This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

Coordination and Turnout in Large Elections*

Daniel Diermeier and Jan A. Van Mieghem

This Version: March 17, 2008

Abstract

Competitive elections are frequently decided by which side can generate larger turnout on polling day. When decided on whether to turnout voters are assumed to balance the cost of going to the polls with the prospect of making a difference in the election outcome. We present a stochastic model of turnout where voters receive information about current turnout propensities of the voting populus through opinion polls. Voters then decide on whether to vote or not by maximizing their expected utility based on the (potentially noisy) polling information as well as the costs and beneftis of participation. We prove the existence of a unique limiting distribution for the process and show that even in large electorates substantial expected turnout is possible if voting factions are sufficiently similar in size. A key requirement for substantial turnout is that polls never provide precise feedback on the current state of the electorate. The effect of noise, however, is non-monotonic: no noise or too much noise results in vanishing turnout, while moderate noise may result in substantial turnout. Our model also can account for known empirical regularities about turnout identified in the political science literature.

1 Introduction

In contrast to theoretical economics,[1] models of bounded rationality and learning have only very recently been adopted in the study of politics.[2]. The study of mass elections in particular seems to offer a natural application of models of bounded rationality as the costs and benefits to voters are low (e.g. Aldrich 1993, Niemi and Weisberg 1993).

*Both authors are at the Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208-2009; d-diermeier@kellogg.northwestern.edu and Van-Mieghem@kellogg.northwestern.edu.

[1] For overviews see e.g. Fudenberg and Levine (1998), Young (1998), Blume (1997).

[2] For example, Bendor, Diermeier, and Ting (2003), Conley, Toosi and Wooders (2001), Dhillon and Demichelis (2002), Fey (1997), and Sieg and Schulz (1995).

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

To investigate the potential fruitfulness of this approach we apply the model to perhaps the most famous anomaly in the study of elections: Anthony Downs' (1957) "paradox of voting" or "turnout anomaly." Simply put, it states that nobody should vote in large electorates when there is even a small cost to voting because each voter's probability to decide an election outcome is vanishingly small. The cost of voting includes not only transportation etc. but also the opportunity cost of spending the time to go to the polls, stand in line and so forth. But, of course, citizens do participate, even in very large electorates.

The turnout anomaly has generated a large literature and many solution attempts.[3] Among the most influential are game-theoretic voting models (Palfrey and Rosenthal 1983, 1985; Myerson 1998). In these models, voters can vote for one of two candidates or stay home. There are two types of citizens with strictly opposed preferences. We will refer to them as Democrats and Republicans. Each voter of a given type strictly prefers the same candidate to win. Elections are decided by majority rule with some tie-breaking provision, such as a coin toss. All members of the winning type earn a payoff or benefit $b > 0$ for winning (whether or not they voted); losers get nothing (payoff = 0). Independent of the outcome, there is an additive and private cost of voting c, where $\frac{b}{2} > c > 0$.[4]

Because the intent to vote for the non-preferred candidate is dominated for each voter by the intent of voting for the preferred candidate, the relevant problem reduces to a *turnout game*[5], which simply involves the binary decision of whether to vote or stay home. Once voting is modeled as a game-theoretic (rather than decision-theoretic) problem, it cannot be an equilibrium for everyone to stay home, for then each voter could unilaterally decide the election by voting instead. Similarly, it cannot be a equilibrium for everybody to vote, unless the two teams are of the exact same size.[6]

It follows that all Nash-equilibria in the turnout problem involve the use of mixed strategies by at least some voters, except if the two teams are of the exact same size. This leads to an abundance of Nash-equilibria, some of them with surprisingly high turnout. However, all equilibria with non-trivial turnout (for example, less than 100 voters in an electorate of millions) in large elections are asymmetric and thus require precise coordination.[7] Because of these highly unrealistic coordination demands subsequent research has regarded these equilibria as implausible (Palfrey and Rosenthal 1985, Myerson 1998). The predominant approach has been to limit the amount of common knowledge present among voters by introducing some

[3] See Aldrich (1993) for an overview.

[4] Palfrey and Rosenthal (1983) also consider the (trivial) cases where $c \geq \frac{b}{2}$, and $c = 0$ as well as different tie-breaking rules.

[5] A turnout game can be interpreted as a competitive public goods game where each group's critical participation threshold is endogenously determined by the turnout-level of the competing group.

[6] To see why, suppose that there is one more Republican than Democrat in the population. If all eligible voters turn out, the Republican candidate wins, but then each Democrat might as well stay home.

[7] In these equilibria, the larger faction is divided into two sub-groups which play different (mixed) strategies. See Myerson (1998) for an instructive example.

form of uncertainty, for example with respect to payoffs (Palfrey and Rosenthal 1985) or the number of players (Myerson 1998). In these modified games all remaining equilibria have vanishing turnout.

Our purpose is to investigate the usefulness of behavioral models in the study of mass elections. We therefore use the same model of voter incentives used in the literature, i.e. we take the game form as given (i.e. as in Palfrey and Rosenthal (1983) or Myerson (1998)), but modify the solution concept from Nash (or Poisson) equilibrium to a stationary distribution of a stochastic dynamic process. So, voters do not receive utility from voting per se (e.g. because they did their "civic duty"), but only from the outcome of an election.[8] The process is specified as follows. In each period a randomly chosen voter observes a poll based on the population's last period vote propensities. Given that polls typically exhibit some sampling noise, the voter uses Bayes' rule to form expectations about the current configuration of play.[9] Based on this expectation, she then chooses a best response. Then again a voter is selected and so forth. This induces a Markov process governed by the best response dynamic and the random selection of voters. While our main purpose is to investigate whether simply modifying the behavioral model will lead to significant turnout, this approach also allows us to explicitly model how voters manage to coordinate their actions through (noisy) opinion polls.

We first prove the existence of a unique limiting distribution for our Markov process for arbitrary levels of information uncertainty (including the no-noise case with perfectly informative polls).[10] That is, voters in turnout games are able to coordinate implicitly through polls. We then investigate the qualitative properties of the limiting distribution. Our main result states that even in large electorates substantial expected turnout (up to 100%!) is possible if the faction sizes are close. A key requirement for substantial turnout is that polls never provide precise feedback on the current state of the electorate. Noisy polling introduces uncertainty about whether an actor is pivotal in determining the outcome of the election. However, if polling information becomes too noisy, turnout again drops to vanishingly small levels consistent with Palfrey and Rosenthal's (1995) and Myerson's (1998) findings. Thus, in contrast to existing game-theoretic models, the amount of uncertainty has a non-monotonic effect: moderate uncertainty may increase participation, while large uncertainty leads to vanishing turnout. Our model also confirms the usual empirical regularities about

[8] The alternative approach of changing the payoff structure of the game while leaving the solution concept alone has a long history beginning with Riker and Ordeshook's (1968) "d-term" to capture civic duty to a recent model of altruism in voting by Feddersen and Sandroni (2002).

[9] In a later section of the paper we also investigate the full information (no noise) case.

[10] The existence of a unique limiting distribution is a non-trivial result in the case of unperturbed best-response dynamics. See Blume (1995). As we will show below the result does not depend on the existence of noisy polls either.

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

	Strategic Uncertainty	**No Strategic Uncertainty**
Behavioral	Our model	e.g., Bendor, Diermeier and Ting (2003) Conley, Toosi and Wooders (2001) Dhillon and Demichelis (2002) Sieg and Schulz (1995)
Classical	-	e.g., Palfrey and Rosenthal (1983) Myerson (1998)

Figure 1 Our behavioral model of voting shows a crucial relationship between noise (in the information that voters receive) and levels of turnout.

turnout.[11] For example, turnout drops as the participation cost or the number of voters increase, or as factions become less equal. Turnout increases in the stakes of an election and the closeness of opinion polls. (Wolfinger and Rosenstone 1980, Hansen, Palfrey, and Rosenthal 1987, Nalebuff and Shachar 1999.)

These results may suggest that by simply relaxing the rationality assumptions on voters we generate turnout. Such a conclusion, however, would be incorrect. The existence of (moderately) noisy polling information is critical. Specifically, we show that in the case of a large, perfectly informed, electorate turnout will be zero unless factions are of exactly equal size. Thus, we recover the no-turnout anomaly in an even more pronounced form. That is, under perfect information our model yields negligibly small turnout fractions for large N, independent of the costs or benefits of participation. The stark contrast with the noisy informative polls shows the importance of uncertainty in turnout models and the subtlety in the effect of uncertainty.

The contribution of our model is thus two-fold, as illustrated by Figure 1:

The first contribution is methodological (and captured by the rows in Figure 1). We use behavioral game theory to study voting rather than classical game theory. There are very few other models that have explored this modeling approach[12]. Given the increasing importance of behavioral modeling in economic

[11] Since these regularities are well-accounted for by game-theoretic models, it is important for an assessment of our approach that our model preserves these theoretical advances.

[12] The ones most closely related to our paper are Dhillon and Demichelis (2002) and Bendor, Diermeier, and Ting (2003). Both papers were completed after the first version of this paper. In contrast to Dhillon and Demichelis our model can account for high turnout. Bendor, Diermeier, and Ting's model also implies high turnout but relies on a different behavioral approach, i.e. aspiration-based learning, and their results are largely computational. While both our model and the Bendor-Diermeier-Ting can account for high turnout, the respective behavioral dynamics and thus each model's predictions are very different. For example, in the Bendor-Diermeier-Ting model, voters may abstain even if they have a weakly dominant strategy to participate.

theory and the strong tradition of behavioral approaches in the empirical study of elections we believe that this approach is well-worth pursuing, especially in the study of voting in large elections.

The second contribution is explanatory (and represented by the columns in Figure 1). We show that polling noise has an important impact on turnout levels. Moreover, this effect is non-monotonic and non-trivially related to the cost-benefit ratio of voting. To the best of our knowledge, this is a completely new insight, not previously reported in neither the empirical nor the theoretical literature on elections. As the Figure indicates it is in principle possible to study polling noise in a classical framework, perhaps by using a global games approach (Morris and Shin 2003).[13] However, it is not clear how a global games framework would capture polling noise. Strategic uncertainty in the global games framework is usually derived from some uncertainty about the fundamentals of the game (e.g. payoffs). But in our case, there is no such uncertainty. Rather, polling noise only refers to uncertainty over the current vector of actors taken by the voters. It is thus an open question how to capture our insight in a global games setting.

Finally, we formally investigate Aldrich's conjecture that voting in large elections does not fully respond to incentives or rational choice as captured by cost-benefit calculations. While a rational-choice model cannot study such conjecture, our model can. Assuming perfectly-informative polls, we consider log-logistic choice (Blume 1993) instead of unperturbed best response. (We call this "action noise" to contrast with "information noise.") We show that little is gained by this modification. While substantial turnout may be expected to occur in the perturbed model, any such turnout is entirely driven by randomness in individual choice behavior, i.e. by voters that participate although their incentives would suggest that they should abstain.

A conclusion summarizes our findings and suggest further avenue for research. All proofs are relegated to the appendix, which also contains large N-approximations as well as computational properties of the model.

2 The Model

Suppose there are two types of voters in a population of size N: Democrats of size N_D and Republicans N_R, where $N = N_D + N_R$, and $N_R \geq N_D > 0$. We use k for individual voters and i, j for types of voters with i and j denoting different types unless otherwise noted. Each voter must choose an action $z \in \{0, 1\}$, where $z = 0$ means "abstaining." The state of the electorate at time t is given by $n^t = (n^t_D, n^t_R)$, where $n^t_i \leq N_i$ is the number of type i that is intending to vote at time t. Superscripts indicating time periods are dropped

This cannot happen in our model. Also, polling information plays no role in the Bendor-Diermeier-Ting model. Both models could easily be tested and compared in laboratory experiments.

[13] We thank an anonymous referee for suggesting this approach.

unless necessary. In the usual fashion we write n_i^{-k} for the number of voters of type i without counting a specific voter k. Similarly we write n^{-k} to denote (n_D^{-k}, n_R) if k is a Democrat and (n_D, n_R^{-k}) if k is a Republican.

We assume the same payoff specification as Palfrey and Rosenthal[14] (1983) and Myerson (1998): Each member of the team that turns out more voters receives a payoff of b while the losers each receive 0. In addition there is a private cost c to participating independent of the election outcome. Ties are decided by a fair coin-toss. Throughout the analysis we assume $0 < c < \frac{b}{2}$.

