Mandate and paternalism: A theory of large elections ☆

Marco Faravelli [a,*], Priscilla Man [a], Randall Walsh [b]

[a] School of Economics, University of Queensland, Australia
[b] Dept. of Economics, University of Pittsburgh & National Bureau of Economic Research, United States

A B S T R A C T

We propose a game theoretic costly voting model of large elections that incorporates the assumption that mandate matters. This innovation is motivated by empirical evidence that US Representatives with larger victory margins on average vote in a more partisan manner. If voters are paternalistic, this new model predicts strictly positive limiting turnout rates as the population grows arbitrarily large. The model also preserves stylized comparative statics results of costly voting models, including the underdog effect and the competition effect. Finally, we develop an innovative computational strategy to solve the model for large, finite, electorates and show that our results are not only qualitatively, but also quantitatively relevant.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

It is natural to think of elections as winner take all contests. Be it a ballot for a member of parliament or a presidential election, we typically believe that the winner will have her own way, while the loser will have no say. Consequently, it is equally customary in voting theory to model elections as winner take all tournaments, where only the candidate who receives the most votes wins the prize. However, in reality elections are more complex and mandate plays a role in shaping the policy implemented by the winner. This notion has received attention in the political science literature and various explanations for the presence of a mandate effect have been put forward.[1] In this paper we argue that the policies adopted by elected politicians are closer to the "center" when their margin of victory is smaller and more "extreme" when their seats are won with a landslide victory. We present empirical evidence from US Congressional voting records in support of this thesis and propose a new theory of large elections that takes mandate into account. While one may be tempted to think of the winner take all assumption as an innocuous simplification, we show that a small deviation from the standard approach generates different and more realistic predictions.

One of the differences generated relates to the "paradox of voting", first identified by Anthony Downs (1957) in the original formulation of his rational choice theory. In a two-candidate winner take all election in which citizens pay a cost to vote, turnout rate is predicted to quickly converge to zero as the electorate grows. The paradox remains whether it is cast in

---

[1] See Smirnov and Fowler (2007) for a survey.

decision theoretic terms, as in Downs, or in game theoretic terms used by recent authors (Ledyard, 1981, 1984; Palfrey and Rosenthal, 1983, 1985). Notably, the assumption that the policy implemented depends on the margin of victory provides a substantial step towards the resolution of the paradox of voting. We show that, under this more realistic assumption, turnout converges to zero at a drastically slower rate than under the winner take all assumption. If we further assume that voters are paternalistic — i.e., they derive spillover benefits from the impact that a policy has on other individuals — our model delivers a strictly positive limiting turnout rate as the size of the electorate grows without bound.[2] In other words, the paradox of voting disappears. Conversely, turnout converges to zero in a winner take all framework even with paternalistic voters.

This positive limiting turnout result is driven by a delicate balance between the significance of one vote on the mandate and the magnitude of the paternalism spillover. The rates at which these two effects converge to zero and infinity, respectively, balance out in equilibrium, prompting a positive proportion of the electorate to vote. We demonstrate that this positive proportion need not be one. We show that, given any policy rule (a mapping from vote share to policy outcome), the limiting turnout is strictly less than one whenever voters' paternalism falls below some threshold.

Further characterizing the equilibrium, we show that under mild symmetry assumptions (so that supporters of neither party are ex-ante more motivated to vote), we obtain the "underdog effect" — that members of the minority party turn out to vote at higher rates than those of the majority; as well as the "competition effect" — that closer elections generate higher turnout.[3] These effects have been well-documented, both in laboratory experiments and empirically in large elections (Levine and Palfrey, 2007; Shachar and Nalebuff, 1999; Blais, 2000). However, theoretical predictions for such effects have been notoriously difficult to derive in game theoretical models of large elections (Taylor and Yildirim, 2010). This is largely due to the fact that most of these models predict zero turnout in large elections, thereby stripping off the possibility for meaningful limiting turnout analysis. Our ability to predict strictly positive turnout in large elections enables us to generate these predictions matching stylized empirical findings in a game theoretical framework without reliance on preference for voting or coordination devices.

The limiting results are of interest not because the population is indeed infinite, but because the limiting case is an approximation of a large election. For this reason we also take a numerical approach to solve models of large, finite elections. This has typically been computationally hard due to the need for computing binomial probabilities. We overcome this difficulty through the concentration properties of binomial distributions (cf., Hoeffding, 1963), which allow us to drastically reduce the number of terms to be computed. On a technical note, given that our computational method is not specific to smooth policy rules, we believe that it might be useful for other researchers in this field. We use this approach to provide the solution to our model for a population of 75 000 voters and for several policy rules and different parameters. This exercise is useful in three ways. First, we show that the limiting turnout is an excellent approximation (to the fourth decimal digit) of the turnout with a finite electorate of this size. Second, the exercise allows us to check our theoretical predictions against empirical observations. We calibrate the policy rule using the aforementioned US Congressional voting data and calculate the predicted turnout using our model. Under plausible preference parameters, the predicted turnout rates fit well with the empirical observations. Third, we highlight the crucial role of relaxing the winner take all assumption. Computing the outcome for a relatively close election (which favors the winner take all rule), we calculate the turnout rate under a proportional policy rule with a low paternalism factor and find that, under a winner take all policy rule with 75 000 voters, the paternalism required to achieve the same turnout rate is 200 trillion times higher.

