

# The Dynamics of Voter Turnout and Party Choice

Nathan A. Collins

Santa Fe Institute

1399 Hyde Park Road, Santa Fe, NM 87501

`nac@santafe.edu`

March 22, 2010

## Abstract

A considerable body of research focuses on why voter turnout changed — specifically, why it declined — in the 1960s and 1970s. Most models of the change focus on factors such as a decline in civic involvement or a shift in the age distribution toward younger citizens who vote less frequently. While these approaches have taught us much about voter turnout, they are puzzling in that none actually focuses on the origins of change. That is, they explain changes in turnout with changes in other factors but never directly explain these other changes. Furthermore, almost all existing models of voter turnout dynamics assume exogenous party choices, although voters' party choices are not fixed in general and this fact likely has consequences for voter turnout. I develop a model of turnout dynamics — one that directly explains changes in turnout — with endogenous party choice. I show that this model explains voter turnout in U.S. presidential elections better than other models, and I reach similar results in the context of California voter registration data. While voter registration is of obvious importance to voter turnout, it receives little attention and its dynamics have never been studied.

## 1 Introduction

Since the decline in voter turnout in the 1960s and 70s, political scientists have worked hard to explain what changed that led to the decline and, more generally, what leads to higher turnout. The resulting models focus on such factors as a decline in civic involvement, a decline in party identification, mobilization efforts, and changes in the age distribution, among other things (e.g.,

Abramson and Aldrich 1982; Ansolabehere and Konisky 2006; Cavanagh 1981; Gerber and Green 2000; Gerber, Green, and Shachar 2003; Highton 1997; Knack and White 2004; Nagler 1991; Rosenstone and Hansen 1993). Strangely, none of these approaches attempts to explain change itself. While we have learned much about voter turnout from these studies, they are essentially static models that explain changes in voter turnout in terms of changes in other variables. Bluntly, these models do not actually explain change itself.

Furthermore, when these models attempt to explain turnout changes, they tend to focus on the 1960s at the expense of other interesting features of voter turnout dynamics, e.g., the post-World War II increase in U.S. turnout and the plateau in the 1970s, 80s, and 90s. Finally, few models attempt to treat party choice and turnout simultaneously — in fact, many formal model of turnout assume fixed party choices (e.g., Diermeier and Van Mieghem 2005; Feddersen and Sandroni 2006; Palfrey and Rosenthal 1983, 1985. Among dynamical models Collins, Kumar, and Bendor (2009) study turnout dynamics, and Achen (2002) and Fiorina (1977) study party identification and party choice dynamics, but no model studies both at the same time. This is also strange, since, for example, the closeness of an election affects voter turnout (e.g., Cox and Munger 1989; Diermeier and Van Mieghem 2005; GROFMAN, COLLET, and GRIFFIN 1998; Matsusaka 1993), and hence party choices may affect turnout.

I present here a model of voter turnout dynamics with endogeneous party choice that addresses *each* of these issues simultaneously. The model starts with the assumption that citizens can be grouped according to whether they voted and what party they voted for. A member of a group has a probability of switching to one of the other choices that depends on whether they voted and, if so, whether their party won. For example, a citizen who votes for a Democrat has some probability of voting for a Republican in the next election and some probability of shirking (staying home), and both probabilities depend on her choice to vote and on whether the Democrat won the election. This approach is inspired by Collins et al. (2009), which treats voter turnout as a fundamentally dynamic quantity and models changes in voter turnout directly rather than appealing to changes in exogeneous variables. (One should not read this as a rejection of the latter approach, which has contributed much to our understanding of voter turnout; rather, it reflects a desire to study a previously ignored aspect of voter turnout dynamics.) A problem with the Collins et al. approach, however, is its flatly wrong assumption that voters never change their party choices. This leads

them to predict that a party's turnout as a fraction of the total population — *party-level turnout* — must either always increase or always decrease, whereas party-level turnout often increases and then decreases or vice-versa, even if we focus only on long-term trends and ignore short-term fluctuations.

The model would not be useful if it was not helpful in explaining data; I show that the new model I present here explains a great deal. I study U.S. presidential election data between 1945 and 2000 and find that the new model with endogenous party choice explains 60 percent of the variance in party-level voter turnout, a substantial fraction considering the magnitude of some turnout changes during this time and a substantial improvement over other models, including the Collins et al. (2009) model. I also study a novel problem, voter registration dynamics, using data from California. Voter registration has not been studied nearly as much as voter turnout, and its dynamics seem never to have been studied. This is odd because of the importance of voter registration in the voting process. I show that the new model explains three-quarters of the variance in party-level voter registration in California between 1945 and 2000, a substantial fraction for any political science model and an improvement over plausible models.

Readers should not overlook the normative implications of this paper. When thinking about the effects of a particular factor in an election, the model and empirical results of this paper show that it is not appropriate to attribute the turnout results in one election simply to, for example, the mobilization efforts, closeness, or demographics of that single election. Rather, we must evaluate voter turnout in the context not only of historical turnout rates but also recent trends. For example, if turnout has been trending upward, another increase in turnout may not be the result of a broader sense of civic involvement or closeness or excitement — it may simply be the continuation of a trend that has little to do with any essentially cross-sectional variable. From an empirical point of view, this issue becomes important any time we study voter turnout cross-sectionally using data from many elections.

The broader purpose of this paper is to draw attention to the role of change in politics and methods for modeling change. Although we often do not think in these terms, change is fundamental to politics. Without changes, there would be no partisan realignment in the U.S. in the middle of the last century, there would be no changes in voting rights, political cycles, etc. In other areas of political science, the shift toward more democracies, the rise of terrorist organizations, and the

development of the European Union are also about change. Thus an aim of this paper is to show that one can model change — and how doing so produces empirically useful results — using voter turnout and registration as one example among many.

In the next section, I develop a new model of voter turnout dynamics with endogeneous party choice. The model predicts that party-level turnout, i.e., the fraction of the total population that votes for a given party may increase for a period of time and then decrease or vice-versa, a phenomenon that turns out to be a key feature of voter turnout data in the United States. I next show that the model fits U.S. presidential election data and California voter registration data between 1945 and 2000 better than one theoretically motivated model (Collins et al. 2009) and at least as well as several highly flexible but atheoretical empirical models.

