

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 voters²⁶ 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

²⁶For 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.

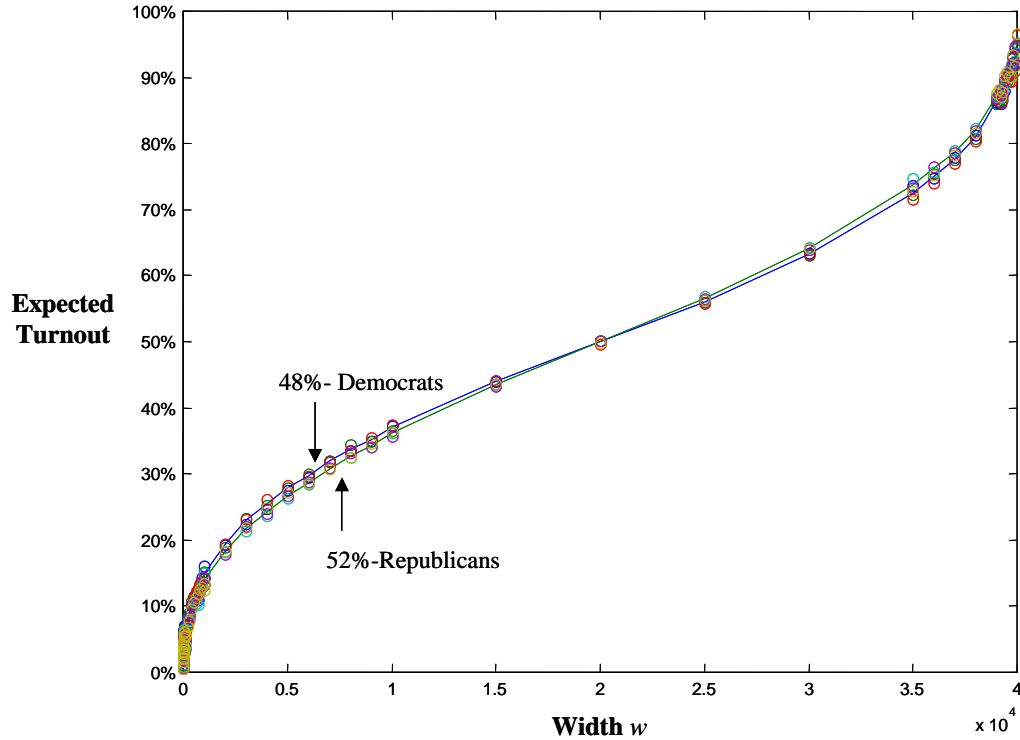


Figure 6 The expected turnout fractions as a function of the width w of the birth zone for an electorate of size 1 million with 48% Democrats and 52% Republicans.

Myerson (1998), which was constructed to demonstrate the strikingly low expected turnout predicted by game-theoretic models. In his case the voting fractions 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 model²⁷ 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-

²⁷The 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.

²⁸Recall 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.

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

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

³⁰It 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.

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

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

population size increases, substantial turnout requires higher noise and lower cost/benefit ratios³³. 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 that³⁵:

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

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 $\xi(N) \leq (1 + \frac{3}{50}N)^{-1}$.

To see how binding this constraint is, consider an electorate with $N = 3$ 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:

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

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

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

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 $N \rightarrow \infty$. 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

³³While 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.

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

³⁵The notation $O(f(x))$ describes the behavior for large x . Formally, $O(f(x))$ denotes any function $g(x)$ such that $\lim_{x \rightarrow \infty} g(x)/f(x) = 1$. 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 ($N \rightarrow \infty$) tends to the probability density function $p(x)$, where $x_i = n_i/N_i$. The density p solves the following partial differential equations:*

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

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 $\tilde{n}(n) = n$. 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

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

³⁷In 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 ($N = N_D + N_R \rightarrow \infty$) 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

³⁸Recall 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 choice³⁹ 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 $\beta \rightarrow \infty$, log-linear choice converges to a distribution that puts positive probability only on best-responses to n_t^{-k} .⁴¹

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

⁴⁰Alternatively, 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.