For a given configuration n, a voter k of type i's payoff can then be summarized in the following matrix.

Payoff Matrix	$n_i^{-k} < n_j - 1$	$n_i^{-k} = n_j - 1$	$n_i^{-k} = n_j$	$n_i^{-k} \geq n_j + 1$
$u_k(z = 0; n^{-k})$	0	0	$b/2$	b
$u_k(z = 1; n^{-k})$	$0 - c$	$b/2 - c$	$b - c$	$b - c$

Note that if any type's faction (not counting k) is behind by more than one vote (column 1) or ahead by at least one vote (column 4), type k's decision on whether to participate is irrelevant for the outcome of an election. In columns 2 and 3, on the other hand, voter k is pivotal.

Rather than specifying the Nash or Poisson equilibria for this payoff specification we define a stochastic process where voters adjust their actions in response to the current agent configuration. The process consists of a *selection rule* and an *action rule*. According to our selection rule, in each period t one specific agent out of N is randomly chosen with probability $1/N$.[15] That agent will choose an action according to the immediate expected return given her expectations about current play. In the next period, again a player is chosen at random (with replacement), and so forth. The selection probabilities are denoted as follows. It is convenient to group the agents by type: a voter of faction i that currently chooses action z is referred to as type (i, z). The probability that the randomly chosen agent is of type (i, z) is denoted by p_{iz}. For example, $p_{D0} = (N_D - n_D)/N$.

Agents condition their behavior on the current configuration of play in the population. Voters do not observe the participation decision of all other voters, but receive their information about current voting behavior from noisy opinion polls. To capture this intuition we assume that each selected voter observes a noisy signal $\widetilde{n}$ of n, written as $\widetilde{n}(n)$.[16] Published polls, for example, typically include a polling error of 3%,

[14] We only consider their model where ties are broken by a fair coin-toss.

[15] For simplicity, we assume that revisions are made each period. All results, however, continue to hold in continuous time when the time between revisions is exponentially distributed. Clearly, the time between revisions can be arbitrarily small, so that the limiting distribution can be reached "quickly" in real time.

[16] Given that we assume the fractions N_D and N_R are know, the noisy signal $\widetilde{n}$ of n is equivalent to a poll reporting the percentages of members of each party intending to vote.

which roughly means that $n = \widetilde{n} \pm 3\%$. Given $\widetilde{n}$, an agent updates his beliefs about the state of the system and chooses a best response given the signal.[17]

Formally, our stochastic model defines a discrete-time, discrete-state Markov process: we have a family of random variables $\{n^t : t \in \mathbb{N}\}$ where n^t assumes values on the state space $S_d \times S_r$ and where $S_d = \{0, 1, 2, \dots, N_D\}$ and $S_r = \{0, 1, 2, \dots, N_R\}$. Given our stationary selection rule and since signals are only a function of n, and not explicitly of time, we have a Markov chain with stationary transition probabilities, which are summarized in a transition matrix P. Because at most one player can change her action in a given period the Markov chain is a two-dimensional *birth-death process*. A "birth" corresponds to an agent changing her action from abstention to participation, while a "death" corresponds to a voting agent deciding now to stay home.

The transition matrix P is completely defined by the selection rule and the action rule, which specifies the probability that a selected agent chooses a given action. To derive the action probabilities for given a noisy signal $\widetilde{n}$, agents now must *estimate* the true state n given the polling information $\widetilde{n}$ and then, based on that information, decide whether to vote. It follows immediately that voters will only vote if they expect to be pivotal. Formally, the expected utility for a type $(i, 0)$ is:

$$\begin{aligned} \mathrm{E}u_i(z &= 0|\widetilde{n}(n), \text{type } (i,0)) = \frac{b}{2}\Pr(n_i = n_j|\widetilde{n}(n)) + b\Pr(n_i \geq n_j + 1|\widetilde{n}(n)), \\ \mathrm{E}u_i(z &= 1|\widetilde{n}(n), \text{type } (i,0)) = \frac{b}{2}\Pr(n_i = n_j - 1|\widetilde{n}(n)) + b\Pr(n_i \geq n_j|\widetilde{n}(n)) - c. \end{aligned}$$

Hence, voting $(z = 1)$ is preferred, iff:

$$\mathrm{E}u_i(z = 0|\widetilde{n}(n), \text{type } (i,0)) \leq \mathrm{E}u_i(z = 1|\widetilde{n}(n), \text{type } (i,0)),$$

or:

$$\text{type } (i,0) \text{ votes } \Leftrightarrow \frac{b}{2}\Pr(n_i = n_j - 1|\widetilde{n}(n)) + \frac{b}{2}\Pr(n_i = n_j|\widetilde{n}(n)) \geq c.$$

Thus, a type-0 voter participates if she expects to create a tie or victory. Similarly, for type $(i, 1)$:

$$\text{type } (i,1) \text{ votes } \Leftrightarrow \frac{b}{2}\Pr(n_i = n_j|\widetilde{n}(n)) + \frac{b}{2}\Pr(n_i = n_j + 1|\widetilde{n}(n)) \geq c,$$

who votes if she expects to sustain a tie or a victory.

We can now partition the rectangular state-space into transition zones: the *birth-zone* for type i is the set of states where a type i agent finds it optimal to participate. In other words, for any state in the birth-zone a type i agent is pivotal and the cost/benefit ratio is sufficiently small. The *death-zone* for type i, on the

[17] While in reality the sampling errors are normally distributed, we will assume a simpler setting of uniformly distributed noise. This is without loss of generality: it yields simple formulas, while the insights extend to normally distributed noise.

other hand, is the state-space subset where it is optimal for a type i agent to stay home. Given that only the cost-benefit *ratio* impacts the decisions, it is convenient to denote the ratio c/b by ξ, where $0 < \xi < \frac{1}{2}$. Formally then:

$$\text{Birth-zone } i: \text{ type } (i,0) \text{ votes} \quad \Leftrightarrow \quad \Pr(n_i = n_j - 1|\widetilde{n}(n)) + \Pr(n_i = n_j|\widetilde{n}(n)) \geq 2\xi.$$

$$\text{Death-zone } i: \text{ type } (i,1) \text{ abstains} \quad \Leftrightarrow \quad \Pr(n_i = n_j|\widetilde{n}(n)) + \Pr(n_i = n_j + 1|\widetilde{n}(n)) < 2\xi.$$

Defining $\Delta n = n_D - n_R$, yields the *general pivot equations*:

$$\begin{aligned} &\text{Type } D \text{ birth} \Leftrightarrow \Pr\left(\Delta n \in \{-1,0\}|\widetilde{n}(n)\right) \geq 2\xi, && \text{Type } R \text{ birth} \Leftrightarrow \Pr\left(\Delta n \in \{0,1\}|\widetilde{n}(n)\right) \geq 2\xi, \\ &\text{Type } D \text{ death} \Leftrightarrow \Pr\left(\Delta n \in \{0,1\}|\widetilde{n}(n)\right) < 2\xi, && \text{Type } R \text{ death} \Leftrightarrow \Pr\left(\Delta n \in \{-1,0\}|\widetilde{n}(n)\right) < 2\xi. \end{aligned} \tag{1}$$

Clearly, the exact form of the birth and death zones depends on how noisy the poll is.

Ultimately, our goal is to study the long-run behavior of our process to estimate turnout fractions. In our model all states communicate so that the Markov chain is regular (Taylor and Karlin 1994; p.171) and hence has a unique limiting distribution denoted by the vector π, where

$$\pi_n = \lim_{t\to\infty} \Pr\{n^t = n|n^0\},$$

and $\pi_n > 0$ for all $n \in S$ and is independent of the starting state n^0. It can easily be shown that π is the unique distribution that solves $\pi = \pi P$. These equations are called the *global balance equations* because, rearranging $\pi_i = \sum_j \pi_j P_{ji}$, yields

$$(1 - P_{ii})\,\pi_i = \sum_{j\neq i} \pi_j P_{ji},$$

which can be interpreted as saying that the probability "flow" out of state i must equal the probability flow into state i. To study turnout we must characterize the limiting distribution π. This requires specifying the transition matrix P and solving $\pi = \pi P$ for π.[18] Like the birth and death zones, P depends on the details of the polling technology.

3 Discussion of Modeling Assumptions

Since our model differs significantly from game-theoretic models of elections it may be useful to discuss some of the basic modeling assumptions and how they relate to game-theoretic models.

[18] We are slightly abusing notation here. Note that there are N states in the Markov chain and we represent the limiting distribution in a $1 \times N$ vector indexed by n. Here n is a scalar that enumerates all states, whereas anywhere else in the text, n is the pair (n_D, n_R). (To stay consistent, we could represent π as a matrix but that would turn the matrix equation $\pi = \pi P$ into a tensor equation.)

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

In contrast to game-theoretic approaches our voters are only boundedly rational. First, there is no assumption about common knowledge concerning the game or other players' rationality. Second, voters are assumed to be "myopic," i.e., they do not take into account that their current action may have some (small) influence on future decisions by themselves or by other actors. On the other hand, we do assume that voters are able to use Bayes' rule to update their beliefs about the current state of the system and that they choose best responses based on these beliefs.

This approach may leave us exposed to criticisms both from adherents and opponents of (classical) game-theory. On the one hand one may argue that voters do take into account future consequences of their actions and thus will respond to polls strategically. Of course, if such foresight were perfect we would be back in the case of fully rational agents, now playing a stochastic game. While this may be an interesting project in its own right, the point of our paper is to investigate the case of boundedly rational agents.[19] On the other hand one may argue, that the rationality requirements on voters assumed by our model are still far too demanding. Specifically, we assume that a boundedly rational voter can use Bayes' rule to compute her probability of being pivotal every time she updates her intention to vote[20]. While this certainly is a weakness of the model, it is worth noting that *any* (Markovian) model of belief formation could be substituted to capture biases in updating probabilities, e.g. the tendency of experimental subjects to over-estimate their probability of being pivotal (Tversky and Quattrone 1988). Of course, such non-Bayesian assumptions would lead to even higher turnout. In this sense our goal is to consider the "hardest" case for generating turnout. Similarly, other decision rules could be used instead of best-response, such as the case of randomly perturbed best-response investigated below.[21]

Other objections may be voiced towards our selection rule and the frequency and availability of polls. One may argue that it is not plausible to assume that in each period only one voter is selected to change his vote. However, given that the discrete periods can be chosen to be arbitrarily small and random, the time scale is arbitrary. The sequential selection assumes that voters may change their ballots as many times as they wish before a given date, but only one voter can cast her ballot at a given time. The model assumption that polls are updated after every casted vote and made available to all voters is a more drastic departure from existing elections where polling data is provided less frequently. Justifications of our model can take two forms. First, our model is an stylized version of the recent tendency to provide more frequent polls.

[19] Alternatively, one may wish to consider models of limited foresight. In this case, our model would constitute the base-line case of no foresight.

[20] Given that voters are picked at random, on average, a given agent only updates her intention after N polls (and not after each poll).

[21] For a rather different approach consider Bendor, Diermeier, and Ting (2003). They investigate a generalization of the Bush-Mosteller model of stimulus response learning as applied to the turnout game.

(Daily polling updates during the 2004 elections were not uncommon.) This model then is an analytically tractable simplification of reality. The second justification is more forward looking to a perhaps not too-distant future where polls may be held electronically on a computer network in a decentralized manner.[22] Consider for example the following instance of our model where agents vote in an election using the internet. Whenever a voter wishes to cast her ballot and logs into the site, she automatically is informed about the current vote totals as calculated by the computer.[23] Polls are calculated instantaneously. Because voters act independently in a decentralized manner, their votes will be cast asynchronously and in random order. Such a scenario would almost literally implement our model. It also means that the predictions of our model are directly testable in laboratory setting. Therefore, literally interpreted, our model faithfully describes online voting with random log-in. As a model of "real world" elections it assumes that there are sufficiently many polls.

