Institutional Persistence and Change:
An Evolutionary Approach

Sung-Ha Hwang[*]

Suresh Naidu[‡] Samuel Bowles[†]

February 28, 2013

Abstract

We propose an evolutionary theory of institutional persistence and change. Many institutional changes are decentralized, with a large number of private actors informally adopting new practices that are later confirmed by changes in formal governance structures. For example, land tenure norms, conventional crop shares, inheritance practices, and traditional property rights all are informal institutions, or conventions, that persist for long periods of time and sometimes experience rapid changes due to social conflict, despite the absence of government interventions. We model these types of institutional transitions by extending evolutionary game theory to incorporate social conflict and endogenous social classes in a "bottom-up" evolutionary contract game. The driving mechanism in our model comes from small probabilities of agents engaging in collective action, leading to some contracts being selected over others in the long-run. We show that non-risk dominant contractual conventions will be highly persistent under a stochastic evolutionary dynamic when the size of the poorer class is relatively large, as it is more difficult for them to generate enough collective action to upset the unequal equilibrium. We then show that this same result obtains when we endogenize class sizes using an intergenerational mobility dynamic, with the result that societies with more barriers to mobility will also have more inequality between rich and poor. Finally, we extend the model to allow the rate of collective action to increase in the inequality of a contract, and introduce a government motivated to support the long-term interest of one of the groups, identifying the conditions under which redistribution implemented by non-democratic states.

JEL codes: D02 (Institutions), D3 (Distribution), C73 (Stochastic and Dynamic Games; Evolutionary Games)

1   Introduction

Economic institutions such as labor relations and land tenure systems often persist over centuries, while transitions among these institutions may occur abruptly. A large recent literature has shown that institutional differences persist, causing long term economic and social effects (for surveys see Nunn (2009) and Acemoglu et al. (2005))). Banerjee and Iyer (2005), for example, show that the informal institutions surrounding the *Zamindari* land tenure system in India persisted long after the formal institution was abolished. Other research has found that informal institutions such as contracted crop shares or unequal gender norms exhibit substantial long-run hysteresis, even in the face of large changes in technology and agricultural fundamentals (Bardhan, 1984; Young and Burke, 2001b; Alesina et al., 2011). This paper presents an evolutionary model of social conflict and equilibrium selection that captures many features of this kind of informal institutional persistence.

Our model and main results are intuitive. We consider two classes or groups playing a 2x2 contract game with 2 pure-strategy Nash equilibria. Empirically, it is known that people sporadically engage in acts that are individually irrational but would be in the interest of the group were they to be widely adopted. We reformulate the stochastic "mutations" in evolutionary game theory to be small, i.i.d, probabilities of agents refusing the status-quo and acting suboptimally in the group interest, and examine stochastic stability of Nash equilibria as this probability of individual collective action goes to 0. If a population is large, however, it will take much more of this idiosyncratic behavior to generate enough collective action to change a status-quo institution.[1] Thus large classes will be disadvantaged in the long-run. When class composition is endogenized, we find the plausible extension that high barriers to class mobility result in a larger class of poor relative to rich, and so inefficient and unequal institutional equilibria are stochastically stable. This generates an empirical relationship between intergenerational mobility, efficiency, and cross-sectional inequality that is driven by the relative difficulty of sufficient collective action in a disadvantaged large population. Further extensions yield familiar results in this new model, where even non-democratic governments may redistribute in order to deter collective action.

Evolutionary game theory has been used extensively for modeling and selecting among

---

[1]As will be shown, this is not due to $\frac{1}{N}$ incentive problems, which depend on functional form assumptions, but instead overcoming the minimum fraction of non-best-response play required to induce the other side of the game to change their best-response.

equilibria that emerge from decentralized interactions in a large population of low-rationality agents (Foster and Young, 1990a; Kandori et al., 1993a). It is particularly well suited to model the evolution of norms, culture, and contracts; institutions that are not maintained by the state (Bowles, 2004a). Often *de facto* changes in practices are only confirmed later by changes in formal governance and structures (North, 1981). In many countries and periods, this decentralized and informal aspect of institutional transition is evident in, for example, labor contracts, changes in conventional crop shares, shifts in inheritance practices, and changes in economic relationships between men and women. Evolutionary models are well-suited to model these types of institutions. For example, Young (1998b) shows that conventional contracts that are selected in an evolutionary dynamic optimize a Rawlsian criterion, in that they maximize the payoff of the worst-off. However, this literature lacks three features that characterize many long-run historical processes: 1) intentional play to benefit group interest, rather than random deviations from best-response strategies, 2) endogenous population composition, and 3) the presence of large institutions, such as the state, that internalize and react to these long-run dynamics.

Another approach, based in political economy, models institutional persistence and change as the outcome of bargaining between representative agents of a small number of economic groups of fixed size. Acemoglu and Robinson (2008), for example, consider a model in which formal political institutions change in response to shocks to political power while the economic institutions may or may not persist. In contrast to this and other political economy models, in our approach, institutions are not directly chosen, but rather emerge as the largely unintended consequence of individual actions of large numbers of imperfectly rational agents, none of whom is powerful enough to choose an institution for the entire society (David, 1985; Greif, 1994; Young, 1998b).

We share with the political economy literature a perspective that emphasizes the intentional pursuit of group objectives and social conflict as key ingredients in a theory of institutional persistence and change. This differs from the evolutionary approach to institutional innovation and change in which the observed institutions are the outcome of a process of path-dependent adaptation with random experimentation. However, we also believe that the strong aggregation and rationality assumptions made by the political economy perspective are not well-suited for modelling the diffusion and evolution of informal institutions.

For concreteness, we follow the evolutionary bargaining literature(Binmore et al., 2003)

and study the emergence and persistence of contracts that govern the size of the joint surplus and its distribution between two classes, and we identify conditions under which efficient and/or egalitarian contracts are likely to emerge and to persist. We represent these institutions as conventions between such discrete classes of economic actors as employers and workers, or landlords and sharecroppers. Conventions that are less likely to be upset by the idiosyncratic collective action of the side that has more to gain from a transition will be stochastically stable in our model, and because the collective action problem is harder to solve in larger groups, asymmetric population sizes will have a large effect on equilibrium selection.

Our model extends stochastic evolutionary game theory in a number of ways: first by restricting "mutations" to be strategies that would improve the payoff of the agent were they to become an equilibrium, building on our work in Naidu et al. (2010). These are individual shocks to engage in actions that are in the collective interest of ones' group, should sufficiently many others do the same, and can be interpreted as shocks to social preferences, class-consciousness, beliefs, or identity, as long as they induce individuals to deviate from the best-response strategy in the pursuit of group interests. We have in mind rejections of the terms of the status quo contract by either side, such as lockouts, legal prosecutions, land evictions, strikes, slave revolts, and urban food riots. This brings a social conflict interpretation to what are normally understood as "mutations" in the evolutionary games literature. Second, after establishing that relative class size matters for equilibrium selection, we endogenize the composition of each class by explicitly modelling intergenerational mobility, which allows for mutual determination of both steady-state population sizes and stable institutions. An insight from evolutionary game theory is that the structure of interactions (for example network structure) and population sizes have direct impacts on stabilizing various equilibria, and our approach uses an intergenerational mobility dynamic to endogenize this. Third, we allow the rate of idiosyncratic collective action to vary with the degree of inequality, and examine the extent of redistribution that a non-democratic far-sighted government would optimally choose.

Our modifications allows us to introduce political economy considerations into evolutionary models of institutions, as we can combine the social conflict over institutions emphasized by the former with the explicit dynamics and stability of equilibria emphasized by the latter. Like both the political economy and evolutionary approaches, our model identifies

conditions under which inefficient economic institutions persist in the long run. But in contrast to the approach pioneered by North and Weingast (1989) and Acemoglu and Robinson (2006, 2008), commitment problems play no role in explaining inefficient institutions in our approach. Rather, an inefficient institution may persist due to coordination failures in an evolutionary game with noise, as in a large literature started by Foster and Young (1990a).[2] But in contrast to this literature, the idiosyncratic shocks to agent behavior are not random mistakes, but instead shocks to the willingness to engage in collective action, instead of best-responding to the status-quo equilibrium (as in Naidu et al. (2010) and Bowles (2004a)). This yields two contrasts with the evolutionary game theory literature: 1) an inefficient institution may endure even when it implements highly unequal outcomes, and 2) the equilibrium favoring the smaller class is more likely to be stochastically stable.

By integrating evolutionary and political economy models of institutions, our approach integrates a number of insights. First, following Young (1998b), we interpret risk dominance in a contract game as a measure of the relative equality and efficiency of an equilibrium contract. Second, we show how class size together with individual collective action shocks changes which Nash equilibrium is stochastically stable, a result that we showed in a different context in Naidu et al. (2010). Third, we integrate best-response evolutionary dynamics with intergenerational mobility to endogenize class size, and we find that whether the risk-dominant equilibrium is stochastically stable or not depends on the degree of class mobility. Our model produces a very general "Great Gatsby Curve", where intergenerational mobility is inversely related to cross-sectional inequality. Finally, we provide novel microfoundations for aggregate shocks to *de facto* political power, arising from idiosyncratic shocks in decentralized interactions, and give conditions under which a far-sighted non-democratic government would still choose to redistribute income.

2 Decentralized Transitions: Europe and South Africa

Empirically, the distinction between the political economy and evolutionary approaches is evident in two very different cases of the demise of European feudalism. Consistent with the political economy approach, the emancipation of Russia's serfs by Tsar Alexander II in 1863 was a deliberate choice to implement a new set of institutions resulting from bargaining

---

[2]See also Young (1993c); Kandori et al. (1993a); Axtell et al. (2001); Ellison (1993).

within Russia's elite (Blum, 1971). In contrast, the demise of English serfdom was not the result of explicit bargaining among social groups. The historian E.B. Fryde (Fryde, 1996, pp. 6) writes:

> throughout the 1380s and long beyond them....the servile velleins refused with ever increasing persistence to accept the implications of serfdom, .. In this atmosphere of frequent local disorder and of continuous tension between lords and tenants, the direct exploitation of domanial estates would largely disappear from England in the fifty years after the [1381] Great Revolt.

Similarly, in France protracted agrarian conflict culminated in the 1789 peasant rebellions and forced local lords to abandon many of their feudal privileges well before any legislation was passed. "Peasant uprisings kept rural France on the legislative agenda and drowned out the tendencies to silence on seigneurial rights that characterized much of the nobility" (Markoff, 1996, pp. 509). The abolition of seigniorial dues by the Estates General in 1789 arguably confirmed the new order, it did not introduce it. Instead, a series of uncoordinated actions by dispersed peasants, each taking the grievances of the entire group as their own, induced the aristocratic class to change the terms of agricultural labor.

South Africa's transition to democracy provides a direct contrast between the evolutionary and political economy approaches. Acemoglu and Robinson (2006, pp. 13) write that "the basic structure of apartheid was unaltered" until "De Klerk concluded that the best hope for his people was to negotiate a settlement from a position of strength". For Acemoglu and Robinson, South Africa's new institutions were introduced as the result of the formal constitutional negotiations beginning in 1990. Consistent with their view that economic institutions will change only after the political institutions change owing to commitment failures, they conclude that the change in economic institutions resulted from the introduction of a new political system. However, our reading of the historical evidence is that fundamental changes in economic practices and hence *de facto* economic institutions predate De Klerk's rise to prominence in the National Party, and are more plausibly seen as the cause of the subsequent political transition, rather than its consequence.

The labor market aspects of South African apartheid were a convention regulating the patterns of racial inequality that had existed throughout most of South Africa's recorded history and had been formalized in the early 20th century and strengthened in the aftermath

of World War II. For white business owners, the convention might be expressed: Offer only low wages for menial work to blacks. For black workers the convention was: Offer one's labor at low wages, do not demand access to skilled employment. These actions represented mutual best responses: As long as (almost) all white employers adhered to their side of the convention, the black workers' best response was to adhere to their aspect of the convention, and conversely.

The power of apartheid labor market conventions is suggested by the fact that real wages of black gold miners did not rise between 1910 and 1970, despite periodic labor shortages on the mines and a many-fold increase in productivity(Wilson, 1972). But a series of strikes beginning in the early 1970s and burgeoning after the mid-1980s with the organization of the Congress of South African Trade Unions (COSATU) signaled a rejection of apartheid by increasing numbers of black workers. The refusal of Soweto students to attend classes taught in Afrikaans and the ensuing 1976 uprising returned civil disobedience to levels not experienced since the anti-pass law demonstrations a decade and a half earlier, including the one at Sharpeville at which 69 protesters had been killed by police. The acceleration of urban protests loosely coordinated by the United Democratic Front (UDF), contributed to what whites came to call the "ungovernability" of the country and its businesses. Figure 1 depicts these trends.

Many business leaders concluded that adherence to the apartheid convention was no longer a best response, leading them independently to alter their labor relations, raising real wages and promoting black workers. An executive of the Anglo American Corporation, South Africa's largest, commented: "...in the business community we were extremely concerned about the long-run ability to do business..." (Wood, 2003, pp. 171). Starting in the mid 1980s, the Corporation developed new policies for 'managing political uncertainty' and to address worker grievances, even granting workers a half day off to celebrate the Soweto uprising. In September 1985, Anglo American's Gavin Relly led several business leaders on a clandestine "trek" to Lusaka to seek common ground with African National Congress leaders in exile. In 1986 the Federated Chamber of Industries issued a business charter with this explanation: "the business community has accepted that far reaching political reforms have to [be] introduced to normalize the environment in which they do business." An official of the Chamber of Mines described the situation in 1987 (Wood, 2003, pp. 169)

The political situation in the country was really dismal and we knew that we

were going to have one mother of a wage negotiation. And that the issue wasn't what level of increases we negotiated; the issue was do we survive or not? Will there, after this negotiation, still be such a thing as managerial prerogative. Who controls the mines, really? That was what it would boil down to.

In addition to conceding many of their black employees' workplace demands, business-led pressure for political reforms mounted, joined by reform advocates from the government's intelligence services, churches and others. Late in 1989, four years after the state of emergency had been declared in response to the strike wave and urban unrest, F. W. de Klerk replaced the intransigent P. W. Botha as State President. In 1990 he lifted the ban on the African National Congress, the South African Communist Party and other anti-apartheid organizations, and released Nelson Mandela from prison. Mandela was elected president in South Africa's first democratic election in 1994.

