# Coordination in Turnout Games*

Daniel Diermeier and Jan A. Van Mieghem

November 17, 2000

**Abstract**

We present a stochastic model of coordination in turnout games. In each period a randomly selected voter receives information about current play through noisy polls and then, based on this information, forms expectations about the current configuration of play and chooses a best response. 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 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. The model's predictions are also consistent with the usual empirical regularities about turnout. We then derive continuum approximation results for large electorates using a partial differential formulation and apply the results to the case of perfectly informative polls. We show that under (perturbed) best response voters are able to spontaneously coordinate their actions on a single state.

## 1 Introduction

Anthony Downs' (1957) "paradox of voting" or "turnout problem" constitutes perhaps the most famous anomaly for rational choice models of politics. 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. But citizens do participate, even in very large electorates.

The turnout problem has generated a large literature and many solution attempts.[1] Among the most influential are game-theoretic voting models (Palfrey and Rosenthal 1983, 1985; Myerson 1998). Here, 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 plurality rule with some tie-breaking provision, such as a coin toss. All members of the winning faction 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 $\frac{b}{2} > c > 0$.[2]

Because voting for the non-preferred candidate is dominated for each voter by voting for the preferred candidate, the relevant problem reduces to a *turnout game*, 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.[3]

It follows that all Nash-equilibria in the turnout problem involve the use of mixed strategies by at least some voters. This leads to an abundance of Nash-equilibria, some of them with surprisingly high turnout. However, all equilibria with non-trivial turnout in large elections are asymmetric and thus require precise

*Both authors are at the Department of Managerial Economics and Decision Sciences (MEDS), Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2009; d-diermeier@kellogg.nwu.edu and VanMieghem@kellogg.nwu.edu.

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

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

[3] 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.

coordination.[4] Subsequent research has regarded these equilibria as problematic (Palfrey and Rosenthal 1985, Myerson 1998). The predominant approach has been to limit the amount of common knowledge present among voters by introducing some 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.

We propose a different approach that allows us to explicitly model coordination in voting games. Voters are assumed to receive information about current play through (noisy) opinion polls. Noise may result from either sampling error or from the fact that polling numbers are usually reported with finite precision (typically two or three digits). Specifically, in each period a randomly chosen voter observes a poll based on the population's last period voting behavior. Given that polls typically exhibit some sampling noise, the voter uses Bayes' rule to form expectations about the current configuration of play. Based on this expectation, she then chooses a best response action. 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.[5]

In part, this model is motivated by our recent paper (Diermeier and Van Mieghem (2000)) that shows how spontaneous coordination on asymmetric states is possible in a discrete public good model.[6] Moreover, the model's limiting distribution exhibits qualitatively different features from the mixed strategy equilibria in the game theoretic models. Since in turnout games all equilibria involve the use of mixed strategies, a stochastic approach may generate new insights into the turnout problem. In particular, it allows us to study polls as coordination devices and their consequences for turnout and election outcomes.

As in Diermeier and Van Mieghem (2000) we prove the existence of a unique limiting distribution. We then investigate the qualitative properties of the limiting distribution. Surprisingly, we find 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. If polling information becomes too noisy, however, turnout again drops to vanishingly small levels consistent with Palfrey and Rosenthal's (1995) and Myerson's (1998) findings. Thus, in contrast to 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 turnout. Turnout, for example, 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).

While we characterize the necessary and sufficient conditions of the limiting distribution, its computational complexity prevents us from deriving a closed form solution unless factions are of approximately equal size. This suggests the use of computational methods. In an appendix we present some computational properties of the model. In particular, we derive a recursive formulation of state probabilities that reduces the model's computational complexity from quadratic to linear.

Additional insights can be derived by using a continuum approximation for large electorates with $N$ players. Such an approach is justified given that the intended domain of our model are large elections. Specifically, we derive a partial differential equation representation of the limiting distribution that shows that elections must be close. Moreover, the closeness depends on the accuracy of election polls. We then apply our large $N$ results to the special case of perfectly informative polls, which corresponds to a Markov process with best response dynamics (Blume 1995) and log-logistic dynamics (Blume 1993). Best response dynamics are usually difficult to analyze due to the lack of ergodicity. In our model, however, the properties of the general model are preserved. In particular, a unique limiting distribution exists. Moreover, for large $N$, voters are able to implicitly coordinate their actions. That is, we show that for large $N$ the model uniquely predicts vanishing turnout independently of the costs or benefits of voting, unless factions are of exactly equal size where coordination occurs on the full turnout state.

To further investigate this "spontaneous coordination" result, we consider perturbed best response dy-

[4] In these equilibria, the larger faction is devided into two sub-groups that play different (mixed) strategies. See Palfrey and Rosenthal (1983) for details and Myerson (1998) for an instructive example.