Finally, the model assumes that payoffs are constant over the duration of the stochastic process. This is not reasonable if we consider a sequence of elections with different candidates. The model is thus better interpreted as capturing one election with many opinion polls. Since voters are assumed to be myopic, it does not matter whether voters cast their ballot in an actual election or are asked how they would vote "if the election were held today." Following this interpretation, our model assumes that the candidates' positions are fixed over a sequence of opinion polls. While this may be objectionable in the case of office-motivated candidates it seems entirely reasonable if the alternatives are e.g. referenda.[24]

Our analysis seeks to characterize the limiting distribution of the process. While (classical) game-theory uses an equilibrium approach as its predictive concept, the prediction of our model is a probability distribution which is unique, as we will show. (While our model thus does not face a problem of equilibrium multiplicity, its predictions are probabilistic.) In analogy to comparative statics analysis we can now vary the parameters of the model, e.g. the cost of participation, relative faction sizes etc., to study the effects on the limiting distribution and (hopefully) account for known empirical regularities.

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

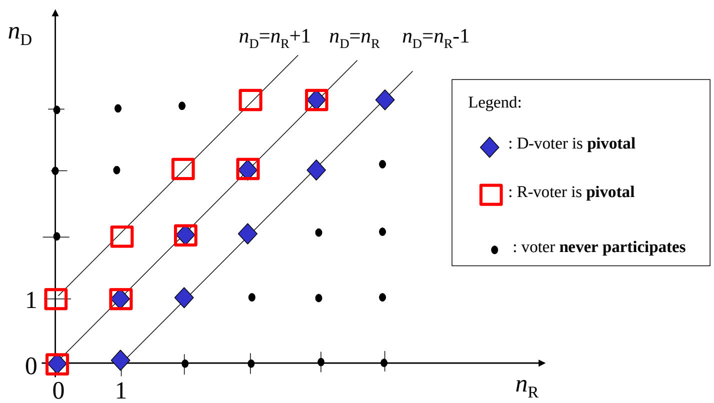

Figure 2 In the absence of noise, only states on the three diagonal lines may make a voter pivotal and induce her to participate in the election.

4 Example

To illustrate the model, first consider an environment without noise. Imagine a Democrat who observes a perfectly-informative poll $n = (n_D, n_R)$. The Democrat will vote and participate in the election if and only if she can either break the tie ($n_D = n_R$) or make the tie ($n_D = n_R - 1$) and if the cost-benefit ratio c/b is less than $\frac{1}{2}$. If $c/b \leq 1/2$, the states where a Democrat will participate are represented by a diamond in Figure 2 and fall on two adjacent diagonal lines. The Democrat will never participate when the poll falls outside those two diagonal lines. A similar reasoning applies to a Republican, his pivoting states being represented by squares in Figure 2. The end result being that only states in the strip $n_D = n_R \pm 1$, which is the birth-zone, may lead an agent to participate.

Now consider the case that the poll is noisy so that an agent only observes the noisy signal $\widetilde{n}$ of n. Agents must now estimate the true state n given the polling information $\widetilde{n}$ and then, based on that information, decide whether to vote. Clearly, that inference depends on the statistics of the noise. For concreteness and analytic tractability, assume a simple additive noise model so that $n = \widetilde{n} + \epsilon$ where the noise ϵ is uniformly distributed over a square-grid of size $[-\varepsilon, \varepsilon]^2$. (Obviously, ϵ is an integer random variable given that both

[22] In this application one may prefer the continuous time formulation discussed above where the waiting time between log-ins is exponentially distributed.

[23] Whether this information is only for a sub-sample of the ballots cast and whether it is rounded would determine the level of noise in this online poll.

[24] A more complicated model would also allow candidates to modify their positions in response to polls. For a game-theoretic analysis of this approach see Ledyard (1984).

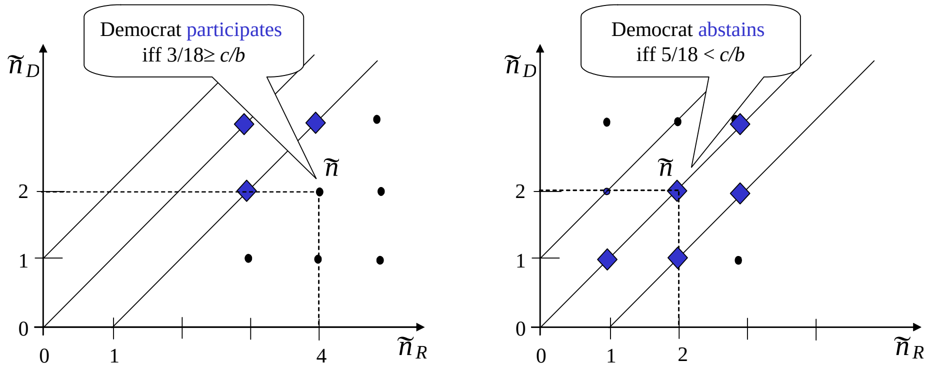

Figure 3 The presence of noise in the polls may increase (left) or decrease (right) turnout relative to the case without noise.

the true and the noisy states are integers.) If $\varepsilon = 1$, this means that

$$\widetilde{n}(n) = (n_D + \epsilon_D, n_R + \epsilon_R) \qquad \text{with probability } p_\varepsilon = \frac{1}{9} \qquad \forall \epsilon_i \in \{-1, 0, 1\}.$$

Bayes' rules is now particularly simple and yields the distribution of the inferred state from the noisy poll:

$$n(\widetilde{n}) = (\widetilde{n}_D - \epsilon_D, \widetilde{n}_R - \epsilon_R) \qquad \text{with probability } p_\varepsilon = \frac{1}{9} \qquad \forall \epsilon_i \in \{-1, 0, 1\}.$$

Again consider a Democrat but this time she observes a noisy poll $\widetilde{n} = (\widetilde{n}_D, \widetilde{n}_R)$. The Democrat will vote and participate in the election if and only if her expected reward exceeds her cost. Imagine that she observes the poll $\widetilde{n} = (2, 4)$. If the poll were perfectly informative the democrat would abstain. After all, adding her vote would still lead to a Republican victory. But if the poll has uniform additive noise with $\varepsilon = 1$, this polling result may imply three pivotal events (indicated by diamonds): $\Pr(n_D = n_R | \widetilde{n} = (2, 4)) = \frac{1}{9}$ and $\Pr(n_D = n_R - 1 | \widetilde{n} = (2, 4)) = \frac{2}{9}$. According to the general pivot equations (1), the Democrat will participate if $c/b \leq \frac{3}{18}$. This shows that, relative to the case without noise, the presence of noise can increase turnout (and widen the birth-zone strip) provided the cost-benefit ratio is small enough. (Note for participation to be optimal the cost-benefit ratio must be *lower* than under complete information. If the cost-benefit ratio is sufficiently low, however, then with uncertainty there are more states where a voter expects to be pivotal.)

In general, however, the impact of noise is ambiguous. Imagine our Democrat observes the poll $\widetilde{n} = (2, 2)$ as illustrated in the right of Figure 3. If the poll were perfectly informative the democrat would participate if $c/b < 1/2$. But if the poll has uniform additive noise with $\varepsilon = 1$, this polling result may imply five pivotal events (indicated by diamonds): $\Pr(n_D = n_R | \widetilde{n} = (2, 4)) = \frac{3}{9}$ and $\Pr(n_D = n_R - 1 | \widetilde{n} = (2, 4)) = \frac{2}{9}$. The general pivot equations show that the Democrat now will abstain if $c/b > \frac{5}{18}$. This shows that, relative to the case without noise, the presence of noise can also decrease turnout (and narrow the birth-zone strip) provided the cost-benefit ratio is sufficiently high ($\frac{5}{18} < c/b < \frac{1}{2}$ in our example).

The next section will show that depending on the cost-benefit ratio c/b and the noise level ϵ, the birth-zone, and thus turnout, may increase *or* decrease.

5 General Results

Consider the additive noise model $n = \widetilde{n} + \epsilon$, where the noise ϵ is uniformly distributed over a square-grid of size $[-\varepsilon, \varepsilon]^2$. (Given that n and are integers, ϵ is a discrete random variable and ε is a positive integer.) As illustrated by the example, the advantage of the additive model is that Bayes' rule becomes analytically tractable, especially so with uniformly distributed noise. Specifically[25]:

$$\widetilde{n}(n) = (n_D + \epsilon_D, n_R + \epsilon_R) \qquad \text{with probability } p_\varepsilon = \frac{1}{(1+2\varepsilon)^2} \qquad \forall \epsilon_i \in \{-\varepsilon, -\varepsilon+1, \ldots, \varepsilon\}. \tag{2}$$

Equivalently, inverting:

$$n(\widetilde{n}) = (\widetilde{n}_D - \epsilon_D, \widetilde{n}_R - \epsilon_R) \qquad \text{with probability } p_\varepsilon \qquad \forall \epsilon_i \in \{-\varepsilon, -\varepsilon+1, \ldots, \varepsilon\}.$$

To fix ideas consider a type-$(i, 0)$ voter and a given noisy poll $\tilde{n}$. The agent now must estimate the true state n given the polling information $\widetilde{n}$ and then, based on that information, decide whether to vote. She will vote if and only if

$$\Pr(n_i = n_j - 1|\widetilde{n}) + \Pr(n_i = n_j|\widetilde{n}) \geq 2\xi.$$

Recall our example where a voter receives a signal of the form $\tilde{n}_i = \tilde{n}_j$. Given uniform noise with $\epsilon = 1$ this implies $\Pr(n_i = n_j|\tilde{n}_i = \tilde{n}_j) = \frac{3}{9}$ and $\Pr(n_i = n_j - 1|\tilde{n}_i = \tilde{n}_j) = \frac{2}{9}$. Hence, the voter will participate if $\frac{5}{18} \geq \xi$.

Note for participation to be optimal the cost-benefit ratio ξ must be *lower* than under complete information (because $\Pr(n_i = n_j - 1|\widetilde{n}) + \Pr(n_i = n_j|\widetilde{n}) < 1$). If the cost-benefit ratio is sufficiently low, however, then with uncertainty there are more events where a voter expects to be pivotal. For example, in the case of an erroneous poll at $\tilde{n}_i = \tilde{n}_j + 1$ a voter will still participate if $\frac{5}{18} \geq \xi$. So depending on the cost-benefit ratio ξ and the level of noise, as indicated by ε, the birth-zone, and thus turnout, may increase *or* decrease, as shown by the following proposition:

Proposition 1 *With uniform additive polling noise over* $[-\varepsilon, \varepsilon]^2$*, the pivot equation* $\Pr(\Delta n(\widetilde{n}) \in \{-1, 0\}|\widetilde{n}) \geq$

[25] To be more realistic, one may add the boundary condition $\widetilde{n} \geq 0$. While this would slightly change the estimates of n near the boundary of the state space, it does not alter any of our conclusions.

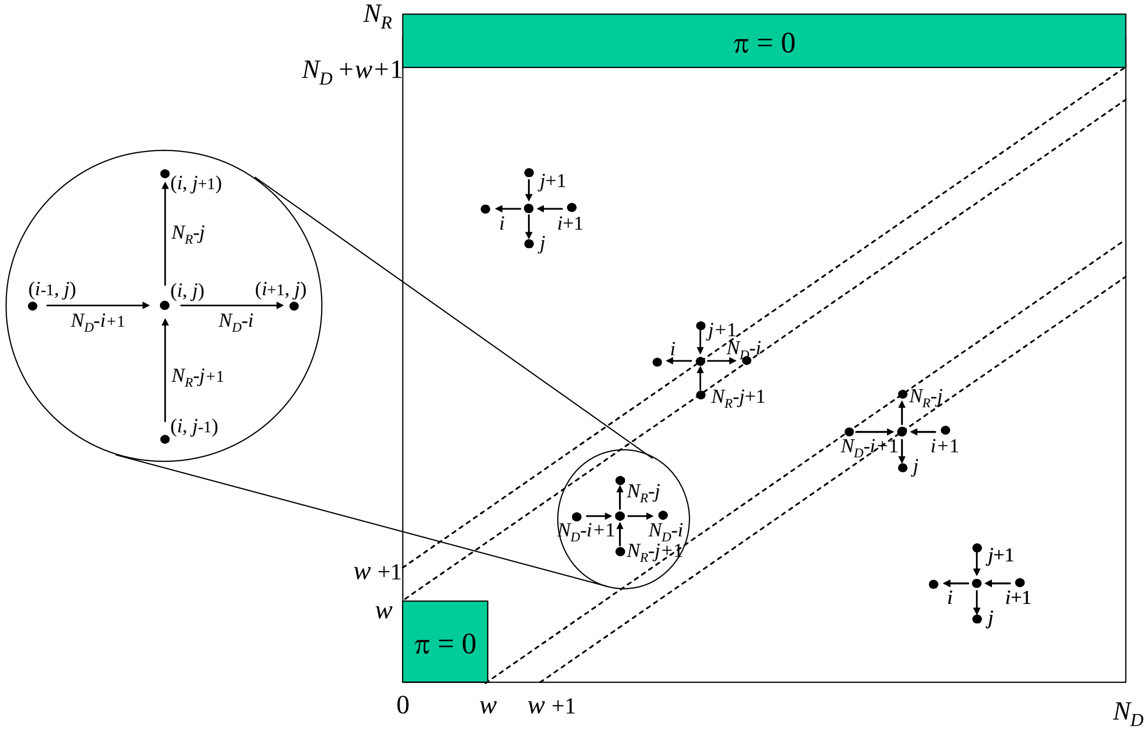

Figure 4 The state space is partitioned in a "birth zone" around the diagonal and two death zones. The specific state transition probabilities (relative to N) are shown, with those within the birth strip magnified.