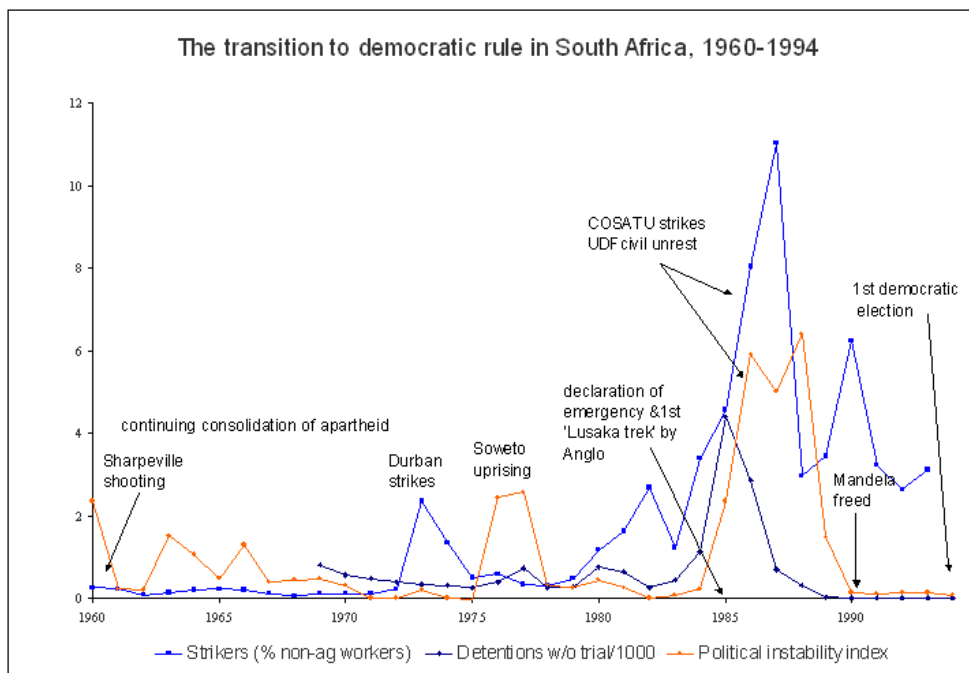Figure 1: Political and economic disturbances in South Africa, 1960-1994 (Sources: Strikers: Statistics South Africa; Detentions: Institute of Race Relations, Yearbooks; Political Instability: Fedderke, De Kadt, and Luiz 2001)

Note the following about this process. First, the concession of best-responding businesses to the collective action of black workers occurred well before and constituted one of the

causes of the political transition. The redistribution of economic resources thus predated and contributed to the redistribution of political resources. Second, the process of transition was extremely abrupt, bringing to an end in less than a decade de facto class and race relations that had endured for a century. Third, while trade unions, 'civics' (community organizations), and other groups were involved in the rent strikes, student stay aways, and strikes against employers, the rejection of apartheid was highly decentralized and only loosely coordinated prior to the 1990 unbanning of the ANC, whose leadership had spent the previous decade either abroad or in prison. We now model an abstract transition process with these general features.

3 Institutional Equilibrium Selection

3.1 Contracts

Contracts differ both in the kinds of incentives that they provide and the distribution of the joint surplus that they implement. To illustrate the kinds of contracts among which decentralized selection may take place, suppose the Bs (who will be column players in our game matrix) are landowners or employers while As (who will be row players) are tenants or workers. Contract E is an equal but inefficient sharecropping or profit-sharing contract and contract U is an unequal but efficient fixed rental or wage contract. We suppose the unequal contract U gives joint surplus $\rho > 2$, of which a share θ goes to the As and the remainder $(1-\theta)$ goes to the Bs, where $0 \leq \theta \leq 1/2$. The equal contract E gives joint payoff 2 divided equally between the A and B players. We furthermore assume that $\theta\rho < 1$. Then from $\rho > 2$, $1 < (1-\theta)\rho$ holds, which means that Bs prefer the unequal contract U while As prefer the equal contract E. We can represent the payoffs from a contract as a 2×2 game matrix, as in Table 1. In the Appendix, we provide a model of optimal contracting that microfounds this reduced-form representation of the payoffs from alternative contract arrangements. Observe that the unique mixed strategy Nash equilibrium of the game defined above is defined as (p^*, q^*) with $p^* := 1/(1 + (1-\theta)\rho)$ and $q^* := 1/(1 + \theta\rho)$.

	U	E
U	$\theta\rho, (1-\theta)\rho$	0,0
E	0,0	1,1

Table 1: Payoffs in the Contract Game. A gets the row payoffs, B gets the column payoffs. Note that because $\theta\rho < 1 < (1-\theta)\rho$ so Bs strictly prefer the unequal contract (U, U) while As strictly prefer the equal contract (E, E).

3.2 Dynamics

We consider a population of agents of size $N = N_A + N_B$, with $\gamma := N_A/N$ the fraction of population that are of class A. The dynamic governing contractual offers is myopic best-response with inertia. Each period, all players are matched with every member of the other class to play the contract game in Table 1. Thus their payoffs depend on the distribution of strategies in the other population. Each period, agents play last period's strategy, U or E, with probability $1 - \nu$ or revise their strategy with probability ν. If they revise their strategy and do not experience a collective action shock, they play the best-response to last-period's distribution of strategies. Each agent in the pair proposes one of two contracts (termed U or E) governing the distribution of the surplus (e.g. union recognition, crop-shares, or land tenure norms). If they fail to coordinate on a contract, both get 0, reflecting the fact that, like the South African employers and workers in Section 2, agents are bargaining over a discrete institution, agreement on which is necessary for the production of a surplus, rather than simply over a divisible surplus.

This setting specifies a stochastic dynamical system, where the states represent the numbers of agents in each population playing U, the unequal strategy. Thus, the state space is given by $\square = \{(x, y) : x = 1, \cdots, N_A, \ y = 1, \cdots, N_B\}$ where each x and y is the number of agents in the A and B population, respectively, who are playing strategy U. We define the best-response rule of a A (respectively B) agent as follows:

$$BR_A^U(y) = \begin{cases} 1 & \text{if} \quad y > N_B q^* \\ 0 & \text{if} \quad y < N_B q^* \end{cases} \quad, \quad BR_B^U(x) = \begin{cases} 1 & \text{if} \quad x > N_A p^* \\ 0 & \text{if} \quad x < N_A p^* \end{cases} \quad, \quad (1)$$

where we suppose that $y \neq N_B q^*$ and $x \neq N_A p^*$ for simplicity. We also set $BR_A^E := 1 - BR_A^U$ and $BR_B^E := 1 - BR_B^U$. Thus, $BR_\alpha^i = 1$ means that an agent in α population chooses strategy

i as a best response. This defines a strategy revision process as follows:

$$X_{t+1} = X_t + \alpha_t^A BR_A^U(Y_t) - \beta_t^A BR_A^E(X_t) \tag{2}$$

$$Y_{t+1} = Y_t + \alpha_t^B BR_B^U(X_t) - \beta_t^B BR_B^E(X_t) \tag{3}$$

where $\alpha_t^A \sim Bin(N_A - X_t, \nu)$, $\alpha_t^B \sim Bin(N_B - Y_t, \nu)$ and $\beta_t^A \sim Bin(X_t, \nu)$, $\beta_t^B \sim Bin(Y_t, \nu)$ are all *i.i.d.* binomial random variables with probability of success being ν. This dynamic is simply a generalization of the best-response dynamic in Kandori et al. (1993a) to asymmetric games. In Appendix, we provide a detailed microscopic dynamic which yields equations (2) and (3) as an aggregate mean dynamic. Intuitively, the equations (2) and (3) can be interpreted as the usual "input and output model" as follows (See Hofbauer and Sigmund (2003)). In the second term of equation (2), the random variable α_t^A describes the event that agents in the A population using strategy E, $(N - X_t)$, revise her strategy with probability v and the term $BR_A^U(Y_t)$ gives the probability that the chosen individuals switch to strategy U (this counts "switch in" numbers). Similarly, in the third term of equation (2) the random variable β_t^A describes the event that agents in the A population using strategy U (X_t) revise her strategy with probability v and the term $BR_A^E(X_t)$ gives the probability that the chosen individuals switch to strategy E (this counts "switch out" numbers). Equation (3) can be interpreted similarly. We call the dynamic defined by (2) and (3) unperturbed process. Under this setting, it is easy to see that the strategy revision process admits two absorbing classes, namely all coordinate to (U, U) or (E, E).

We now add a perturbation to this dynamic. Suppose that when agents revise their strategies, they play a non-best response with probability ϵ if the status-quo contract is not their preferred contract. When A players deviate, they play contract E, while when B players deviate, they choose contract U. This formulation of the perturbations is the key difference between our model and the standard stochastic evolutionary game theory models that have ϵ being the probability of playing a randomly chosen strategy. By contrast, our model has ϵ as the probability of engaging in collective action by playing the strategy that would be best for that sub-population were it to be played by both sub-populations in equilibrium. We describe the stochastic process more fully and apply it to a more general class of bargaining games in Naidu et al. (2010). We think of these as forms of decentralized social conflict, where one actor incurs a cost by playing a strategy which is not a best-response, but would yield

a higher payoff were it to become an equilibrium. Thus we consider our "collective action shocks" as a reduced form way of incorporating activities such as strikes and lockouts, legal prosecutions, land invasions and evictions. The non-best-response play can be considered collective action because it uses the strategy that would yield a higher payoff for the group, were all agents to play it, the shocks are independent across individuals[3], so it is only when a large enough set of agents is simultaneously perturbed that the equilibrium changes.

This perturbation modify the underlying dynamic (2) and (3) as follows:

$$X_{t+1} = X_t + \alpha_t^A BR_A^U(Y_t) - \beta_t^A BR_A^E(Y_t) - \omega_A BR_A^U(Y_t) \tag{4}$$

$$Y_{t+1} = Y_t + \alpha_t^B BR_B^U(X_t) - \beta_t^B BR_B^E(Y_t) + \omega_B BR_B^C(Y_t) \tag{5}$$

where $\omega_A \sim Bin(N_1, v\epsilon)$ and $\omega \sim Bin(N_2, v\epsilon)$ and all the other terms are as defined above. In equation (4), the random variable ω_A describe the event that the idiosyncratic players are chosen from the A population with the probability of selecting being ϵ. The chosen idiosyncratic players will play strategy E (a favorable strategy to the A population) when the best response is to play an unfavorable strategy, namely U. Similar interpretation is possible for the last term in equation (5). While similar to the structure of mutations in Kandori et al. (1993a) and Young (1993a), our dynamic differs in that the mutations are unidirectional for each population, so that the A population only idiosyncratically *reduces* its play of the unequal contract while the B population only idiosyncratically *increases* its play of the unequal contract. This is the sense in which our mutations are not mistakes, but instead strategies that would improve the payoff of the entire class were they to become an equilibrium.

This dynamic can define a perturbed best response dynamics which is irreducible (See Naidu et al. (2010)). Thus there exists a unique stationary distribution, μ_ϵ, for the perturbed dynamic Since the original unperturbed process admits two absorbing states, we are interested in which absorbing state has positive mass in the ergodic distribution μ when the probability of idiosyncratic play goes to 0. Such a state, called stochastically stable state, can be found by taking $\epsilon \to 0$ for μ_ϵ; more precisely, the stochastic stable state is defined to be a state in which the limiting distribution of $\lim_{\epsilon \to 0} \mu_\epsilon$ put a positive weight (Foster and Young, 1990b).

---

[3]In reality, collective action shocks are likely correlated across individuals, but we abstract from that here, and discuss it in the conclusion.

3.3 Institutional Equilibrium Selection

Now we determine which Nash equilibrium is stochastically stable. To do this we call the state in which every agent plays (U, U) (or (E, E)) U convention (or E convention). It is known that the stochastic stability can be studied determining the number of idiosyncratic players it takes to upset each convention(Young, 1993a; Kandori et al., 1993a; Young, 1998a). These numbers are called resistances or the costs of transition from one convention to the other in the literature. To compute these numbers, first suppose that the status quo convention is (U, U), the unequal convention that favors the Bs. If sufficiently many idiosyncratically playing As demand contract E rather than the status quo contract U, best responding Bs will switch to offering contract E. By letting p be the fraction of idiosyncratic players in the A population, we see that the B population agents will play E as their best responses if $(1 - p)(1 - \theta)\rho < p$. Thus the minimum number of As deviating from the status quo to induce a switch from contract U to contract E, $r(U, E)$, is given by the first equation in (6). Similarly, the corresponding resistance for a B-induced transition from the U contract to the E contract is given by the second equation in (6).

$$r(U, E) = \left\lceil N_A \frac{(1 - \theta)\rho}{1 + (1 - \theta)\rho} \right\rceil, \quad r(E, U) = \left\lceil N_B \frac{1}{1 + \theta\rho} \right\rceil \tag{6}$$

where $\lceil t \rceil$ denotes the least integer that is greater than or equal to t.

Then the stochastically stable state is the state i in which $r(i, j) > r(j, i)$, a state which requires more non-best-response play to escape. If $r(i, j) > r(j, i)$ holds, expected waiting time before a transition out of i to j will exceed that of the reverse transition, so that the population will spend more than half of the time near the convention given by i (Ellison, 1993, 2000; Beggs, 2005). Specifically the expected time to exit from the convention U, $\mathbb{E}[W_U]$, and the expected time to exit from the convention E, $\mathbb{E}[W_E]$, are respectively given by

$$\mathbb{E}[W_U] \approx \epsilon^{-r(U,E)}, \quad \mathbb{E}[W_E] \approx \epsilon^{-r(E,U)}.$$

These resistances differ from those in the standard perturbed Markov process models in which the resistances that drive transitions are identified by letting the degree of *unintentional* idiosyncratic behavior to zero; so transitions are induced by the idiosyncratic play of that group for which the least number are required to induce the best responders in the other

14

group to switch strategies (Binmore et al., 2003). By contrast our resistances are the least number of *intentional* idiosyncratic plays, or collective action shocks, required to induce a transition by those who would benefit should a transition occur. In the contract game under the standard model, it is always the case that the number of idiosyncratic plays required to induce a transition is least for members of the sub-population that stands to lose from the transition, because inducing best responders in the opposing sub-population to switch to a contract that they prefer requires fewer idiosyncratic players than inducing a switch to a worse contract. This is why in the standard model with random errors transitions are always induced by those who lose as a result. In our model transitions are induced by those who stand to gain, as agents do not 'experiment' with contracts under which they would be worse off. Thus the resistances that drive the two processes (intentional or unintentional) are always different: resistances in the standard perturbed Markov process model are always less than one half, while ours are greater than one half.