[5] For a behavioral model of trail-and-error learning in turnout games see Bendor, Diermeier and Ting (2000).

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

namics using a log-logistic formulation (Blume 1993).[7] In general, perturbed best response models (even in the limit of vanishing noise) may have very different properties from unperturbed best response models (e.g. Blume 1997, Young 1998). We show that in the long-run expected turnout fraction may be positive for large $N$, but that any such turnout is the direct consequence of noise in the individual choice process. In the case of vanishing noise we recover negligibly small turnout fractions in the limit. The loglogistic model also demonstrates how spontaneous coordination in elections with equal faction sizes ($N_D = N_R$) depends on a critical threshold $\beta^C$ of the uncertainty parameter $\beta$. For high levels of noise ($1/\beta > 1/\beta^C$), noise prevails and voters coordinate on turnouts of less that 50%. As the noise level drops below $1/\beta^C$, however, a second coordinated outcome is possible where all voters turn out. As uncertainty drops even further, the noisy low-turnout coordinated outcome disappears and only the full turnout equilibrium remains, in accordance with deterministic best-response.

In summary, 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.

# 2 The Model

There are two factions 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 factions with $i$ and $j$ denoting different factions unless otherwise noted. Each voter must choose from an action $z \in \{0, 1\}$, where $z = 0$ means "abstaining." The state of the electorate at time $t$ is given by $n^t = (n_D^t, n_R^t)$, where $n_i^t \leq N_i$ is the number of type $i$ that is intending to vote at time $t$. Superscripts indicating time periods are dropped unless necessary. In the usual fashion we write $n_i^{-k}$ for the number of voters of type $i$ without counting $k$. 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[8] (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 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 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, voters are 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 $1/N$.[9] 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

[7] For technical reasons researchers have overwhelmingly prefered the use of perturbed best-response dynamics because it guarantees the existence of a unique limiting distributions parametrized by the stochastic disturbance (e.g. Blume 1993, Blume 1995, Kandori, Mailath, Rob 1993, and Young 1993).

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

[9] 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.

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 $\widetilde{n}(n)$. Published polls, for example, typically include a polling error of 3%, which roughly means that $n = \widetilde{n} \pm 3\%$. Given $\widetilde{n}$, an agent updates her beliefs about the state of the system and chooses a best response given the signal.[10]

Formally, our stochastic model defines a discrete time, discrete state Markov process: we have a family of random variables $\{X^t : t \in \mathbb{N}\}$ where $X^t$ assumes values on the state space $S_d \times S_r$ and where $S_d = \{0, 1, 2, \ldots, N_D\}$ and $S_r = \{0, 1, 2, \ldots, N_R\}$. 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 vote only 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 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 $\xi$, where $0 < \xi < \frac{1}{2}$. Formally then:

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

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

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

Clearly, the exact form of the birth and death zones depends on how true states $n$ are reported as a noisy poll. In the remainder we will investigate various polling structures and characterizes the corresponding transition matrix $P$ and birth and death zones.

[10] 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.

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 column vector $\pi$, where

$$\pi_j = \lim_{t\to\infty} \Pr\{X^t = j | X^0 = i\}, \tag{2}$$

and $\pi_j > 0$ for all $j \in S$. It can easily be shown that $\pi$ 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

$$(1 - P_{ii})\,\pi_i = \sum_{j\neq i} \pi_j P_{ji}, \tag{3}$$

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 $\pi$. This requires specifying the transition matrix $P$ and solving $\pi = \pi P$ for $\pi$. Like the birth and death zones, $P$ depends on the details of the polling technology, which we study next.

# 3 General Results

We will consider a simple noise model where noise is uniformly distributed over a square-grid of size $[-\varepsilon, \varepsilon]^2$. As mentioned above, this is computationally convenient, but without loss of generality. Specifically[11]:

$$\widetilde{n}(n) = (n_D + \epsilon_D, n_R + \epsilon_R) \qquad \text{with probability } p_\varepsilon = \frac{1}{(1+2\varepsilon)^2} \qquad \forall \epsilon_i \in \{-\varepsilon, -\varepsilon+1, \ldots, \varepsilon\}. \tag{4}$$

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

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

Suppose the 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 that the necessary cost-benefit ratio $\xi$ must be *lower* than under complete information. 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 parameters $c$, $b$, and $\epsilon$ turnout may increase or decrease. That is, the effect of polling noise is that the birth zone may increase at the expense of the death zone:

**Proposition 1** *With uniform polling noise over the square $[-\varepsilon, \varepsilon]^2$, the pivot equation $P(\Delta n(\widetilde{n}) \in \{-1, 0\} | \widetilde{n}) \geq 2\xi$ defines a birth-zone of "width" $w(\xi, \varepsilon)$:*

$$\begin{array}{ll} \textit{Type D birth} \Leftrightarrow \Delta\widetilde{n} \in [-w-1, w] & \textit{Type R birth} \Leftrightarrow \Delta\widetilde{n} \in [-w, w+1] \\ \textit{Type D death} \Leftrightarrow \Delta\widetilde{n} \notin [-w, w+1] & \textit{Type R death} \Leftrightarrow \Delta\widetilde{n} \notin [-w-1, w], \end{array} \tag{5}$$

*where*

$$w(\xi, \varepsilon) = \lfloor 2\varepsilon - \xi(1+2\varepsilon)^2 + \tfrac{1}{2} \rfloor. \tag{6}$$

[11] To be more accurate, 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.