## 2 A Turnout Model with Endogeneous Party Choice

In this section, I present a model of voter turnout dynamics with endogeneous party choice. Doing so addresses a key shortcoming of models of voter turnout that fix party choices (Collins et al. 2009; Diermeier and Van Mieghem 2005; Myerson 1998; Palfrey and Rosenthal 1983, 1985) and for the first time allows us to study the dynamics of party choice and voter turnout simultaneously, something other dynamical models of turnout and party choice fail to do (Achen 2002; Collins et al. 2009; Fiorina 1977). I cover mathematical details of the model in supplementary material and focus here on the conceptual motivations for the model, which I refer to for clarity as the *new model*.

The new model assumes that citizens decide whether to vote and, if so, which of two parties  $A$  and  $B$  to vote for in a series of elections  $t = 1, 2, 3, \dots$ . The set of possible decisions identifies three groups: those who vote for  $A$ , those who vote for  $B$ , and those who shirk, i.e., stay home. I label these groups  $A$ ,  $B$ , and  $S$ , respectively. Citizens may change their decisions between elections; doing so corresponds to moving between two groups. For example, a person who votes for party  $A$  in one election may switch to staying home in the next; this corresponds to switching from the  $A$ -voting group to the  $S$  group.

I assume that the probability of moving from one group to another depends only on whether a citizen voted and, if so, whether the party she voted for won the election. For these purposes, it is useful to relabel the groups as *winners*, i.e., winning voters, *losers*, and *shirkers*. The probability

of moving from one group to another is called a *failure rate* (see below for an explanation of this terminology). A failure rate  $f$  is, in the large-population limit, the fraction of a group who make a given set of choices who move to another group. Thus,  $f_{ws}$  and  $f_{ls}$  are the fractions of winners and losers who switch to shirking in the next election, respectively;  $f_{sw}$  and  $f_{sl}$  are the fractions of shirkers who switch to voting for the winning and losing party, respectively; and  $f_{wl}$  and  $f_{lw}$  are the fraction of winners who change to the losing party and vice versa in the next election. Under the assumption that  $A$  wins a given election, this framework leads to the model in Figure 1.

These two ideas — three groups of citizens, each corresponding to a particular choice, and probabilities of moving between these groups — are the key conceptual ideas in the model. This framework is distinct from past efforts to understand voter turnout, almost none of which approach turnout in terms of flows between groups within the whole population, and almost all of which focus either on individuals, as is the case in all cross-sectional empirical research, or two groups of citizens (voters and non-voters, possibly stratified by party), as is the case in formal models of turnout.

One way — though not the only way — to interpret the new model is in terms of satisficing (Simon 1955). Satisficing is the idea that people stick with a decision until it fails to satisfy them (hence the term failure rates). We can understand the decision to switch from one of the groups above to another in terms of satisficing. Winning voters may be dissatisfied because of the effort expended voting — “winning didn’t make a big enough difference to me to make voting worthwhile” — in which case they would not vote in the next election; or the party or candidate may dissatisfy them, in which case they may switch parties. Losing voters may be dissatisfied with voting itself, in which case they may stay home in the future, or with their party for losing, in which case they may switch parties to join the winning group. Finally, those who stayed home may be dissatisfied with the winning party and hence vote for the opposing party, or they may simply be dissatisfied with not voting — perhaps because of social pressure to vote — and vote for either party in the next election. Note again that satisficing is not the only possible interpretation; although I have not explored it, an interpretation in terms of subjective utility maximization may be possible.

[Figure 1 about here.]

To describe the model mathematically, we need to talk about the flows between groups, which

we can determine in terms of the failure rates. Let  $v_{A,t}$ ,  $v_{B,t}$ , and  $v_{S,t}$  be the fractions of the population who vote for  $A$ , vote for  $B$ , and shirk at election  $t$ . Suppose  $A$  wins (the equations are similar if  $B$  wins). Turnout evolves as

$$\begin{aligned} v_{A,t+1} - v_{A,t} &= -(f_{ws} + f_{wl})v_{A,t} + f_{lw}v_{B,t} + f_{sw}v_{S,t} \\ v_{B,t+1} - v_{B,t} &= f_{wl}v_{A,t} - (f_{ls} + f_{lw})v_{B,t} + f_{sl}v_{S,t} \\ v_{S,t+1} - v_{S,t} &= f_{ws}v_{A,t} + f_{ls}v_{B,t} - (f_{sw} + f_{sl})v_{S,t}. \end{aligned}$$

Each term in these equations represents a flow from one subgroup to another. For example, in the first equation, the term  $-(f_{ws} + f_{wl})v_{A,t}$  represents the total flow out of the  $A$ -voting population and into the shirking and  $B$ -voting populations. Likewise,  $f_{lw}v_{B,t}$  represents the flow from the  $B$ -voting population to the  $A$ -voting population, and  $f_{sw}v_{S,t}$  represents the flow from the shirking population to the  $A$ -voting population. Note that since  $v_{A,B,S}$  are turnout fractions,  $v_A + v_B + v_S = 1$ , and we can eliminate one equation:

$$\begin{aligned} v_{A,t+1} - v_{A,t} &= f_{sw} - (f_{ws} + f_{wl} + f_{sw})v_{A,t} + (f_{lw} - f_{sw})v_{B,t} \\ v_{B,t+1} - v_{B,t} &= f_{sl} + (f_{wl} - f_{sl})v_{A,t} - (f_{ls} + f_{lw} + f_{sl})v_{B,t}. \end{aligned} \tag{1}$$

There are two properties of this model that are important conceptually. First, the model that Equations 1 represent has the property that the change in one party's turnout between elections  $t$  and  $t + 1$  depends on *both* parties' turnout levels at election  $t$ . As a result, party-level turnout can change nonmonotonically, e.g., it may increase to a maximum and then decline as time progresses. I want to provide some intuition for this result. Consider Figure 1, and suppose that the net flow out of the  $A$ -voter population into the  $B$ -voter population is slow compared to the flow from the shirking population into the  $A$ -voter population. Then, the  $A$  voter population will build up quickly, but the population of shirkers will soon deplete, so that this flow will dry up quickly. Once that happens, the slower flow out of the  $A$ -voter population takes over, starting a slow but steady decline. One can think of this in terms of buckets of water: we pour the  $S$  bucket into  $A$ , but  $A$  leaks into  $B$ . Then,  $A$  fills up quickly, but once it is full it slowly drains away. Figure 2a presents an example of nonmonotonic changes in voter turnout at the party level.