4   Persistence

We can now investigate how the level of equality and efficiency of a contract, together with the relative class sizes, affects the persistence of the associated convention. Efficiency is measured by the level of the joint surplus, that is, 2 in the equal contract and ρ in the unequal contract, while the level of equality in the unequal contract is measured by the share of the surplus received by the least well off group, θ. Ignoring the integer consideration and setting $r(U, E) = r(E, U)$ from (4) and (5) gives the characteristics of unequal contracts such that the population would spend approximately half of the time at the unequal and half at the egalitarian contract. We let γ be the fraction of the A class in the population; i.e., $\gamma = N_A/N$ and $\gamma^* = \gamma^*(\theta, \rho)$ be the critical fraction satisfying

$$\gamma^* \frac{(1 - \theta)\rho}{1 + (1 - \theta)\rho} = (1 - \gamma^*) \frac{1}{1 + \theta\rho}. \tag{7}$$

It is simple to check that if $\gamma = 1/2$, the stochastically stable state is risk-dominant. In the 2x2 contract game, this will be the contract that maximizes the product of the payoffs of the two classes, namely $\rho^2(1 - \theta)\theta$ for convention U and 1 for convention E. Thus, if

$\rho^2(1-\theta)\theta > 1$ then $r(U, E) > r(E, U)$, and U will be selected. The reverse inequality implies that E is selected. We can generalize these results to the case where the class sizes differ. Our key result is that unequal and inefficient contracts that are not risk dominant will be selected if the class suffering the inequality is sufficiently large relative to the favored class.

Proposition 4.1. *For the dynamic process we have the following results:*

1. *If $\gamma > \gamma^*$ then U is the stochastically stable state.*

2. *If U is risk-dominant (i.e., $\rho^2(1-\theta)\theta < 1$), then $\gamma^* < 1/2$ and if E is risk-dominant (i.e., $\rho^2(1-\theta)\theta > 1$), then $\gamma^* > 1/2$.*

3. *$d\gamma^*(\theta, \rho)/d\theta < 0$.*

4. *$d\gamma^*(\theta, \rho)/d\rho < 0$.*

Proof. See Appendix.

Proposition 4.1 shows that for a given γ, there exists a locus of inequality and efficiency levels $\gamma^*(\theta, \rho)$ that satisfies equation (7), so that if $\gamma > \gamma^*$ the unequal contract becomes stochastically stable. The intuition is not the incentive-based logic stressed by the literature on collective action inspired by Olson (1965); Esteban and Ray (2001); nor is it related to the fact that excess supply of a factor of production may disadvantage its 'owners' in markets. Rather the advantage of small size arises simply because smaller groups are more likely to experience realizations of idiosyncratic play large enough to induce a transition, as long as the rate of idiosyncratic play is less than the critical fraction of idiosyncratic players required to induce a transition (which we assume throughout, given that the relevant resistances in our model are always greater than one-half). By contrast, the standard evolutionary dynamic will have the opposite prediction in this class of games, where a larger relative class size for the As favors the equal contract (Naidu et al., 2010).

If the unequal contract is risk-dominant, then this can occur even if the A class is smaller than the B class (Proposition 4.1, 2). If contract U is not risk-dominant, hence contract E is risk-dominant, then stochastic stability requires that the numer of As be larger than the B-population (Proposition 4.1, 2). As the total surplus of contract U shrinks, it takes a larger and larger relative population of As to maintain the stochastic stability of contract U.

Similarly, as the inequality of contract U increases, so that the As receive less and less of the surplus, it takes a larger relative population size of the As for the unequal contract to be stochastically stable. Now suppose that the degree of inequality is greatest; i.e., $\theta = 0$. In this case, if $\gamma > (1+\rho)/(1+2\rho)$ the resistance of the transition from the equal to the unequal contract $(r(E, U))$ will be less than the resistance of the transition from the unequal to the equal contract $(r(U, E))$, so the unequal contract will be selected even if contract U offers nothing to the A class. This occurs because in a population all of whom are best responding by playing contract U favored by the Bs if all of the Bs idiosyncratically select their preferred (unequal) contract, the average payoff to the As of persisting with their preferred contract (E) is zero, so they will (weakly) best respond by conceding to the Bs and playing U. In order for the As to induce the Bs to concede to a switch from a contract in which they receive the entire surplus to the equal contract, it is not necessary for all the As to deviate; just a fraction $\rho/(1+\rho)$ of them will be sufficient. But if γ is sufficiently large, this required number of deviating As will exceed the critical number of deviating Bs to induce a shift in the opposite direction, namely $(1 - \gamma)$, so the unequal contract will be selected. Thus the equilibrium selection process favors smaller classes.

5   Intergenerational Mobility

The assumption that class sizes are given may now be relaxed. We augment the dynamics of the strategy updating process, in which agents choose strategy of demanding contracts U or E, with a class mobility process in which the relative size of the A and B class endogenously changes depending on the payoffs of agents in each of the two classes. It is important to see if the result relating asymmetric class sizes to stochastic stability of the unequal contract remains true when class sizes are plausibly endogenized in an explicit dynamic process.

In the class mobility dynamic, we suppose that each generation randomly matches into mating pairs, has offspring, and then dies. We assume that the probability of a child becoming a member of the B class is increasing in parents' joint income. This barrier to class mobility could arise because class membership requires that one undertake a project with a minimum size, for example inheriting capital goods sufficient to employ an economically viable team of workers or the amount needed to acquire the educational credentials and social connections necessary to be an elite member. We suppose members of the less well off

class are credit constrained. Those who inherit less than this amount become members of the A class. In the resulting model, then, the stochastically stable contract and the relative sizes of the two classes will be jointly determined.

To simplify the analysis, we suppose that class mobility occurs only when both A and B classes plays either one of conventions. However, all that is needed for our results is that the strategy updating dynamic is sufficiently fast compared to the class mobility dynamic. Also, to exclude the uninteresting case in which one of the class sizes is zero, we suppose that there is at least one agent in each class throughout in every state. Under the joint process of strategy revision and class mobility, the state is described by a triplet: the numbers of A population agents choosing U (x), the number of B population agents choosing E (y) and the size of A population (n_A). Thus the state space Δ is given by

$$\Delta := \{(x, y, n_A) : 1 \leq n_A \leq n - 1, \ 0 \leq x \leq n_A, \ 0 \leq y \leq N - n_A\}.$$

To be concrete, we suppose the following joint strategy updating and class mobility dynamics:

- Each period, every agent plays last period's strategy with probability $1 - v$, a best response strategy with probability $v(1-\epsilon)$, and an idiosyncratic strategy with probability $v\epsilon$.

- Next if the state corresponds to either of conventions (i.e., all agents in both populations play either (U, U) or (E, E); $(x, y, n_A) = (0, 0, k)$ or $(k, N-k, k)$ for some k), then agents who have not update their strategy are matched into mating with agents from the other population and give birth to two offspring, and then die.

Changes in the sizes of the two classes will occur when an offspring of a B parent has insufficient wealth to retain its parent's upper class status, or when a child of an A has sufficient wealth to become a B. The class mobility in general could depend on four things: the degree of class assortment in parenting, the inheritance rules in force (primogeniture or equal inheritance, for example), the minimal inheritance required for membership in the upper class, and the incomes of the two parents. We assume equal inheritance to the two offspring of each couple and abstract from marital assortment as it will not affect the resulting equilibria in our model. We assume that when parents belong to the same class, the two offspring retain the parents' class membership, the incomes of two Bs always being sufficient

for both offspring to become Bs and the incomes to two As never being sufficient to allow their two offspring to become Bs.

To define the expected income of the cross-class couple, first observe that when the state of the system is (x, y, n_A), x/n_A fraction and $1 - x/n_A$ fraction of the A population play strategy U and E and $y/(N - n_A)$ fraction and $1 - y/(N - n_A)$ fraction of the B population play strategy U and E, respectively. Thus, the expected income of the A agent is given by

$$y_A(x, y, n_A) := \frac{x}{n_A} \frac{y}{N - n_A} \theta \rho + (1 - \frac{x}{n_A})(1 - \frac{y}{N - x_A})1. \tag{8}$$

Similarly, the expected income of the B agent is given by

$$y_B(x, y, n_A) := \frac{x}{n_A} \frac{y}{N - n_A} \theta(1 - \rho) + (1 - \frac{x}{n_A})(1 - \frac{y}{N - x_A})1 \tag{9}$$

Since the class couple will be formed either by A agents' matching with the $(N - n_A)/N$ fraction of the B agents in the population or B agents' matching with the n_A/N fraction of the A agents in the population, the expected income of the cross class couple is

$$y_c(x, y, n_A) := \frac{N - n_A}{N} y_A(x, y, n_A) + \frac{n_A}{N} y_B(x, y, n_A). \tag{10}$$

To capture the relationship between parental wealth and class mobility, we suppose that a measure of the barrier to class mobility y_b is given by

$$y_b(n_A) := (1 - \frac{n_A}{N})\bar{y} \tag{11}$$

for some positive $\bar{y}$, called a baseline income barrier. The idea behind (11) is that it is more difficult for the A population to move up to the B population when the B population is numerous. While this specific assumption is needed to guarantee interior class sizes, the general premise that mobility becomes harder as the size of the rich class increases seems intuitive, and data to support this is presented in Turchin (2003).

Then we suppose that within the class dynamics the matched cross couples give birth to two B children if $y_c(x, y, n_A) > y_b(n_A)$. Thus if we let G be a function describing the net

increase in the number of agents in the A population when the class dynamic occurs, then

$$G(x, y, n_A) := \begin{cases} 1 & \text{if} \quad y_c(x, y, n_A) < y_b(n_A) \\ 0 & \text{if} \quad y_c(x, y, n_A) = y_b(n_A) \\ -1 & \text{if} \quad y_c(x, y, n_A) > y_b(n_A) \end{cases}$$

To describe the changes in the number of agents in the A population during the class dynamic, we recall that a randomly chosen agent from n_A A agents who has not revised the strategy with probability $1 - v$ will be matched with an B agent whose fraction is $1 - n_A/N$. Thus, we infer that there will be $Bin(n_A, (1 - v)(1 - n_A/N))$ cross class couples formed by the A agents and similarly that there will be $Bin(N - n_A, (1 - v)n_A/N)$ cross class couples formed by the B agents. In fact, we obtain the following aggregate mean class dynamic (12) from the microscopic model of class dynamics in the Appendix (See Hwang, Katsoulakis, and Rey-Bellet, 2013 for aggregation):

$$N_{t+1}^A = N_t^A + \chi_t G(X_t, Y_t, N_t^A). \tag{12}$$

where $\chi_t \sim (Bin(N - N_t^A, (1 - v)\frac{N_t^A}{n}) + Bin(N_t^A, (1 - v)(1 - \frac{N_t^A}{n})))$. The class dynamics also modify the strategy revision dynamics of (4) and (5) slightly. This is because that an agent in the population is now decribed by the choice of strategy U or E and the membership of A class or B class and new born children's strategies depend on the process of class dynamics. We present the detailed equations in the Appendix (See equations (24) and (25)).

To study the equilibrium selection problem in the joint dynamic of strategy revision and class defined by (12), (??), and (??), we first need to identify the absorbing states for the unperturbed system. To do this, observe that from the strategy updating dynamics the absorbing states only involve one of conventions; i.e., the absorbing states are of the form: $(x, y, n_A) = (0, 0, k)$ or $(k, N-k, k)$ for some k. Consider the case, $(X_t, Y_t, N_t^A) = (k, N-k, k)$ first. In this case, from (8), (9), and (10) we have

$$y_c(X_t, Y_t, N_t^A) = \frac{N - k}{N}\theta\rho + \frac{k}{N}\theta(1 - \rho).$$

Since $y_c(X_t, Y_t, N_t^A)$ is increasing in k and y_b in (11) is decreasing in k, if $\bar{y}$ is greater than $\theta\rho$, there exists a unique $k_U^* = k_U^*(\bar{y}, \theta, \rho)$ such that $y_c(k_U^*, N - k_U^*, k_U^*) = y_b(k_U^*)$ (See Panel
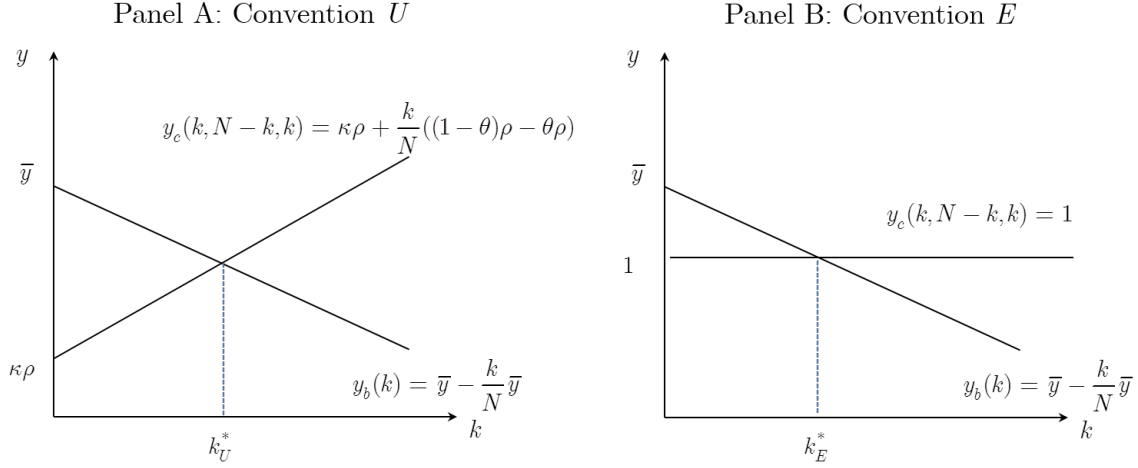
A in Figure 2). Then it follows that

Panel A: Convention U Panel B: Convention E

Figure 2: **Determination of the critical values k_U^* and k_E^*.** The panels show the determination of the critical values, k_U^* and k_E^*, in the class mobility dynamics

$$y_c(k, N - k, k) < y_b(k) \text{ for } k < k_U^* \text{ and } y_c(k, N - k, k) > y_b(k) \text{ for } k > k_U^*. \quad (13)$$

We suppose that k_U^* is an integer for simplicity. Alternatively, we can add the state $(k_U^*, N - k_U^*, k_U^*)$ to the state space Ξ as in Binmore et al. (2003).   Then the equation, $y_c(k_U^*, N - k_U^*, k_U^*) = y_b(k_U^*)$ and inequalities (13) implies that

$$N_{t+1}^A > N_t^A \text{ if } k < k_U^*, \;\; N_{t+1}^A = N_t^A \text{ if } k = k_U^*, \;\; N_{t+1}^A < N_t^A \;\; \text{if } k > k_U^* \quad (14)$$

for the class dynamic equation (12) and this shows that $(k_U^*, N - k_U^*, k_U^*)$ is a unique absorbing state of the form $(k, N - k, k)$ (See Panel A in Figure 2).Similarly, for the case of $(X_t, Y_t, N_t^A) = (0, 0, k)$, it is easy to see that $(0, 0, k_E^*)$ is the unique absorbing state of the form $(0, 0, k)$. We record this observation in the following lemma, where we note that $\bar{y} > 1$ implies $\bar{y} > \theta\rho$.