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

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

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

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

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):

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

Thus:

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

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

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

■

We can summarize the transition probabilities at each state in the transition matrix $P$, which is graphically represented in Figure 1. The up and right transitions inside the strip around the diagonal represent births, while the down or left transitions are deaths. The proposition summarizes the joint impact of relative cost-benefit $\xi = c/b$ and the noise level $\varepsilon$ on pivot probabilities into one parameter: the width $w$ of the birth zone.

The transition probabilities directly imply that 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 $N_R \le N_D + w + 1$, then the state $(n_D, n_R)$ is absorbing. Proposition 2 summarizes these findings:

**Proposition 2** *A unique limiting distribution exists and solves $\pi = \pi P$. Moreover, $\pi$ is zero at all states $n = (n_D, n_R)$ such that $(n_D, n_R) < (w, w)$ or $n_R > N_D + w + 1$. If $N_R \le N_D + w + 1$, 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 $w$. As an example, we simulated the expected turnout in our model for various values of $w$ for a total electorate of 1 million voters[12] with 48% democrats and 52% republicans. The

[12] For such large electorates the limiting distribution $\pi$ can no longer be calculated exactly (the linear system $\pi = \pi P$ has $480 \times 520$ million unknowns). The time dynamics of the Markov chain, however, can easily by 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 models computational properties are discussed in an appendix.

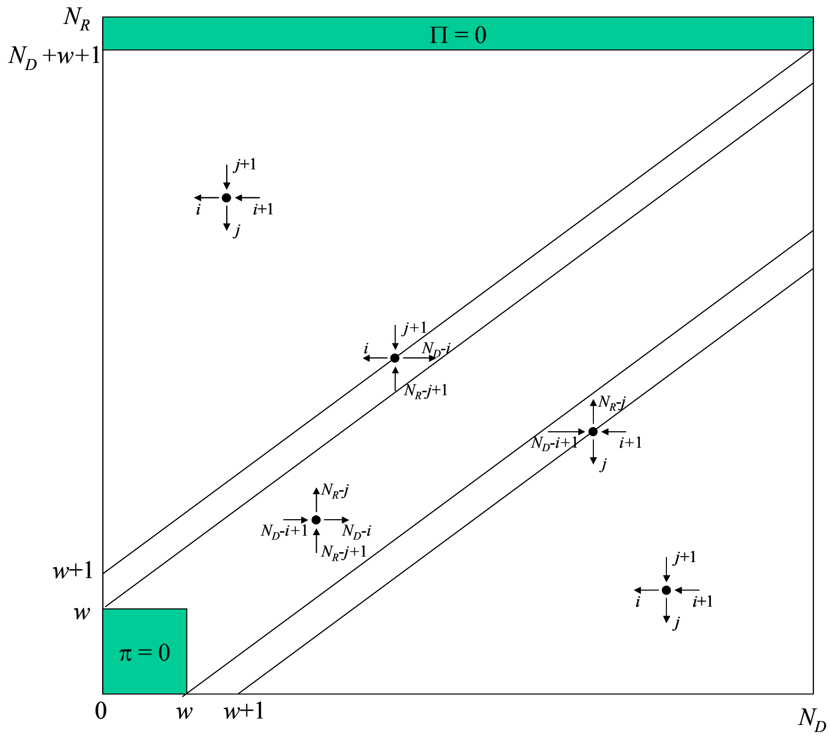


Figure 1: The state space is partitioned in a "birth zone" around the diagonal and two death zones. The specific state transition probabilities (relative to $N$) are shown.

expected turnout together with sample points are reported in Figure 2. 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 $w$.)

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 $\xi$ and $\epsilon$ to assess the effects of the cost/benefit ratio and uncertainty on expected turnout.

