of Agricultural Economics* 57(1): 1–10.

Alchian, Armen Albert, and William R. Allen (1969). *Exchange and Production: Theory in Use* (3rd ed.). Belmont, CA: Wadsworth Publishing Company.

Alesina, Alberto, and Eliana La Ferrara (2000). "Participation in Heterogeneous Communities." *Quarterly Journal of Economics* 115(3): 847–904.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky (2019). "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *Quarterly Journal of Economics* 134(3): 1557–1626.

Allen, Robert C. (2001). "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War." *Explorations in Economic History* 38(4): 411–447.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2017). "Global Inequality Dynamics: New Findings from WID.world." *American Economic Review* 107(5): 404–09.

Andersen, Steffen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A. List (2011). "Stakes Matter in Ultimatum Games." *American Economic Review* 101(7): 3427–39.

Anderson, Christopher M., and Louis Putterman (2006). "Do Non-strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism." *Games and Economic Behavior* 54(1): 1–24.

Andreoni, James, and Laura K. Gee (2012). "Gun for Hire: Delegated Enforcement and Peer Punishment in Public Goods Provision."

Journal of Public Economics 96(11-12): 1036–1046.

Andreoni, James, and John Miller (2002). "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70(2): 737–753.

Ariely, Dan, Ximena Garcia-Rada, Katrin Gödker, Lars Hornuf, and Heather Mann (2019). "The Impact of Two Different Economic Systems on Dishonesty." *European Journal of Political Economy* 59(4): 179–195.

Arrow, Kenneth J. (1951). "An Extension of the Basic Theorems of Classical Welfare Economics." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 507–532. The Regents of the University of California.

Arrow, Kenneth J. (1971). "Political and Economic Evaluation of Social Effects and Externalities." In Michael D. Intriligator (Ed.), *Frontiers of Quantitative Economics*, 3–23. North Holland.

Arrow, Kenneth J., and Gerard Debreu (1954). "Existence of an Equilibrium for a Competitive Economy." *Econometrica* 22(3): 265–290.

Arrow, Kenneth J., and Frank Hahn (1971). *General Competitive Analysis*. San Francisco, CA: Holden-Day.

Ashton, John K. (2003). "Capital Utilisation and Scale in the English and Welsh Water Industry." *The Service Industries Journal* 23(5): 137–149.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, Tariq Raza, and Eric Verhoogen (2017). "On the Origins and Development of Pakistan's Soccer-Ball Cluster." *The World Bank Economic Review* 30 (Supplement 1): S34–S41.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen (2017). "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan." *Quarterly Journal of Economics* 132(3): 1101–1164.

Attanasio, Orazio, and Elena Pastorino (2018). "Nonlinear Pricing in Village Economies." NBER Working Paper 21718, National Bureau of Economic Research: 7.

Autor, David H., David Dorn, and Gordon H. Hanson (2016). "The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade." *Annual Review of Economics* 8(1): 205–240.

Axtell, Robert, and Omar Guerrero (2022). *Dynamics of Firms: Data, Theories, and Models*. Cambridge, MA: MIT Press.

Azar, José, Emiliano Huet-Vaughn, Ioana Marinescu, Bledi Taska, and Till von Wachter (2019). "Minimum Wage Employment Effects and Labor Market Concentration." NBER Working Paper 26101, National Bureau of Economic Research.

Baker, George P., and Thomas N. Hubbard (2004). "Contractibility and Asset Ownership: On-Board Computers and Governance in US Trucking." *Quarterly Journal of Economics* 119(4): 1443–1479.

Baland, Jean-Marie, Pranab K. Bardhan, and Samuel Bowles (2006). *Inequality, Cooperation and Environmental Sustainability*. Princeton, NJ: Princeton University Press.

Ball, Sheryl, Catherine C. Eckel, and Maria Heracleous (2010). "Risk Aversion and Physical Prowess: Prediction, Choice and Bias." *Journal of Risk and Uncertainty* 41(3): 167–193.

Banerjee, Abhijit V., and Esther Duflo (2010). "Giving Credit Where It Is Due." *Journal of Economic Perspectives* 24(3): 61–80.

Banerjee, Abhijit V. and Esther Duflo (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York, NY: Public Affairs.

Banerjee, Abhijit V., Paul J. Gertler, and Maitreesh Ghatak (2002). "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy* 110(2): 239–280.

Banerjee, Nirmala (2006). "A Survey of Occupations and Livelihoods of Households in West Bengal." Technical report, mimeo Sachetana, Kolkata.

Bar-On, Yinon M, Rob Phillips, and Ron Milo (2018). "The Biomass Distribution on Earth." *Proceedings of the National Academy of Sciences* 115(25): 6506–6511.

Bardhan, Pranab K. (1984). *Land, Labor, and Rural Poverty: Essays in Development Economics*. New York, NY: Columbia University Press.

Bardhan, Pranab K., and Jeff Dayton-Johnson (2002). "Unequal Irrigators: Heterogeneity and Commons Management in Large-scale Multivariate Research." In Elinor Ostrom, Thomas Dietz, Nives Dolsçak, Paul C Stern, Susan Stonich, and Elke U. Weber (Eds.), *The Drama of the Commons*, 87–112. Washington, DC: National Academies Press.

Bardhan, Pranab K., and Jeff Dayton-Johnson (2007). "Inequality and the Governance of Water Resources in Mexico and South India." In Jean- Marie Baland, Pranab K. Bardhan, and Samuel Bowles (Eds.), *Inequality, Cooperation and Environmental Sustainability*, 97–129.

Barkai, Simcha (2020). "Declining Labor and Capital Shares." *The Journal of Finance* 75(5): 2421–2463.

Barone, Enrico (1935). "The Ministry of Production in the Collectivist State." In Friedrich August Hayek (Ed.), *Collectivist Economic Planning*, 245–290. Abingdon, UK: Routledge.

Bassanini, Andrea, and Andrea Garnero (2013). "Dismissal Protection and Worker Flows in OECD Countries: Evidence from Cross-Country/Cross-Industry Data." *Labour Economics* 21(2): 25–41.

Becker, Gary S (1998). *Accounting for Tastes*. Cambridge, MA: Harvard University Press.

Belkin, Douglas (2002). "Boston Firefighters Sick—Or Tired of Working." *Boston Globe*, January 8, 3rd ed., B1.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen (2018). "Who Becomes an Inventor in America? The Importance of Exposure to Innovation." *Quarterly Journal of Economics* 134(2): 648.

Bellas, Carl J. (1972). *Industrial Democracy and the Worker-Owned Firm: A Study of Twenty-One Plywood Companies in the Pacific Northwest*. New York, NY: Irvington Publishers.

Bentham, Jeremy (1996 [1789]). *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation*. Oxford, UK: Clarendon Press.

Bernstein, Lisa (1992). "Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry." *The Journal of Legal Studies* 21(1): 115–157.

Bessen, James, and Eric Maskin (2009). "Sequential Innovation, Patents, and Imitation." *The RAND Journal of Economics* 40(4): 611–635.

Bharti, Nitin Kumar (2018). "Wealth Inequality, Class and Caste in India, 1961-2012." *WID.World Working Paper* 14.

Bittlingmayer, George (1982). "Decreasing Average Cost and Competition: A New Look at the Addyston Pipe Case." *The Journal of Law and Economics* 25(2): 201–229.

Black, Jane, David de Meza, and David Jeffreys (1996). "House Prices, the Supply of Collateral, and the Enterprise Economy." *The Economic Journal* 106(434): 60–75.