⁴¹For 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 ($N = N_D + N_R \rightarrow \infty$) 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 $\beta \rightarrow \infty$, 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 $\beta \rightarrow \infty$) we recover the best response model with zero turnout (and “close” elections in the sense that n_D/N and $n_R/N \rightarrow 0$ 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.

⁴²This already holds for finite N .

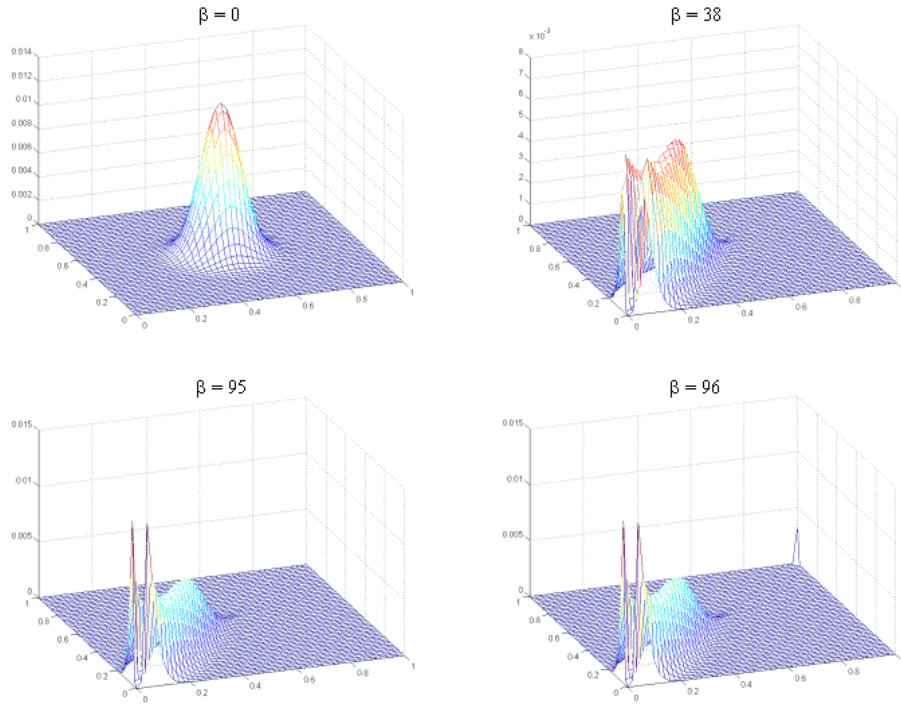


Figure 7 The limiting turnout density π for the log-logistic model as a function of the two turnout fractions and the parameter β for an example with $N_D = N_R = 50$ and cost-benefit ratio $\xi = 0.1$.

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

⁴³Note that in the turnout game the two limits ($t \rightarrow \infty$) and ($\beta \rightarrow \infty$) 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

⁴⁴The 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.

⁴⁵In 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.

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):

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

Thus:

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

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 $w \leq 2\epsilon$ and

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

■

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

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

and associated maximal width is:

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

Similarly, w reaches its minimal value 0 when

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

or when

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

■

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:

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

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:

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

Hence, for $N \rightarrow \infty$, 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

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

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 $N \rightarrow \infty$, we have:

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

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 $N \rightarrow \infty$, 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 $\theta = 45^\circ$, 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:

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

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

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

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 $\{i, i + 1, \dots, j\}$ if $i < j$ and $i : j = \emptyset$ otherwise.:

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

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

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

⁴⁶A 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 $N_i \simeq 89$.

⁴⁷Hence, 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 $u_j, u_{j+1}, \dots, u_{N_D}$. 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:

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

The lower strip boundary yields:

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

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,

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

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 I. 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.

⁴⁸The 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 Shabar. 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.