Lemma 5.1. *Suppose that $\bar{y} > 1$. Then there are two absorbing states, $(x, y, n^A) = (k_U^*, N - k_U^*, k_U^*)$ and $(x, y, n^A) = (0, 0, k_E^*)$, for the unperturbed dynamic of the joint process.*

To compute the resistances for the transitions between these two absorbing states, we need to first determine resistances for the transitions between states, $r(i, j)$ for $i, j \in \Xi$
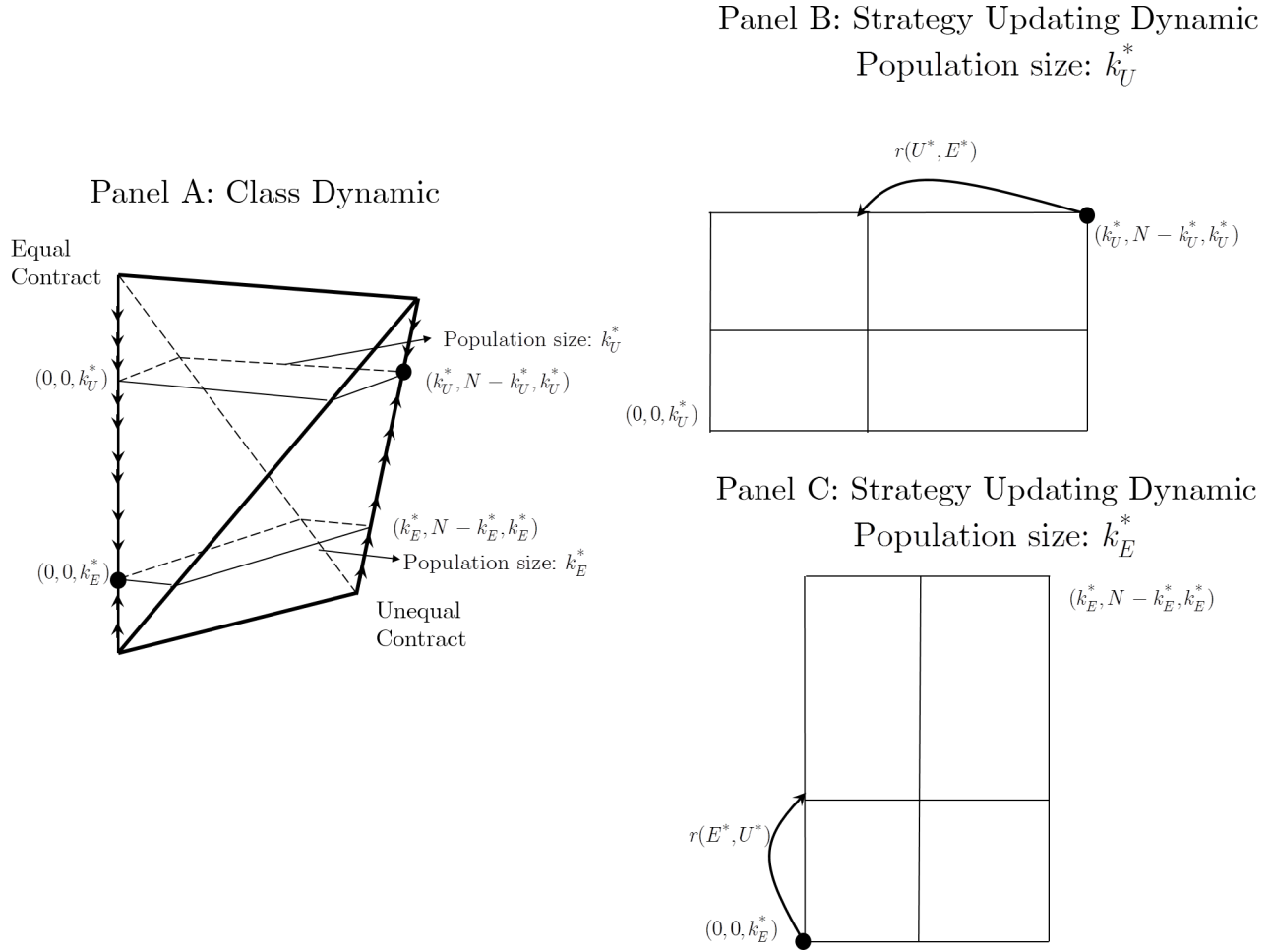
Panel B: Strategy Updating Dynamic
Population size: k_U^*

$r(U^*, E^*)$

$(k_U^*, N - k_U^*, k_U^*)$

Panel A: Class Dynamic

Equal
Contract

Population size: k_U^*

$(0, 0, k_U^*)$

$(k_U^*, N - k_U^*, k_U^*)$

$(0, 0, k_U^*)$

Panel C: Strategy Updating Dynamic
Population size: k_E^*

$(k_E^*, N - k_E^*, k_E^*)$

$(k_E^*, N - k_E^*, k_E^*)$

Population size: k_E^*

$(0, 0, k_E^*)$

Unequal
Contract

$r(E^*, U^*)$

$(0, 0, k_E^*)$

Figure 3: **Class dynamic and strategy updating dynamic**. Panel A shows that possible transitions between two absorbing states in a diagram of the state space (See the Appendix for the picture of the state space). Panels B, C show the strategy updating dynamic when the size of the A population is k_U^* or k_E^*.

(See Young (1998)). We note that the resistances between states for a given size of the A population is the same as ones in the previous section (equation (6)). Also, for the states belonging to $\{(k, N - k, k) : k = 1, \cdots, N - 1\}$, since the unperturbed system always can reach $(k_U^*, N - k_U^*, k_U^*)$ (from (14)) and cannot escape from $(k_U^*, N - k_U^*, k_U^*)$ under the class dynamic, $r(i, (k_U^*, N - k_U^*, k_U^*)) = 0$ and $r((k_U^*, N - k_U^*, k_U^*), i) = \infty$ for all $i \in \{(k, N - k, k) : k = 1, \cdots, N - 1\}$. Similarly we find that $r(i, (0, 0, k_E^*)) = 0$ and $r((0, 0, k_E^*), i) = \infty$ for all $i \in \{(0, 0, k) : k = 1, \cdots, N - 1\}$ (See Panel 1 in Figure 3 and other resistances in the Appendix).

Using these observations, the resistances between two absorbing states, $(k_U^*, N - k_U^*, k_U^*)$

and $(0, 0, k_E^*)$, are computed as follows (See Panels 2, 3 in Figure 3).

Lemma 5.2. *Suppose that* $\bar{y} > 1$. *Let* $U^* = (k_U^*, N - k_U^*, k_U^*)$ *and* $E^* = (0, 0, k_E^*)$. *The resistances between* U^* *and* E^* *are*

$$r(U^*, E^*) = \left\lceil k_U^*(\bar{y}, \theta, \rho) \frac{(1-\theta)\rho}{1 + (1-\theta)\rho} \right\rceil, \quad r(E^*, U^*) = \left\lceil k_E^*(\bar{y}, \theta, \rho) \frac{1}{1 + \theta\rho} \right\rceil, \quad (15)$$

Proof. See Appendix.

Since the class sizes are different in each absorbing state in the joint dynamics, resistances in (15) are the modified versions of (6) by endogenizing the relative class fractions. [The resistances of (15) generalize (6) by showing how the class size depends on income barrier and clarifying another mechanism through which inequality and productivity affect the resistance of absorbing states.]

Thus we are able to explore the effects of exogenous changes in $\bar{y}$, θ, and ρ on the stochastically stable contract, the equilibrium class sizes, and hence on the income inequality between members of the two classes. Intuitively, we would expect that as the barrier to mobility increased (higher $\bar{y}$): a) the A class would be more numerous in equilibrium and that as a result b) the population would spend a larger fraction of the time at the unequal contract. Both consequences of an increase in $\bar{y}$ would be to increase the income difference between the two classes. Proposition 5.3 shows that these intuitions are correct.

Proposition 5.3. *Suppose that* $\bar{y} > 1$. *Then, there exists* $\bar{y}^* = \bar{y}^*(\theta, \rho)$ *such that, for all* $\bar{y} > \bar{y}^*$, *we have* $U^* = (k_U^*, N - k_U^*, k_U^*)$ *as the stochastically stable state.*

Proof. See Appendix.

An implication of Proposition 5.3 is that the risk dominant contract will not be selected if the cost of vertical class mobility is sufficiently high. An increase in ρ lowers k_U^* as it increases the income of the cross-class couple ($dk_U^*/d\rho < 0$), thereby facilitating mobility out of the A class, reducing the equilibrium number of As and thus favoring them. In contrast to the exogenous population size model, however, the effect on equilibrium selection is ambiguous, as an increase in the productivity of the unequal contract ρ (with no change in the equal contract) will also increase the fraction of idiosyncratically playing As necessary to induce

the best responding Bs to abandon the unequal contract, $(1-\theta)\rho/(1+(1-\theta)\rho)$, so the sign of $dr(U^*, E^*)/d\rho$ is ambiguous. However, a proportional increase in the productivity of both contracts, for example, scaling up the payoff matrix in Table 1 by some $\omega > 1$, does not affect the fraction of each class whose idiosyncratic play is sufficient to induce a transition. In this case the only effect of an increase in ρ is, assuming $\bar{y}$ fixed as above, to reduce the equilibrium size of the A population in both contracts, favoring the As and unambiguously increasing the fraction of time spent at the more equal convention.

Our model's evolutionary dynamic thus yields the "Great Gatsby Curve", where low intergenerational mobility is correlated with high cross-sectional inequality. In our model, this occurs because large populations require many more deviant players to induce the other side to change behavior, and unequal contracts make it harder for cross-class couples to send their children into the wealthier class. So high barriers to mobility make the population of the poor larger, making collective action harder, and so unequal contracts are more likely to be stochastically stable. Evidence for this correlation is abundant across a wide variety of historical contexts: from pre-industrial populations (Mulder et al., 2009) to the modern day (Corak, 2012), suggesting that it is general phenomenon unrelated to particular technological or political settings, and therefore worth replicating in our relatively abstract model.

This section shows that the results from the previous section about which contracts are persistent are robust to endogenizing the class sizes. High barriers to mobility create asymmetric class sizes, which increases the total income of the Bs, as there are now many As to interact with. It is also harder for the As to generate enough idiosyncratic deviance to tip the equilibrium to one that is favorable to them, and so a high barrier to mobility will favor an unequal contract. Even if one starts from equal population sizes and at the egalitarian contract, the intergenerational transmission dynamic will eventually produce few elites and many poor, and this will make it easier for the Bs to obtain their preferred convention.

6   Population Size vs Interaction Structure

We have so far focused on exploring the role that asymmetric population size, in an environment of idiosyncratic collective action shocks, has on the stochastic stability of the unequal contract. But another interpretation of our model is that it is the structure of interactions between two equally sized populations changes the stochastically stable equilibrium. It is

well known that interaction structure (modelled as a network) can alter many stochastic stability results; in this section we lay out the relationship between interaction structure and population size in our model. Essentially, instead of each population best-responding to the play of the full population of the other type, we can instead think of each population observing only a sample of the other population. Then, if the sample sizes differ between the two populations, the population with a larger sample will have an advantage, even if the population sizes are identical. Because it is sampling noise that drives institutional transitions, populations whose play is not completely observed have the same advantage as small populations whose play is fully observed.

We have in mind arguments laid out in other social sciences. Early industrial capitalism, for example, agglomerated workers in large establishments facilitating collective action. By contrast, earlier class systems, according to Gellner (1983), were characterized by "laterally separated petty communities of the lay members of society" speaking different dialects or even languages, presided over by a culturally and linguistically homogeneous class. Economic relations in such societies often took the form of patron-client relationships that endured over generations with little mobility of the clients among the patrons (Blau, 1964; Fafchamps, 1992; Platteau, 1995).

The patron client relationship will support a very different dynamic from the relationship of employee to employer in the modern labor market. The reason is that these two institutions affect the information available to agents when they adopt best responses.

Suppose that when adopting a best response the members of the two classes do not know the entire distribution of play in the previous period. As before, players are randomly matched to play with members of the other population, but now members of a given class know the distribution of play in only a subset of the other population. While we could in principle investigate heterogeneous sets of opposing play known by each agent, we simplify dramatically and focus only on the case where each agent knows the distribution of play of a fraction of the opposing class. As know the play of a fraction of Bs given by η_A and Bs know the distribution of play in a fraction η_B of As, where $\eta_A, \eta_B \leq 1$. Pre-capitalist agrarian institutions, in Gellner's view, entailed $\eta_A < \eta_B$, for the upper class communicated readily amongst themselves and therefore had information about the recent play of a large segment of the less well-off class. The geographical, cultural and linguistic isolation of the As, by contrast, militated against information sharing beyond ones local community.

The advantage enjoyed by the Bs is not that a given B-patron may engage the A-clients of other Bs. Rather, by drawing information from a larger sample of As, the B's less noisy signal of the distribution of play reduces the likelihood that their myopic best response will overreact to the chance occurrence of a high level of idiosyncratic play among their particular A-clients. Since the introduction of η_A and η_B modifies the best response rules as follows

$$BR_A^U(y) = \begin{cases} 1 & \text{if } y > \eta_A N q^* \\ 0 & \text{if } y < \eta_A N q^* \end{cases}, \quad BR_B^U(x) = \begin{cases} 1 & \text{if } x > \eta_B N p^* \\ 0 & \text{if } x < \eta_B N p^* \end{cases},$$

the new resistances $r(E,U)$ and $r(U,E)$ are given by

$$r(U,E) = \left\lceil \eta_B N \frac{(1-\theta)\rho}{1 + (1-\theta)\rho} \right\rceil, \quad r(E,U) = \left\lceil \eta_A N \frac{1}{1 + \theta\rho} \right\rceil \tag{16}$$

Instead of population sizes, we now have asymmetric 'scope of vision' parameters in the resistances (η_A, η_B) mean that more idiosyncratic players are required to induce a concession by the best responding members of the population that has more information.

If η_A is small, then it takes only a few idiosyncratic plays by the subset of Bs in a given A's 'scope of vision' to convince the best responding A to concede to the unequal contract. As is evident from (16), a decrease in η_A will reduces $r(E,U)$ and the inequality $r(U,E) > r(E,U)$ is more likely to be satisfied and hence, the unequal contract will persist for a long time.

Proposition 6.1. *Suppose that $\bar{y} > 1$ and $\eta_B = 1$. Then there exists a $\eta_A^* > 0$ such that for all $\eta_A < \eta_A^*$, the unequal contract is stochastically stable.*

If we redefine $\gamma = \eta_A/(\eta_A + \eta_B)$, Proposition 4.1 will obtain in this model. While we do not explore endogenous population structure (only population sizes) dynamics in this paper, future work could endogenize the η_i's along similar lines. This framework could potentially be extended to incorporate results from the literature on stochastic games on graphs to relate more complex network properties to the stochastically stable equilibrium (Blume, 1995b; Ellison, 1993; Hojman and Szeidl, 2006).