2ξ *defines a birth-zone of "width"* $w(\xi, \varepsilon)$*:*

$$\textit{Type D birth} \Leftrightarrow \Delta\widetilde{n} \in [-w-1, w] \quad \textit{Type R birth} \Leftrightarrow \Delta\widetilde{n} \in [-w, w+1]$$

$$\textit{Type D death} \Leftrightarrow \Delta\widetilde{n} \notin [-w, w+1] \quad \textit{Type R death} \Leftrightarrow \Delta\widetilde{n} \notin [-w-1, w],$$

where

$$w(\xi, \varepsilon) = \lfloor 2\varepsilon - \xi(1+2\varepsilon)^2 + \tfrac{1}{2} \rfloor. \tag{3}$$

The proposition shows that the impact of the three model parameters–cost c, benefit b, and level of noise ε–on the equilibrium turnout statistics can be captured by one single parameter w. Moreover, this parameter has a direct graphical interpretation: it is the width of the strip in the state space where a randomly selected voter will choose to participate in the election. The width of this birth-zone is clearly represented in Figure 4. The figure also shows the transition probabilities at each state in the transition matrix P by arrows. The up and right transitions inside the strip around the diagonal represent births, while the down or left transitions are deaths. Consider, for example, the magnified state (i, j) in the birth zone. Recall that this means that currently i Democrats and j Republicans are intending to vote. Now select an agent at random. With probability $(N_D - i)/N$ this randomly selected agent is a Democrat who is changing his decision from abstain to participate, representing a birth from state (i, j) to $(i+1, j)$. Similarly, with probability $(N_R - j)/N$ the randomly selected agent is a Republican changing his decision from abstain to participate, representing a birth in the vertical direction. Considering the states $(i-1, j)$ and $(i, j-1)$ similarly yield the birth probabilities $(N_D - (i-1))/N$ and $(N_R - (j-1))/N$. (The figure only shows the transition probabilities relative to N.)

The transition probabilities also directly show that all states communicate so that the Markov chain is regular and a unique limiting distribution π exists. Furthermore, the states $n_R > N_D + w + 1$ and the states $(n_D, n_R) < (w, w)$ are transient and the limiting distribution is zero for those states. The latter is consistent with the fact that $(0, 0)$ can never be a Nash equilibrium in a game-theoretic turnout model. In addition, if $N_R \leq N_D + w + 1$, then the state (N_D, N_R) is absorbing and everyone will vote (with probability one). Proposition 2 summarizes these findings:

Proposition 2 *A unique limiting distribution exists and solves $\pi = \pi P$. Moreover, π is zero at the states $(n_D, n_R) < (w, w)$ and $n_R > N_D + w + 1$. If $N_R \leq N_D + w + 1$, then $\pi(N_D, N_R) = 1$ so that the expected turnout is* $(100\%, 100\%)$.

Clearly, the limiting expected turnout is minimal when the birth zone is minimal $(w = 0)$ but is monotonically increasing in the width of w. Given that the width parameter w captures all system dynamics for fixed faction sizes, it only remains to analyze how w changes as function of ξ and ϵ to assess the effects of the cost/benefit ratio and uncertainty on expected turnout.

Corollary 1 *The birth width w decreases linearly in the cost-benefit ratio of voting ξ, but is concave in the level of noise ε. Specifically, turnout is minimal ($w = 0$) for large cost-benefit ratios (if $\xi \geq \overline{\xi}(\varepsilon)$), in the absence of noise or with a large amount of noise (if $\varepsilon = 0$ or if $\varepsilon \geq \overline{\varepsilon}(\xi)$), while turnout is maximal ($w^* = \varepsilon$) for the intermediate level of noise $\varepsilon^*(\xi)$, where*

$$\begin{aligned} \overline{\xi}(\varepsilon) &= \frac{4\varepsilon + 1}{2\left(1 + 2\varepsilon\right)^2}, \\ \overline{\varepsilon}(\xi) &= \frac{1 - 2\xi + \sqrt{(1 - 2\xi)}}{4\xi}, \\ \varepsilon^*(\xi) &= \frac{1 - 2\xi}{4\xi}. \end{aligned}$$

The corollary is summarized in Figure 5, which shows a contour map of the parameter w that summarizes the impact of both cost/benefit (ξ) and noise (ε). Recall, high values of the width w imply high turnout. Somewhat surprisingly, while small levels of noise increase turnout, high levels of noise decrease it. This finding is in contrast with the results from game-theoretic models where the introduction of uncertainty destroys the high-turnout equilibria (Palfrey and Rosenthal 1985, Myerson 1998).

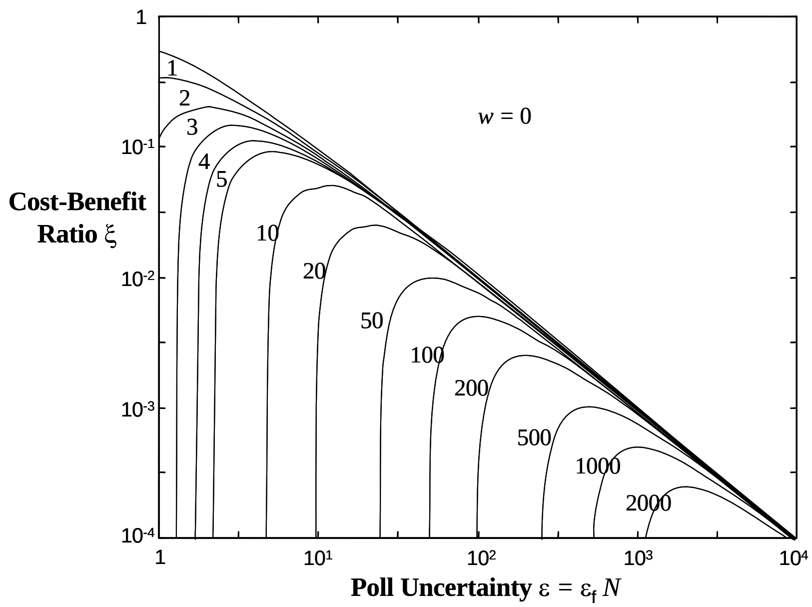

Figure 5 Contour graph of the width parameter w as a function of the cost-benefit ratio $\xi = c/b$ and the poll uncertainty $\varepsilon = \varepsilon_f N$.

6 Discussion of General Results

Proposition 2 constitutes perhaps the most striking result of our analysis: our behavioral approach can account for significant, even universal, turnout. Moreover, this finding is not limited to a knife-edge case, but holds for a range of parameter values as long as factions are "close" in size, depending on w. To see how expected turnout depend on w consider the following numerical example. We simulated the expected turnout in our model for various values of w for a total electorate of 1 million voters[26] with 48% democrats and 52% republicans. The expected turnout together with sample points are reported in Figure 6. Clearly turnout increases rapidly for small values of w, after which growth slows to an almost linear rate, until it picks up again as we approach $w = 40000$. At $w = 40,000$, we have that $N_R = N_D + w$ so that turnout is 100% for both parties. (Interestingly, the smaller party has larger proportional turnout for $w < 20,000$, while the reverse is true for larger w.)

While it is reassuring that the model can support high turnout, one may also be interested in assessing the quantitative improvement over existing models. To assess this magnitude, consider an example by

[26] For such large electorates the limiting distribution π can no longer be calculated exactly (the linear system $\pi = \pi P$ has 480×520 million unknowns). The time dynamics of the Markov chain, however, can easily be simulated. For each w, we simulated three sample paths, each with 10 million time periods. As the graph shows, the simulation error is remarkably small. The computational properties of the model are discussed in an appendix.

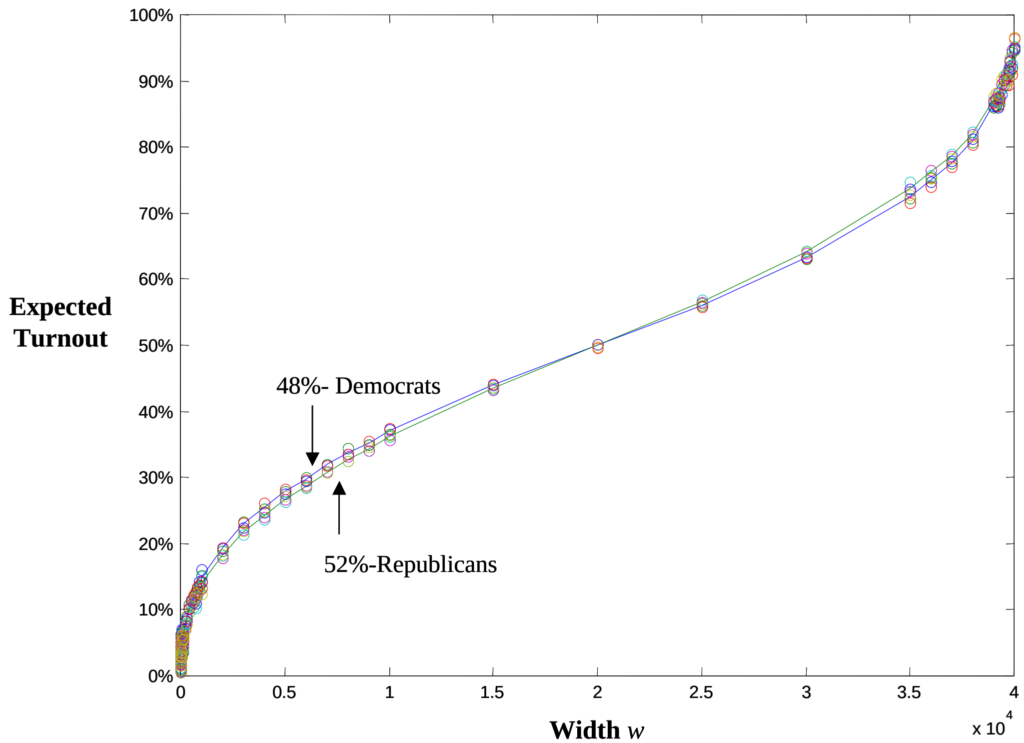

Figure 6 The expected turnout fractions as a function of the width w of the birth zone for an electorate of size 1 million with 48% Democrats and 52% Republicans.

Myerson (1998), which was constructed to demonstrate the strikingly low expected turnout predicted by game-theoretic models. In his case the voting factions are assumed to be very dissimilar ($N_D = 1$ million and $N_R = 2$ million with a cost-benefit ratio of 0.05). Myerson shows that in the unique Poisson voting equilibrium expected total turnout is about 64(!). For our model, highest expected turnout occurs for largest value of w. The corollary and Figure 5 show that the highest w for $c/b = 0.05$ is $w^* = 10$ for a rather low polling noise level of $\varepsilon^*(\xi) = 9.5$, which corresponds to an polling noise level of about 0.001%. Nevertheless, such little amount of noise is critical and results in an expected turnout in our model[27] of $(1.05\% \pm 0.26\%, 0.53 \pm 0.12\%)$, which means that about ten thousand voters of each party are expected to vote.[28]

One possible interpretation of this discrepancy is suggested by the stochastic assumptions. In our ap-

[27] The expected turnout was obtained through dynamic simulation of 10 sample paths, each simulated during 20 million time periods. We report averages together with 95% confidence intervals.