Blanchflower, David G., and Andrew J. Oswald (1994). *The Wage Curve*. Cambridge, MA: MIT Press.

Blanchflower, David G., and Andrew J. Oswald (1998). "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1): 26–60.

Blanchflower, David G., and Andrew J. Oswald (2008). "Hypertension and Happiness Across Nations." *Journal of Health Economics* 27(2): 218–233.

Blinder, Alan, Elie R.D. Canetti, David E. Lebow, and Jeremy B. Rudd (1998). *Asking about Prices: A New Approach to Understanding Price Stickiness*. New York, NY: Russell Sage Foundation.

Bonnefoy, Pascale (2020). " 'An End to the Chapter of Dictatorship': Chileans Vote to Draft a New Constitution." *New York Times*, October 25. https://www.nytimes.

com/2020/10/25/world/americas/chile-constitution-plebiscite.html

Booth, Alison L., and Patrick Nolen (2012). "Gender Differences in Risk Behaviour: Does Nurture Matter?" *The Economic Journal* 122(558): F56–F78.

Bowles, Samuel (1985). "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models." *American Economic Review* 75(1): 16–36.

Bowles, Samuel (1991). "The Reserve Army Effect on Wages in a Labour Discipline Model: US, 1954-1987." In T. Mizoguchi (Ed.), *Making Economies More Efficient and More Equitable: Factors Determining Income Distribution*, 385–406. Oxford: Oxford University Press.

Bowles, Samuel, and Wendy Carlin (2020a). "The Coming Battle for the COVID-19 Narrative." *Vox EU*, April 10. https://voxeu.org/article/coming-battle-covid-19-narrative.

Bowles, Samuel, and Wendy Carlin (2020b). "Shrinking Capitalism." AEA *Papers and Proceedings* 110(1): 372–77.

Bowles, Samuel, Wendy Carlin, and M Stevens (2017). *The Economy: Economics for a Changing World*. Oxford, UK: Oxford University Press.

Bowles, Samuel, and Herbert Gintis (2002). "Homo reciprocans." *Nature* 415(6868): 125–127.

Bowles, Samuel, and Arjun Jayadev (2007). "Garrison America." *The Economists' Voice* 4(2): 1–7.

Bowles, Samuel, and Arjun Jayadev (2014). "One Nation under Guard." *New York Times*, February 15. https://opinionator.blogs.nytimes.com/2014/02/15/one-nation-under-guard/.

Bowles, Samuel, and Sandra Polania-Reyes (2012). "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature* 50(2): 368–425.

Brody, Richard (2009). "The Women." *The New Yorker*, November 11. https://www.newyorker.com/culture/richard-brody/the-women.

Brown, Martin, Armin Falk, and Ernst Fehr (2004). "Relational Contracts and the Nature of Market Interactions." *Econometrica* 72(3): 747–780.

Buchanan, James M. Jr. (1986). "The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1986." *NobelPrize.org*. Accessed March 29, 2021, https://www.nobelprize.org/prizes/economic-sciences/1986/buchanan/lecture/.

Buchanan, James M. Jr., and Gordon Tullock (1962). *The Calculus of Consent*. Ann Arbor, MI: University of Michigan Press.

Burdin, Gabriel, Simon D. Halliday, and Fabio Landini (2018). "The Hidden Benefits of Abstaining from Control." *Journal of Economic Behavior & Organization* 147(3): 1–12.

Burdin, Gabriel, Simon D. Halliday, and Fabio Landini (2020). "Why Using Technology to Spy on Home-working Employees May Be a Bad Idea." *LSE Business Review*, June 17. https://blogs.lse.ac.uk/businessreview/2020/06/17/why-using-technology-to-spy-on-home-working-employees-may-be-a-bad-idea/.

Camerer, Colin F. (2011). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Cameron, Lisa A. (1999). "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia." *Economic Inquiry* 37(1): 47–59.

Campbell, John Y., and N. Gregory Mankiw (1989). "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence." NBER *Macroeconomics Annual* 4(1): 185–216.

Card, David, and Alan B. Krueger (1994). "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84(4): 772–793.

Cardenas, Juan-Camilo, Toh-Kyeong Ahn, and Elinor Ostrom (2004). "Communication and Co-operation in a Common-Pool Resource

Dilemma: A Field Experiment." In Steffen Huck (Ed.), *Advances in Understanding Strategic Behaviour*, 258–286. Heidelberg, Netherlands: Springer.

Carpenter, Jeffrey, Stephen Burks, and Eric Verhoogen (2005). "Comparing Students to Workers: The Effects of Social Framing on Behavior in Distribution Games." *Research in Experimental Economics* 10(1): 261–290.

Carpenter, Jeffrey, Eric Verhoogen, and Stephen Burks (2005). "The Effect of Stakes in Distribution Experiments." *Economics Letters* 86(3): 393–398.

Carroll, Lewis (2000 [1865]). *The Annotated Alice: The Definitive Edition of "Alice's Adventures in Wonderland" and "Through the Looking-Glass" by Lewis Carroll*. New York: Norton.

Caves, Douglas W., Laurits R. Christensen, and Michael W. Tretheway (1984). "Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ." *The RAND Journal of Economics* 15(4): 471–489.

Cecchetti, Stephen, and Schoenholtz Kermit L. (2017). "Looking Back: The Financial Crisis Began 10 Years Ago This Week." *Money, Banking and Financial Markets*, August 7. https://www.moneyandbanking.com/commentary/2017/8/6/looking-back-the-financial-crisis-began-10-years-ago-this-week.

Chomo, Grace, and Michael J. Ferrantino (2000). *NAFTA Environmental Impacts on North American Fisheries*. Washington, DC: US International Trade Commission.

Christensen, Laurits R., and William H. Greene (1976). "Economies of Scale in US Electric Power Generation." *Journal of Political Economy* 84(4): 655–676.

Clark, W. A. V. (1991). "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model." *Demography* 28(1): 1–19.

Coase, Ronald H. (1937). "The Nature of the Firm." *Economica* 4(16): 368–405.

Coase, Ronald H. (1960). "The Problem of Social Cost." *Journal of Law and Economics* 3(1): 1–44.

Coase, Ronald H. (1988). *The Firm, the Market, and the Law*. Chicago, IL: University of Chicago Press.

Coase, Ronald H. (1992). "The Institutional Structure of Production." *American Economic Review* 82(4): 713–719.

Cockburn, Iain M., and Rebecca M. Henderson (2001). "Scale and Scope in Drug Development: Unpacking the Advantages of Size in Pharmaceutical Research." *Journal of Health Economics* 20(6): 1033–1057.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe (2016). "Using Big Data to Estimate Consumer Surplus: The Case of Uber." NBER Working Paper 22627, National Bureau of Economic Research.

Cohn, Alain, Michel André Maréchal, David Tannenbaum, and Christian Lukas Zünd (2019). "Civic Honesty Around the Globe." *Science* 365(6448): 70–73.

Colchero, M. Arantxa, Barry M. Popkin, Juan A. Rivera, and Shu Wen Ng (2016). "Beverage Purchases from Stores in Mexico under the Excise Tax on Sugar Sweetened Beverages: Observational Study." *British Medical Journal* 352.

Congressional Budget Office (2020). "The Distribution of Household Income, 2017—Data and Supplemental Information: Data Underlying Exhibits."

Coppage v. Kansas (1915). "Pitney, Mahlon and Supreme Court of the United States, 236 U.S. 1."