**Corollary 3** *The birth width $w$ decreases linearly in the cost-benefit ratio of voting $\xi$, but is concave in the level of noise $\varepsilon$. 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*

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

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

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

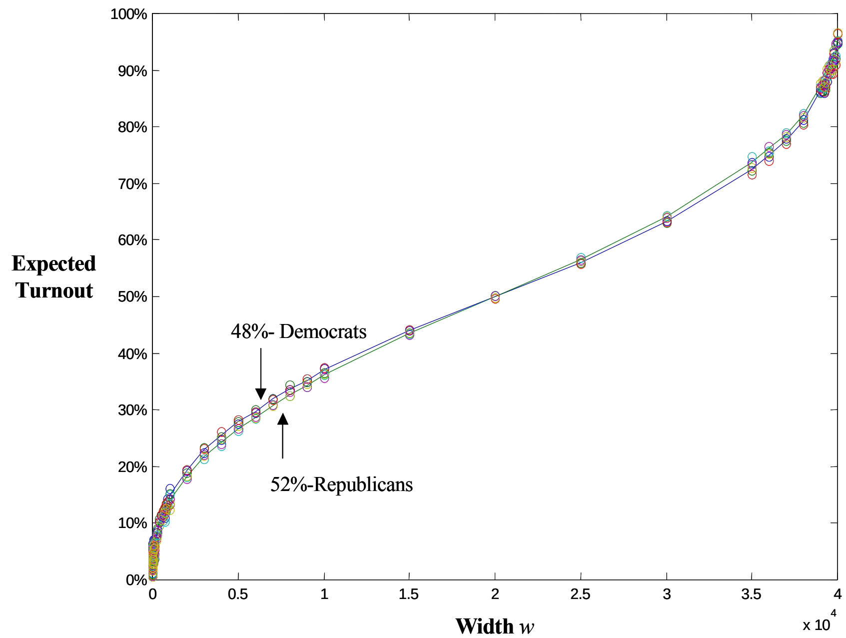


Figure 2: The expected turnout fractions as a function of the width $w$ of the birth zone for an electorate of size 1 million with 48% Democrats and 52% Republicans.

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

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

■

The corollary is summarized in Figure 3, which shows a contour map of the parameter $w$ that summarizes the impact of both cost/benefit ($\xi$) and noise ($\varepsilon$). Recall that 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). Moreover, our model can account for some the well-known empirical regularities of turnout with respect to the costs and potential benefits of voting (Hansen, Palfrey, and Rosenthal 1987, Wolfinger and Rosenstone 1980). Turnout decreases in the cost of participation (because $w$ decreases), but increases in the stakes of the election (because $w$ increases)[13], and of course, the closeness of the race as reported in the opinion poll (Nalebuff and Shachar 1999).[14]

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

[14] It is worth pointing out that turnout may be substanstially higher if voters decide 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 driving to the polls, then our model suggests that participation in all election 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.

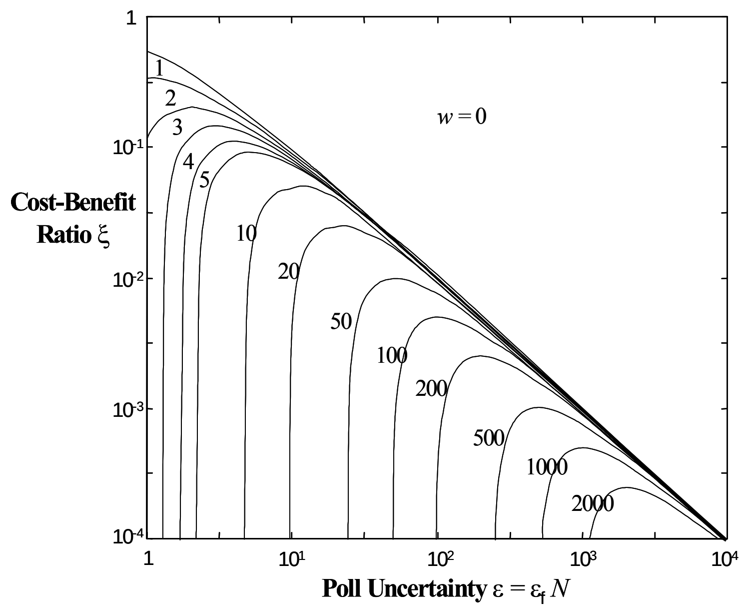


Figure 3: The width parameter $w$ as a function of $\xi = c/b$ and $\varepsilon = \varepsilon_f N$.

To see the critical effect of polling noise, consider an example by Myerson (1998), which was constructed to demonstrate the strikingly low expected turnout in large elections with costly voting with random faction sizes $N_D$ and $N_R$. In Myerson's model, it can be easily shown that the expected turnout of Democrats and Republicans in large electorates is roughly the same at approximately $1/(4\pi\xi^2)$. In his example, 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) so that in the unique Poisson voting equilibrium expected turnout equals about 32 voters of either party. For our model, highest expected turnout occurs for largest value of $w$. The corollary and Figure 3 show that the highest $w$ for $\xi = 0.05$ 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[15] 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.