Second, the equations above assumes that party  $A$  always wins, but in fact this equation may lead to a situation of *competitive elections*, in which each party frequently switches from winner to loser. Figure 2b depicts such a situation. If we are content with using computer simulations,

then it is fine to leave well-enough alone, but for the sake of understanding the model better, we would like to have a relatively simple expression for voter turnout as a function of time. To achieve this goal, I focus on the continuous-time limit of the model (cf. Collins et al. 2009).<sup>1</sup> It is vitally important to understand that the continuous-time model is qualitatively and quantitatively nearly identical to the discrete-time model.

[Figure 2 about here.]

In what follows, I present the continuous-time differential equations corresponding to Equation (1) (and the cases where  $B$  wins and where  $A$  and  $B$  are in a series of competitive elections), and I describe the solutions to these equations. My focus will be on developing intuition for these equations; I deal with the technical details in supplementary material. Because I wish to focus this paper on empirical questions — namely, how well does this model fit real turnout data in relative and absolute terms — I will not address certain issues, such as the steady-state turnout levels and the relationship between the underlying failure rates and the directly estimable equation parameters. The reason for this is that these parameters can not be estimated very well from the available data, and in any case there is no compelling theory of what these parameters should be and therefore no hypothesis to test.

*Turnout Equations.* In continuous time, Equation (1) becomes

$$\begin{aligned}\frac{dv_A}{dt} &= f_{sw} - (f_{ws} + f_{wl} + f_{sw})v_A(t) + (f_{lw} - f_{sw})v_B(t) \\ \frac{dv_B}{dt} &= f_{sl} + (f_{wl} - f_{sl})v_A(t) - (f_{ls} + f_{lw} + f_{sl})v_B(t).\end{aligned}\tag{2}$$

Recall that this equation covers the case in which  $A$  wins for some interval of time. The intuition is fairly straightforward. Equation (1) describes a difference in turnout between two elections. In continuous time, such a difference becomes a derivative. To describe cases in which  $B$  wins for some interval of time, we can just swap the labels  $A$  and  $B$  in Equation (2), i.e., change the role of  $A$  from winner to loser and similarly for  $B$ .

The more complicated case is when  $A$  and  $B$  are engaged in a series of competitive elections. In such elections, the two turnout levels remain close and parties go back and forth from winning to

---

<sup>1</sup>In addition to its mathematical tractability, it is also surprisingly easier to fit a continuous-time model to data than a discrete-time model. This observation has to do with the problem of writing the discrete-time model's solutions in an easily-estimable form.

losing, a process Collins et al. (2009) dubbed *chattering*. (See Figure 2b.) In the continuous-time limit, chattering is equivalent to tied turnout, so that the two parties' turnout levels are the same for as long as the tie persists. Thus,  $v_A(t) = v_B(t)$  for some interval of time and, by symmetry, each party must win half of the elections and lose the other half. (The symmetry argument, while correct, may be made more precise; see the supplementary material.) This means that the differential equation for  $v(t) = v_A(t) = v_B(t)$  is just the average of the two in Equation (2):

$$\frac{dv_A}{dt} = \frac{dv_B}{dt} = \frac{1}{2}(f_{sw} + f_{sl}) - \frac{1}{2}(f_{ws} + 2f_{sw} + 2f_{sl} + f_{ls})v_A(t). \quad (3)$$

*Solutions.* The evolution of turnout under the assumption that  $A$  always wins (i.e., the solutions of Equation (2)) are sums of exponential functions:

$$\begin{aligned} v_A(t) &= v_w + a_+ e^{-\lambda_+ t} + a_- e^{-\lambda_- t} \\ v_B(t) &= v_l + b_+ e^{-\lambda_+ t} + b_- e^{-\lambda_- t}, \end{aligned} \quad (4)$$

where the time constants  $\lambda_{\pm}$  are

$$\lambda_{\pm} = \frac{1}{2}(f_{ls} + f_{lw} + f_{sl} + f_{sw} + f_{wl} + f_{ws}) \pm \left[ (f_{ls} + f_{lw} + f_{sl} + f_{sw} + f_{wl} + f_{ws})^2 - 4[(f_{sl} + f_{sw})(f_{lw} + f_{wl}) + (f_{lw} + f_{sl})f_{ws} + f_{ls}(f_{sw} + f_{wl} + f_{ws})] \right]^{1/2};$$

the coefficients  $a_{\pm}$  and  $b_{\pm}$  are

$$\begin{aligned} a_{\pm} &= \frac{\lambda_{\pm} - f_{sl} - f_{lw} + f_{ls}}{f_{sl} - f_{wl}} b_{\pm} \\ b_{\pm} &= \pm \frac{(f_{wl} - f_{sl})(v_A(0) - v_w) + (\lambda_{\mp} - f_{sl} - f_{lw} + f_{ls})(v_B(0) - v_l)}{\lambda_- - \lambda_+}; \end{aligned}$$

and the steady states  $v_w$  and  $v_l$  are

$$\begin{aligned} v_w &= \frac{f_{ls}f_{sw} + f_{sl}f_{lw} + f_{lw}f_{sw}}{(f_{sl} + f_{sw})(f_{lw} + f_{wl}) + (f_{lw} + f_{sl})f_{ws} + f_{ls}(f_{sw} + f_{wl} + f_{ws})} \\ v_l &= \frac{f_{ws}f_{sl} + f_{sw}f_{wl} + f_{wl}f_{sl}}{(f_{sl} + f_{sw})(f_{lw} + f_{wl}) + (f_{lw} + f_{sl})f_{ws} + f_{ls}(f_{sw} + f_{wl} + f_{ws})}. \end{aligned}$$

Equation (4) — and in particular its functional form — is the key result of this paper, since it establishes the main prediction I will test. This prediction, that party-level turnout evolves over time as the sum of the same two exponential functions, distinguishes the model from Collins et al. (2009), the only other formal model of long-term voter turnout dynamics, as well as other plausible empirical models. The fact that in the model turnout evolves as the sum of two exponentials reflects



the fact that citizens can move between three groups and that  $v_A + v_B + v_S = 1$ . That is, the fact that we are dealing with turnout fractions places a restriction on the patterns of flow from one group to another so that there are two modes of change rather than three. Note that I present the parameters of Equation (4) for completeness; they turn out to be difficult to estimate precisely and are therefore not of particular interest for the mainly empirical purposes of this paper.