The game theoretic literature on costly voting is extensive.[4] Nonetheless, the prediction that turnout converges to zero as the size of the electorate grows has remained typical.[5] Part of the literature has attempted to overcome this result by assuming a preference for the act of voting itself (Riker and Ordeshook, 1968). We maintain that voting is instrumental. Another part of the literature assume the existence of some coordination mechanism (such as a leader who can mobilize supporters) that effectively reduces the voting game to one with a small number of players (Harsanyi, 1977, 1992; Morton, 1991; Shachar and Nalebuff, 1999). One notable example is the "ethical voter" model by Feddersen and Sandroni (2006), in which a proportion of the voters receive a benefit for following the voting strategy that would be adopted by a social planner whose party preference is aligned with hers.[6] In this strand of literature, voters within the same party do not interact strategically among themselves. Our model departs from this setting by assuming that all voters interact with each other strategically, thereby allowing voters within the same party to free-ride on each other.

Two concurrent and independent papers by Evren (2012) and Myatt (2012) tackle the paradox of voting by introducing uncertainty to a winner take all election with paternalistic voters.[7] Evren introduces uncertainty over the proportion of other-regarding voters. Myatt, on the other hand, introduces uncertainty over the beliefs on voters' party alliance. These

---

[2] To be precise, by "strictly positive limiting turnout rate" we mean that the proportion of potential voters who turn out is bounded away from zero — as opposed to the case where the absolute number of voters is bounded away from zero, but the proportion goes to zero.

[3] The terms underdog effect and competition effect are introduced by Levine and Palfrey (2007).

[4] See, for instance, Campbell (1999), Börgers (2004), Goeree and Grosser (2007), Krasa and Polborn (2009), Krishna and Morgan (2012). The empirical implications of the pivotal voter model in small elections have been examined by Coate et al. (2008). There are also information aggregation voting models that analyze voters' participation. See McMurray (2013).

[5] E.g., see Taylor and Yildirim (2010).

[6] Coate and Conlin (2004) use a version of this model to explain turnout in the Texas liquor referenda.

[7] Paternalism also appears in Feddersen and Sandroni (2006) in the sense that an ethical voter projects her preferred election outcome onto the entire measure of voters. Myatt (2012) does not incorporate paternalism in his basic model, but his Proposition 7 considers paternalistic voters.
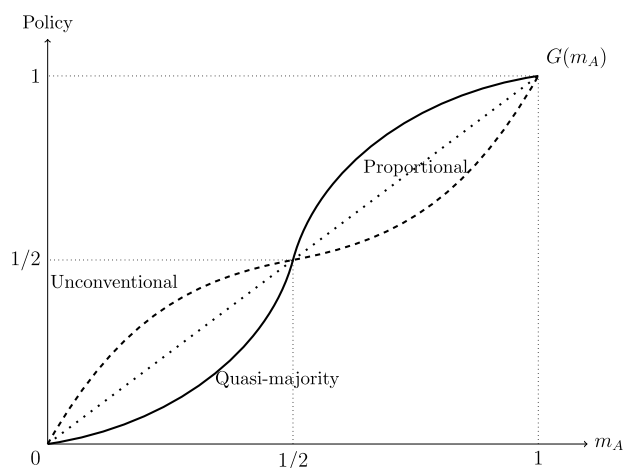
Fig. 1. Examples of smooth policy rule.

different forms of uncertainty "smooth" the expected utility of each voter in the sense of slowing the rate at which the benefit from voting approaches zero, just as mandate does in our model. Thus Evren's, Myatt's and our models present three different (and mutually compatible) approaches to resolve the paradox of voting. It is noteworthy that our model does not assume any uncertainty over a standard costly voting model, suggesting that significant election turnout can be compatible with models without additional uncertainty.