The discrepancy, we believe, can be explained by the stochastic assumptions: Myerson assumes a best-response model with only uncertainty in the faction sizes. Our model assumes certainty in faction sizes but introduces polling noise in a dynamic model. Our results indicate that polling noise has a much more pronounced impact on turnout than faction size uncertainty. Indeed, consider an example with the same cost-benefit ratio of $\xi = 0.05$, but more plausible faction sizes: $N_D = 480,000$ and $N_R = 520,000$. Turnout for this example in Myerson's model remains virtually unchanged[16] at about 64 voters (or .0064%), similar to the earlier example with 3 million voters. In our model, more similar faction sizes have a substantial impact and with $w = 10$, expected turnout now is about (4.40%, 4.07%), as reflected in Figure 2. This means about twenty-one thousand voters of each party are expected to show up. Finally, if the cost-benefit ratio were to decrease fifty fold to $\xi = 0.0001$, the highest width increases to $w^* = 2500$ and expected turnout in our model[17] increases to about 20%, as shown in Figure 2. So, a fifty fold drop in the cost/benefit ratio roughly corresponded to a five fold increase in turnout percentages. This highlights the subtle yet crucial impact of polling uncertainty, rather than faction uncertainty, in turnout models.

The example also confirms and highlights the impact of the cost-benefit ratio in voting turnout models. As

[15] 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.

[16] The reason is that Myerson assumes that $N_D$ and $N_R$ are Poisson random variables, whose standard deviation is the square root of the mean. Hence, with a mean of about 500,000, the standard deviation is only 707. Hence, whether $N_R - N_D$ is 1 million as in the first example, or 40,000, as in the second, statistically the faction sizes are clearly not equal so that expected turnout is not much affected.

[17] Myerson's large population approximation does no longer apply in this low cost-benefit region.

population size increases, substantial turnout requires higher noise and lower cost/benefit ratios. Substantial turnout is possible if the stakes in large elections are substantially higher than in small elections.[18] If, for example voters care more about Senatorial than a House election, then although the district size is larger in the former election turnout may still be substantial. The critical question then is how fast the cost/benefit must change in $N$ to ensure non-vanishing turnout. The corollary allows us to formally analyze the relationship between $\xi$ and $N$. First note that the answer depends on the noise level. Consider a fixed relative amount of noise, i.e. $\varepsilon_f = \varepsilon/N$ is constant. Then, we have that[19]:

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

so that a substantial turnout with large population size requires that $\xi$ decreases inversely in $N$. For example, for a polling noise level $\varepsilon_f = 3\%$, substantial turnout requires that $\xi(N) \leq \left(1 + \frac{3}{50}N\right)^{-1}$.

# 4 Results for Large Electorates

The intended domain of applications for our model are large elections. Therefore, it seems appropriate to investigate a continuum approximation for large population size $N$. Consider the fractional state descriptor:

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

Clearly, 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 $0 \leq w_f < \Delta\alpha = \alpha_R - \alpha_D = \frac{N_R - N_D}{N}$, so that from before, we 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 4** *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:*

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

*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.*

**Proof**: 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:

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

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

[19] 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 $x$, $O(f(x)) \simeq f(x)$.

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

$$\alpha_D (1 - (x_D - \varepsilon_D)) \pi(x - \varepsilon_D e_D) + \alpha_R (1 - (x_R - \varepsilon_R)) \pi(x - \varepsilon_R e_R) = (\alpha_D(1 - x_D) + \alpha_R(1 - x_R)) \pi(x). \tag{7}$$

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 $p_i$, (7) is equivalent up to $o(\frac{1}{N})$ for large $N$ to:

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

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

$$\text{PDE}_1(x) : (1 - x_D)\frac{\partial p}{\partial x_D} + (1 - x_R)\frac{\partial p}{\partial x_R} = 2p. \tag{8}$$

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:

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

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

$$\text{PDE}_2(x) : x_D \frac{\partial p}{\partial x_D} + x_R \frac{\partial p}{\partial x_R} = -2p, \tag{9}$$

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

Given that an interior extremum would require $\frac{\partial p}{\partial x_i} = 0$, the two PDE's directly yield:

**Corollary 5** *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. How close depends on $w$.

# 5 Perfectly Informative Polls

The PDEs' boundary conditions are too complex to derive a closed form solution for the general case. However, we can characterize 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.[20]

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

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

$$\begin{array}{ll} \text{Type } D \text{ birth} \Leftrightarrow \Delta n \in \{-1, 0\} & \text{Type } R \text{ birth} \Leftrightarrow \Delta n \in \{0, 1\} \\ \text{Type } D \text{ death} \Leftrightarrow \Delta n \notin \{0, 1\} & \text{Type } R \text{ death} \Leftrightarrow \Delta n \notin \{-1, 0\}. \end{array}$$