Substantively, this interpretation of the model thus suggests another possible reason for the trend in many countries over the past 2 centuries towards a reduction in the relative incomes of the well off (Piketty 2005). The geographic, industrial, and occupational mobility characteristic of modern labor markets (coupled with the spread of literacy and greater ease

of communication) made workers less responsive to the demands of a small number of local employers, as they knew about the offers of employers outside their local area. The effect would be to raise η_A and thus to destabilize highly unequal contracts.[4]

7 Inequality, the Rate of Idiosyncratic Play, and Redistributive Politics

Economic inequality may enhance the frequency of deviant play by the less well off group by providing additional motives and opportunities to challenge the status quo contract (Scott 1976, Moore 1978, Wood 2003). To capture this insight we make the rate of idiosyncratic play state-dependent, and study the response of a far-sighted government that on behalf of the myopic Bs seeks to deter a transition to the egalitarian state.

Bergin and Lipman (1996) show that, if one allows ϵ to vary arbitrarily as a function of the state, (x, y), then one can choose a function that selects any recurrent class of the unperturbed process as the stochastically stable state. But what error functions are empirically plausible? We would like to capture the idea that idiosyncratic play by the poorer As will be greater in highly unequal societies. To study this, we modify our baseline model (i.e., exogenous class size model) in Section 3 as follows. With some abuse of notation, in this section we incorporate a state-dependent probability, $\epsilon(x, y)$, of idiosyncratic play into the model with

$$\epsilon(x, y) = \epsilon^{\phi(\lambda(\pi_B(x,y) - \pi_A(x,y)))}, \tag{17}$$

where we recall that π_A and π_B are expected payoffs for classes A and B from the underlying game, given by

$$\pi_A(x, y) := \frac{x}{N_A} \frac{y}{N_B} \theta \rho + \frac{N_A - x}{N_A} \frac{N_B - y}{N_B}, \ \ \pi_B(x, y) := \frac{x}{N_A} \frac{y}{N_B} (1 - \theta)\rho + \frac{N_A - x}{N_A} \frac{N_B - y}{N_B}.$$

Here ϕ is a decreasing function and $\lambda > 0$ captures the extent to which inequality increases

---

[4]Similarly, male-female interactions may be differentially structured in traditional patriarchies, despite numerical parity. Men who can publically circulate and fraternize with other men have an informational advantage vis-a-vis women confined to domestic roles and family networks. By observing the behavior of many women, male strategies will not be altered in response to idiosyncratic play of a few women, resulting in the persistent of unequal gender norms. The formation of information sharing networks within groups of women, e.g. via employment outside of the household, would then give women a larger share of the intrahousehold pie.

idiosyncratic play. Sociological conditions favoring rejection of unequal contracts as well as religious or other cultural influences that make economic inequality illegitimate will increase λ. In the equal contract, the B class idiosyncratically plays at rate ϵ, since $\pi_B(0,0) = \pi_A(0,0) = 1$, and in the unequal contract, the A class plays at a rate $\epsilon^{\phi(\lambda((1-\theta)\rho-\theta\rho))}$, which is clearly decreasing in θ, so an decrease in income inequality reduces the rate of idiosyncratic play. To simplify the analysis from now on, we adopt $\phi(t) := 1/(1+t)$, however our results do not depend on the particular choice of this functional form. We show that in Proposition 4.1 the sufficiently numerous A population may induce the unequal institution to persist. However, when the idiosyncratic play rate is dependent on the payoffs, this may not be the case.

Proposition 7.1. *There exists some $\overline{\lambda}$ such that for any $\lambda > \overline{\lambda}$ the equal contract is selected.*

    Proof. See Appendix.

    This result is simple, but it allows us to incorporate a plausible relationship between inequality and institutional persistence; namely that inequality increases the level of non-best-response play for the worse-off group (the As), destabilizing the unequal contract. In addition, this feature in our model allows us to endogenize a politically chosen level of redistribution at the unequal contract.

    To study this question more precisely, we introduce a forward-looking government that may seek to stabilize the status quo contract. We focus on government redistribution of income as a device to reduce idiosyncratic play in the unequal state, thus prolonging the contract preferred by the upper class. Here we consider the reduced strategy revision processes $\{s_t\}_t$ whose two states consist of U convention and E convention and transition probabilities are based on the resistances between the two states (expression 18). Suppose that the government implements a tax rate τ in each state to maximize the weighted sum of the present discounted value of the two groups period payoffs subject to the dynamics governed by the stochastic strategy revision process, whose transition probabilities $P(s, B, \tau)$, the probability of transition from s to a state in B, are indexed by the control parameter τ : i.e., $P(s, B, \tau)$ is given by

$$
\begin{array}{c|cc}
 & U & E \\
\hline
U & 1 - \epsilon^{r_{UE}(\tau)} & \epsilon^{r_{UE}(\tau)} \\
E & \epsilon^{r_{EU}(\tau)} & 1 - \epsilon^{r_{EU}(\tau)}
\end{array} \tag{18}
$$

Since the after tax income for the A and B populations are $\theta\rho + \tau(1-\theta)\rho$ and $(1-\tau)(1-\theta)\rho$, respectively, the modified resistances $r_{UE}(\tau)$ and $r_{EU}(\tau)$ in (18) are

$$
\begin{aligned}
r_{UE}(\tau) \quad &: \quad = \left[N_A \frac{(1-\tau)(1-\theta)\rho}{1 + (1-\tau)(1-\theta)\rho} \frac{1}{1 + \lambda((1-\tau)(1-\theta)\rho - \theta\rho - \tau(1-\theta)\rho)} \right] \quad (19) \\
r_{EU}(\tau) \quad &: \quad = \left[N_B \frac{1}{1 + \theta\rho + \tau(1-\theta)\rho} \right].
\end{aligned}
$$

We parameterize the weight put on the payoff of the rich by a parameter $\xi \in [0,1]$ in its per-period objective function. If we follow Acemoglu and Robinson (2006) and model democracy as taxation chosen by the median voter, then, since $\gamma > 1/2$, democracy would be equivalent to $\xi = 0$, and non-democracy would correspond to $\xi = 1$. In this latter case, the state faces a trade-off similar in spirit to that in Acemoglu and Robinson (2006): the state weighs the costs of the tax on the per period income of the Bs against the effect of reduced income inequality on the transition probabilities from U contract to E contract. To study this possibility we suppose that $\xi > 1/2$. Then, the per-period return function to the government $u(s,\tau)$ is given as follows:

$$
u(s,\tau) := \begin{cases} (1-\xi)(\theta\rho + \tau(1-\theta)\rho) + \xi(1-\tau)(1-\theta)\rho & \text{if} \quad s = U \\ 1 & \text{if} \quad s = E \end{cases}
$$

We note that when $\bar{\tau} = ((1-\theta)\rho - \theta\rho)/2(1-\theta)\rho$, the payoffs for the A and B classes in U contract are the same; i.e.,

$$
(1-\xi)(\theta\rho + \bar{\tau}(1-\theta)\rho) + \xi(1-\bar{\tau})(1-\theta)\rho = \frac{\rho}{2}.
$$

Thus we consider $\bar{\tau}$ to be the maximum possible tax rate that the government can impose. Then the government maximizes

$$
\max_{\{\tau_t\}_{t=1}^{\infty}} \mathbb{E}_0\left[\sum_{t=0}^{\infty} \beta^t u(s_t, \tau_t) \right] \quad (20)
$$

such that $\tau_t \in [0, \bar{\tau}]$ for all t and $0 < \beta < 1$ and s_t is the Markov process starting from s.

The corresponding Bellman equation for (20) is given as follows:

$$v(s) = \max_{\tau}\{u(s,\tau) + \beta \int v(s')P(s,ds',\tau)\}.$$

More explicitly for $s = U, E$, we have the following two equations:

$$v(U) = \max_{\tau \in [0,1]}\{u(U,\tau) + \beta(v(U)P(U,U,\tau) + v(E)P(U,E,\tau))\} \quad (21)$$

$$v(E) = \max_{\tau \in [0,1]}\{1 + \beta(v(U)P(E,U,\tau) + v(E)P(E,E,\tau))\} \quad (22)$$

Then by solving equations (21) and (22), we find the optimal tax policy for a given state $s = U, E$. We denote by τ_E and τ_U the optimal tax rates at state E and state U.

We first consider the tax rate in the E contract state, given by (22). Since the after tax income for B's is always greater than the after tax income for A's, the per-period return to the government in U contract is always greater than the per-period return in E contract; i.e., $u(U,\tau) = \rho/2 \geq 1$ for all $\tau \leq \bar{\tau}$. Thus we have $u(s_t, \tau_t) \geq 1$ for all t and this implies that $v(U)$ and $v(E)$ are greater than $1/(1-\beta)$. In particular, it is easy to see that $v(E) \geq 1/(1-\beta)$ implies that $v(U) \geq v(E)$ (see Lemma 1 in the Appendix). This shows that the value of U contract is always greater than E contract; $v(U) \geq v(E)$. Then from (18) and (19), we see that an increase in τ in E contract state reduces the minimum number of B agents inducing transition to U contract, thus makes transition to U contract more likely. Thus from $v(U) > v(E)$ and equation (22), it is easy to see that the government sets the highest possible tax rate, namely $\tau_E = \bar{\tau}$.

Next we consider the optimal tax rate τ_U in the U contract state. In this case, observe that because the government put more weight on the B's payoffs ($\xi > 1/2$), an increase in the tax will reduce the per-period return to the government ($du(U,\tau)/d\tau < 0$). To study the effect on the transition probabilities $P(U,E,\tau)$, we first ignore the effect of inequality (i.e., $\lambda = 0$). Then similarly to the tax effect in the E contract state, an increase in the tax will reduces the minimum number of A's agent inducing transition to E contract (the first term in (19)) and makes transition to E contract more likely. However, when the idiosyncratic rate is state-dependent and increasing in income inequality, an increase in τ in the U contract state may reduce the degree of inequality by redistribution and thus increase the minimum number of A's agent necessary to induce a transition to E, thus delaying the transition

to the E contract. The second effect stabilizing the unequal contract dominates the first destabilizing effect, when the responsiveness of idiosyncratic play to inequality is sufficiently sensitive (λ is sufficiently high). This observation leads to the following proposition.

Proposition 7.2. *There exists $\bar{\lambda}$ such that for $\lambda > \bar{\lambda}$ such that $P(U, E, \tau)$ is nondecreasing in τ.*

Proof. See Appendix.

How is the optimal tax determined in U contract? From the discussion above, when λ is sufficiently high, an increase in the tax rate reduces the per-period return to government, but raises the life-long return by reducing the transition probability from U contract to E contract. This precisely shows the trade-off the government faces between the per-period loss by a higher level of redistribution and the long term gain by delaying the transition from U contract to E contract. As Proposition 7.3 shows, when λ is sufficiently high and A's are sufficiently numerous, the government would set the positive tax rate adopting redistribution policy today in the favor of the future gains.

Proposition 7.3. *Suppose that λ and N_A are sufficiently large. Then $\tau_U > 0$.*

Proof. See Appendix.

If inequality generates more idiosyncratic play by the worse-off party, one would expect far-sighted governments acting on behalf of the long term interests of the B class to subject the well to do to redistributive taxation. This proposition thus lies in the spirit of Acemoglu and Robinson (2006), where far-sighted elites choose partial reform in order to prevent (or delay) the transition to an equilibrium that is even more egalitarian. As in Acemoglu and Robinson (2006) the paradigmatic historical case on non-democratic redistribution of this type is Bismark's expansion of the welfare state (without the franchise) in response to the German workers movement in the late 19th century. The novelty is that we obtain this result in an evolutionary model, with individual shocks to collective action, rather than aggregate shocks to the bargaining power of representative agents.

8 Conclusion

By synthesizing the collective conflict between groups stressed by political economy and the decentralized and stochastic aspects emphasized in evolutionary approaches we hope to match several stylized facts about *de facto* institutional transitions. For example, long periods of decentralized conflict that occasionally generate large changes in conventional practices, the coexistence of large barriers to intergenerational mobility and unequal distribution of income, and the willingness of non-democratic governments to redistribute in the face of widespread social unrest.

One limitation of our framework is that we have assumed that deviations from best-response are independent across individuals. In reality, individual members of each class may choose to act in unison, whether best responding or playing idiosyncratically. Leaders and organizations may have a role in coordinating strategies, as in Acemoglu and Jackson (2011), where the play of "prominent" agents affect the pattern of play in all future periods. The entire membership of a trade union may decide to work under the current contact, or to refuse to do so. Where members of such organizations may commit themselves to acting in unison, dynamics are affected in two ways. First the effective size of the class is reduced to the number of autonomously acting entities. The effect is to increase the fraction of time spent governed by the contract favored by the affected class. By embedding an explicit model of collective action as a public goods game in the above dynamic and assuming that some agents are permanently other regarding, we could generate more adequate behavioral foundations for idiosyncratic play as we have modelled it. We leave this for future work.[5]

Another limitation is that for reasonable updating processes, group sizes, and rates of idiosyncratic play, the waiting times for transitions from one basin of attraction to another are extraordinarily long, certainly surpassing historically relevant time spans (see Kreindler and Young (2011b) for a discussion). However, the above and other extensions can dramatically accelerate the dynamic process, yielding transitions over historically relevant time scales. First, most populations (nations, ethno-linguistic units) are composed of smaller groups of frequently interacting members. Because groups are of quite variable size, the process may be considerably accelerated because the transition times will depend not on the mean group

---

[5]The result is that non best response play is correlated, with deviance from the status quo contract being largely absent when the number of potentially deviant players is insufficient to induce a transition (Kuran, 1991; Bowles, 2004a).

size but on the size of the smallest groups. Second, chance events affect the payoff structures as well as the behaviors of the members of the population, occasionally greatly reducing the size of the basin of attraction of the status quo convention. These effects in conjunction with non best-response play will accelerate the process of transition. Third, there are generally far more than two feasible conventions, and some of them may be adjacent (that is, the resistances among them are small.) A population may traverse a large portion of the state space by means of a series of transitions among adjacent conventions. Fourth conformism and collective action will reduce the effective numbers of players and tend to bunch deviant play, resulting in more frequent transitions.

We think that this model illuminates the dynamics of highly decentralized popular unrest and elite response during the French Revolution (Markoff, 1996; Soboul, 1964; Rudé, 1972), the U.S. civil rights (McAdam, 1986) and labor (Freeman, 1998) movements, as well as the fall of apartheid that we gave as motivation. By specifying historically plausible dynamics of institutional change, we are able to account for properties of contracts that persist. While under some conditions efficient and egalitarian institutions are stable in the long-run, highly unequal and inefficient institutions may outlast (in an evolutionary sense) more egalitarian and efficient institutions if the barriers to upward class mobility are sufficiently great. We also think the approach taken in this model could be extended to incorporate network structure, imitation and conformism, and many other important features determining the diffusion and persistence of economic institutions.