Aggregate turnout in the case that  $A$  wins for all  $t$  in some interval is

$$v(t) = v_w + v_l + (a_+ + b_+)e^{-\lambda_+ t} + (a_- + b_-)e^{-\lambda_- t},$$

This is significant because it has the same functional form as in the Collins et al. (2009) model, i.e., a sum of two exponential functions. Therefore, distinguishing the Collins et al. and new models using aggregate turnout data requires distinguishing parameters, which the authors found to have large standard errors because of a fairly small amount of data. Aggregate turnout is therefore not likely to be useful for comparing the two models.

I now consider tied turnout. Because of the additional constraint on tied turnout that  $v_A = v_B$ , tied turnout evolution (the solutions of Equation (3)) may have only one time constant:

$$v_A(t) = v_B(t) = v_{wl} + (v_A(0) - v_{wl})e^{-\lambda t}, \quad (5)$$

where  $v_{wl} = (f_{sw} + f_{sl})/(f_{ws} + 2f_{sw} + 2f_{sl} + f_{ls})$  and  $\lambda = (f_{ws} + 2f_{sw} + f_{ls} + 2f_{sl})$ .

Figure 3 presents some examples of turnout dynamics that can result. In first example, turnout ties after about 15 elections and remains tied forever. In the second example, one party always dominates. Note that in both cases party-level turnout may change nonmonotonically, a property that is not possible in Collins et al.'s turnout dynamics model, in which each party's turnout fraction evolves as a single exponential function of time. We will see that nonmonotonicity is a property the data require. In the new model, such nonmonotonic changes are a direct consequence of the coupling between the two parties' turnout levels.

[Figure 3 about here.]

### 3 Empirical Analysis

I address two main questions in this section. First, does the new model fit real data better than other models, and in particular does it fit better than the Collins et al. (2009) model, the only other

formal model of voter turnout dynamics? In other words, is a party’s turnout level — the fraction of the population who votes for that party’s candidate — better explained by a sum of exponential functions of time, or is a single exponential sufficient? More generally, is the sum-of-exponentials form a better description of the data than other plausible empirical models? Second, can the model’s applicability be extended to a new domain, voter registration? The choice to register with a party is similar to the choice to vote for a party, but no one has tested a model of registration dynamics.

I begin this section by discussing the empirical models I test and the standards I use to judge them. I then discuss the data sets — one covering U.S. presidential elections and one covering California voter registration — I use to test the models and the reasons I selected them. Finally, I present results of the analysis.

## Models, Comparisons, and Data

*Empirical Models.* The primary empirical goal of this paper is to determine whether the new model fits real election data better than the original other models. In order to do so, I now develop empirical models of party-level voter turnout, starting with the new model. The full new model is given by the solutions in Equations (4) and (5). However, these general solutions are fairly difficult to fit to data because of the possibility that ties may begin and end. As a practical matter, however, one can approximate each party’s turnout as the sum of two exponential functions:

$$\begin{aligned} v_A(t) &= a + a_+ e^{-\lambda_+ t} + a_- e^{-\lambda_- t} \\ v_B(t) &= b + b_+ e^{-\lambda_+ t} + b_- e^{-\lambda_- t}, \end{aligned} \tag{6}$$

where  $a$ ,  $b$ ,  $a_{\pm}$ , and  $b_{\pm}$  are parameters to be estimated. The approximation works because the differential equation for tied turnout is the mean of Equations (2), which cover turnout evolution when one party wins for some period of time. Therefore, to first order, tied turnout evolves as the average of the winner and loser’s turnout levels under the assumption that one party wins and one party loses each election. This approximation works best when the winner’s and loser’s turnout would remain close in Equations (2), which will happen, for example, if both turnout levels are already near the steady-state levels, as is the case in Figure 3a.

The other extant formal model of voter turnout dynamics is Collins et al. (2009). Under this model (with a similar approximation covering the tied-turnout case), each party’s turnout evolves

a single exponential function,

$$\begin{aligned} v_A(t) &= a_0 + a_1 e^{-\lambda_a t} \\ v_B(t) &= b_0 + b_1 e^{-\lambda_b t}, \end{aligned} \tag{7}$$

Note that this model nests within new model, Equation (6).

Although there are no theoretical justifications for other models, it is worth considering two standard empirical models of dynamics. First, I consider a polynomial model:

$$\begin{aligned} v_A(t) &= \sum_{n=0}^N a_n t^n \\ v_B(t) &= \sum_{n=0}^N b_n t^n, \end{aligned} \tag{8}$$

where  $T$  is the total number of years covered by the data. Second, I consider a Fourier or sum-of-sines model:

$$\begin{aligned} v_A(t) &= \sum_{n=0}^N a_n \sin(2\pi n t / T) \\ v_B(t) &= \sum_{n=0}^N b_n \sin(2\pi n t / T), \end{aligned} \tag{9}$$

where  $T$  is the total number of years covered by the data. In each of the preceding three models,  $a_n$  and  $b_n$  are parameters to be estimated.

*Comparisons and Statistical Tests.* I now turn to the matter of which comparisons to focus on and what standards to use in these comparisons. First, as I discussed earlier, the functional forms of the aggregate turnout functions for the new and Collins et al. models are the same, so we can not differentiate the models using aggregate turnout data. Therefore, I compare the two models using party-level data only.

Second, both the polynomial and Fourier models are extremely flexible in a way that the other two are not, and in fact the polynomial and Fourier models can be excellent approximations to the either of the other models. To see this, note that because they are linear in their parameters, the polynomial and Fourier models with  $N$  parameters will fit any  $N$  data points perfectly. The corresponding flexibility also means that they may be used to approximate any other function, and as we add parameters —  $N = 8$  in the tests that follow — the approximation becomes better and better. Neither the new model nor the Collins et al. model have this property, since they are not linear in their parameters.