Corkery, Michael, and Jessica Silver-Greenberg (2014). "Miss a Payment? Good Luck Moving That Car." *The New York Times*, September 24. https://www.cnbc.com/2014/09/25/miss-a-payment-good-luck-moving-that-car.html.

Cortada, James W. (2019). *IBM: The Rise and Fall and Reinvention of a Global Icon*. Cambridge, MA: MIT Press.

Coviello, Decio, Erika Deserranno, and Nicola Persico (2020). "Minimum Wage and

Individual Worker Productivity: Evidence from a Large US Retailer." Technical report.

Craig, Ben, and John Pencavel (1992). "The Behavior of Worker Cooperatives: The Plywood Companies of The Pacific Northwest." *American Economic Review* 81(5): 1083–1105.

Craig, Ben, John Pencavel, Henry S. Farber, and Alan B. Krueger (1995). "Participation and Productivity: A Comparison of Worker Cooperatives and Conventional Firms in the Plywood Industry." *Brookings Papers on Economic Activity. Microeconomics* 1995(1): 121–174.

Credit Suisse (2020). "Global Wealth Databook." Credit Suisse Research Institute.

Cronon, William (2009). *Nature's Metropolis: Chicago and the Great West*. New York, NY: Norton.

Dawson, Philip J., and Lionel J. Hubbard (1987). "Management and Size Economies in the England and Wales Dairy Sector." *Journal of Agricultural Economics* 38(1): 27–38.

Debreu, Gerard (1951). "The Coefficient of Resource Utilization." *Econometrica* 19(3): 273–292.

Debreu, Gérard (1984). "La Supériorité de l'Économie Libérale Est Incontestable et Mathématiquement Démontrée." *Le Figaro* (10 March): 80.

de Groot, Hans, Walter W. McMahon, and J. Fredericks Volkwein (1991). "The Cost Structure of American Research Universities." *The Review of Economics and Statistics* 73(3): 424–431.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020). "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics* 135(2): 561–644.

de V. Cavalcanti, Tiago V., and Jose Tavares (2008). "Assessing the "Engines of Liberation": Home Appliances and Female Labor Force Participation." *The Review of Economics and Statistics* 90(1): 81–88.

Dhar, Tirtha, Jean-Paul Chavas, Ronald W. Cotterill, and Brian W. Gould (2005). "An Econometric Analysis of Brand-Level Strategic Pricing Between Coca-Cola Company and PepsiCo." *Journal of Economics & Management Strategy* 14(4): 905–931.

Diamond, Rebecca, Tim McQuade, and Franklin Qian (2018). "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco." NBER Working Paper 24181, National Bureau of Economic Research: 4.

Domar, Evsey D., and Richard A. Musgrave (1944). "Proportional Income Taxation and Risk-Taking." *Quarterly Journal of Economics* 58(3): 388–422.

Dow, Gregory K. (2003). *Governing the Firm: Workers' Control in Theory and Practice*. Cambridge, UK: Cambridge University Press.

Dow, Gregory K., and Louis Putterman (2000). "Why Capital Suppliers (Usually) Hire Workers: What We Know and What We Need to Know." *Journal of Economic Behavior & Organization* 43(3): 319–336.

Dube, Arindrajit, T. William Lester, and Michael Reich (2016). "Minimum Wage Shocks, Employment Flows, and Labor Market Frictions." *Journal of Labor Economics* 34(3): 663–704.

Dupas, Pascaline, and Edward Miguel (2017). "Impacts and Determinants of Health Levels in Low-Income Countries." In Esther Duflo and Abhijit V. Banerjee (Eds.), *Handbook of Economic Field Experiments*, 3–93. Elsevier.

Durkheim, Émile (1893). *De la Division du Travail Social: Étude sur l'Organisation des Sociétés Supérieures*. Paris: F. Alcan.

Economist, The (2007). "Intelligent Design." *The Economist*. Accessed March 29. https://www.economist.com/finance-and-economics/2007/10/18/intelligent-design.

Economist, The (2009). "The Bankruptcy of General Motors: A Giant Falls." *The Economist*. Accessed March 29, 2021, https://www.economist.com/briefing/2009/06/04/a-giant-falls.

Edgeworth, Francis Ysidro (1881). *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. Abingdon, UK: Kegan Paul.

Editorial Board (1987). "The Right Minimum Wage: $0.00." *The New York Times*, January 14. https://www.nytimes.com/1987/01/14/opinion/the-right-minimum-wage-0.00.html.

Engel, Christoph (2011). "Dictator Games: A Meta Study." *Experimental Economics* 14(4): 583–610.

Ertan, Aytekin, Stefan Lewellen, and Jacob K Thomas (2018). "The Long-Run Average Cost Puzzle." Available at SSRN.

Falk, Armin, and James J. Heckman (2009). "Lab Experiments Are a Major Source of Knowledge in the Social Sciences." *Science* 326(5952): 535–538.

Falk, Armin, and Michael Kosfeld (2006). "The Hidden Costs of Control." *American Economic Review* 96(5): 1611–1630.

Farber, Henry S. (2005). "What Do We Know about Job Loss in the United States? Evidence from the Displaced Workers Survey, 1984–2004." *Federal Reserve Bank of Chicago, Economic Perspectives* 29(2): 13–28.

Farrell, Joseph (1987). "Information and the Coase Theorem." *Journal of Economic Perspectives* 1(2): 113–129.

Fehr, Ernst, and Urs Fischbacher (2002). "Why Social Preferences Matter—the Impact of Non-selfish Motives on Competition, Cooperation, and Incentives." *The Economic Journal* 112(478): C1–C33.

Fehr, Ernst, and Simon Gächter (2000a). "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90(4): 980–994.

Fehr, Ernst, and Simon Gächter (2000b). "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14(3): 159–181.

Fehr, Ernst, and Simon Gächter (2002). "Altruistic Punishment in Humans." *Nature* 415(6868): 137–140.

Fehr, Ernst, and Andreas Leibbrandt (2011). "A Field Study on Cooperativeness and Impatience in the Tragedy of the Commons." *Journal of Public Economics* 95(9-10): 1144–1155.

Fehr, Ernst, and John A. List (2004, 09) "The Hidden Costs and Returns of Incentives-Trust and Trustworthiness among CEOs." *Journal of the European Economic Association* 2(5): 743–771.

Fehr, Ernst, and Bettina Rockenbach (2003). "Detrimental Effects of Sanctions on Human Altruism." *Nature* 422(6928): 137–140.

Feldstein, Martin S., Ted Halstead, and N. Gregory Mankiw (2017). "A Conservative Case for Climate Action." *New York Times*, February 8. https://www.nytimes.com/2017/02/08/opinion/a-conservative-case-for-climate-action.html.

Feynman, Richard (1988). "Horizon: The Pleasure of Finding Things Out." *British Broadcasting Corporation*, March 27. Audio, 13:10. https://www.bbc.co.uk/programmes/p018dvyg.

Foley, Duncan K. (1994). "A Statistical Equilibrium Theory of Markets." *Journal of Economic Theory* 62(2): 322.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton (1994). "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior* 6(3): 347–69.

Frank, Kenneth T., Brian Petrie, Jae S. Choi, and William C. Leggett (2005). "Trophic Cascades in a Formerly Cod-Dominated Ecosystem." *Science* 308(5728): 1621–1623.

Fremstad, Anders and Mark Paul (2019). "The Impact of a Carbon Tax on Inequality." *Ecological Economics* 163(1): 88–97.