[28] Recall from Figure 6 that in the more realistic case of similarly-sized factions substantial turnout is possible even for considerably larger noise terms. The point of the example is to demonstrate that even in a case designed to show vanishing turnout, adding noise can substantially increase expected participation.

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

proach, randomness is introduced through noisy polls, not through uncertainty about parameters of the game form. Moreover, Myerson assumes a Poisson structure to model uncertainty, which has a large relative amount of uncertainty (coefficient of variation = 1). Such large variability would drive w down to zero in our model, resulting in minimal turnout. This highlights the subtle yet crucial impact of uncertainty in turnout models. Note, however, that our approach also replaces Nash equilibrium by the limiting distribution of our stochastic model. While it would be desirable to separate these two dimensions and, perhaps, construct a (classical) game-theoretic model with noisy polls, conceptually it is not clear how one could capture noisy polls in (Nash or Bayesian) equilibrium. Below we take one step in this direction and discuss the case of perfectly informative polls. This establishes a direct comparison between Palfrey and Rosenthal (1983) and our approach. One of the key insights of this comparison is the critical impact of noisy polls who serve as the "catalyst" for generating substantial turnout.

While much of the discussion of the turnout anomaly has focused on its troubling point prediction of vanishing turnout, the comparative static properties of game-theoretic models have explained empirical regularities remarkably well. (See Palfrey and Rosenthal (1983) and Hansen, Palfrey, and Rosenthal (1987) for a discussion.) It is thus important that our model can account for these regularities equally well. This is indeed the case. Turnout decreases in the cost of participation (because w decreases), but increases in the stakes of the election (because w increases)[29], and of course, the closeness of the race as reported in the opinion poll (Hansen, Palfrey, and Rosenthal 1987, Wolfinger and Rosenstone 1980, Nalebuff and Shachar 1999).[30] Note that in contrast to some game-theoretic models (e.g. Palfrey and Rosenthal) these predictions are probabilistic, but unique.

Much of the theoretical work on elections (including turnout) has relied on limit arguments as N goes to infinity. While such an approach seems justified given the intended application to large elections, it is important to know whether large turnout can occur for large electorates even if "in the limit" it vanishes. In other words, if substantial turnout can occur for even 100 million voters then a result of vanishing turnout in the limit is much less problematic.[31] However, such an analysis is usually absent in game-theoretic models.[32]

In our model turnout depends on the relationship between noise and the cost-benefit ratio. As the

[29] Participation in national elections is higher than in state or local elections.

[30] It is worth pointing out that turnout may be substantially higher if voters vote on many elections simultaneously. For example, in presidential elections voters also vote on House elections, and perhaps on Senate elections, referenda etc. If the marginal cost of filling out an additional ballot is small compared to the cost of going to the polls, then our model suggests that participation in all elections may be driven by the election with the largest w, leading to substantially larger turnout. We like to thank Ken Shepsle for suggesting this conjecture.

[31] We wish to thank John Ledyard for suggesting this interpretation.

[32] An exception is Hansen, Palfrey, and Rosenthal (1987).

population size increases, substantial turnout requires higher noise and lower cost/benefit ratios[33]. The critical question then is how fast the cost/benefit changes in N.[34] The corollary allows us to answer this question. Consider a fixed relative amount of noise, i.e. $\varepsilon_f = \varepsilon/N$ is constant. Then, we have that[35]:

$$\overline{\xi}(\varepsilon_f) = \frac{4\varepsilon_f N + 1}{2\left(1 + 2\varepsilon_f N\right)^2} = O(\frac{1}{2\varepsilon_f N}),$$

so that a substantial turnout with large population size requires that the cost-benefit ratio ξ decreases inversely in N. That is, e.g. for a polling noise level $\varepsilon_f = 3\%$, substantial turnout requires that $\xi(N) \leq \left(1 + \frac{3}{50}N\right)^{-1}$.

To see how binding this constraint is, consider an electorate with $N = 3\,\text{million}$. In this example a polling error of 1% yields $\varepsilon = 3 \times 10^4$. Hence, the cost/benefit ratio should be less than 10^{-4} to yield values of w substantially larger than 1, which is required for substantial turnout.

7 The Impact of Polling Noise in Large Electorates

The intended domain of applications for our model certainly is large elections. Therefore, we now derive a continuum approximation for large population size N which greatly simplifies the analysis and allows us to investigate the role of noisy polls in large electorates.

Consider the fractional state descriptor:

$$x_i = \frac{n_i}{N_i} \qquad \text{and} \qquad \alpha_i = \frac{N_i}{N} \qquad \text{and} \qquad w_f = \frac{w}{N}.$$

The state space for x is a discrete grid or subset of the unit square. The birth zones become, slightly abusing notation,

$$S(w_f) = \{x \in [0,1]^2 : x_i N_i \in \{0, 1, ..., N_i\} \text{ and } \alpha_R x_R - \alpha_D x_D \in [-w_f - \frac{1}{N}, w_f]\}.$$

In this section, we consider the approximation where x is considered a continuous state variable on the unit square, which formally obtains as the limit for $N \to \infty$. Similarly, we denote the continuous extension of $\pi(s)$ by $p(x)$. To avoid trivialities, we assume $0 \leq w_f < \Delta\alpha = \alpha_R - \alpha_D = \frac{N_R - N_D}{N}$. Hence, as above, we

[33] While polling noise and N are easily determinable, the measurement of c and b is a difficult, perhaps insolvable, empirical problem. In their study of Oregon school board referenda, Hansen, Palfrey and Rosenthal (1987) structurally estimate the cost of participation. Of course, that estimate critically depends on the underlying game-theoretic model of turnout.

[34] Alternatively, substantial turnout is possible if the stakes in large elections are substantially higher than in small elections.

[35] The notation $O(f(x))$ describes the behavior for large x. Formally, $O(f(x))$ denotes any function $g(x)$ such that $\lim_{x\to\infty} g(x)/f(x) = 1$. Informally, it means that for large x, $O(f(x)) \simeq f(x)$.

know that:

$$\begin{aligned} x &\in [0, w_f]^2 \text{ is transient } \Rightarrow p(x) = 0, \\ x_R &> \frac{\alpha_D + w_f}{\alpha_R} \text{ is transient } \Rightarrow p(x) = 0. \end{aligned}$$

Proposition 3 *The limiting distribution $\pi(n)$ for large population sizes ($N \to \infty$) tends to the probability density function $p(x)$, where $x_i = n_i/N_i$. The density p solves the following partial differential equations:*

$$\begin{aligned} \textit{inside the birth strip, p solves } PDE_1(x) &: (1 - x_D)\frac{\partial p}{\partial x_D} + (1 - x_R)\frac{\partial p}{\partial x_R} = 2p, \\ \textit{inside the death zone, p solves } PDE_2(x) &: x_D\frac{\partial p}{\partial x_D} + x_R\frac{\partial p}{\partial x_R} = -2p. \end{aligned}$$

Thus, $p(1 - x)$ is homogeneous of degree -2 inside the birth strip and $p(x)$ is homogeneous of degree -2 in the death zone.

The PDEs' boundary conditions are too complex to derive a closed form solution for the general case. The PDE formulation, however, does yield additional insight on the most likely turnout and on the impact of noise.

The most likely turnout correspond to the state where the probability density p reaches on extremum. Given that an interior extremum requires $\frac{\partial p}{\partial x_i} = 0$, the PDEs directly yield the following corollary:

Corollary 2 *In the large population limit, the limiting density p cannot attain an extremum in the interior of the birth or death zones. Hence, the most likely outcome must be on either the upper or lower strip boundary $\alpha_R x_R - \alpha_D x_D = \pm w_f$.*

So, elections must be close: the model predicts that the most likely turnout in large electorates is $n_D = n_R \pm w$, regardless whether the two factions are of similar size ($N_D \approx N_R$) or not ($N_D \gg N_R$).[36] The cost-benefit ratio and the level of noise determine (via the width parameter w) how close the elections will be. (The prediction that elections will be "close within noise tolerances" is in agreement with the example shown in Figure 6, even though it concerns a finite population size.)

Besides predicting the most likely turnout, the continuum approximation also allows us to investigate the impact of noise in large electorates. To contrast the results under strategic uncertainty (Corollary 2), consider the special case of perfectly informative polls where $\widetilde{n}(n) = n$. This case corresponds to the minimal width birth-zone: $w = 0$. So, the pivot probabilities are either one or zero. Notice that this case corresponds to Blume's (1995) best-response dynamic as applied to the turnout game.[37]

[36] This result is in stark contrast with the outcome that would obtain under mandatory voting!

[37] In general, the analysis of best-response dynamics even in simple 2×2 games may be highly non-trivial. See Blume (1995) for details.

Given that $0 < 2\xi < 1$, the general pivot equations simplify to

$$\text{Type } D \text{ birth} \Leftrightarrow \Delta n \in \{-1, 0\} \quad \text{Type } R \text{ birth} \Leftrightarrow \Delta n \in \{0, 1\}$$
$$\text{Type } D \text{ death} \Leftrightarrow \Delta n \notin \{0, 1\} \quad \text{Type } R \text{ death} \Leftrightarrow \Delta n \notin \{-1, 0\}.$$

Notice that the pivot equations are independent of ξ. This corresponds to the following matrix

Best-Response Action Probabilities	$n_i < n_j - 1$	$n_i = n_j - 1$	$n_i = n_j$	$n_i = n_j + 1$	$n_i > n_j + 1$
Type $(i, 0)$: $z = 0$	1	0	0	1	1
Type $(i, 0)$: $z = 1$	0	1	1	0	0
Type $(i, 1)$: $z = 0$	1	1	0	0	1
Type $(i, 1)$: $z = 1$	0	0	1	1	0

Even though the action rule is deterministic, the selection rule induces stochasticity in the state transitions. De-conditioning on types through the selection rule allows us to map the best response action probabilities into the state transition probability matrix yields:

Best-Response Transition Matrix	to $(n_i + 1, n_j)$	$(n_i - 1, n_j)$	$(n_i, n_j + 1)$	$(n_i, n_j - 1)$	n
from n with $n_i < n_j - 1$	0	p_{i1}	0	p_{j1}	$1 - p_{i1} - p_{j1}$
from n with $n_i = n_j - 1$	p_{i0}	p_{i1}	0	0	$1 - p_{i0} - p_{i1}$
from n with $n_i = n_j$	p_{i0}	0	p_{j0}	0	$1 - p_{i0} - p_{j0}$
from n with $n_i = n_j + 1$	0	0	p_{j0}	p_{j1}	$1 - p_{j0} - p_{j1}$
from n with $n_i > n_j + 1$	0	p_{i1}	0	p_{j1}	$1 - p_{i1} - p_{j1}$

The unique limiting distribution π can now be found by solving the linear system of equations $\pi = P\pi$ given by the global balance equations. Applying Proposition 3 we can show the following:

Proposition 4 *As the size of the electorate grows ($N = N_D + N_R \to \infty$) while the fractions $\alpha_i = N_i/N$ remain constant, the limiting distribution of turnout fractions with a perfectly informative poll converges to zero everywhere except for a probability-one mass point at* (0%, 0%) *if $N_D \neq N_R$ or at* (100%, 100%) *if $N_D = N_R$.*

Thus, in the absence of noise, voters in large electorates will (almost surely) coordinate on a state with zero turnout level, unless we have the knife-edge case of *exactly equal* factions.[38] This result obtains in the absence of uncertainty and is purely driven by the explicit coordination device. Hence, we recover the vanishing turnout result *even if voters act in a myopic fashion.* Moreover, the implication of vanishing turnout occurs in an even sharper form since in contrast to the multiplicity of equilibria in the Palfrey-Rosenthal model, the prediction is unique. In addition, in our model there is no analogue to the mixed

[38] Recall that in the case of exactly equal factions there is a Nash-equilibrium in pure strategies with full turnout.

strategy equilibria in the game-theoretic model or the asymmetric high-turnout equilibria found in Palfrey and Rosenthal (1983). We thus conclude that simply shifting from a fully rational to a boundedly rational model cannot resolve the turnout problem in large electorates whereas introducing (moderate) polling noise can.