For these reasons, I will consider it evidence for the new model if, using party-level turnout data, (1) the new model fits the data better than the original Collins et al. (2009) model and (2) the new model fits the data at least as well as polynomial and Fourier models with the same number of parameters. Since the Collins et al. model is nested within the new model, these two can be compared using a likelihood ratio test. The polynomial and Fourier models are not nested within the new model, so I use Vuong tests (Vuong 1989) to compare these models. The Vuong statistic is similar to a  $z$ -statistic, i.e., the difference between two parameters divided by the variance of this difference. To compute a Vuong statistic, one divides the difference between the log-likelihoods for two models by the difference's variance, which one constructs using individual data points' contributions to the likelihood. Like the  $z$ -statistic, the Vuong statistic is asymptotically normally distributed.

*Data.* Data one uses to test the model must meet several requirements. First, the data must be from a democratic political system in which two parties garner almost all the votes for some period of time, and as a practical matter this means that the model applies to the United States only. Other systems in which two parties dominate nonetheless have substantial third party turnout, e.g., in the United Kingdom third parties routinely reach 10 percent of the vote and sometimes much more. Thus, it is inappropriate to apply an explicitly-two party model to the United Kingdom and other such countries, which, unfortunately for this analysis, includes most democracies.

Second, it must be plausible that the model is stationary over the period of time considered, i.e., we must believe that the failure rates  $f$  are fairly constant over this time period. I therefore focus on the period between 1945 and 2000. 1945 signalled the end of a war, U.S. involvement in which was precipitated by an attack on U.S. soil involving significant casualties, preceded by a major economic downturn lasting more than a decade. 2000 signalled the beginning of a period of economic decline — according to Federal Reserve statistics, it was the first time since at least 1945 when household net worth declined<sup>2</sup> — followed shortly by the first attack on U.S. soil since 1941 with significant (civilian) casualties, an attack that precipitated wars in two theaters. Another candidate for an endpoint year is 1973, which Collins et al. (2009) used because it was the beginning of the OPEC

---

<sup>2</sup><http://www.federalreserve.gov/RELEASES/z1/Current/data.htm>, accessed December 7, 2009. Curiously, the Fed actually reports statistics on households and nonprofits rather than households alone, although it is unlikely that households experienced an economic upturn while nonprofits did not.

oil crisis. I disagree with that choice because, while gas lines were long, there was not an economic downturn of the same nature as that of the 1930s or 2000s, and the United States had not been attacked on its own soil.

I focus on two data sets that meet the requirements. The first is U.S. presidential election data, available from the Clerk of the House <sup>3</sup>. This data comprises information on vote totals for each party between 1920 and 2008. I construct turnout fractions using voting-age population, data on which is available from the U.S. Census Bureau <sup>4</sup>. I excluded 1964 from the data because there is a large and temporary spike in Democratic voting and a corresponding decline in Republican voting resulting no doubt from Kennedy's assassination the previous year. This does not seem to have had a lasting impact on turnout trends since in 1968 and 1972 there are substantial increases in Republican voting. Unfortunately, this is a difficult assertion to test. If we include 1964, Democratic turnout in particular is all over the place. Second, if we focus on the period 1964-2000, turnout mainly declines and all the models fit quite well, i.e., this period does not discriminate well between models. For the sake of full disclosure, I will discuss tests that include 1964 as well; the results are quantitatively different but reflect much the same pattern as the results that exclude 1964.

For comparison, I also report results of fitting the model to McDonald and Popkin's (2001) data, which computes voter turnout using an estimate of the voting-eligible — rather than voting-age — population between 1948 and 1996. While the parameter estimates differ, the main conclusions do not differ between voting-age and voting-eligible turnout data.

The second dataset is novel. The California Secretary of State maintains partisan registration data dating back to 1910 (Available at [http://www.sos.ca.gov/elections/elections\\_u.htm](http://www.sos.ca.gov/elections/elections_u.htm), accessed December 8, 2009). This dataset provides registration totals for the two major parties as well as the total voting-eligible population in every election since 1910. I compute registration fractions as the party's number of registered voters divided by the size of the voting-eligible population. For simplicity, I exclude special elections, e.g., the recent gubernatorial recall election, and focus on data from congressional and presidential elections only.

Partisan voter registration has not received much attention, and its dynamics seemingly even less. Most studies involving registration focus on it only as it pertains to voter turnout or some

---

<sup>3</sup>[http://clerk.house.gov/member\\_info/electionInfo/index.html](http://clerk.house.gov/member_info/electionInfo/index.html), accessed November 19, 2009

<sup>4</sup><http://www.census.gov/popest/archives/>, accessed November 19, 2009 and December 4, 2009

other issue (e.g., Ansolabehere and Konisky 2006; Highton 1997; Knack and White 2004; Nagler 1991), and there are none that concern registration dynamics (in the sense of this paper). These dynamics are, however, interesting because registration is typically, at least, a prerequisite for voter turnout and therefore registration dynamics are likely an important component of voter turnout dynamics. Furthermore, they form a more stable indication of voters' overall sentiments toward the parties than election returns, which are subject to greater fluctuations due to the high cost of changing one's registration relative to changing one's party choice in an election.

## Results

*U.S. presidential elections.* The first test concerns U.S. presidential elections between 1948 and 2000. I present the data and the (estimated) new model in Figure 4a, and I present results of likelihood ratio tests and Vuong tests in the first column of Table 1. These results indicate that (1) the new model performs much better than the Collins et al. (2009) model (log likelihood ratio 13.85,  $p < 10^{-8}$ ) and (2) the new model performs marginally better than both the polynomial or Fourier models (Vuong statistics 1.00 and 1.25,  $p = 0.16$  and  $0.11$ , respectively). To give a sense of the absolute fit of the models, I computed  $R^2$  values. For the new model,  $R^2 = 0.59$  compared with 0.31 for the Collins et al. model, 0.49 for the polynomial model, and 0.45 for the Fourier model. By the standards set out above, the new model is a better fit to U.S. presidential election data than other models.

The results are essentially unchanged if we instead estimate the models using the voting-eligible population to determine turnout rates (McDonald and Popkin 2001). The new model again performs better than the Collins et al. (2009) model (LLR 10.65,  $p < 10^{-8}$ ) and performs as well as the polynomial model (Vuong statistic 0.65,  $p = 0.26$ ) and Fourier model (Vuong statistic 0.73,  $p = 0.23$ ). The overall fit for the new model is also better than the other models:  $R^2 = 0.46$  compared with 0.39, 0.35, and 0.17 for the polynomial, Fourier, and Collins et al. models, respectively. Thus, using the voting-age or voting-eligible population yields the same conclusion: the new model with endogenous party choice performs better than the Collins et al. model and slightly better than the (highly flexible and theoretically unmotivated) polynomial and Fourier models.