Frey, Bruno S., and Iris Bohnet (1995). "Institutions Affect Fairness: Experimental Investigations." *Journal of Institutional and Theoretical Economics* 151(2): 286–303.

Frey, Bruno S., and Iris Bohnet (1996). "Cooperation, Communication and Communitarianism: An Experimental Approach." *Journal of Political Philosophy* 4(4): 322–336.

Friedlaender, Ann F., Clifford Winston, and Kung Wang (1983). "Costs, Technology, and Productivity in the US Automobile Industry." *The Bell Journal of Economics* 14(1): 1–20.

Friedman, Milton (1953). *Essays in Positive Economics*. Chicago, IL: University of Chicago Press.

Friedman, Milton (1974). *There's No Such Thing as a Free Lunch*. Chicago, IL: Open Court.

Furman, Jason, and Peter R Orszag "Slower Productivity and Higher Inequality: Are They Related?" *Peterson Institute for International Economics Working Paper 2018*.

Galbraith, John Kenneth (1967). *The New Industrial State*. Boston: Houghton Mifflin.

García-Barrio, Raúl, and Luis García-Barrios (1990). "Environmental and Technological Degradation in Peasant Agriculture: A Consequence of Development in Mexico." *World Development* 18(11): 1569–1585.

Gates, Bill (2018). "At the time I was in college, this was basically how the global economy worked." Twitter Post, August 21, 1:09pm.

Gates, Guilbert, Jack Ewing, Karl Russell, and Derek Watkins (2017). "How Volkswagen's 'Defeat Devices' Worked." *The New York Times*, March 16. https://www.nytimes.com/interactive/2015/business/international/vw-diesel-emissions-scandal-explained.html.

Gauthier, David (1986). *Morals by Agreement*. New York, NY: Oxford University Press.

Gibbard, Allan (1973). "Manipulation of Voting Schemes: A General Result." *Econometrica* 41(4): 587–601.

Giri, Jacques (1983). *Le Sahel Demain: Catastrophe ou Renaissance?* Paris, France: Karthala Editions.

Gneezy, Uri, Andreas Leibbrandt, and John A. List (2016). "Ode to the Sea: Workplace Organizations and Norms of Cooperation." *The Economic Journal* 126(595): 1856–1883.

Gneezy, Uri, and Aldo Rustichini (2000a). "A Fine is a Price." *The Journal of Legal Studies* 29(1): 1–17.

Gneezy, Uri, and Aldo Rustichini (2000b). "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115(3): 791–810.

Government of India (2020). "Period Labor Force Survey." Accessed April 1, 2021, http://mospi.nic.in/sites/default/files/publication_reports/Annual_HL.pdf.

Grabowski, Richard (2002). "East Asia, Land Reform and Economic Development." *Canadian Journal of Development Studies* 23(1): 105–126.

Grant, Charles, Christos Koulovatianos, Alexander Michaelides, and Mario Padula (2010). "Evidence on the Insurance Effect of Redistributive Taxation." *The Review of Economics and Statistics* 92(4): 965–973.

Greenberger, Scott S. (2003). "Sick Day Abuses Focus of Fire Talks." *Boston Globe*: September 17, B7.

Greif, Avner (1994). "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102(5): 912–950.

Griliches, Zvi, and Vidar Ringstad (1971). *Economies of Scale and the Form of the Production Function: An Econometric Study of Norwegian Manufacturing Establishment Data*. Amsterdam, Netherlands: North Holland.

Gross, David B., and Nicholas S. Souleles (2002). "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data." *Quarterly Journal of Economics* 117(1): 149–185.

Grossman, Sanford J., and Oliver D. Hart (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4): 691–719.

Grubs, Matt (2017). "Sugar Tax Fails." *Santa Fe Reporter*, Accessed August 4, 2020. https://www.sfreporter.com/news/2017/05/02/sugar-tax-fails/.

Guiso, Luigi, Tullio Jappelli, and Daniele Terlizzese (1996). "Income Risk, Borrowing Constraints, and Portfolio Choice." *American Economic Review* 86(1): 158–172.

Güth, Werner, Rolf Scmittberger, and Bernd Schwarze (1982). "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior and Organization* 3(4): 367–88.

Hall, Robert E. (1988). "The Relation between Price and Marginal Cost in US Industry." *Journal of Political Economy* 96(5): 921–947.

Hall, Robert E., and Charles I. Jones (1999). "Why Do Some Countries Produce So Much More Output Per Worker than Others?" *Quarterly Journal of Economics* 114(1): 83–116.

Hardin, Garrett (1968). "The Tragedy of the Commons." *Science* 162(3859): 1243–1248.

Hart, Oliver (1995). *Firms, Contracts, and Financial Structure* Oxford, UK: Clarendon Press.

Hausmann, Ricardo, and Dani Rodrik (2003). "Economic Development as Self-Discovery." *Journal of Development Economics* 72(2): 603–633.

Hayek, Friedrich August (1945). "The Use of Knowledge in Society." *American Economic Review* 35(4): 519–530.

Hayek, Friedrich August (1975). "The Pretence of Knowledge." *The Swedish Journal of Economics* 77(4): 433–442.

Hayek, Friedrich August (2014 [1948]). "The Meaning of Competition." In Bruce Caldwell (Ed.), *The Market and Other Orders*, The Collected Works of F. A. Hayek, 105–119. Chicago, IL: University of Chicago Press.

Hemingway, Ernest (1961). "The Snows of Kilimanjaro." In *The Snows of Kilimanjaro, and Other Stories*. New York: Scribner.

Henderson, Rebecca M. (2000). "Drug Industry Mergers Won't Necessarily Benefit R&D." *Research Technology Management* 43(4): 10.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin F. Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath (2001). "In Search of *Homo Economicus*: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review* 91(2): 73–78.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010a). "Most people are not WEIRD." *Nature* 466(7302): 29.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010b). "The WEIRDest People in the World?" *Behavioral and Brain Sciences* 33(2-3): 61–83.

Henrich, Robert, Robert Boyd, Samuel Bowles, Colin Camerer F., Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton, and David Tracer (2005). " 'Economic man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-scale Societies." *Behavioral and Brain Sciences* 28(6): 795–855.

Herndon, William H., and Jesse William Weik (1889). *Herndon's Lincoln: The True Story of a Great Life*. Chicago, IL: Belford Clarke.

Hill, Andrew, Junzheng Wang, Jacob Levi, Katie Heath, and Joseph Fortunak (2020). "Minimum Costs to Manufacture New Treatments for COVID-19." *Journal of Virus Eradication* 6(2): 61.

Himansu, Lanjouw, Peter and Nicholas Stern (2018). *How Lives Change: Palanpur, India, and Development Economics*. Oxford, UK: Oxford University Press.

Hobbes, Thomas (1968 [1651]). *Leviathan*. London: Penguin.

Hofferth, Sandra L., Sarah M. Flood, and Matthew Sobek (2013). "American Time Use Survey Data Extract System: Version 2.4" [machine-readable database]. Maryland Population Research Center, University of Maryland, College Park, Maryland, and Minnesota Population Center, University of Minnesota.

Hoffman, Elizabeth, Kevin A. McCabe, and Vernon L. Smith (1996). "On Expectations and the Monetary Stakes in Ultimatum Games." *International Journal of Game Theory* 25(3): 289–301.

Holden v. Hardy (1898). Supreme Court of the United States. U.S. Reports, 169 U.S. 366.R.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen (1994). "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." *Journal of Political Economy* 102(1): 53–75.