9 Appendix

9.1 Microfoundations for the Contract Game

In this section we model the contract E as a share contract that is egalitarian but inefficient, and the contract U as a fixed payment contract that divides the surplus unequally but produces efficiently. The share in contract E maximizes the employer/landowner's profits subject to the tenant/worker's incentive compatibility constraint for the supply of labor, while the rent or wage in the fixed payment contract is determined by the reservation position of the As. Under both contracts, hours of labor, L, produce output, q, according to $q = f(L)$, where f is a concave, increasing production function satisfying the Inada conditions. A's utility varies with income y and hours worked: $V(y, L) = y - h(L)$. The employer/landowner's (B's) opportunity cost of holding the land is k_c. A's utility-maximizing labor supply under either contract is $L(s), L' > 0$ where s is the share of the residual output retained by A and is equal to 1 in the fixed rental/wage contract and $s \in (0, 1)$ in the share contract. Under the rental/wage contract the A (as residual claimant) works $L(1)$ hours, so total output is $f(L(1))$. Subtracting from this the disutility of the A's labor $h(L(1))$ and the opportunity cost of the land, the joint surplus is $f(L(1)) - h(L(1)) - k_c$.

Under the share contract B's profits of $(1 - s)f(L(s))$ are maximized at a share $s^* < 1$, under which terms A works $L(s^*)$ hours, yielding a total output of $f(L(s^*))$ and a joint surplus of $f(L(s^*)) - h(L(s^*)) - k_c$. Define k_c^* such that $(1-s^*)f(L(s^*)) - k_c^* = s^* f(L(s^*)) - h(L(s^*))$, so that the share contract equally divides the surplus. As expected, the joint surplus under the fixed rental contract is larger, reflecting its superior incentives. To ensure that both contracts are Pareto optima, so that the interests of the classes are opposed, we assume the bargaining power of the Bs in the U contract to be such that the rent, R^* is fixed at $R^* > f(L(1)) - s^* f(L(s^*)) - h(L(1)) + h(L(s^*))$, so that As are worse off in the fixed rent contract.

We can also define $D \equiv (1 - s^*)f(L(s^*)) - k_c^*$, and divide all the payoffs by D. Now by definition of k_c^*, the normalized joint surplus produced under sharecropping is 2, and $\frac{1}{2}$ is the share that the tenant receives. Also define $\rho = \frac{f(L(1)) - h(L(1)) - k_c^*}{D} > 2$ as the joint surplus produced under the rental contract with $\theta = \frac{f(L(1)) - h(L(1)) - R^*}{\rho} < .5$ being the share received by the tenant. This gives the normalized payoffs in the contract game in Table 1.

9.2   Aggregating Individual Dynamics

In this section, we provide a microscopic model of strategy updating and endogenous population size in Section 5. Then the simplified version of this model will yield the models in Section 4. Suppose that there are N agents, $x_1, x_2, \cdots, x_N$. Each agent is identified by two pieces of traits – the strategy used for the contract game (either U or E) in Section 3 and the membership of groups (A class or B class). We use the notations $\sigma(x_i) \in \{0, 1\}$ and $\gamma(x_i) \in \{0, 1\}$, meaning that

$$
\begin{aligned}
\sigma(x_i) &= 1 : \text{agent } x_i \text{ uses } U \text{ strategy in the underlying game; } \sigma(x_i) = 0 \text{ otherwise} \\
\gamma(x_i) &= 1 : \text{agent } x_i \text{ belongs to } A \text{ class; } \gamma(x_i) = 0 \text{ otherwise.}
\end{aligned}
$$

Then the state space for the microscopic model is

$$
\Xi := \left\{ \begin{pmatrix} \sigma \\ \gamma \end{pmatrix} := \begin{pmatrix} \sigma(x_1), \sigma(x_2), \cdots, \sigma(x_N) \\ \gamma(x_1), \gamma(x_2), \cdots, \gamma(x_N) \end{pmatrix} : \sigma(x_i) = 0, 1, \ \gamma(x_i) = 0, 1 \right\} = (\{0, 1\} \times \{0, 1\})^{\{1, \cdots, N\}}.
$$

Then the number of agents in the A population using U, X, the number of agents in the B population using U, Y, and the number of the A population, N^A are

$$
X = \sum_{i=1}^{N} \sigma(x_i) \gamma(x_i), \quad Y = \sum_{i=1}^{N} \sigma(x_i)(1 - \gamma(x_i)), \quad N^A = \sum_{i=1}^{N} \gamma(x_i).
$$

We recall the following best response rules for determining the strategy choice :

$$
BR_A^U(y, N^R) = \begin{cases} 1 & \text{if } y > (N - N^R)q^* \\ 0 & \text{if } y < (N - N^R)q^* \end{cases}, \quad BR_A^E(y, N^R) := 1 - BR_A^U(y, N^R)
$$

$$
BR_B^U(x, N^R) = \begin{cases} 1 & \text{if } x > N^R p^* \\ 0 & \text{if } x < N^R p^* \end{cases}, \quad BR_B^E(x, N^R) := 1 - BR_B^U(x, N^R)
$$

Here $BR_A^U(y, N^R) = 1$ and $BR_B^U(x, N^R) = 1$ mean that the best responses are choosing U.

The events of strategy updating and inter class movement which are decribed in Section 5 can be specifically modeled as follows:

- At the start of each period, every agent $x_1, \cdots, x_N$ receives a strategy revision oppor-

tunity independently with probability v. Thus the random variable, c_i, determining whether agent i receives the strategy revision opportunity is the Bernoulli random variable with the probability of success being v, denoted $Ber(v)$. Here the Bernoulli random variable c_i takes 1 when agent i receives the revision opportunity and takes 0, otherwise. When agents who have received the revision opportunity update their strategies, each agent plays his/her best response with probability $1 - \epsilon$ and play a idiosyncratic strategy with probability ϵ. Then, the random variable, ξ^i, indicating whether agent i plays the idiosynctratic strategy is the Bernoulli random variable with the probability of succee being ϵ, denoted by $Ber(\epsilon)$.

- Next if the state is the one that all agents play either U or E (i.e., $(X_t, Y_t, N_t^A) = (0, 0, k)$ or $(k, N-k, k)$ for some $k = 1, \cdots, N-1$), agents who did not update their strategy (i.e., $c_i = 0$) are matched with another agent from the different population and give birth to childern. Specifically, the event that a chosen A agent can be matched with a B agent can be decribed by a Bernoulli random variable d_1 with the success probability being $(N - n^R)/N$, while the event that a chosen B agent can be matched with an A agent can be decribed by a Bernoulli random variable d_2 with the success probability being n^R/N. Then, the cross couples yield the net increases in agents in the A population according to g function in p. 14 in the text.

This defines the following microscopic stochastic process:

$$\sigma_{t+1}(x_i) = \sigma_t(x_i) \quad + \quad \gamma_t(x_i)c_i((1 - \sigma_t(x_i))BR_A^U - \sigma_t(x_i)BR_A^E - \xi^i BR_A^U)$$
$$+ \quad (1 - \gamma_t(x_i))c_i(((1 - \sigma_t(x_i))BR_B^U - \sigma_t(x_i)BR_B^E + \xi^i BR_B^E)$$

$$\gamma_{t+1}(x_i) = \gamma_t(x_i) + (1 - c_i)\left[\gamma_t(x_i)d_1 + (1 - \gamma_t(x_i))d_2\right]G(X_t, Y_t, N_t^A),$$

where $i = 1, \cdots, N$, and G is defined in the text.

We aggregate the microscopic processes as follows. We denote the Binomial random variable with repetition n and success probability p, by $\mathcal{B}(n, p)$. First we have

$$
\begin{aligned}
N_{t+1}^A &= \sum_{i=1}^{N} \gamma_t(x_i) + (\sum_{i=1}^{N}(1 - c_i)\gamma_t(x_i)d_1 + \sum_{i=1}^{N}(1 - c_i)(1 - \gamma_t(x_i))d_2)G(X_t, Y_t, N_t^R) \\
&= N_t^A + (\mathcal{B}(N_t^A, (1 - v)(1 - \frac{N_t^A}{N})) + \mathcal{B}(N - N_t^A, v\frac{N_t^A}{N}))G(X_t, Y_t, N_t^R),
\end{aligned}
$$

where we note that, for example, $(1-c_i)d_1 \sim Ber((1-v)\frac{N_t^A}{N})$ and

$$\sum_{i=1}^{N} \gamma_t(x_i)(1-c_i)d_1 \sim \sum_{\{i:\gamma_t(x_i)=1\}} (1-c_i)d_1 \sim \sum_{j}^{\sum_i \gamma_t(x_i)} (1-c_i)d_1 \sim \mathcal{B}(N_t^A, (1-v)(1-\frac{N_t^A}{N})).$$

Similiarly,

$$X_{t+1} = \sum_{i=1}^{N} \gamma_{t+1}(x_i)\sigma_t(x_i) + \sum_{i=1}^{N} \gamma_{t+1}(x_i)\gamma_t(x_i)c_i[(1-\sigma_t(x_i))BR_A^U - \sigma_t(x_i)BR_A^E - \xi^i BR_A^U]$$

$$+ \sum_{i=1}^{N} \gamma_{t+1}(x_i)(1-\gamma_t(x_i))c_i[(1-\sigma_t(x_i))BR_B^U - \sigma_t(x_i)BR_B^E + \xi^i BR_B^E]$$

$$= : I + II + III.$$

For the first term (I), we find

$$I = \sum_{i=1}^{N} \gamma_{t+1}(x_i)\sigma_t(x_i) = \sum_{i=1}^{N} \gamma_t(x_i)\sigma_t(x_i)$$

$$+ (\sum_{i=1}^{N}(1-c_i)\gamma_t(x_i)d_1\sigma_t(x_i) + \sum_{i=1}^{N}(1-c_i)\gamma_t(x_i)d_2\sigma_t(x_i))G(X_t, Y_t, N_t^A)$$

$$= X_t + (\mathcal{B}(X_t, (1-v)\frac{N-N_t^A}{N}) + \mathcal{B}(Y_t, (1-v)\frac{N_t^A}{N}))G(X_t, Y_t, N_t^A)$$

and for (II), if $c_i = 0$, $\gamma_{t+1}(x_i) = \gamma_t(x_i)$ and $\gamma_t^2(x_i) = \gamma_t(x_i)$. Thus

$$II = \sum_{i=1}^{N} \gamma_t(x_i)c_i((1-\sigma_t(x_i))BR_A^U - \sigma_t(x_i)BR_A^E - \xi^i BR_A^U)$$

$$= BR_A^U \mathcal{B}(N_t^A - X_t, v) - BR_A^E \mathcal{B}(X, v) - BR_A^U \mathcal{B}(N_t^A, v\epsilon)$$

For (III), since $\gamma_t(x_i)(1-\gamma_t(x_i)) = 0$, we have $(III) = 0$. We then aggregate similarly and find the following aggregated equations:

$$N_{t+1}^A = N_t^A + (\mathcal{B}(N - N_t^A, (1-v)\frac{N_t^A}{N}) + \mathcal{B}(N_t^A, (1-v)\frac{N-N_t^A}{N}))G \qquad (23)$$
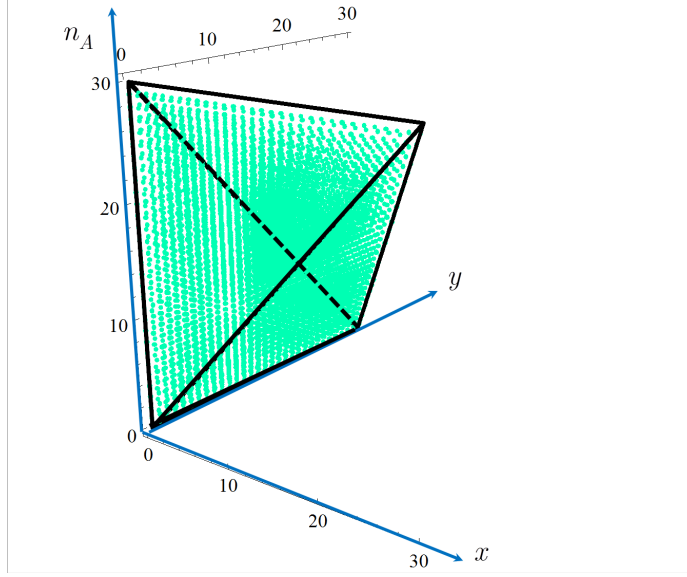
Figure 4: **State space for the aggregate population dynamic**

$$X_{t+1} = X_t + (\mathcal{B}(Y_t, (1-v)\frac{N_t^A}{N}) + \mathcal{B}(X_t, (1-v)\frac{N - N_t^A}{N}))G(X_t, Y_t, N_t^A) \quad (24)$$
$$+ BR_A^U(Y_t, N_t^A)\alpha_t^A - BR_A^E(Y_t, N_t^R)\beta_t^B - BR_A^U(Y_t, N_t^R)\omega_t^A$$

$$Y_{t+1} = Y_t - (\mathcal{B}(Y_t, (1-v)\frac{N_t^R}{N})F_1 + \mathcal{B}(X_t, (1-v)\frac{N - N_t^R}{N}))G(X_t, Y_t, N_t^A) \quad (25)$$
$$+ BR_B^U(X_t, N_t^A)\alpha_t^B - BR_B^E(X_t, N_t^R)\beta_t^B + BR_B^E(X_t, N_t^R)\omega_t^B$$

where

$$\alpha_t^A \sim \mathcal{B}(N_t^A - X_t, v), \beta_t^A \sim \mathcal{B}(X_t, v), \ \omega_t^A \sim \mathcal{B}(N_t^R, v\epsilon)$$
$$\alpha_t^B \sim \mathcal{B}(N - N_t^A - Y_t, v), \beta_t^B \sim \mathcal{B}(Y_t, v), \omega_t^B \sim \mathcal{B}(N - N_t^A, v\epsilon).$$

We show the figure of the state space of the aggregate population process.

9.3 Determination of Resistances for the Process with Endogenous Class Mobility

We first recall that a resistance for the transitions between states, $r(i, j)$, is defined to be a number such that $0 < \lim P^\epsilon(i, j)/\epsilon^{r(i,j)} < \infty$, where $i, j \in \Delta$ and P^ϵ is the transition prob-

ability for the perturbed joint dynamics (See Young (1993)). Then from the specifications of the population and strategy updating dynamics, we find

$$
\begin{aligned}
r((k, N-k, k), (k_U^*, N-k_U^*, k_U^*)) &= 0 \ \text{ for } k = 1, \cdots, N-1 \\
r((0, 0, k), (0, 0, k_E^*)) &= 0 \ \text{ for } k = 1, \cdots, N-1 \\
r((x, y, k), (x', y', k)) &= [x - x']_+ BR_A^U(x, y, k) + [x' - x]_\infty BR_A^E(x, y, k) \\
&\quad + [y - y']_\infty BR_B^U(x, y, k) + [y' - y]_+ BR_B^E(x, y, k)
\end{aligned}
$$

where $[t]_+ = t$ if $t > 0$, $= 0$ otherwise and $[t]_\infty = \infty$ if $t > 0$, $= 0$ otherwise and other $r(i, j)$s are infinity.