8 The Impact of Action Noise versus Incentives

In a recent survey paper Aldrich (1993) has suggested that voting does not fully respond to incentives or rational choice as captured by cost-benefit calculations. In a traditional rational choice model this distinction cannot be modeled. Using a stochastic approach, however, we can investigate this concern by looking at a model where actions are driven both by randomness and by incentives. For concreteness, we assume that polls are perfectly-informative but that actions are subject both to "action noise" and incentives.

We introduce a parameter $\beta \geq 0$ that measures the relative impact of incentives versus randomness to actions, where larger values of β mean that incentives become more important in an agent's voting decisions. Specifically, consider the case of log-logistic choice[39] and let $p^{\beta}(z|n_t^{-k})$ denote the conditional probability that in period $t+1$ agent k will play action z given that the current configuration of play is n_t. Then the log-linear choice rule is given by:

$$p^{\beta}(z|n_t^{-k}) = \frac{\exp[\beta u(z; n_t^{-k})]}{\sum_{z' \in Z} \exp[\beta u(z'; n_t^{-k})]},$$

It is equivalent to the assumption that the pair-wise probability ratios of choosing actions are proportional to the respective pay-off differences.[40] A low β corresponds to the case where a participation decision is not much influenced by the incentives specified in the model, in agreement with Aldrich's suggestion. For $\beta = 0$ choice is completely random. That is, for all possible configurations, a voter will play each action with probability 1/2. As β increases, the utility differences become more important in determining a voter's decision. For $\beta \to \infty$, log-linear choice converges to a distribution that puts positive probability only on best-responses to n_t^{-k}.[41]

[39] See Blume (1997), and Young (1998) for overviews of alternative choice models.

[40] Alternatively, this rule can be interpreted as a random utility model (e.g. McFadden 1973). In the latter interpretation, rather than specifying that agents have fixed incentives, utilities are assumed to vary randomly according to a given probability distribution with a fixed mean. Given these incentives agents choose optimal actions. This interpretation is equally suitable for a model of voting, since the (perceived) benefits and costs of participating may well vary substantially over time. Turnout is notoriously affected by bad weather, for instance.

[41] For different (i.e. technical) reasons, the existing literature has used perturbed best response as the action rule (Foster and Young 1990, Blume 1993, Kandori, Mailath, and Rob 1993, Young 1993). Using perturbed best response ensures the

In the log-logistic model the action probabilities are given by the following matrix:

Log-Logistic Action Probabilities	$n_i < n_j - 1$	$n_i = n_j - 1$	$n_i = n_j$	$n_i = n_j + 1$	$n_i > n_j + 1$
Type $(i,0)$: $z = 0$	$\frac{1}{1+e^{-\beta c}}$	$\frac{1}{1+e^{\beta(0.5b-c)}}$	$\frac{1}{1+e^{\beta(0.5b-c)}}$	$\frac{1}{1+e^{-\beta c}}$	$\frac{1}{1+e^{-\beta c}}$
Type $(i,0)$: $z = 1$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$
Type $(i,1)$: $z = 0$	$\frac{1}{1+e^{-\beta c}}$	$\frac{1}{1+e^{-\beta c}}$	$\frac{1}{1+e^{\beta(0.5b-c)}}$	$\frac{1}{1+e^{\beta(0.5b-c)}}$	$\frac{1}{1+e^{-\beta c}}$
Type $(i,1)$: $z = 1$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}}$

Mapping these action probabilities into the state transition probability matrix yields:

Log-Logistic Transition Matrix	to (n_i+1, n_j)	(n_i-1, n_j)	(n_i, n_j+1)	(n_i, n_j-1)	n
from n with $n_i < n_j - 1$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$	$\frac{1}{1+e^{-\beta c}} p_{i1}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$	$\frac{1}{1+e^{-\beta c}} p_{j1}$	$1-\sum P(n,n')$
from n with $n_i = n_j - 1$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{i0}$	$\frac{1}{1+e^{-\beta c}} p_{i1}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$	$\frac{1}{1+e^{\beta(0.5b-c)}} p_{j1}$	$1-\sum P(n,n')$
from n with $n_i = n_j$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{i0}$	$\frac{1}{1+e^{\beta(0.5b-c)}} p_{i1}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{j0}$	$\frac{1}{1+e^{\beta(0.5b-c)}} p_{j1}$	$1-\sum P(n,n')$
from n with $n_i = n_j + 1$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$	$\frac{1}{1+e^{\beta(0.5b-c)}} p_{i1}$	$\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{j0}$	$\frac{1}{1+e^{-\beta c}} p_{j1}$	$1-\sum P(n,n')$
from n with $n_i > n_j + 1$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$	$\frac{1}{1+e^{-\beta c}} p_{i1}$	$\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$	$\frac{1}{1+e^{-\beta c}} p_{j1}$	$1-\sum P(n,n')$

We can then show:

Proposition 5 *As the size of the electorate grows ($N = N_D + N_R \to \infty$) while the fractions $\alpha_i = N_i/N$ remain constant with $N_D \neq N_R$, the limiting distribution of turnout fractions for the log-logistic model with perfectly-informative polls converges to zero everywhere except for a probability-one mass point at $(x_0\%, x_0\%)$ where*

$$x_0 = \frac{e^{-\beta c}}{1 + e^{-\beta c}}.$$

The Proposition shows that the equilibrium turnout distribution is stochastically decreasing in c. Also, the equilibrium distribution converges to the best-response limiting distribution for $\beta \to \infty$, regardless of c.[42] The main substantive conclusion from Proposition 5 is that adopting an approach in line with Aldrich's conjecture where voter decisions are not driven by incentives would not alter any of our conclusions. While substantial participation may occur in the perturbed model, any such participation is driven by the random perturbations of the best response correspondence, i.e., by those agents that vote although their (unperturbed) incentives would suggest to abstain. This explains why this proposition predicts turnout of equal fractions $x_D = x_R = x_0$ for any finite level of β. In the absence of action noise (as $\beta \to \infty$) we recover the best response model with zero turnout (and "close" elections in the sense that n_D/N and $n_R/N \to 0$

existence of a unique limiting distribution in generic games. Since we derive a unique limiting distribution even in the case of (unperturbed) best-response this technical assumption is in general unnecessary for our model.

[42] This already holds for finite N.

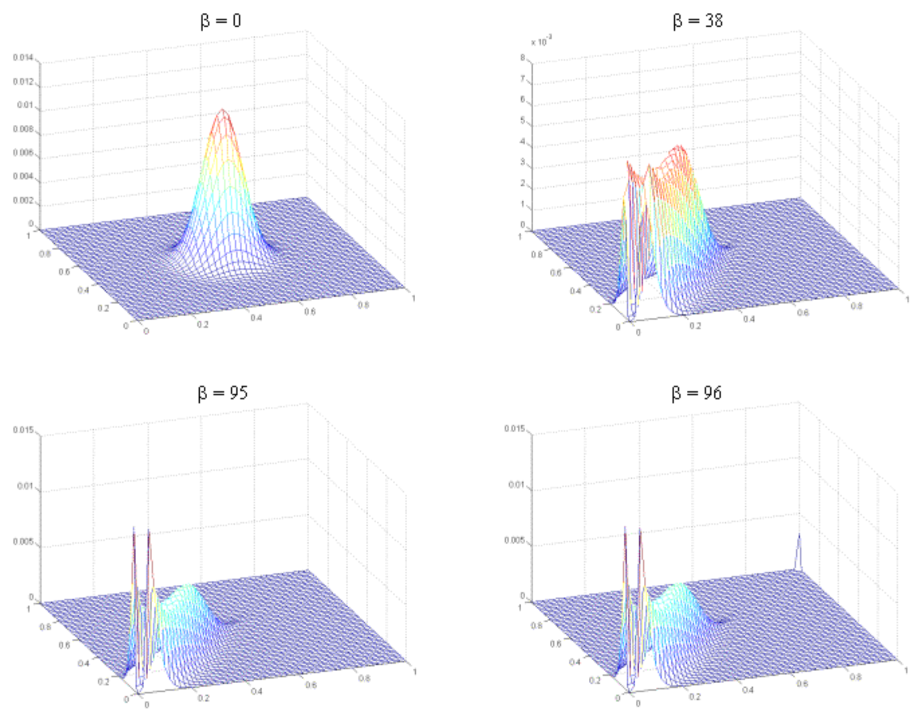

Figure 7 The limiting turnout density π for the log-logistic model as a function of the two turnout fractions and the parameter β for an example with $N_D = N_R = 50$ and cost-benefit ratio $\xi = 0.1$.

in agreement with Corollary 2 and Proposition 4).[43] Thus, while action noise may lead to higher turnout than polling noise, the key insight from the (unperturbed) best response model is still valid: Without either action or polling noise, expected participation in a best-response model is negligible if $N_D \neq N_R$.

In addition to providing a robustness check to the incentives vs. action noise argument, the log-logistic formulation allows us to study the "spontaneous" coordination through polls that may happen if $N_D = N_R$. Consider Figure 7, which shows the limiting distribution of turnout in an example with $N_D = N_R = 50$ for various values of β. The minimal value $\beta = 0$ corresponds to pure random choice and the turnout distribution thus is a Gaussian mountain with maximum turnout likelihood at (25,25) = (50%, 50%). As β increases, behavior is more and more driven by the incentives given by the game form. Recall that without noise the unique best-response turnout is (50, 50)=(100%, 100%) with probability 1. So, we may expect a convergence to universal turnout for vanishing noise. This, however, is *not* the case. Rather, the dynamics as a function of β are non-linear: as β increases, voters coordinate on *smaller* turnouts, which are consistent

[43]Note that in the turnout game the two limits ($t \to \infty$) and ($\beta \to \infty$) *are* interchangeable. This is *not* the case even in closely related games such as the discrete public goods game (Diermeier and Van Mieghem 2000), let alone general games (e.g. Blume 1997).

with *unequal* faction sizes. In the case of $\beta = 38$, there are three most likely turnout states: the most likely noise-induced state is 14 people from each party, with probability 0.68%, and two small turnout states: 4 Democrats and zero Republicans, or the reverse (0, 4), each also with probability 0.68%. As β increases, the two small turnout states become the most likely outcome at even lower turnout. At $\beta = 95$, for example, 2 voters of one party and zero of the other are the two most likely states with probability 1.47%, while the symmetric 14 people state still has probability of 0.68%. But at a critical value β^C between 95 and 96, suddenly spontaneous coordination at the (100%,100%) outcome becomes possible: for $\beta = 95$, the state (50,50)=(100%,100%) has probability 10^{-17}, whereas for $\beta = 96$ that state has probability 0.32%!

This phenomenon is reminiscent of the well-known phase transitions in theoretical physics.[44] For low β noise prevails, while at lower β two low-turnout states that are each other's mirror image (or differing only in "spin") are equally likely. At β^C, two phases can be in equilibrium: the low turnout phase (with two most-likely states, differing only in "spin") and the full turnout phase (with one most likely state). Finally, as temperature drops further, the low turnout phase becomes less likely, and ultimately, the full turnout phase prevails with probability 100%.[45]

9 Conclusion

We have proposed a new methodology to study coordination in voting games. As in game-theoretic models, the voters' incentives are given by a normal form. As in stochastic learning models, however, voters adjust their voting behavior in response to polling information about the current state of the electorate.

The model is applied to turnout games (Palfrey and Rosenthal 1983, 1985) where we investigate how noisy opinion polls may serve as coordination devices. Voters coordinate in both noisy and perfectly informative polls, under the assumption of both perturbed and unperturbed best response. We characterize the effect of uncertainty, induced either through information coarseness or sampling error, on turnout. We show that the effect of noise is non-monotonic: some uncertainty is necessary for non-zero participation levels, but too much uncertainty again leads to vanishing turnout. Using large-N approximations we then show that

[44] The threshold $1/\beta^C$ plays a role similar to the Curie temperature in models of spontaneous magnetization, i.e. magnetization (an ordered state) in the absence of any external magnetic field. Once the temperature drops below a critical threshold (the Curie temperature) the system suddenly switches to a magnetized state. This analogy can be made precise by the use of Ising models (e.g. Blume 1993). Ising models are isomorphic to infinite lattice games where each node "plays" a 2x2 coordination game with its immediate neighbors. The case of pure coordination with $x > 0$ on the diagonals and 0 everywhere else then corresponds to the case of spontaneous magnetization.