If I include 1964 in the analysis of voting-age population turnout, the new model still comes out on top but less dramatically so.  $R^2$  for the new model is 0.40 compared with 0.23, 0.38, and

0.34 for the Collins et al., polynomial, and Fourier models, respectively. If we focus on the period 1964-2000 — under the argument that the failure rates changed in 1964 —  $R^2$  for the new model is 0.69 compared with 0.65, 0.69, and 0.62 for the Collins et al., polynomial, and Fourier models, respectively. Between 1948 and 2000, the new model still explains the data better than the Collins et al. model and at least as well as the polynomial and Fourier models. Between 1964 and 2000, the model does not differentiate well between the models but still performs somewhat better than the polynomial and Fourier models. Thus, while less dramatically than when I exclude 1964, these data still provide overall support for the new model.

[Figure 4 about here.]

*California Voter Registration.* I next consider California voter registration data between 1946 and 2000. Note that I included 1964 in this analysis. While the assassination of President Kennedy clearly had an impact on election returns in 1964, it is less obvious that it would have an impact on registration because it probably did not change citizens' fundamental partisan dispositions. Examining the data bears out this conclusion: unlike the presidential voting data, there is no sudden and dramatic change in registration in 1964.

I present the data and fitted new model in Figure 4b. The results of the analysis appear in the second column of Table 1. These results show that, again, the new model fits the data better than the Collins et al. (2009) model (LLR statistic 69.57,  $p < 10^{-8}$ ) as well as the polynomial and Fourier models (Vuong statistics  $-0.37$  and  $-0.82$ ,  $p = 0.64$  and  $0.57$ , respectively). For the new model,  $R^2 = 0.74$  compared with 0.73 for the polynomial model, 0.72 for the Fourier model, and 0.70 for the Collins et al. model.

[Table 1 about here.]

The standard I set out above for the new model with endogenous party choice was that it perform better than the original Collins et al. (2009) model and that it do at least as well as polynomial and Fourier models with the same number of parameters. The new model tested with California voter registration data meets both requirements. Tested with U.S. presidential election data, the model meets the first and somewhat exceeds the second requirement, performing better than the Collins et al. model and marginally better than both the polynomial and Fourier models

in the more than five-decade period between 1948 and 2000. The new model is, therefore, the most successful model of long-term turnout dynamics yet.

## 4 Conclusion

The aim of this paper was to develop and test a model of the dynamics of voter turnout with endogenous party choice and, in so doing, focus attention on the problem of directly explaining change in political behavior and politics more generally. The model I developed assumes that we can divide the population into subgroups of shirkers and voters for each of the two major parties and that flows between these groups explain the dynamics we observe in voter turnout. The present paper is the first to model the long-term time dynamics of party choice and voter turnout at the same time. I tested the model against the Collins et al. (2009) model and two sensible although theoretically unmotivated empirical models, a polynomial model and a Fourier or sum-of-sines model. The Collins et al. model is important because it makes a prediction about how the fraction of the population voting for a given party changes over time, and this constitutes a sort of null prediction about party choice, i.e., that people make the same party choices every time they vote. I found that the new model explained turnout and registration data better than the Collins et al. model and as well as and sometimes better than the other two, which, combined with the great flexibility and lack of theoretical content of the polynomial and Fourier models, provides substantial support for the new model.

There are a number of factors that I have not addressed, e.g., mobilization efforts. As with most studies of turnout, these are usually considered in a single election (e.g., Gerber and Green 2000), but it is plausible that mobilization efforts have their own long-term dynamics, e.g., mobilization efforts may increase when a party loses a series of elections or is engaged in competitive elections. In terms of the broader aim of this paper, such an effort focuses attention on how variables that have been treated in an exogeneous and cross-sectional fashion play a role in shaping long-term changes in political behavior. Other similar issues include economic dynamics and changes in the age distribution. I hope to model these and other aspects in the future.

Finally, the broader point of this paper is to show how a dynamical approach — an approach focused on change as a variable of interest — can provide valuable explanations for phenomenon



of interest to political scientists. In this paper, a simple approach based on examining how people move between groups — voters for different parties and shirkers — explained upwards of half of the variance in voter turnout and three-quarters of the variance in voter registration over time. By social science standards, the model presented here explains a great deal, and I have not even included mobilization or demographic effects, let alone differences between individual citizens (as cross-sectional studies do). As I have already pointed out, these factors are important, but so too is the lesson learned here: that dynamics are interesting in their own right and that we can learn a great deal by studying them directly.

## References

- Paul R. Abramson and John H. Aldrich. The decline in electoral participation in america. *American Political Science Review*, 76:502–521, 1982.
- Christopher Achen. Parental socialization and rational party identification. *Political Behavior*, 24(2):151–170, 2002.
- Stephen Ansolabehere and David M. Konisky. The introduction of voter registration and its effect on turnout. *Political Analysis*, 14:83–100, 2006.
- Thomas E. Cavanagh. Changes in american voter turnout, 1964-1976. *Political Science Quarterly*, 96:53–65, 1981.
- Nathan A. Collins, Sunil Kumar, and Jonathan Bendor. The adaptive dynamics of turnout. *Journal of Politics*, 71(2):457–472, 2009.
- Gary W. Cox and Michael C. Munger. Closeness, expenditures, and turnout in the 1982 u.s. house elections. *American Political Science Review*, 83:217–231, 1989.
- Daniel Diermeier and Jan A. Van Mieghem. Coordination and turnout in large electorates. Type-script, 2005.
- Timothy Feddersen and Alvaro Sandroni. A theory of participation in elections. *American Economic Review*, 96:1271–1282, 2006.