Huberman, Michael (2004). "Working Hours of the World Unite? New International Evidence of Worktime, 1870–1913." *The Journal of Economic History* 64(4): 964–1001.

Hume, David (1967 [1742]). *A Treatise of Human Nature: Reprinted from the Original Ed. in Three Volumes and Ed., with an Analytical*

Index by L. A. Selby-Bigge. Oxford, UK: Clarendon Press.

Ibn Battuta (2013 [1929]). *Travels in Asia and Africa*. Abingdon, UK: Routledge.

Initiative on Global Markets (IGM) Forum (2015, 22 September). "$15 Minimum Wage." Chicago Booth, University of Chicago.

Inklaar, Robert, Herman de Jong, Jutta Bolt, and Jan Luiten van Zanden (2018). "Rebasing Maddison: New Income Comparisons and the Shape of Long-Run Economic Development." *GGDC Research Memorandum* 174: 1–67.

Isaacson, Walter (2014). *The Innovators: How a Group of Inventors, Hackers, Geniuses and Geeks Created the Digital Revolution*. New York, NY: Simon and Schuster.

İşcan, Talan B. (2018). "Redistributive Land Reform and Structural Change in Japan, South Korea, and Taiwan." *American Journal of Agricultural Economics* 100(3): 732–761.

Iversen, Torben, and David Soskice (2001). "An Asset Theory of Social Policy Preferences." *American Political Science Review* 95(4): 875–893.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan (1993). "Earnings Losses of Displaced Workers." *American Economic Review* 83(4): 685–709.

Jaffe, William (1972). "Pareto Translated: A Review Article." *Journal of Economic Literature* 10(4): 1190–1201.

Jayadev, Arjun, and Samuel Bowles (2006). "Guard Labor." *Journal of development Economics* 79(2): 328–348.

Jensen, Jørgen Dejgård, and Sinne Smed (2013). "The Danish Tax on Saturated Fat—Short Run Effects on Consumption, Substitution Patterns and Consumer Prices of Fats." *Food Policy* 42(1): 18–31.

Jensen, Robert T., and Nolan H. Miller (2008). "Giffen Behavior and Subsistence Consumption." *American Economic Review* 98(4): 1553–77.

Jobs, Steve (2005). " 'You've got to find what you love,' Jobs says." Commencement Speech for Stanford University, 12 June, https://news.stanford.edu/2005/06/14/jobs-061505/.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986). "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* 76(4): 728–741.

Kahneman, Daniel, and Alan B. Krueger (2006). "Developments in the Measurement of Subjective Well-Being." *Journal of Economic perspectives* 20(1): 3–24.

Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur Stone (2004). "Toward National Well-being Accounts." *American Economic Review* 94(2): 429–434.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin (1997). "Back to Bentham? Explorations of Experienced Utility." *Quarterly Journal of Economics* 112(2): 375–406.

Kapur, Devesh (2002). "The Causes and Consequences of India's IT Boom." *India Review* 1(2): 91–110.

Kavlak, Goksin, James McNerney, and Jessika E. Trancik (2018). "Evaluating the Causes of Cost Reduction in Photovoltaic Modules." *Energy Policy* 123(1): 700–710.

Keynes, John Maynard (1921). *A Treatise on Probability*. New York, NY: Macmillan.

Keynes, John Maynard (1923). *A Tract on Monetary Reform*. New York, NY: Macmillan.

Keynes, John Neville (1897). *The Scope and Method of Political Economy*. New York, NY: Macmillan.

Khan, Lina M. (2016). "Amazon's Antitrust Paradox." *Yale Law Journal* 126(3): 710.

King County Recorder's Office (1929). "Deeds." *Seattle Civil Rights and Labor History Project* 1450: 348.

Klette, Tor Jakob (1999). "Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data." *The Journal of Industrial Economics* 47(4): 451–476.

Kohn, Melvin L. (1989). *Class and Conformity: A Study in Values*. Chicago, IL: University of Chicago Press.

Kohn, Melvin L. (1990). "Unresolved Issues in the Relationship between Work and Personality." *Contemporary Sociological Studies* 3(1): 27–46.

Kohn, Melvin L., and Carmi Schooler (1983). *Work and Personality: An Inquiry into the Impact of Social Stratification*. Norwood, NJ: Ablex Pub. Corp.

Kollock, Peter (1998). "Transforming Social Dilemmas: Group Identity and Co-operation." In Peter Danielson (Ed.), *Modeling Rationality, Morality, and Evolution*, 185–209. Oxford, UK: Oxford University Press.

Kollock, Peter, and Jodi O'Brien (1992). "The Social Construction of Exchange." In Edward J. Lawler and Barry. Markovsky (Eds.), *Advances in Group Processes* 89–112 Greenwich, CT: JAI Press.

Koshal, Rajindar K., and Manjulika Koshal (1999). "Economies of Scale and Scope in Higher Education: A Case of Comprehensive Universities." *Economics of Education Review* 18(2): 269–277.

Krueger, Alan B., and Orley Ashenfelter (2018). "Theory and Evidence on Employer Collusion in the Franchise Sector." Working Paper 24831, National Bureau of Economic Research.

Krueger, Alan B., and David A. Schkade (2008). "The Reliability of Subjective Well-Being Measures." *Journal of Public Economics* 92(8-9): 1833–1845.

Laband, David N., and Bernard F. Lentz (2003). "New Estimates of Economies of Scale and Scope in Higher Education." *Southern Economic Journal* 70(1): 172–183.

Laffont, Jean-Jacques, and Mohamed Salah Matoussi (1995). "Moral Hazard, Financial Constraints, and Sharecropping in El Oulja." *The Review of Economic Studies* 62(3): 381–399.

Lafond, Francois, Diana Seave Greenwald, and J Doyne Farmer (2020). "Can Stimulating Demand Drive Costs Down? World War II as a Natural Experiment." Available at SSRN.

Lanjouw, Peter, and Nicholas Stern (1998). *Economic Development in Palanpur over Five Decades*. Oxford, UK: Oxford University Press.

Lazear, Edward P. (2000). "Performance Pay and Productivity." *American Economic Review* 90(5): 1346–1361.

Lazear, Edward P., Kathryn L. Shaw, and Christopher Stanton (2016). "Making Do with Less: Working Harder during Recessions." *Journal of Labor Economics* 34(S1): S333–S360.

Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research." In John H. Kagel and Alvin E. Roth (Eds.), *The Handbook of Experimental Economics*, 111–194. Princeton, NJ: Princeton University Press.

Leontief, Wassily (1983). "National Perspective: The Definition of Problems and Oopportunities." In *The Long-Term Impact of Technology on Employment and Unemployment*. A *National Academy of Engineering Symposium*, 3–8. Washington, D.C.: National Academy Press.

Leopold, Aldo (2020 [1949]). A *Sand County Almanac, and Sketches Here and There*. Oxford, UK: Oxford University Press.

Lepper, Mark R., David Greene, and Richard E. Nisbett (1973). "Undermining Children's Intrinsic Interest with Extrinsic Reward: A Test of the "Overjustification" Hypothesis.." *Journal of Personality and Social Psychology* 28(1): 129.

Levitt, Stephen D., and John List (2007). "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives* 21(2): 153–174.

Lewin, Robert (2019). "Railroads Can Teach Utilities How to Stop Causing Fires." *Homeland Security Today*, November 15. https://www.hstoday.us/subject-matter-areas/infrastructure-security/railroads-can-teach-utilities-how-to-stop-causing-fires/.

Lewis, W. Arthur (2000). "Economic Development with Unlimited Supplies of Labour." *The Manchester School* 22(2): 139–191.