[45] In the case of $N_D \neq N_R$ this "phase transition" does not occur. Rather, more and more probability weight is put on the low turnout states.

unless there is some uncertainty about polling information turnout will be vanishingly small. Thus, merely assuming bounded rationality does not resolve the turnout problem.

Overall our results indicate a potentially important role for stochastic models in voting models, especially if coordination is an important characteristic of the strategic problem faced by voters. This suggests other applications of the model in voting games, for example in the case of multi-candidate elections or under different electoral rules. Eventually, such application may also include candidates as strategic actors.

This paper has been published in Mathematical and Computer Modelling 48 (2008) 1478-1496

A Proofs

Proof of Proposition 1: From the general pivot equations (1) we have

$$\Pr(\Delta n(\tilde{n}) \in \{-1, 0\}|\tilde{n}) = \Pr(\tilde{n}_D + \epsilon_D - \tilde{n}_R - \epsilon_R \in \{-1, 0\}|\tilde{n})$$

Defining $\Delta\epsilon = \epsilon_D - \epsilon_R$ yields

$$\Pr(\tilde{n}_D + \epsilon_D - \tilde{n}_R - \epsilon_R \in \{-1, 0\}|\tilde{n}) = \Pr(\Delta\epsilon \in \{-\Delta\tilde{n} - 1, -\Delta\tilde{n}\}|\tilde{n})$$

Denote $p_{\Delta\epsilon}(z) = \Pr(\Delta\epsilon = z)$. Given that $\Delta\epsilon$ is a sum of two random variables, its distribution is the convolution so that, using the indicator function $1\{\cdot\}$ ($1\{A\} = 1$ if A, otherwise 0):

$$\begin{aligned}
p_{\Delta\epsilon}(z) &= \sum_y \Pr(\epsilon_R = y - z)\Pr(\epsilon_D = y) \\
&= \sum_y p_\varepsilon 1\{-\varepsilon \leq y - z \leq \varepsilon\} p_\varepsilon 1\{-\varepsilon \leq y \leq \varepsilon\} \\
&= \sum_y p_\varepsilon^2 1\{\max(-\varepsilon + z, -\varepsilon) \leq y \leq \min(\varepsilon + z, \varepsilon)\} \\
&= \begin{cases} 0 & \text{if } |z| > 2\varepsilon \\ p_\varepsilon^2 \left(\min(\varepsilon + z, \varepsilon) - \max(-\varepsilon + z, -\varepsilon) + 1\right) & \text{if } |z| \leq 2\varepsilon, \end{cases} \\
&= \begin{cases} 0 & \text{if } |z| > 2\varepsilon, \\ \frac{2\varepsilon - |z| + 1}{(1+2\varepsilon)^2} & \text{if } |z| \leq 2\varepsilon. \end{cases}
\end{aligned}$$

Thus:

$$\Pr(\Delta\epsilon \in \{-\Delta\tilde{n} - 1, -\Delta\tilde{n}\}|\tilde{n}) = \begin{cases} 0 & \text{if } \Delta\tilde{n} \notin [-2\varepsilon - 1, 2\varepsilon], \\ \frac{1}{(1+2\varepsilon)^2} & \text{if } \Delta\tilde{n} \in \{-2\varepsilon - 1, 2\varepsilon\} \\ \frac{4\varepsilon - 2|\Delta\tilde{n}| + 2 - sign(\Delta\tilde{n})}{(1+2\varepsilon)^2} & \text{otherwise.} \end{cases}$$

Now, Type D birth $\Leftrightarrow P(\Delta n \in \{-1, 0\}|\tilde{n}) \geq 2\xi$, which is equivalent to $\Delta\tilde{n} \in [-w - 1, w]$, where $w \leq 2\varepsilon$ and

$$\begin{aligned}
w &= \max\left\{ i \in \{0, 1, \ldots, 2\varepsilon\} \text{ such that } \frac{4\varepsilon - 2i + 2 - 1}{(1+2\varepsilon)^2} \geq 2\xi \text{ and } \frac{4\varepsilon - 2(i+1) + 2 + 1}{(1+2\varepsilon)^2} \geq 2\xi \right\}, \\
&= \max\left\{ i \in \{0, 1, \ldots, 2\varepsilon\} : 4\varepsilon - 2\xi(1+2\varepsilon)^2 + 1 \geq 2i \right\} \\
&= \lfloor 2\varepsilon - \xi(1+2\varepsilon)^2 + \tfrac{1}{2} \rfloor.
\end{aligned}$$

■

Proof of Corollary 1: Clearly, $w(\xi, \varepsilon)$ is jointly concave in $\xi = c/b$ and ε, and for each c is maximal for (neglecting integrality restrictions):

$$\frac{\partial w}{\partial \varepsilon} = 2 - 2\xi(1+2\varepsilon)2 = 0 \Leftrightarrow \varepsilon^*(\xi) = \frac{1 - 2\xi}{4\xi},$$

and associated maximal width is:

$$w_{\max}(\xi) = w(\xi, \varepsilon^*(\xi)) = \lfloor 2\frac{1-2\xi}{4\xi} - \xi(1+2\frac{1-2\xi}{4\xi})^2 + \tfrac{1}{2}\rfloor = \frac{1-2\xi}{4\xi} = \varepsilon^*(\xi).$$

Similarly, w reaches its minimal value 0 when

$$2\varepsilon - \xi(1+2\varepsilon)^2 + \tfrac{1}{2} = 0 \Leftrightarrow \varepsilon \geq \overline{\varepsilon}(\xi) = \frac{1-2\xi+\sqrt{(1-2\xi)}}{4\xi},$$

or when

$$\xi \geq \overline{\xi}(\varepsilon) = \frac{4\varepsilon+1}{2\,(1+2\varepsilon)^2}.$$

■

Proof of Proposition 3: Denote by e_i a unit vector on the i-axis and let $\varepsilon_i = \frac{1}{N_i} = \frac{1}{\alpha_i N}$. For a state x inside the birth zone, we only have births:

$$x \to x + \varepsilon_i e_i \text{ w.p. } p_{i0} = \frac{N_i - n_i}{N} = \frac{N_i}{N}(1-x_i) = \alpha_i(1-x_i).$$

The limiting distribution $\pi(x)$ solves the global balance equations $\pi = \pi P$, which inside the birth zone thus reduce to:

$$\alpha_D\,(1-(x_D-\varepsilon_D))\,\pi(x-\varepsilon_D e_D) + \alpha_R\,(1-(x_R-\varepsilon_R))\,\pi(x-\varepsilon_R e_R) = (\alpha_D(1-x_D)+\alpha_R(1-x_R))\,\pi(x). \quad (4)$$

Now, consider the continuum approximation $p(x)$ of $\pi(x)$ by using a first-order Taylor expansion: $\pi(x - \varepsilon_i e_i) = p(x) - \varepsilon_i \frac{\partial p}{\partial x_i} + o(\varepsilon_i)$. Denoting $\frac{\partial p}{\partial x_i}$ by p_i, (4) is equivalent up to $o(\frac{1}{N})$ for large N to:

$$\begin{aligned} \alpha_D\,(1-x_D+\varepsilon_D)\,(p-p_D\varepsilon_D)) + \alpha_R\,(1-x_R+\varepsilon_R)\,(p-p_R\varepsilon_R)) - (\alpha_D(1-x_D)+\alpha_R(1-x_R))\,p &= 0 \\ \Leftrightarrow -\alpha_D p_D \varepsilon_D + \alpha_D x_D p_D \varepsilon_D + \alpha_D \varepsilon_D p - \alpha_D p_D \varepsilon_D^2 - \alpha_R p_R \varepsilon_R + \alpha_R x_R p_R \varepsilon_R + \alpha_R \varepsilon_R p - \alpha_R p_R \varepsilon_R^2 &= 0 \end{aligned}$$

Recall that $\alpha_i \varepsilon_i = 1/N$, so that this equality is equivalent to:

$$\begin{aligned} &\Leftrightarrow \quad -p_D + x_D p_D + p - p_D\varepsilon_D - p_R + x_R p_R + p - p_R \varepsilon_R = 0 \\ &\Leftrightarrow \quad (1-x_D+\varepsilon_D)p_D + (1-x_R+\varepsilon_R)p_R - 2p = 0 \end{aligned}$$

Hence, for $N \to \infty$, we have:

$$\text{PDE}_1(x) : (1-x_D)\frac{\partial p}{\partial x_D} + (1-x_R)\frac{\partial p}{\partial x_R} = 2p. \quad (5)$$

Changing variables $u_i = 1 - x_i$, we get:

$$\text{PDE}_1(u) : u_D \frac{\partial p}{\partial u_D} + u_R \frac{\partial p}{\partial u_R} = -2p,$$

with general solution: $p(u)$ is homogeneous of degree -2. If x is outside the birth strip, we only have deaths so that

$$x \to x - \varepsilon_i e_i \text{ w.p. } p_{i0} = \frac{n_i}{N} = \frac{N_i}{N} x_i = \alpha_i x_i.$$

The limiting distribution in the death zone solves:

$$\alpha_D (x_D + \varepsilon_D) \pi(x + \varepsilon_D e_D) + \alpha_R (x_R + \varepsilon_R) \pi(x + \varepsilon_R e_R) = (\alpha_D x_D + \alpha_R x R) \pi(x).$$

Similar to before, for $N \to \infty$, we have:

$$\text{PDE}_2(x) : x_D \frac{\partial p}{\partial x_D} + x_R \frac{\partial p}{\partial x_R} = -2p, \tag{6}$$

with general solution: $p(x)$ is homogeneous of degree -2. ■

Proof of Proposition 4: With perfect information, we know that $w = 0$. Using our fractional state descriptor $x_i = \frac{n_i}{N_i}$, the type i birth-zone in the scaled state space are the two lines $\alpha_R x_R - \alpha_D x_D \in [-\frac{1}{N}, 0]$. Clearly, as $N \to \infty$, both type's birth zones reduce to the line $\alpha_R x_R - \alpha_D x_D = 0$. First consider the case $N_D \neq N_R$. Anywhere outside that birth-line, the continuum approximation $p(x)$ is homogeneous of degree -2. Thus, in polar coordination $p(x_1, x_2) = p(r\cos\theta, r\sin\theta) = r^{-2} p(\cos\theta, \sin\theta)$, which means that p has a pole of order -2 at the origin. Because p must be integrable, it must be that $p(\cos\theta, \sin\theta) = 0$ for all θ. By extension, p is zero in the interior of the death zone, which yields that p has a mass point (Dirac impulse) of measure 1 at the origin $x = (0, 0)$. In the special case where $N_D = N_R$, we have that $\alpha_D = \alpha_R = \frac{1}{2}$ and our earlier argument must exclude the angle $\theta = 45^o$, which corresponds to the birth line. Indeed, we know that for $N_D = N_R$ (even for small values of N) we have a Dirac impulse of measure 1 at $x = (1, 1)$ because that state is absorbing for any value of N (thus also in the limit).■

Proof of Proposition 5: Set $\gamma = \frac{1}{1+e^{-\beta c}}$. Analogous to the derivation of the continuum approximation earlier, we have that the drifts at any state $x = (n_D/N_D, n_R/N_R)$ in the death zone are:

$$\begin{aligned}
x &\to x + \varepsilon_D e_D \text{ w.p. } (1-\gamma) \frac{N_D - n_D}{N} = (1-\gamma) \frac{N_D - n_D}{N_D} \frac{N_D}{N} = (1-\gamma) \alpha_D (1 - x_D), \\
x &\to x - \varepsilon_D e_D \text{ w.p. } \gamma \frac{n_D}{N} = \gamma \alpha_D x_D, \\
x &\to x + \varepsilon_R e_R \text{ w.p. } (1-\gamma) \frac{N_R - n_R}{N} = (1-\gamma) \alpha_R (1 - x_R), \\
x &\to x - \varepsilon_R e_R \text{ w.p. } \gamma \frac{n_R}{N} = \gamma \alpha_R x_R.
\end{aligned}$$