Notice that the pivot equations are independent of $\xi$. 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 | $n_i + 1$ | $n_i - 1$ | $n_j + 1$ | $n_j - 1$ | $n$ |
|---|---|---|---|---|---|
| $n_i < n_j - 1$ | 0 | $p_{i1}$ | 0 | $p_{j1}$ | $1 - p_{i1} - p_{j1}$ |
| $n_i = n_j - 1$ | $p_{i0}$ | $p_{i1}$ | 0 | 0 | $1 - p_{i0} - p_{i1}$ |
| $n_i = n_j$ | $p_{i0}$ | 0 | $p_{j0}$ | 0 | $1 - p_{i0} - p_{j0}$ |
| $n_i = n_j + 1$ | 0 | 0 | $p_{j0}$ | $p_{j1}$ | $1 - p_{j0} - p_{j1}$ |
| $n_i > n_j + 1$ | 0 | $p_{i1}$ | 0 | $p_{j1}$ | $1 - p_{i1} - p_{j1}$ |

The unique limiting distribution $\pi$ can now be found by solving the linear system of equations $\pi = P\pi$ given by the global balance equations. Applying Proposition (4) we can show the following:

**Proposition 6** *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 Dirac impulse at* $(0\%, 0\%)$ *if $N_D \neq N_R$ or at* $(100\%, 100\%)$ *if $N_D = N_R$.*

Proof: 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 $\theta$. By extension, $p$ is zero in the interior of the death zone, which yields that $p$ has a 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).■

For large $N$, voters will (almost surely) coordinate on a state with zero turnout level, unless we have the knife-edge case of *exactly equal* factions.[21] There is no analogue to the mixed strategy equilibria in the game-theoretic model or the asymmetric high-turnout equilibria found in Palfrey and Rosenthal (1983).[22] Moreover, in contrast to the multiplicity of equilibria in that model, the prediction is unique. Note that the result obtains in the absence of any uncertainty or noise. It is purely driven by the explicit coordination device.

For technical reasons most of the 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 existence of a unique limiting distribution. Rather than characterizing the limiting distribution directly researchers have focussed on the case of arbitrarily small noise. The critical notion here is that

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

[22] Compare this result to the discrete public good game analyzed in Diermeier and Van Mieghem (2000) where asymmetric states with high participation may be the most likely long-run states.

of stochastically stable state (Foster and Young 1990). Intuitively, a state is stochastically stable if its assigned limiting probability is strictly positive for vanishing noise. These are states that are most likely to be observed over the long run if perturbations from best response behavior are arbitrarily small. Technically, this approach amounts to a double-limit argument. First time is driven to infinity to find the limiting distribution parametrized by the noise term, and then noise is driven to zero to identify the stochastically stable states. The advantage of this approach is that in many games the stochastically stable states can be characterized directly without having to specify the limiting distribution first.[23]

In our case the stable state short-cut is not available. Rather, we need to explicitly derive the limiting distribution. Following Blume (1993) we consider the case of log-logistic choice.[24]. Let $p^{\beta}(z|X^t_{-k})$ denote the conditional probability that in period $t+1$ agent $k$ will play action $z$ given that the current configuration of play (not including $k$) is $X^t_{-k}$. Then the log-linear choice rule is given by:

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

It is equivalent to the assumption that the pair-wise probability ratios of choosing actions are proportional to the respective pay-off differences. This rule can either be interpreted as "perturbed" decision making (e.g. Blume 1993) or 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 particularly suitable for a model of voting since the (perceived) benefits and costs of participating may well vary substantially over time.[25] Nothing in our model presupposes a particular interpretation of the log-linear rule. All we assume is that the agents' behavioral regularities can be captured by it. The parameter $\beta$ may be interpreted as the degree to which choices respond to the incentives in the model. For $\beta = 0$ choice is completely random. That is, for all possible configurations, $k$ will play each action with probability 1/2. For $\beta \to \infty$, log-linear choice converges to a distribution that puts positive probability only on best-responses to $X^t_{-k}$.

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 | $n_i + 1$ | $n_i - 1$ | $n_j + 1$ | $n_j - 1$ |
|---|---|---|---|---|
| $n_i < n_j - 1$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$ | $\frac{1}{1+e^{-\beta c}} p_{i1}$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$ | $\frac{1}{1+e^{-\beta c}} p_{j1}$ |
| $n_i = n_j - 1$ | $\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{i0}$ | $\frac{1}{1+e^{-\beta c}} p_{i1}$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$ | $\frac{1}{1+e^{\beta(0.5b-c)}} p_{j1}$ |
| $n_i = n_j$ | $\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{i0}$ | $\frac{1}{1+e^{\beta(0.5b-c)}} p_{i1}$ | $\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{j0}$ | $\frac{1}{1+e^{\beta(0.5b-c)}} p_{j1}$ |
| $n_i = n_j + 1$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$ | $\frac{1}{1+e^{\beta(0.5b-c)}} p_{i1}$ | $\frac{e^{\beta(0.5b-c)}}{1+e^{\beta(0.5b-c)}} p_{j0}$ | $\frac{1}{1+e^{-\beta c}} p_{j1}$ |
| $n_i > n_j + 1$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{i0}$ | $\frac{1}{1+e^{-\beta c}} p_{i1}$ | $\frac{e^{-\beta c}}{1+e^{-\beta c}} p_{j0}$ | $\frac{1}{1+e^{-\beta c}} p_{j1}$ |

and the probability of staying in state $n$ equals 1 minus the above probabilities of leaving that state. We can then show:

[23] These two limits are in general *not* interchangeable, even in closely related games, such as a discrete public good game (Blume 1995, Diermeier and Van Mieghem 2000). So, a model with vanishing noise may have very different properties from a model with unperturbed best response.

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

[25] For example, turnout may be affected by bad weather.

**Proposition 7** *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 for the log-logistic model for any $\beta$ converge to a Dirac impulse at*

$$x_D = x_R = \frac{e^{-\beta c}}{1 + e^{-\beta c}}. \tag{10}$$

Proof: 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:

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

Using the continuum approximation $p(x)$ for $\pi$ 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, integrality implies that $p$ must be zero everywhere except at $u_i = 0$, where it thus must have a Dirac impulse of measure 1.■

Notice that the equilibrium distribution is stochastically decreasing in $c$. Also, the equilibrium distribution converges to the best-response limiting distribution for $\beta \to \infty$, regardless of $c$.[26] 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. Indeed as $\beta \to \infty$ we recover the best response model with zero turnout.[27] The key insight from the (unperturbed) best response model is still valid: If information about current voting behavior is perfect, voters almost surely coordinate (in the limit as $\beta \to \infty$) on a single state. Moreover, expected participation is small whenever voting factions are not of exactly the same size.

In addition to providing a robustness check of best-response dynamics, the log-logistic formulation allows us to study "spontaneous" coordination through polls in more detail. Consider Figure 4, which shows the limiting distribution of turnout in an example with $N_D = N_R = 50$ for various values of $\beta$. The minimal value $\beta = 0$ corresponds to pure random choice. The turnout distribution thus is a Gaussian mountain with maximum turnout likelihood at (25,25) = (50%, 50%). As $\beta$ 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 $\beta$ are non-linear: as $\beta$ increases, voters coordinate on *smaller* turnouts, which are consistent 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 $\beta$ 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%. At a critical value $\beta^C$ between 95 and 96, however, spontaneous coordination at the (100%,100%)

[26] This already holds for finite $N$.

[27] So, in the turnout participation game the two limits ($t \to \infty$) and ($\beta \to \infty$) *are* interchangeable.