Lipsey, Richard G., and Kelvin Lancaster (1956). "The General Theory of Second Best." *The Review of Economic Studies* 24(1): 11–32.

Liu, Elaine M., and Sharon Xuejing Zuo (2019). "Measuring the Impact of Interaction between Children of a Matrilineal and a

Patriarchal Culture on Gender Differences in Risk Aversion." *Proceedings of the National Academy of Sciences* 116(14): 6713–6719.

Loewenstein, George F., Leigh Thompson, and Max H. Bazerman (1989). "Social Utility and Decision Making in Interpersonal Contexts." *Journal of Personality and Social Psychology* 57(3): 426.

Lucas, Richard E., Andrew E. Clark, Yannis Georgellis, and Ed Diener (2003). "Reexamining Adaptation and the Set Point Model of Happiness: Reactions to Changes in Marital Status." *Journal of personality and social psychology* 84(3): 527.

Makowski, Louis, and Joseph M. Ostroy (2001). "Perfect Competition and the Creativity of the Market." *Journal of Economic Literature* 39(2): 479–535.

Maliranta, Mika, Niku Määttänen, and Vesa Vihriälä (2012). "Are the Nordic Countries Really Less Innovative than the US?" *Vox EU*, Accessed March 31. https://voxeu.org/article/nordic-innovation-cuddly-capitalism-really-less-innovative.

Markovits, Daniel (2020). "How McKinsey Destroyed the Middle Class." *The Atlantic*, February 3. https://www.theatlantic.com/ideas/archive/2020/02/how-mckinsey-destroyed-middle-class/605878/.

Marshall, Alfred ([1890] 1920). *Principles of Economics* (8th ed.). London, UK: MacMillan and Company.

Martín, Juan Carlos, and Augusto Voltes-Dorta (2011). "The Econometric Estimation of Airports' Cost Function." *Transportation Research Part B: Methodological* 45(1): 112–127.

Marx, Karl (1967 [1867]). *Capital: A Critique of Political Economy (Volume 1), The Process of Capitalist Production*. New York, NY: International Publishers.

Mas-Colell, Andreu (1989). "Capital Theory Paradoxes: Anything Goes." In George R. Feiwel (Ed.), *Joan Robinson and Modern Economic Theory*, 505–520. Heidelberg, Netherlands: Springer.

Maskin, Eric (2008). "Biography of Eric Maskin." In Karl Grandin (Ed.), *The Nobel Prizes*. Nobel Media AB 2021.

McDonald, Rory, and Suresh Kotha (2015). "Boeing 787: Manufacturing a Dream." *Harvard Business School Case Study* 615-048

McKenzie, David (2017). "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition." *American Economic Review* 107(8): 2278–2307.

Mian, Atif, and Amir Sufi (2014). *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again*. Chicago, IL: University Of Chicago Press.

Mihet, Roxana, and Thomas Philippon (2019). "The Economics of Big Data and Artificial Intelligence." In *Disruptive Innovation in Business and Finance in the Digital World*, Volume 20 of *International Finance Review*, 29–43. Bingley, UK: Emerald Publishing Limited.

Milgrom, Paul, and John Roberts (1990). "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review* 80(3): 511–528.

Mill, John Stuart (1844). *Essays on Some Unsettled Questions of Political Economy*. London, UK: John W. Parker.

Mill, John Stuart (1848). *The Collected Works of John Stuart Mill, Volume II: The Principles of Political Economy with Some of Their Applications to Social Philosophy (Books I-II)*. Online Library of Liberty.

Mill, John Stuart (1998 [1861]). *Utilitarianism*. Oxford, UK: Oxford University Press.

Mill, John Stuart (2002 [1859]). *On Liberty*. New York, NY: Dover Publications.

Miravete, Eugenio J., Katja Seim, and Jeff Thurk (2020). "One Markup to Rule Them All: Taxation by Liquor Pricing Regulation." *American Economic Journal: Microeconomics* 12(1): 1–41.

Moser, Petra (2012). "Innovation without Patents: Evidence from World's Fairs." *The Journal of Law & Economics* 55(1): 43–74.

Moser, Petra (2016). "Patents and Innovation in Economic History." *Annual Review of Economics* 8(1): 241–258.

Moyer, Justin Wm. (2020). "Racist housing covenants haunt property records across the country. New laws make them easier to remove." *The Washington Post*, 22 October. Accessible at: https://www.washingtonpost.com/local/racist-housing-covenants/2020/10/21/9d262738-0261-11eb-8879-7663b816bfa5_story.html Access 1 April 2021.

Munier, Bertrand, and Costin Zaharia (2002). "High Stakes and Acceptance Behavior in Ultimatum Bargaining." *Theory and Decision* 53(3): 187–207.

Murphy, Jessica (2017). "Can Cod Comeback Keep a Canadian Fishery Afloat?" *BBC News*, June 22. https://www.bbc.com/news/world-us-canada-40252481.

Myerson, Roger B., and Mark A. Satterthwaite (1983). "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 29(2): 265–281.

Nasar, Sylvia (1998). *A Beautiful Mind: The Life of Mathematical Genius and Nobel Laureate John Nash*. New York, NY: Simon & Schuster.

Nash, John F. (1950). "Equilibrium Points in n-person Games." *Proceedings of the National Academy of Sciences* 36(1): 48–49.

Nash, John F. (1951). "Non-cooperative Games." *Annals of Mathematics* 54(2): 286–295.

Nerlove, Marc (1963). "Returns to Scale in Electricity Supply." In Carl Christ (Ed.), *Measurement in Economies: Studies in Mathematical Economies and Econometrics in Memory of Yehuda Granfeld*. Redwood City, CA: Stanford University Press.

Nixon, Richard, Nikita Khrushchev, US Embassy, and Soviet Union Moscow (1959). "The Kitchen Debate–Transcript." *CIA Library* 24.

Nobel Prize Committee (2007). "Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2007: Mechanism Design Theory." Accessed May 13, 2019, https://www.nobelprize.org/uploads/2018/06/advanced-economicsciences2007.pdf.

Nordhaus, William D. (1996). "Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not." In Timothy F. Bresnahan and Robert J. Gordon (Eds.), *The Economics of New Goods* 27–70 Chicago, IL: University of Chicago Press.

OECD (2014). "Gender, Institutions and Development Database 2014 (GIDDB): Unpaid Care Work 2014." Accessed 1 April 2021, https://stats.oecd.org/Index.aspx?DataSetCode=GIDDB2014.

OECD data (2020). "Purchasing Power Parities (PPP)." data.oecd.org Accessed 1 April 2021, https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.

Office for National Statistics, UK (2021). "Unemployment Rate (Aged 16 and Over, Seasonally Adjusted)." Accessed April 1, 2021, https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsx/lms.

Oh, Seung-Yun, Yongjin Park, and Samuel Bowles (2012). "Veblen Effects, Political Representation, and the Reduction in Working Time Over the 20th Century." *Journal of Economic Behavior & Organization* 83(2): 218–242.

Okun, Arthur (1975). *Equality and Efficiency, the Big Tradeoff*. Brookings Institution Press.

Ostrom, Elinor (2015 [1990]). *Governing the Commons*. Cambridge, UK: Cambridge University Press.

Ostrom, Elinor, Joanna Burger, Christopher B. Field, Richard B. Norgaard, and David Policansky (1999). "Revisiting the Commons: Local Lessons, Global Challenges." *Science* 284(5412): 278–282.