- Morris P. Fiorina. An outline for a model of party choice. *American Journal of Political Science*, 21:601–625, 1977.
- Alan S. Gerber and Donald P. Green. The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review*, 94:653–663, 2000.
- Alan S. Gerber, Donald P. Green, and Ron Shachar. Voting may be habit-forming: Evidence from a randomized field experiment. *American Journal of Political Science*, 47:540–550, 2003.
- BERNARD GROFMAN, CHRISTIAN COLLET, and ROBERT GRIFFIN. Analyzing the turnout-competition link with aggregate cross-sectional data. *Public Choice*, 95:233–246, 1998.
- Benjamin Highton. Easy registration and voter turnout. *Journal of Politics*, 59:565–575, 1997.
- Stephen Knack and James White. Election day registration and turnout inequality. *Political Behavior*, 22:29–44, 2004.
- John G. Matsusaka. Election closeness and voter turnout: Evidence from california ballot propositions. *Public Choice*, 76:313–334, 1993.
- Michael P. McDonald and Samuel L. Popkin. The myth of the vanishing voter. *American Political Science Review*, 95:963–974, 2001.
- Roger Myerson. Population uncertainty and poisson games. *International Journal of Game Theory*, 27:375–392, 1998.
- Jonathan Nagler. The effect of registration laws and education on U.S. voter turnout. *American Political Science Review*, 85:1393–1405, 1991.
- Thomas Palfrey and Howard Rosenthal. A strategic calculus of voting. *Public Choice*, 41:7–53, 1983.
- Thomas Palfrey and Howard Rosenthal. Voter participation and strategic uncertainty. *American Political Science Review*, 79:62–78, 1985.
- Steven J. Rosenstone and John Mark Hansen. *Mobilization, Participation, and Democracy in America*. Macmillan Publishing Company, New York, 1993.

Herbert A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69: 99–118, 1955.

Quang H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333, 1989.

COMPARISON OF EMPIRICAL MODELS.		
Test	U.S. Presidential Elections	Cal. Voter Registration
New vs. Collins et al. (LR)	13.85 (0.00)	69.57 (0.00)
New vs. Poly. (Vuong)	1.00 (0.16)	-0.37 (0.64)
New vs. Fourier (Vuong)	1.25 (0.11)	-0.82 (0.57)

Table 1: Results of statistical comparisons between the new, Collins et al. (2009), polynomial, and Fourier models using party-level and aggregate turnout in U.S. presidential elections between 1948 and 2000, inclusive. Each entry is either a likelihood ratio or Vuong statistic followed by its  $p$ -value in parentheses. Note that, using the methods described above, the new and Collins et al. models' predictions about aggregate turnout are indistinguishable.

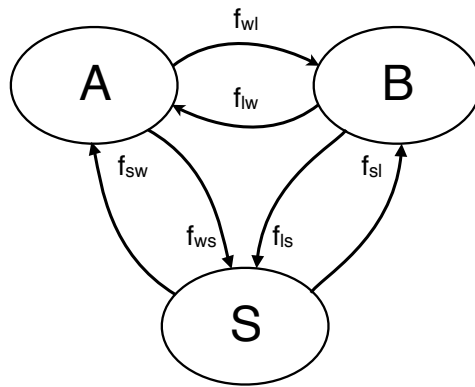
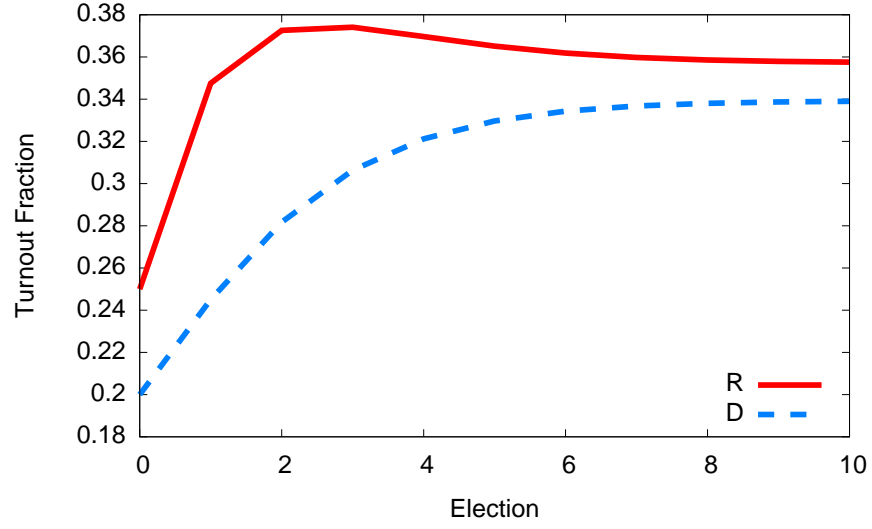
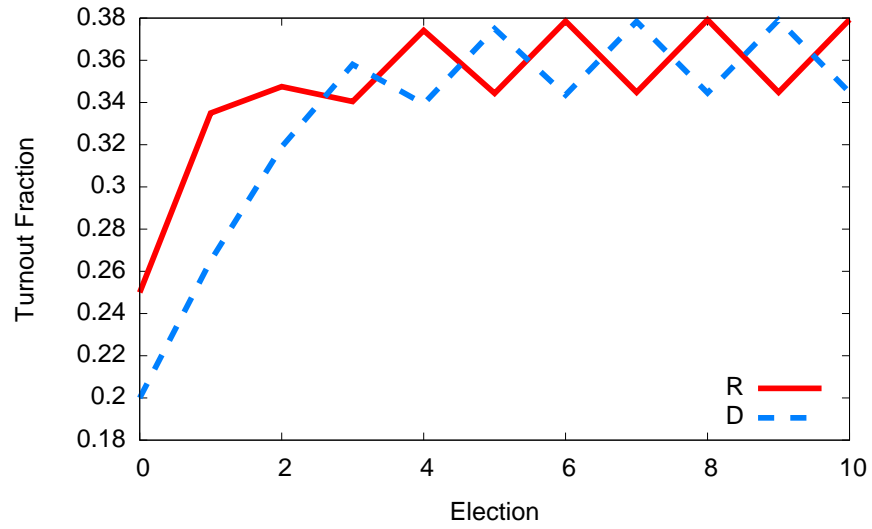


Figure 1: A depiction of the model with endogenous party choice.