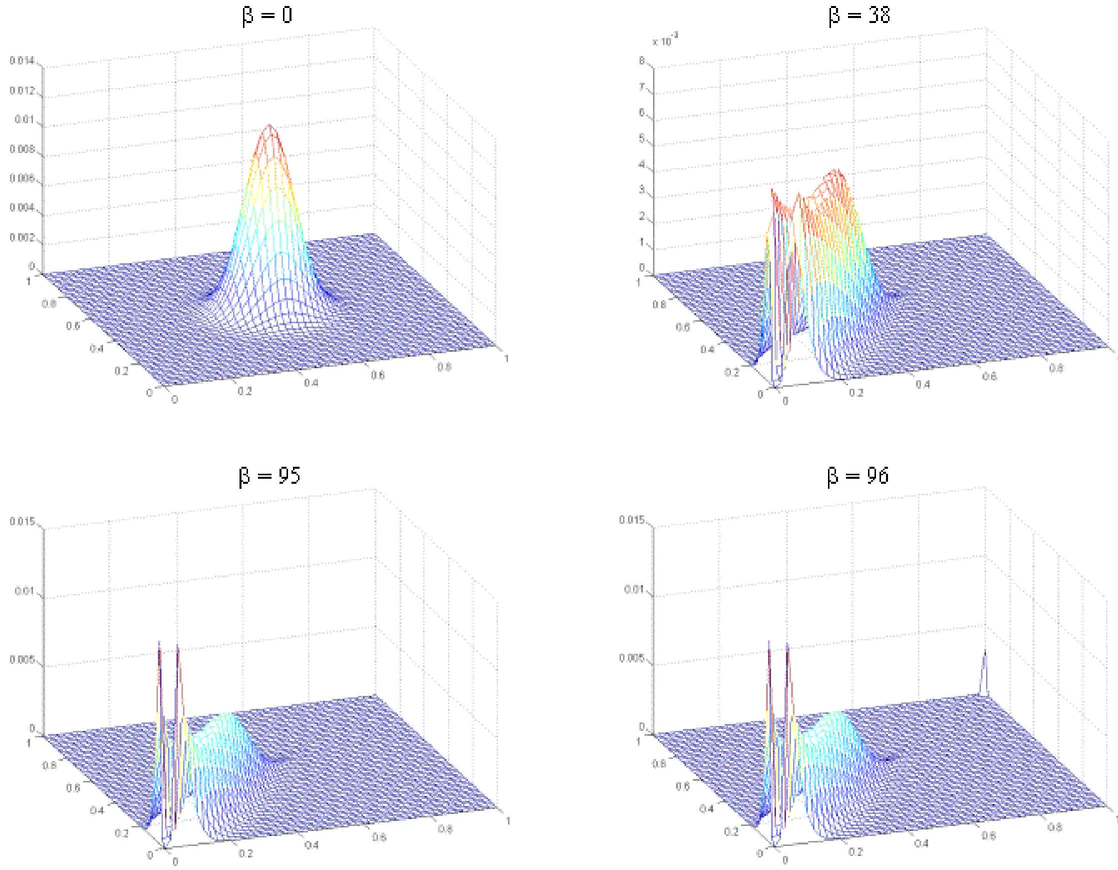


Figure 4: The limiting turnout density $\pi$ as a function of the two turnout fractions and the parameter $\beta$ for an example with $N_D = N_R = 50$ and cost-benefit ratio $\xi = 0.1$.

outcome suddenly 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.[28] For low $\beta$ noise prevails, while at higher $\beta$ two low-turnout states that are each other's mirror image (or differing only in "spin") are equally likely. During a $\beta$-interval starting at $\beta^C$, two "phases", i.e the low turnout phase (with two most-likely states, differing only in "spin") and the full turnout phase (with one most likely state), can be balanced. Finally, as $\beta$ increases beyond $\beta^C$, the low turnout phase becomes less likely, and ultimately, the full turnout phase prevails with probability 100%.[29]

[28]The threshold $1/\beta^C$ plays a role similar to the Curie temperature in models of "spontaneous magnetization," which is 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 2×2 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.

[29]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. So, at least from a stochastic point of view, the system is better behaved in the case where there is *no* Nash equilibrium in pure strategies.

# 6 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 demonstrate how voters can spontaneously coordinate their actions through polls.

Overall our results indicate a potentially important role for stochastic models in voting environments, 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, the model should also include candidates as strategic actors.

# 7 Appendix - 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 $\pi$. 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, $\pi$ 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.[30]

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.[31]

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 $j > \overline{j} := N_D + w$:

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

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 $\pi$ 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

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

[30] A simple personal computer with 128MB of RAM can solve a linear system with a few thousand unknowns. Hence, with $N_1 N_2 \simeq 4000$, one solves exactly for populations $N_i \simeq 200$.

[31] Hence, using this recursive formulation our simple personal computer can solve populations of size $N_i \simeq 2000$ exactly.

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:

$$\begin{aligned}
(N_R - j + 1)\pi(i, j-1) + (j+1)\pi(i, j+1) &= N_D \pi(i, j), \\
\Leftrightarrow (N_R - i - w)\pi(i, i+w) + (i+w+2)\pi(i, i+w+2) &= N_D \pi(i, i+w+1) \\
\Leftrightarrow (N_R - i - w)\left[\sum_{j=0}^{i-1} U_{ij}^{+} u_j + \sum_{j=w+1}^{i-1} L_{ij}^{+} l_j\right] + (i+w+2)\sum_{j=i+1}^{N_1} U_{ij} u_j &= N_D u_i.
\end{aligned} \tag{11}$$

The lower strip boundary yields:

$$\begin{aligned}
(N_D - i + 1)\pi(i-1, j) + (i+1)\pi(i+1, j) &= N_R \pi(i, j), \\
\Leftrightarrow (N_D - i + 1)\pi(i-1, i-w-1) + (i+1)\pi(i+1, i-w-1) &= N_R \pi(i, i-w-1) \\
\Leftrightarrow (N_D - i + 1)\left[\sum_{j=0}^{i-2} U_{i-1,j}^{-} u_j + \sum_{j=w+1}^{i-2} L_{i-1,j}^{-} l_j\right] + (i+1)\sum_{j=i+2}^{N_1} L_{i+1,j} l_j &= N_R l_i.
\end{aligned} \tag{12}$$

Equations (11)–(12) 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.[32] This technique exploits the ergodic properties of the process, i.e., the fact that $\pi_j$ 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 $\pi$ 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.

[32] 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!

# 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. 2000. "A Behavioral Model of Turnout". Mimeo. Graduate School of Business. Stanford University.

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

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

[5] 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.

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

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

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

[9] 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.

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

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

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

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

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

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

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

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

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

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