Oswald, Andrew J., and Stephen Wu (2010). "Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the USA." *Science* 327(5965): 576–579.

Ouchi, William G. (1980). "Markets, Bureaucracies, and Clans." *Administrative Science Quarterly* 25(1): 129–141.

Our World in Data (2018). "Divergence of GDP Per Capita among Latecomers to the Capitalist Revolution (1800-2016)."

Pareto, Vilfredo (1971). "Manuale di Economia Politica (Societá Editrice Libraria, Milano, Italy, 1906). English translation: V. Pareto,

Manual of Political Economy, trans. A. S. Schwier, 1971.

Pareto, Vilfredo (2014 [1906]). *Manual of Political Economy: A Critical and Variorum Edition*. Oxford: Oxford University Press.

Parrott, James A., and Michael Reich (2018). "An Earnings Standard for New York City's App-based Drivers: Economic Analysis and Policy Assessment." Technical report, Center for New York City Affairs.

Pencavel, John (2002). *Worker Participation: Lessons from Worker Co-ops of the Pacific Northwest*. New York, NY: Russell Sage Foundation.

Perese, Kevin (2010). "Input-Output Model Analysis: Pricing Carbon Dioxide Emissions." Tax Analysis Division, Congressional Budget Office Working Paper Series, Washington, DC.

Perez-Truglia, Ricardo (2020). "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment." *American Economic Review* 110(4): 1019–54.

Peters, Max S., Klaus D. Timmerhaus, and Ronald E. West (1991). "Cost Estimation." In *Plant Design and Economics for Chemical Engineers*, 150–215. New York, NY: McGraw-Hill.

Pizer, William A., and Steven Sexton (2019). "The Distributional Impacts of Energy Taxes." *Review of Environmental Economics and Policy* 13(1): 104–123.

Platteau, Jean-Philippe, and Erika Seki (2001). "Community Arrangements to Overcome Market Failures: Pooling Groups in Japanese Fisheries." In Masahiko Aoki and Yujiro Hayami (Eds.), *Communities and Markets in Economic Development*, 343–403. Oxford, UK: Oxford University Press.

Powell, Robert (2017). "Q&A: What Options Do I Have on an Underwater Mortgage?" *USA Today*, April 20. https://www.usatoday.com/story/money/personalfinance/columnist/2017/04/20/q-what-options-do-have-underwater-mortgage/100631398/.

Prendergast, Canice (1999). "The Provision of Incentives in Firms." *Journal of Economic Literature* 37(1): 7–63.

Raff, Daniel M.G. (1988). "Wage Determination Theory and the Five-Dollar Day at Ford." *Journal of Economic History* 48(2): 387–399.

Ransom, Roger L., and Richard Sutch (2001). *One Kind of Freedom: The Economic Consequences of Emancipation*. Cambridge, UK: Cambridge University Press.

Rawls, John (2009 [1971]). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Richards, John F (1995). *The Mughal Empire*. Cambridge, MA: Cambridge University Press.

Robbins, Lionel (1932). *An Essay on the Nature and Significance of Economic Science*. London, UK: Macmillan and Company.

Robinson, Joan (1962). *Economic Philosophy*. Rutgers, NJ: Transaction Publishers.

Roemer, John E. (1988). *Free to Lose: An Introduction to Marxist Economic Philosophy*. London, UK: Radius.

Roine, Jesper, and Daniel Waldenström (2015). "Long-Run Trends in the Distribution of Income and Wealth." *Handbook of Income Distribution* 2(1): 469–592.

Rose, George A., and Sherrylynn Rowe (2015). "Northern Cod Comeback." *Canadian Journal of Fisheries and Aquatic Sciences* 72(12): 1789–1798.

Rosenzweig, Mark R., and Kenneth I. Wolpin (1993). "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India." *Journal of Political Economy* 101(2): 223–244.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir (1991). "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." *American Economic Review* 81(5): 1068–1095.

Ruffini, Krista (2020). "Do Higher Wages Save Consumer Lives? Evidence from Nursing Homes and Minimum Wage Reforms." In *9th Annual Conference of the American Society of Health Economists*. ASHECON: 4.

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld (2010). "Conditional Cooperation and

Costly Monitoring Explain Success in Forest Commons Management." *Science* 330(6006): 961–965.

S & P Dow Jones Indices LLC (2020). "S & P/Case-Shiller NY-New York Home Price Index [NYXRNSA]."

Sally, David (1995). "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992." *Rationality and Society* 7(1): 58–92.

Salomons, Anna, and David H. Autor (2017). "Robocalypse Now: Does Productivity Growth Threaten Employment?" In *Economics of Artificial Intelligence*. Cambridge, MA: National Bureau of Economic Research conference proceedings.

Samuelson, Paul A. (1948). *Economics, an Introductory Analysis*. New York, NY: McGraw Hill.

Samuelson, Paul A. (1957). "Wages and Interest: A Modern Dissection of Marxian Economic Models." *American Economic Review* 47(6): 884–912.

Satterthwaite, Mark Allen (1975). "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10(2): 187–217.

Schelling, Thomas C. (1980). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Schmelz, Katrin (2021). "Enforcement May Crowd Out Voluntary Support for COVID-19 Policies, Especially Where Trust in Government Is Weak and in a Liberal Society." *Proceedings of the National Academy of Sciences* 118(1).

Schonfeld, Erick (2011). "Nobody Predicted The iPad's Growth. Nobody.." *TechCrunch*, January 19. https://techcrunch.com/2011/01/19/nobody-predicted-ipad-growth/.

Schumpeter, Joseph A. (1954). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Cambridge, MA: Harvard University Press.

Schumpeter, Joseph A. (1942). *Capitalism, Socialism and Democracy*. New York, NY: George Allen and Unwin.

Schumpeter, Joseph A. (1954). *History of Economic Analysis*. London, UK: Psychology Press.

Sen, Amartya (1985). "The Moral Standing of the Market." *Social Philosophy and Policy* 2(2): 1–19.

Sethi, Rajiv, and Rohini Somanathan (2009). "Racial Inequality and Segregation Measures: Some Evidence from the 2000 Census." *The Review of Black Political Economy* 36(2): 79–91.

Shapiro, Carl, and Joseph E. Stiglitz (1984). "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74(3): 433–444.

Shearer, Bruce (2004). "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *The Review of Economic Studies* 71(2): 513–534.

Sherman, Senator John (1890). "Act of July 2, 1890 (Sherman Anti-Trust Act)." In *Enrolled Acts and Resolutions of Congress*, 1789-1992, Volume Record group 11. Washington, DC: General Records of the United States Government; National archives.

Schultz, T. W. (1980). Nobel Lecture: The Economics of Being Poor. *Journal of Political Economy* 88(4): 639.

Silver-Greenberg, Jessica (2014). "New York Prosecutors Charge Payday Loan Firms With Usury." *The New York Times*, August 11. https://dealbook.nytimes.com/2014/08/11/new-york-prosecutors-charge-payday-lenders-with-usury/.

Simon, Herbert A. (1951). "A Formal Theory of the Employment Relationship." *Econometrica* 19(3): 293–305.

Simon, Herbert A. (1991). "Organizations and Markets." *Journal of Economic Perspectives* 5(2): 25–44.

Slonim, Robert, and Alvin E. Roth (1998). "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic." *Econometrica* 66(3): 569–596.

Smale, Stephen (1976). "Exchange Processes with Price Adjustment." *Journal of Mathematical Economics* 3(3): 211–226.