10 Proofs

10.1 Proof of 4.1

Proof. The proof of the first claim is in the text. By solving (6) for γ, we find

$$
\gamma^* = \frac{1 + (1 - \theta)\rho}{1 + 2(1 - \theta)\rho + \rho^2(1 - \theta)\theta} \tag{26}
$$

For the second claim, differentiating with respect to ρ and θ yields:

$$
\frac{d\gamma^*}{d\rho} = -\frac{\rho\left(-1 + \rho + \rho^2(\theta - 1)^2 - 2\rho\theta\right)}{\left(-1 + 2\rho(-1 + \theta) + \rho^2(\theta - 1)\theta\right)^2} < 0 \tag{27}
$$

The negative sign follows from the fact that $2\rho\theta + 1 < (\rho(1 - \theta))^2 + \rho$.

For the third claim, differentiate γ^* with respect to θ to get

$$
\frac{d\gamma^*}{d\theta} = \frac{(1 - \theta)\left(-1 - 2\rho\theta - \rho^2(1 - \theta)\theta\right)}{\left(-1 + 2\rho(\theta - 1) + \rho^2(\theta - 1)\theta\right)^2} < 0 \tag{28}
$$

since $\theta < 1/2$ implies $(1 - \theta) > 0$. The fourth claim follows from substituting $\rho^2(1 - \theta)\theta < 1$ into equation (7). $\qquad\square$

10.2 Proof of 5.3

Proof. First from equation (15) , we explicitly find

$$k_U^* = N\frac{\bar{y} - \theta\rho}{\bar{y} - \theta\rho + (1-\theta)\rho}, \qquad k_E^* = N\frac{\bar{y} - 1}{\bar{y}}.$$

We let $\alpha = \theta\rho$ and $\beta = (1-\theta)\rho$ and define a function

$$f(y) := \frac{(y - \alpha)/(y - \alpha + \beta)}{(y - 1)/y} \tag{29}$$

and if we can find y_0 such that for all $y > y_0$, $f'(y) > 0$, then we can find $\bar{y}^*$ such that for all $\bar{y} > \bar{y}^*$, $\lceil k_U^*(1 - p^*) \rceil > \lceil k_E^* q^* \rceil$. By taking the derivative of (29), we see that $f'(y) > 0$ if and only if $(\beta - 1)y^2 - 2(\beta - \alpha)y + \alpha(\beta - \alpha) > 0$. Since $\alpha < 1 < \beta$, if we choose y_0 such that $y_0 > 2(\beta - \alpha)/(\beta - 1)$, then for all $y > y_0$, we have $f'(y) > 0$. ☐

10.3 Proof of 7.1.

Proof. The new idiosyncratic play rate given by (17) modifies the resistance expressions in (6) as follows:

$$r(U, E) = \left\lceil N_A \frac{(1-\theta)\rho}{1 + (1-\theta)\rho} \frac{1}{1 + \lambda((1-\theta)\rho - \theta\rho)} \right\rceil, \quad r(E, U) = \left\lceil N_B \frac{1}{1 + \theta\rho} \right\rceil$$

Then, since $r(U, E)$ is decreasing with respect to λ, the result follows. ☐

10.4 Proof of 7.2

Proof. We set

$$\varphi_1(\tau) := \frac{\tilde{b}(\tau)}{1 + \tilde{b}(\tau)}, \qquad \varphi_{2,\lambda}(\tau) := \frac{1}{1 + \lambda(\tilde{b}(\tau) - \tilde{a}(\tau))},$$

where $\tilde{a}(\tau) := a + b\tau$ and $\tilde{b}(\tau) = (1 - \tau)b$. Then by ignoring the integer problem, we find

$$sign(\frac{d}{d\tau} P(U, E, \tau)) = -sign(\frac{\varphi_1'(\tau)}{\varphi_1(\tau)} + \frac{\varphi_{2,\lambda}'(\tau)}{\varphi_{2,\lambda}(\tau)})$$

where

$$\frac{\varphi_1'(\tau)}{\varphi_1(\tau)} = -\frac{b}{(1 + \tilde{b}(\tau))\tilde{b}(\tau)}, \quad \frac{\varphi_{2,\lambda}'(\tau)}{\varphi_{2,\lambda}(\tau)} = \frac{2b\lambda}{1 + \lambda(\tilde{b}(\tau) - \tilde{a}(\tau))}.$$

Then since $\varphi_1'(\tau)/\varphi_1(\tau)$ is decreasing in τ and $\varphi_{2,\lambda}'(\tau)/\varphi_{2,\lambda}(\tau)$ is increasing in τ, we have

$$\frac{\varphi_1'(\tau)}{\varphi_1(\tau)} + \frac{\varphi_{2,\lambda}'(\tau)}{\varphi_{2,\lambda}(\tau)} > -\frac{b}{(1+\tilde{b}(\bar{\tau}))\tilde{b}(\bar{\tau})} + \frac{2b\lambda}{1+\lambda(\tilde{b}(0)-\tilde{a}(0))} = -\frac{2}{2+a+b}\frac{2b}{a+b} + \frac{2b\lambda}{1+\lambda(b-a)}.$$

By choosing

$$\bar{\lambda} := \frac{\alpha}{2b-(b-a)\alpha}, \quad \text{where } \alpha = \frac{2}{2+a+b}\frac{2b}{a+b},$$

we obtain the above claim. $\qquad\square$

10.5 Proof of 7.3

We begin with a lemma:

Lemma 1. Suppose that $v(E) \geq 1/(1-\beta)$. Then $v(E) \leq v(U)$.

Proof. Let τ_E be the optimal choice in (22) and let $q(\tau) := P(E,U,\tau)$. Then we have

$$1 + \beta v(U)q(\tau_E) + \beta v(E)(1-q(\tau_E)) = v(E)$$
$$\geq 1 + \beta v(E) = 1 + \beta v(E)q(\tau_E) + \beta v(E)(1-q(\tau_E))$$

By rearranging we obtian $v(U) > v(E)$. $\qquad\square$

Proof. We let $p_{\epsilon,\lambda}(\tau) := P(U,E,\tau)$ and define

$$V(\tau) := u(U,\tau) - \beta(v(U)-v(E))P(U,E,\tau) + \beta v(U)$$

and (ignoring the integer problem) find

$$V'(\tau) = (2\xi-1)b + \beta(v(U)-v(E))p_{\epsilon,\lambda}(\tau)(-\ln\epsilon)N_A\varphi_1(\tau)\varphi_2(\tau)\left(\frac{\varphi_1'(\tau)}{\varphi_1(\tau)} + \frac{\varphi_{2,\lambda}'(\tau)}{\varphi_{2,\lambda}(\tau)}\right)$$

Then for sufficiently large λ, we have $\varphi_1'(0)/\varphi_1(0) + \varphi_{2,\lambda}'(0)/\varphi_{2,\lambda}(0) > 0$ as in the above proposition. Thus for sufficiently large N_A we find $V'(0) > 0$ which implies $\tau_U > 0$. $\qquad\square$

References

Acemoglu, Daron and James Robinson, *Economic Origins of Dictatorship and Democracy*, Cambridge University Press, 2006.

_ **and** _ , "Persistence of Power, Elites and Institutions," *The American economic review*, 2008, *98* (1), 267–293.

_ **and Matthew O Jackson**, *History, expectations, and leadership in the evolution of cooperation*, National Bureau of Economic Research, 2011.

_ , **Simon Johnson, and James A Robinson**, "Institutions as a fundamental cause of long-run growth," *Handbook of economic growth*, 2005, *1*, 385–472.

Adler, Glenn and Eddie Webster, "Challenging Transition Theory: The Labour Movement, Radical Reform and the Transition to Democracy in South Africa," *Politics and Society*, 1995, *23* (1), 75–106.

Aldous, David, "Stopping Times and Tightness," *Annals of Probability*, 1978, *6*, 335–340.

Alesina, Alberto, Paola. Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," Technical Report, National Bureau of Economic Research 2011.

Anderlini, Luca and Antonella Ianni, "Path Dependence and Learning from Neighbors," *Games and Economic Behavior*, 1996, *13*, 141–177.

Apesteguia, Jose, Steffen Huck, and Jorg Oechssler, "Imitation - theory and experimental evidence," *Journal of Economic Theory*, 2007, *136*, 217–235.

_ , _ , _ , **and Simon Weidenholzer**, "Imitation and the evolution of Walrasian behavior: Theoretically fragile but behaviorally robust," *Journal of Economic Theory*, 2010, *145*, 1603–1617.

Aronson, Donald G., "Density-Dependent Interaction-Diffusion Systems," in Warren E. Stewart, Willis H. Ray, and Charles C. Conley, eds., *Dynamics and Modeling of Reactive Systems*, Academic Press, 1980, pp. 161–176.

Axtell, Robert L., Joshua M. Epstein, and H. Peyton Young, "The Emergence of Classes in a Multi Agent Bargaining Model," in Steven Durlauf and H. Peyton Young, eds., *Social Dynamics*, Cambridge: MIT Press, 2001, pp. 191–211.

Banerjee, A. and L. Iyer, "History, institutions, and economic performance: the legacy of colonial land tenure systems in India," *The American economic review*, 2005, *95* (4),

1190–1213.

Banerjee, Abhijit and Andrew Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, *101* (2), 274–298.

Bardhan, Pranab, *Land, Labor, and Rural Poverty: Essays in Development Economics*, Columbia University Press, 1984.

Bates, Peter W., Paul C. Fife, Xiaofeng Ren, and Xuefeng Wang, "Traveling waves in a convolution model for phase transitions," *Archive for Rational Mechanics and Analysis*, 1997, *138*, 105–136.

Beggs, A. W., "Waiting times and equilibrium selection," *Economic Theory*, 2005, *25*, 599–628.

Benabou, Roland, "Unequal Societies: Income Distribution and the Social Contract," *The American Economic Review*, 2000, *90* (1), 96–129.

Benaim, Michel and Jorgen W. Weibull, "Deterministic approximation of stochastic evolution in games," *Econometrica*, 2003, *71*, 873–903.

Bergin, James and Bart Lipman, "Evolution with State-Dependent Mutations," *Econometrica: Journal of the Econometric Society*, 1996, pp. 943–956.

_ **and Dan Bernhardt**, "Cooperation Through Imitation," *Games and Economic Behavior*, 2009, *67*, 976–988.

Betzig, Laura, *Despotism and Differential Reproduction : a Darwinian View of History.*, New York: Aldine Pub, 1986.

Bhowmik, Samir Kumar, "Numerical Approximation of a Nonlinear Partial Integro-Differential Equation." PhD dissertation, Heriot-Watt University 2008.

Billingsley, Patrick, *Convergence of Probability Measures*, New York: Wiley, 1968.

Binmore, K., L. Samuelson, and P. Young, "Equilibrium selection in bargaining models," *Games and economic behavior*, 2003, *45* (2), 296–328.

Bisin, Alberto and Thierry Verdier, ""Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits," *The Quarterly Journal of Economics*, 2000, *115* (3), 955–988.

_ **and Thierry Verider**, "The Economics of Cultural Transmission and Socialization," in "Handbook of Social Economics," Elsevier, 2011, pp. 339–416.

Bjornerstedt, Jonas and Jorgen W. Weibull, "Nash equilibrium and evolution by imitation," in Kenneth Arrow and Enrico Colombatto, eds., *Rationality in Economics*,

Macmillan, 1993.

Blau, Peter, *Exchange and Power in Social Life*, New York: John Wiley, 1964.

Blum, Jerome, *Lord and Peasant in Russia*, Princeton University Press, 1971.

Blume, Lawrence E., "The Statistical Mechanics of Strategic Interaction," *Games and economic behavior*, 1993, *5*, 387–424.

＿ , "The Statistical Mechanics of Best-Response Strategy Revision," *Games and Economic Behavior*, 1995, *11*, 111–145.

＿ , "Population Games," in W.Brian Arthur, Steven N. Durlauf, and David A. Lane, eds., *The Economy as a Complex Adaptive System II*, MA: Addison-Wesley, 1997, pp. 425–460.

Blume, L.E., "The statistical mechanics of best-response strategy revision," *Games and Economic Behavior*, 1995, *11* (2), 111–145.

Bowles, Samuel, *Microeconomics: Behavior, Institutions, and Evolution*, Princeton: Princeton University Press, 2004.

＿ , *Microeconomics: Behavior, Institutions and Evolution*, Princeton University Press, 2004.

＿ **and Herbert Gintis**, ""The Inheritance of Inequality"," *The Journal of Economic Perspectives*, 2002, *16* (3), 3–30.

Brenner, Robert, "Agrarian Class Structure and Economic Development in Pre-Industrial Europe," *Past and Present*, 1976, *70*, 30–70.

Brock, William. A. and Steven N. Durlauf, "A formal model of theory choice in science," *Economic Theory*, 1999, *14*, 113–130.

Brock, William A. and Steven N. Durlauf, "Discrete choice with social interactions," *Review of Economic Studies*, 2001, *68*, 235–260.

Carr, Jack and Robert L. Pego, "Metastable Patterns in Solutions of $u_t = \epsilon^2 u_x x - f(u)$," *Communications on pure and applied mathematics*, 1989, *42* (5), 523–576.

Chen, Xinfu, "Existence, Uniqueness, and Asymptotic Stability of Traveling Waves in Nonlocal Evolution Equations," *Advances in Differential Equations*, 1997, *2*, 125–160.

Collet, Pierre and Jean Pierre Eckmann, *Instabilities and Fronts in Extended Systems*, Princeton University Press, 1990.

Comets, Francis, "Nucleation for a Long Range Magnetic Model," *Annales de l'Institut Henri Poincare*, 1987, *23*, 135–178.

＿ **and Theodor Eisele**, "Asymptotic Dynamics, Non-Critical and Critical Fluctuations for a Geometric Long-Range Interacting Model," *Communications in Mathematical Physics*,

1988, *118*, 531–567.

Corak, Miles, "Inequality from Generation to Generation: The United States in Comparison," in Robert Rycroft, ed., *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, ABC-CLIO, 2012.

Darling, Richard and James Norris, "Differential Equation Approximations for Markov Chains," *Probability Survey*, 2008, *5*, 37–79.

David, P.A., "Clio and the Economics of QWERTY," *The American economic review*, 1985, *75* (2), 332–337.

Duclos, Joan Esteban J-Y and Debraj Ray, "Polarization: Concepts, Measurement, Estimation.," *Econometrica*, 2004, *62*, 819–852.

Durlauf, Steven, "A Framework for the Study of Individual Behavior and Social Interactions," *Sociological Methodology*, 2001, *31*, 47 – 87.