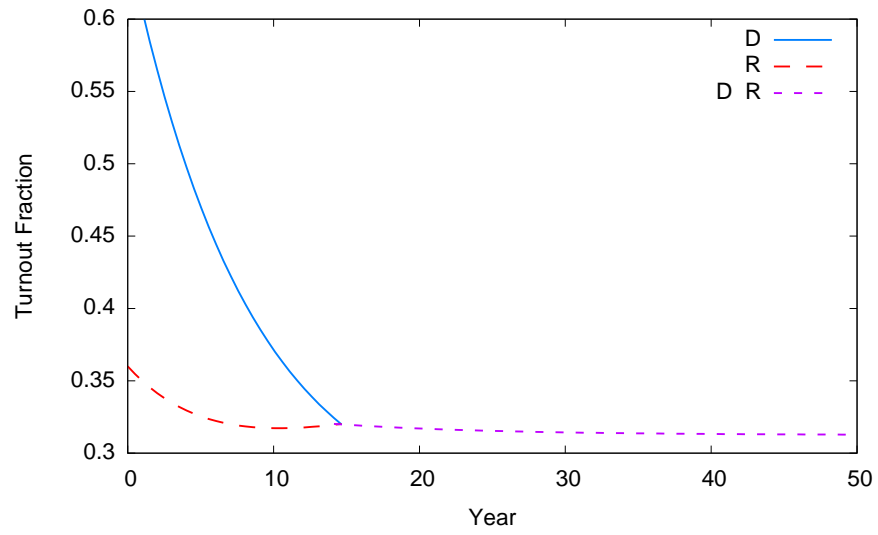


(a)

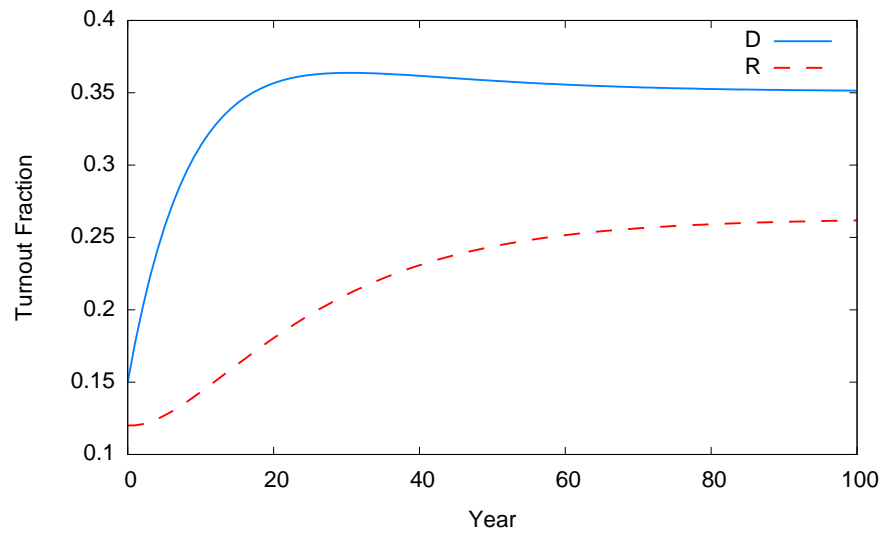


(b)

Figure 2: Examples of turnout dynamics in the model. (a) Party-level turnout can change non-monotonically. (b) Parties can switch from winner to loser rapidly, a situation that in continuous time corresponds to a tie.

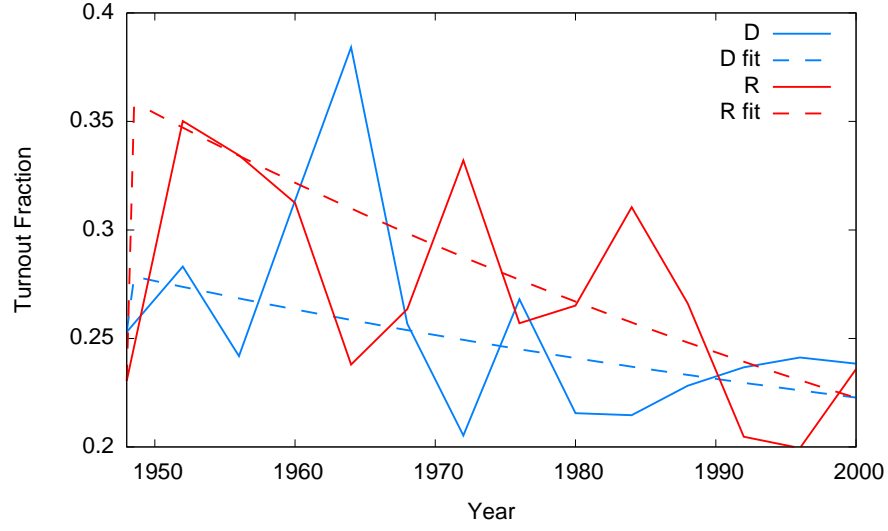


(a)

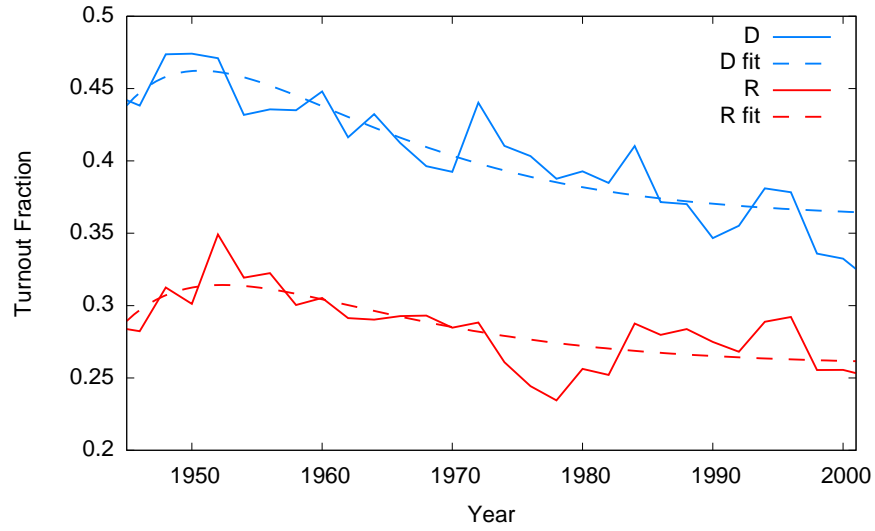


(b)

Figure 3: Examples of turnout evolution under the endogeneous party-choice model.



(a)



(b)

Figure 4: (a) U.S. presidential election data and fitted new model. 1964 data is included in this plot but not in the estimated model because, as the plot shows, it is an outlier associated with the assassination of John Kennedy. (b) California partisan voter registration data and fitted exponential model.