Smed, S. (2012). "Financial Penalties on Foods: The Fat Tax in Denmark." *Nutrition Bulletin* 37(2): 142–147.

Smith, Adam (1976 [1759]). *The Theory of Moral Sentiments*. The Glasgow Edition of the Works and Correspondence of Adam Smith. Oxford, UK: Oxford University Press (Liberty Fund).

Smith, Adam (1982 [1776]). *An Inquiry into the Nature and Causes of the Wealth of Nations*. The Glasgow Edition of the Works of Adam Smith. Oxford, UK: Oxford University Press (Liberty Fund).

Smith, Fred (2013). "Ronald Coase Was the Greatest of the Many Great University of Chicago Economists." *Forbes*, Forbes Magazine, September 3. https://www.forbes.com/sites/fredsmith/2013/09/03/ronald-coase-was-the-greatest-of-the-many-great-university-of-chicago-economists/#3936dfda384f.

Smith, Vernon L., and Arlington W. Williams (1992). "Experimental Market Economics." *Scientific American* 267(6): 116–121.

S&P Dow Jones Indices LLC (2020). "S&P/Case-Shiller U.S. National Home Price Index [CSUSHPINSA]."

Spielmaker, D. M. (2018). "Growing a Nation Historical Timeline." Accessed July 28, 2020. https://www.agclassroom.org/ganarchive/timeline/index.htm.

Starr, Evan, James J Prescott, and Norman Bishara (2019). "Noncompetes in the US Labor Force." *Journal of Law and Economics*

Stewart, James B. (2012). "How Broccoli Landed on Supreme Court Menu." *The New York Times*, June 13. https://www.nytimes.com/2012/06/14/business/how-broccoli-became-a-symbol-in-the-health-care-debate.html.

Stigler, George J., and Gary S. Becker (1977). "De Gustibus Non Est Disputandum." *American Economic review* 67(2): 76–90.

Stiglitz, Joseph E. (1996). *Whither Socialism?* Cambridge, MA: MIT press.

Sugden, Robert (1986). *The Economics of Rights, Co-operation and Welfare*. Oxford, UK: Basil Blackwell.

Taylor, Michael (1987). *The Possibility of Cooperation*. Cambridge, UK: Cambridge University Press.

Teitelbaum, D., and R. Axtell (2005). *Firm Size Dynamics of Industries: Stochastic Growth Processes, Large Fluctuations, and the Population of Firms as a Complex System*. Small business research summary. Washington, DC: SBA Office of Advocacy.

The Editors of Encyclopaedia Britannica (2019). "Léon Walras." *Encyclopaedia Britannica*. Accessed August 20, 2020. https://www.britannica.com/biography/Leon-Walras.

Thunberg, Greta (2019). "Speech for the Climate Action Summit." Climate activist Greta Thunberg, 16, addressed the U.N.'s Climate Action Summit in New York City, https://www.npr.org/2019/09/23/763452863/transcript-greta-thunbergs-speech-at-the-u-n-climate-action-summit.

Tierney, John (2000). "A Tale of Two Fisheries." *New York Times Sunday Magazine* (August 27): 38–43.

Tirole, Jean (2017). *Economics for the Common Good*. Princeton, NJ: Princeton University Press.

Tooze, Adam (2018). *Crashed: How a Decade of Financial Crises Changed the World*. London, UK: Penguin.

Trotsky, Leon (1933). *Soviet Economy in Danger: The Expulsion of Zinoviev*. Pioneer Publishers.

US Bureau of Labor Statistics (2016). "Business Employment Dynamics: Entrepreneurship and the US Economy." US Bureau of Labor Statistics Accessed May 28, 2020. https://www.bls.gov/bdm/entrepreneurship/entrepreneurship.htm.

US Bureau of Labor Statistics (2020a). "Charts Related to the Latest "The Employment Situation" News Release." US Bureau of Labor Statistics Accessed April 1, 2021, https://www.bls.gov/charts/employment-situation/unemployed-27-weeks-or-longer-as-a-percent-of-total-unemployed.htm.

US Bureau of Labor Statistics (2020b, November). "Economic News Release: "Unemployed Persons by Duration of Unemployment"." US Bureau of Labor Statistics. US Bureau of Labor Statistics (2021). "Unemployment Rate [UNRATE]." Accessed March 9, 2021. https://fred.stlouisfed.org/series/UNRATE.

US Census Bureau (2020). "Historical Income Tables: Income Inequality. Table A-4. Selected Measures of Household Income Dispersion." Accessed March 15, 2021. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-inequality.html.

Valentino-DeVries, Jennifer, Jeremy Singer-Vine, and Ashkan Soltani (2012). "Websites Vary Prices, Deals Based on Users' Information." *The Wall Street Journal*, December 24. https://www.wsj.com/articles/SB10001424127887323777204578189391813881534.

Veblen, Thorstein (1934 [1899]). *The Theory of the Leisure Class: An Economic Study of Institutions*. New York, NY: The Modern Library.

Vega, Sara (1999). "Short Term Lending Final Report." *Illinois Department of Financial Institutions*

Vicini, James (2012). "Broccoli, Cellphones and the Obama Healthcare Law." *Reuters*, March 27. https://www.reuters.com/article/us-usa-healthcare-court-broccoli/broccoli-cellphones-and-the-obama-healthcare-law-idUSBRE82Q1B420120327.

Visser, Jelle (2015). "ICTWSS: Database on Institutional Characteristics of Trade Unions, Wage Setting, State Intervention, and Social Pacts in 34 Countries between 1960 and 2014.".

Voltaire ([1734] 1961). "Sur Les Presbyteriens." In Jacques van den Heuvel (Ed.), *Mélanges*, 16–18. Paris, France: Gallimard.

Von Neumann, John, and Oskar Morgenstern (1944). *Game Theory and Economic Behavior*. Princeton, NJ: Princeton University Press.

Weber, Max (2004). "Politics as a Vocation." In David Owen and Tracy B. Strong (Eds.), *The Vocation Lectures*. Indianapolis, IN: Hackett Books.

Weisskopf, Thomas E., Samuel Bowles, and David M. Gordon (1983). "Hearts and Minds: A Social Model of US Productivity Growth." *Brookings Papers on Economic Activity* 1983(2): 381–450.

Winkelmann, Liliana, and Rainer Winkelmann (1998). "Why Are the Unemployed So Unhappy? Evidence from Panel Data." *Economica* 65(257): 1–15.

World Bank Development Indicators (2020). Databank.worldbank.org.

World Health Organization (2018). "Household Air Pollution and Health." World Health Organization Factsheet.

Yakovlev, Evgeny (2018). "Demand for Alcohol Consumption in Russia and Its Implication for Mortality." *American Economic Journal: Applied Economics* 10(1): 106–49.

YouGov (2018). "Daily Survey: Socialism."

Young, H. Peyton, and Mary A. Burke (2001). "Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture." *American Economic Review* 91(3): 559–573.

Zelman, Kathleen M. (2008). "How Many Calories Should You Eat? Based on Gender, Age, and Activity Level." *WebMD*. Accessed 31 March 2021, https://www.webmd.com/diet/features/estimated-calorie-requirement.

Ziegelmeyer, Anthony, Katrin Schmelz, and Matteo Ploner (2012). "Hidden Costs of Control: Four Repetitions and an Extension." *Experimental Economics* 15(2): 323–340.

INDEX

Note: The bold-faced page numbers indicate where the term is defined.

A

B

C

D

E

F

G

H

M

N

O

P

Q

R

S

T

U

V

W

Z