The limiting distribution $\pi(x)$ at any interior death-zone state x solves:

$$\begin{aligned}
&(1-\gamma) \alpha_D (1 - (x_D - \varepsilon_D)) \pi(x - \varepsilon_D e_D) + \gamma \alpha_D (x_D + \varepsilon_D) \pi(x + \varepsilon_D e_D) \\
&+ (1-\gamma) \alpha_R (1 - (x_R - \varepsilon_R)) \pi(x - \varepsilon_R e_R) + \gamma \alpha_R (x_R + \varepsilon_R) \pi(x + \varepsilon_R e_R) \\
&- ((1-\gamma) \alpha_D (1 - x_D) + \gamma \alpha_D x_D + (1-\gamma) \alpha_R (1 - x_R) + \gamma \alpha_R x_R) \pi(x) \quad = \quad 0
\end{aligned}$$

Using the continuum approximation $p(x)$ for π and Taylor's expansion to the first order yields:

$$(1-\gamma-x_D)\frac{\partial p}{\partial x_D}+(1-\gamma-x_R)\frac{\partial p}{\partial x_R}=2p.$$

Hence, p is homogeneous of degree -2 in $u_i = 1-\gamma-x_i$. As before, integrability implies that p must be zero everywhere except at $u_i = 0$, where it thus must have a mass point (Dirac impulse) of measure 1.■

B Computational Properties

Universal turnout is possible if factions are close in size, costs are small or polling noise is moderate. To calculate specific turnout numbers, however, one must solve the general balance equations $\pi = \pi P$ for π. Unfortunately, the derivation of a closed form solution is a very hard problem. This suggests the use of computational methods. From the global balance equations (and the normalization condition) it follows that in principle, π can be solved for exactly by solving a simple system of linear equations. This direct procedure involves $(N_D+1)(N_R+1)$ states and thus unknowns, which, computationally, makes this a viable approach only for relatively small populations.[46]

The balance equations, however, have a sparse structure, as each state only involves its direct neighbors. More importantly, in the death zones it involves only lower states, whereas in the birth zone only higher states are involved. This special structure can be exploited recursively to reduce the "quadratic complexity" of the problem from $(N_D+1)(N_R+1)$ to a "linear" complexity of only $2N_D - w + 1$ unknowns.[47]

This recursive formulation expresses all state probabilities in terms of the upper and lower strip boundary probabilities. We use $i : j$ to denote the set of integers $\{i, i+1, \ldots, j\}$ if $i < j$ and $i : j = \emptyset$ otherwise.:

$$\begin{aligned} u_i &= \pi(i, i+w+1) \qquad \forall i \in 0 : N_D, \\ l_i &= \pi(i, i-w-1) \qquad \forall i \in (w+1) : N_D. \end{aligned}$$

We can write all other $\pi(i,j)$ in terms of u and l as follows. Above the strip, the balance equation

$$(i+1)\pi(i+1,j)+(j+1)\pi(i,j+1)=(i+j)\pi(i,j)$$

can be solved backwards recursively given that $\pi(i,j) = 0$ for $j > \bar{j} := N_D + w$:

$$\pi(i,\bar{j}) = \frac{i+1}{i+\bar{j}}\pi(i+1,\bar{j}) \Rightarrow \pi(i,\bar{j}) = \frac{(i+1)\cdots N_D}{(i+\bar{j})\cdots(N_D+\bar{j})}u_{N_D}.$$

[46] A simple personal computer can solve a linear system with a few thousand unknowns in reasonable time. For example, a PC with 128MB of RAM can store 8000 numbers (assuming IEEE double extended precision, each number requires 16 bytes of storage). Thus, with $N_1 N_2 \simeq 8000$, one solves exactly for populations $N_i \simeq 89$.

[47] Hence, using this recursive formulation our simple personal computer can solve populations of size $N_i \simeq 4000$ exactly.

Now, full backward recursion applies to the upper triangle and specifies $\pi(i,j)$ in terms of $u_j, u_{j+1}, \ldots, u_{N_D}$. Specifically, $\forall i \in 0:(N_D-1)$ we have that

$$\pi(i,i+w+2) = \sum_{j=i+w+2}^{N_1} U_{ij}u_j,$$

Similarly, we solve the lower triangle in terms of l and $\forall i \in (w+1):(N_D-1)$ we have that

$$\pi(i,i-w-2) = \sum_{j=i+1}^{N_D} L_{ij}l_j.$$

Inside the strip, we can solve for all π in terms of both u and l. Indeed, the balance equation inside:

$$(N_D-i+1)\pi(i-1,j) + (N_R-j+1)\pi(i,j-1) = (N-i-j)\pi(i,j),$$

can now be solved by forward recursion. Thus, this also solves for the diagonals one-off the strip boundaries: $\forall i \in 0:(N_D-1)$ we have that

$$\begin{aligned} \pi(i,i+w) &= \sum_{j=0}^{i-1} U^{+}_{ij}u_j + \sum_{j=w+1}^{i-1} L^{+}_{ij}l_j. \\ \pi(i,i-w) &= \sum_{j=0}^{i-1} U^{-}_{ij}u_j + \sum_{j=w+1}^{i-1} L^{-}_{ij}l_j. \end{aligned}$$

Now we only need to solve for the line probabilities u and l, which follow from the balance equations on those lines. Specifically, the upper strip boundary yields:

$$\begin{aligned} (N_R-j+1)\pi(i,j-1) + (j+1)\pi(i,j+1) &= N_D\pi(i,j), \\ \Leftrightarrow (N_R-i-w)\pi(i,i+w) + (i+w+2)\pi(i,i+w+2) &= N_D\pi(i,i+w+1) \\ \Leftrightarrow (N_R-i-w)\left[\sum_{j=0}^{i-1} U^{+}_{ij}u_j + \sum_{j=w+1}^{i-1} L^{+}_{ij}l_j\right] + (i+w+2)\sum_{j=i+1}^{N_1} U_{ij}u_j &= N_D u_i. \end{aligned} \tag{7}$$

The lower strip boundary yields:

$$\begin{aligned} (N_D-i+1)\pi(i-1,j) + (i+1)\pi(i+1,j) &= N_R\pi(i,j), \\ \Leftrightarrow (N_D-i+1)\pi(i-1,i-w-1) + (i+1)\pi(i+1,i-w-1) &= N_R\pi(i,i-w-1) \\ \Leftrightarrow (N_D-i+1)\left[\sum_{j=0}^{i-2} U^{-}_{i-1,j}u_j + \sum_{j=w+1}^{i-2} L^{-}_{i-1,j}l_j\right] + (i+1)\sum_{j=i+2}^{N_1} L_{i+1,j}l_j &= N_R l_i. \end{aligned} \tag{8}$$

Equations (7)–(8) specify the recursive problem formulation. Since it yields a linear system of equations with full coefficient matrix, an analytic closed form solution seems unlikely. Computational complexity, however, is greatly reduced by the recursive formulation, which as a linear system the numeric solution is straightforward to solve.

Nevertheless, even that approach cannot compute electorate sizes of millions. In that case one needs to resort to simulations.[48] This technique exploits the ergodic properties of the process, i.e., the fact that π_j also gives the long-run mean fraction of time that the process occupies state j (e.g., Taylor and Karlin 1994; p.176). Formally,

$$\pi_j = \lim_{m\to\infty} \frac{1}{m} \sum_{\tau=0}^{m-1} \Pr\{X^\tau = j | X^0 = i\}$$

Invoking the fact that the limiting distribution is independent of the starting state, one obtains π by simulation the dynamics for an arbitrarily long period of time, starting from any state at time 0. Of course, for finite time-spans simulations only yield approximate results.

References

[1] Aldrich, John. 1993. "Rational Choice and Turnout." *American Journal of Political Science.* 37(1):246-278.

[2] Bendor, Jonathan, Daniel Diermeier, and Michael M. Ting. 2003. "A Behavioral Model of Turnout." *American Political Science Review* 97(2).

[3] Bergin, James, and Barton l. Lipman. 1996. "Evolution with State-Dependent Mutations." *Econometrica* 64(4):943-956.

[4] Blume, Lawrence E. 1993. "The Statistical Mechanics of Strategic Interaction." *Games and Economic Behavior* 4:387-424.

[5] Blume, Larry. 1995. "The Statistical Mechanics of Best-Response Strategy Revision." *Games and Economic Behavior* 11: 111-45.

[6] Blume, Lawrence E. 1997. "Population Games." In W. Brian Arthur, Steven N. Durlauf, and David A. Lane, eds. *The Economy as an Evolving Complex System II.* Reading: Addison-Wesley.

[7] Conley, J., Toossi, A. and M. Wooders, 2001, "Evolution and Voting: how nature makes us public spirited", Warwick Economic Research Papers, No.601.

[8] Dhillon, Amrita, and Stefano Demichelis. 2002. "Learning in Elections and Voter Turnout". Working Paper. CORE.

[48] The general problem is not computational time, but storage: the coefficient matrix of our recursive formulation is dense so that with $N_i \simeq 1$ million, we need to store 1 trillion numbers!

[9] Diermeier, Daniel, and Jan A.Van Mieghem. 2000. "Spontaneous Collective Action." CMSEMS Discussion Paper No. 1302. Kellogg Graduate School of Management. Northwestern University.

[10] Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

[11] Feddersen, Timothy and Alvaro Sandroni. 2002. "A Theory of Participation in Elections with Ethical Voters." Working paper, available from www.kellogg.northwestern.edu/faculty/fedderse/homepage.

[12] Fey, M. 1997, "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting", *American Political Science Review* 91(1):135-47.

[13] Foster, Dean and H. Peyton Young. 1990. "Stochastic Evolutionary Game Dynamics." *Theoretical Population Biology* 38: 219-232.

[14] Fudenberg, Drew, and David K. Levine. 1998. *The Theory of Learning in Games.* Cambridge, Mass: MIT Press.

[15] Hansen, Steven, Thomas Palfrey, and Howard Rosenthal. 1987. "The Relationship between Constituency Size, and Turnout: Using Game Theory to Estimate the Cost of Voting." *Public Choice* 52:15-33.

[16] Kandori, Michihiro, George Mailath and Raffael Rob. 1993. "Learning, Mutation, and Log-Run Equilibria in Games." *Econometrica* 61: 29-56.

[17] Ledyard, John. 1984. "Pure theory of two-candidate competition." *Public Choice* 44:7-41.

[18] McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." in P. Zarembka, ed., *Frontiers in Econometrics.* New York: Academic Press.

[19] Morris, Steve, and Hyun Song Shin. 2003. "Global Games: Theory and Applications," in M. Dewatripont, L. Hansen and S. Turnovsky, eds..*Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)*, Cambridge, England: Cambridge University Press..

[20] Myerson, Roger. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory* 27:375-392.

[21] Nalebuff, Barry, and Ron Shahar. 1999. "Follow the Leader: Theory and Evidence on Political Participation." *American Economic Review 89*(3)*:525-549.*

[22] Niemi, Richard G., and Herbert F. Weisberg, eds. 1993. *Classics in Voting Behavior.* 3rd. edition. Washington, DC: Congressional Quarterly Press.

[23] Palfrey, Thomas R., and Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice* 41:7-53.

[24] Palfrey, Thomas R., and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *American Political Science Review* 79:62-78.

[25] William H. Riker; and Ordeshook, Peter C. 1968. "A Theory of the Calculus of Voting". *The American Political Science Review* 62 (1): 25-42

[26] Sieg, G. and C.Schulz, 1995, "Evolutionary Dynamics in the voting game", *Public Choice* 85: 157-72.

[27] Taylor, Howard M., and Samuel Karlin. 1994. *An Introduction to Stochastic Modeling*. Second Edition. Boston et al.:Academic Press.

[28] Tversky, A. and G. Quattrone. 1988. "Contrasting Rational and Psychological Analyses of Political Choice." *American Political Science Review* 82:719-36.

[29] Wolfinger, Raymond E., and Rosenstone, Steven. 1980. *Who Votes?* New Haven. Yale University Press.

[30] Young, H. Peyton. 1993. "The Evolution of Conventions." *Econometrica* 61:57-84.

[31] Young, H. Peyton. 1998. *Individual Strategy and Social Structure*. Princeton: Princeton University Press.