Durrett, Rick, "Stochastic Spatial Models," *SIAM Review*, 1999, *41*, 677–718.

_ , *Mutual Invadability Implies Coexistence in Spatial Models*, American Mathematical Society, 2002.

Eisele, Theodor and Richard S. Ellis, "Symmetry Breaking and Random Waves for Magnetic Systems on a Circle," *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1983, *63*, 297–348.

Ellison, Glenn, "Learning, Local Interaction, and Coordination," *Econometrica*, September 1993, *61* (5), 1047–1071.

_ , "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step by Evolution," *Review of Economic Studies*, 2000, *67*, 17–45.

Esteban, Joan and Debraj Ray, "Collective action and the group size paradox," *American political science review*, 2001, *95* (3), 663–672.

Ethier, Stewart N. and Thomas G. Kurtz, *Markov Processes: Characterization and Convergence*, New York: Wiley, 1986.

Fafchamps, Marcel, "Solidarity Networks in Preindustrial Societies: Rational Peasants with a Moral Economy," *Economic Development and Cultural Change*, 1992, *41* (1), 147–74.

Fedderke, J.W., RH de Kadt, and JM Luiz, "Indicators of political liberty, property rights and political instability in South Africa: 1935-97," *International Review of Law and Economics*, 2001, *21*, 103–134.

Fife, Paul C., *Mathematical aspects of reaction and diffusing systems*, Springer-Verlag, 1979.

Fisher, Arthur and Jerrold E. Marsden, "Linearization stability of nonlinear partial differential equations," *Proceedings of Symposia in Pure Mathematics*, 1975, *27*, 219–263.

Foster, D. and H. P. Young, "Stochastic evolutionary game dynamics," *Theoretical Population Biology*, 1990, *38* (2), 219–232.

Foster, Dean and Peyton Young, "Stochastic Evolutonary Game Dynamics," *Theoretical Population Biology*, 1990, *38*, 219–232.

Freeman, Richard, "Spurts in Union Growth: Defining Moments and Social Processes," in A Bordo, C. Goldin, and L. Katz, eds., *Defining Moment: The Great Depression and the American Economy in the 20th Century*, Chicago: University of Chicago Press, 1998.

Freidlin, Mark and Alexander Wentzell, *Random Perturbations of Dynamical Systems*, Springer, 1984.

Fryde, E.B., *Peasants and landlords in later Medieval England, C. 1380-C. 1525*, Stroud, Gloucestershire: Alan Sutton Publishing, 1996.

Fudenberg, Drew and David K. Levine, *The theory of learning in games*, MIT Press, 1998.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *The Review of Economic Studies*, 1993, *60* (1), 35–52.

Gellner, Ernest, *Nations and nationalism* New perspectives on the past., Ithaca: Cornell University Press, 1983.

Greif, A., "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies," *Journal of political economy*, 1994, pp. 912–950.

Harsanyi, John C. and Reinhard Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, 1988.

Helbing, Dirk, "Microscopic Foundation of Stochastic Game Dynamical Equations," in Werner Leinfellner and Eckehart Koehler, eds., *Game Theory, Experience, Rationality*, Dordrechts: Kluwer Academic, 1998, pp. 211–224.

Hofbauer, Josef, "Equilibrium Selection Via Traveling Waves," in Werner Leinfellner and Eckehart Kohler, eds., *Game Theory, Experience, Rationality: Foundations of Social Sciences, Economics and Ethics*, Vienna Circle Institute Yearbook 5, 1997, pp. 245–249.

_ , "The spatially dominant equilibrium of a game," *Annals of Operations Research*, 1999, *89*, 233–251.

_ **and Karl Sigmund**, *Evolutionary Games and Population Dynamics*, Cambridge: Cambridge University Press, 1998.

_ **and** _ , "Evolutionary Game Dynamics," *Bulletin of American Mathematical Society*, 2003, *40*, 479–519.

_ **and William Sandholm**, "Evolution in Games with Randomly Disturbed Payoffs," *Journal of Economic Theory*, 2007, *132*, 47–69.

_ , **Vivian Hutson, and Glenn T. Vickers**, "Travelling Waves for Games in Economics and Biology," *Nonlinear Analysis*, 1997, *30*, 1235–1244.

Hojman, Daniel and Adam Szeidl, "Endogenous Networks, Social Games, and Evolution," *Games and Economic Behavior*, 2006, *55* (1), 112–130.

Hopkins, Ed, "A note on best response dynamics," *Games and Economic Behavior*, 1999, *29*, 138–150.

Horntrop, David J., Markos A. Katsoulakis, and Dionisios G. Vlachos, "Spectral Methods for Mesoscopic Models in Pattern Formation," *Journal of Computational Physics*, 2001, *173*, 361–390.

Hotelling, Harold, "Stability in Competition," *Economic Journal*, 1929, *39* (153), 41–57.

Hutson, Vivian and Glenn T. Vickers, "Travelling Waves and Dominance of ESS's," *Journal of Mathematical Biology*, 1992, *30*, 457–471.

Josephson, Jens and Alexander Matros, "Stochastic imitation in finite games," *Games and Economic Behavior*, 2004, *49*, 244–259.

Kandori, M. G., G. Mailath, and R. Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 1993, *61*, 29–56.

Kandori, Michihiro, George J. Mailath, and Rafael Rob, "Learning,, Mutation, and Long-Run Equilibria in Games," *Econometrica*, 1993, *61*, 29–56.

Katsoulakis, Markos A. and Panagiotis E. Souganidis, "Stochastic Ising Models and Anisotropic Front Propagation," *Journal of Statistical Physics*, 1997, *87*, 63–89.

_ , **Petr Plechac, and Dimitrios K. Tsagkarogiannis**, "Mesoscopic Modeling for Continuous Spin Lattice Systems: Model Problems and Micromagnetic Applications," *Journal of Statistical Physics*, 2005, *119*, 347–389.

Kipnis, Claude and Claudio Landim, *Scaling Limits of Interacting Particle Systems*,

Springer-Verlag, 1999.

Kreindler, Gabriel H. and H. Peyton Young, "Fast convergence in evolutionary equilibrium selection," *University of Oxford working paper*, 2011.

Kreindler, G.E. and H.Peyton Young, "Fast Convergence in Evolutionary Equilibrium Selection," *Economics Series Working Papers*, 2011.

Kuran, T., "Now out of never: The element of surprise in the East European revolution of 1989," *World Politics: A Quarterly Journal of International Relations*, 1991, pp. 7–48.

Kurtz, Thomas G., "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Process," *Journal of Applied Probability*, 1970, *7*, 49–58.

Lebowitz, Joel L. and Oliver Penrose, "Rigorous Treatment of the Van der Waals-Maxwell Theory of the Liquid Vapour Transition," *Journal of Mathematical Physics*, 1966, *7*, 98–113.

Levine, David K. and Wolfgang Pesendorfer, "The Evolution of Cooperation Through Imitation," *Games and Economic Behavior*, 2007, *58*, 293–315.

Ligget, Thomas, *Interacting Particle Systems*, Springer-Verlag, 1985.

Lohmann, S., "The dynamics of informational cascades: The Monday demonstrations in Leipzig, East Germany, 1989-91," *World Politics*, 1994, pp. 42–101.

Markoff, J., *The abolition of feudalism: Peasants, lords, and legislators in the French Revolution*, Pennsylvania State Univ Pr, 1996.

Masi, Anna De, Enza Orlandi, Errico Presutti, and Livio Triolo, "Glauber Evolution with Kac Potentials:I. Mesoscopic and Macroscopic Limits, Interface Dynamics," *Nonlinearity*, 1994, *7*, 633–696.

_ , _ , _ , **and** _ , "Glauber evolution with Kac potentials: III. Spinodal decomposition," *Nonlinearity*, 1996, *9*, 53–114.

_ , _ , _ , **and** _ , "Glauber evolution with Kac potentials: II. Fluctuations," *Nonlinearity*, 1999, *9*, 27–51.

_ , **Pablo Ferrari, and Joel Lebowitz**, "Reaction Diffusion Equations for Interacting Particle Systems," *Journal of Statistical Physics*, 1986, *44*, 589–644.

_ , **Thierry Gobron, and Errico Presutti**, "Traveling Fronts in Non-Local Evolution Equations," *Archive for Rational Mechanics and Analysis*, 1995, *132*, 143–205.

McAdam, D., "Recruitment to high-risk activism: The case of freedom summer," *American Journal of Sociology*, 1986, pp. 64–90.

Mookherjee, Dilip and Debraj Ray, "Contractual Structure and Wealth Accumulation," *The American Economic Review*, 2002, *92* (4), 818–849.

Morris, Stephen, "Contagion," *Review of Economic Studies*, 2000, *67*, 57–78.

Mulder, Monique Borgerhoff, Samuel Bowles, Tom Hertz, Adrian Bell, Jan Beise, Greg Clark, Ila Fazzio, Michael Gurven, Kim Hill, Paul L Hooper et al., "Intergenerational wealth transmission and the dynamics of inequality in small-scale societies," *science*, 2009, *326* (5953), 682–688.

Murray, James D., *Mathematical Biology*, Springer, 1989.

Naidu, Suresh, Samuel Bowles, and Sung-Ha Hwang, "Evolutionary Bargaining with Intentional Idiosyncratic play," *Economics Letters*, 2010, *109* (1), 31–33.

Newman, William I., "Some Exact Solutions to a Nonlinear Diffusion Problem in Population Genetics and Combustion," *Journal of Theoretical Biology*, 1980, *85*, 325–334.

_ , "The Long-Time Behaviour of Solutions to a Nonlinear Diffusion Problem in Population Genetics and Combustion," *Journal of Theoretical Biology*, 1983, *104*, 473–484.

North, Douglass C., *Structure and Change in Economic History*, New York: W. W. Norton and Co., 1981.

North, Douglass C and Barry R Weingast, "Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England," *The journal of economic history*, 1989, *49* (04), 803–832.

Nunn, Nathan, "The Importance of History for Economic Development," *Annual Reviews of Economics*, 2009, *1* (1), 65–92.

Olson, Mancur, *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, MA: Harvard University Press, 1965.

Orlandi, Enza and Livio Triolo, "Traveling Fronts in Nonlocal Models for Phase Separation in External Field," *Proceedings of the Royal Society*, 1997, *127 A*, 823–835.

Passo, Roberta Dal and Piero de Mottoni, "The Heat Equation with a Non-Local Density Dependent Advection Term," *preprint*, 1991.

Piketty, T., "Top income shares in the long run: An overview," *Journal of the European economic association*, 2005, *3* (2-3), 382–392.

Platteau, J. P., "A framework for the analysis of evolving patron-client ties in agrarian economies," *World Development*, 1995, *23* (5), 767–786.

Presutti, Errico, *Scaling Limits in Statistical Mechanics and Microstructures in Contin-*

uum Mechanics, Springer, 2009.

Robson, Arthur J. and Fernando Vega-Redondo, "Efficient equilibrium selection in evolutionary games with random matching," *Journal of Economic Theory*, 1996, *70*, 65–92.

Roemer, John E, "Kantian equilibrium," *The Scandinavian Journal of Economics*, 2010, *112* (1), 1–24.

Rudé, G.F.E., *The crowd in the French Revolution*, Oxford University Press, 1972.

Sandholm, William H., "Pairwise Comparison Dynamics and Evolutionary Foundations for Nash Equilibrium," *Games*, 2010, *1*, 3–17.

_ , *Population Games and Evolutionary Dynamics*, MIT Press, 2010.

Schelling, Thomas C., "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1971, *1*, 143–186.

Schlag, Karl H., "Why imitate, and if so, how?," *Journal of Economic Theory*, 1998, *78*, 130–156.

Scott, James C., *The Moral Economy of the Peasant: Rebellion and Subsistence in Southeast Asia*, New Haven: Yale University Press, 1976.

Sebastian, Florence Goffette-Nago Grauwin and Pablo Jensen, "Dynamic models of residential segregation: an analytical solution," *Journal of Public Economics*, 2012, *96*, 124–141.

Shigesada, Nanako, "Spatial Distribution of Dispersing Animals," *Journal of Mathematical Biology*, 1980, *9*, 85–96.

Smith, John Maynard, *Evolution and the Theory of Games*, Cambridge University Press, 1982.

Smith, Michael J., "The stability of a dynamic model of traffic assignment - an application of a method of Lyapunov," *Transportation Science*, 1984, *18*, 245–252.

Soboul, A., *The Parisian sans-culottes and the French Revolution, 1793-4*, Soc French Studies, 1964.

Sokoloff, Kenneth and Stan Engerman, "Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 2000, *14* (3), 217–232.

Stein, Elias M. and Rami Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.

Szabo, Gyorgy and Gabor Fath, "Evolutionary Games on Graphs," *Physics Report*,

2007, *446*, 97–216.

Trigger, Bruce, *Understanding early civilizations*, Cambridge: Cambridge University Press, 2003.

Turchin, Peter, *Historical dynamics*, Princeton University Press Princeton, 2003.

Turing, Alan, "The chemical basis of morphogenesis," *Philosophical Transactions of the Royal Society B*, 1952, *237*, 37–72.

Vega-Redondo, Fernando, "The evolution of Walrasian behavior," *Econometrica*, 1997, *65*, 375–384.

Vickers, Glenn, Vivian Hutson, and Chris J. Budd, "Spatial Patterns in Population Conflicts," *Journal of Mathematical Biology*, 1993, *31*, 411–430.

Volpert, Aizik I., Vitaly A. Volpert, and Vladimir A. Volpert, *Traveling wave solutions of parabolic systems*, American Mathematics Society, 1994.

Weibull, Jorgen W., *Evolutionary Games Theory*, MA: Cambridge, 1995.

Wilson, F., *Labour in the South African gold mines 1911-1969*, Vol. 6, Cambridge University Press, 1972.

Wood, Elisabeth, *Insurgent Collective Action and Civil War In El Salvador*, Cambridge: Cambridge University Press, 2003.

Wood, Elisabeth Jean, *Forging Democracy From Below: Insurgent Transitions in South Africa and El Salvador*, Cambridge: Cambridge University Press, 2000.

Young, H. Peyton, "The Evolution of Conventions," *Econometrica*, 1993, *61* (1), 57–84.

_ , "The Evolution of Conventions," *Econometrica*, January 1993, *61* (1), 57–84.

_ , "An Evolutionary Model of Bargaining," *Journal of Economic Theory*, February 1993, *59* (1), 145–168.

_ , *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton Univerisity Press, 1998.

_ **and Mary A. Burke**, "Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture," *American Economic Review*, 2001, *91*, 559–573.

_ **and Mary Burke**, "Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture," *American Economic Review*, 2001, *91* (3), 559–573.

Young, H.P., *Individual strategy and social structure: An evolutionary theory of institutions*, Princeton Univ Press, 1998.

_ , "The diffusion of innovations in social networks," *Economy as an evolving complex sys-*

tem, 3, 2006, *3*, 267.