Closest to our work are the papers by Castanheira (2003) and Herrera et al. (2014b) that compare voters' behavior under "power sharing rules" to winner take all elections.[8] Both papers show that, under their particular form of power sharing rule (Castanheira uses a quadratic loss function and a weighted average policy rule between the two parties, with the weights being smooth functions of the mandate; Herrera et al. use a simple proportional rule[9]), the benefit from voting is inversely proportional to the expected number of voters. Both papers conclude from this result that turnout converges to zero at a slower rate under power sharing than in a winner take all election (unless the winner take all election is of the knife-edge case where the two parties are exactly symmetric). Our paper generalizes their result to account for *any* form of mandate consideration. The advantage of this generalization is three-fold. First, when there is no paternalism, our result (see Corollary 1) can be viewed as an extension of the famous zero turnout result of Palfrey and Rosenthal (1985) to a very general class of policy rules. Second, we demonstrate that the limiting underdog and competition effect do not depend on the specific functional form of the electoral rule, nor on the specific distribution of voting costs. Finally, proofs of the convergence rate in both Castanheira and Herrera et al. rely on the functional form they impose. In particular, Herrera et al. utilize a closed form marginal benefit function. Our proof adopts instead a general statistical inequality (Hoeffding, 1963), freeing us from the reliance on specific functional forms. This is a solid technical advancement for the literature on power sharing.

2. Mandate matters: empirical evidence

The notion that the margin of victory matters has been discussed in the political science literature and different explanations have been proposed (see Smirnov and Fowler, 2007). The two leading arguments are that candidates may find an extreme partisan position risky for the next election unless they are supported by a large mandate and/or the margin of victory may signal the position of the center of the electorate, toward which candidates lean. Formally, Razin (2003) proposes a model of "responsive candidates" who infer information on an imperfectly observed common preference shock from the election result, and position themselves optimally given this inference.

We do not intend to offer a theoretical foundation for why elected officials respond to their mandate. However, we do wish to provide an interpretation of this crucial assumption and provide empirical support for the presence of a mandate effect. For simplicity, suppose that the policy position of a given elected official can be mapped into the interval $[0, 1]$ and that members of party A strictly prefer outcomes closer to 1, while members of party B strictly prefer outcomes closer to 0. Assume that, if elected, candidates respond to the margin of victory in terms of their policy positions. Thus, in a two-party election, the proportion of votes received by party A can be mapped into a policy outcome through an increasing function G — which we assume to be smooth. This G function can be viewed as a "policy rule". We plot several possible policy rules against the mandate of party A in Fig. 1.[10] The dotted line presents a "proportional rule" under which the winning

---

[8] Relatedly, Faravelli and Sanchez-Pages (2014) and Kartal (2014) compare welfare under plurality and proportional representation.

[9] Herrera et al. (2014b) also consider Tullock contest success function power sharing rules but do not provide closed form analytical expressions for them.

[10] All three policy rules in Fig. 1 are symmetric around 1/2. This is not necessary for obtaining our positive limiting turnout result. See Section 3 for the precise restrictions on the G function.

Table 1

Estimations of the effect of margin of victory on degree of partisanship, US 105th to 111th Congress[a].

Congress	105th	106th	107th	108th	109th	110th	111th	105th–111th	105th–111th
Victory margin	0.1368***	0.0572**	0.1754***	0.0599**	0.1431***	0.2443***	0.2604***	0.1507***	0.0746***
	(0.0338)	(0.0263)	(0.0343)	(0.0271)	(0.0275)	(0.0257)	(0.0299)	(0.0111)	(0.0095)
Years in office	−0.0008	−0.0010	−0.0007	−0.0013	−0.0001	0.0001	0.0011	−0.0002	0.0049***
	(0.0010)	(0.4261)	(0.3867)	(0.0009)	(0.0009)	(0.0008)	(0.0009)	(0.0004)	(0.0006)
Candidate FE	·	·	·	·	·	·	·	No	Yes
Congress FE	·	·	·	·	·	·	·	Yes	Yes
Marginal effect[b]	0.2011	0.1060	0.2452	0.1082	0.2479	0.4338	0.3981	0.2473	0.1224
N	423	431	428	430	431	424	432	2999	2999
No. legislators	423	431	428	430	431	424	432	753	753
Partisanship									
Mean	0.4253	0.4426	0.4496	0.4661	0.4793	0.4881	0.5017	0.4647	0.4647
S.D.	0.1498	0.1551	0.1778	0.1439	0.1445	0.1465	0.1705	0.1578	0.1578
Victory margin									
Mean	0.3033	0.4264	0.3953	0.4253	0.4070	0.3603	0.3740	0.3847	0.3847
S.D.	0.2201	0.2876	0.2487	0.2597	0.2503	0.2602	0.2608	0.2590	0.2590

[a] The dependent variable in each regression is the degree of partisanship which is constructed using Poole and Rosenthal's DW-Nominate Scores (see text). Numbers in parentheses are standard errors. *, ** and *** indicate statistical significance at 10%, 5% and 1% levels, respectively.

[b] Measures the change in partisanship (expressed in standard deviations) associated with a one standard deviation increase in the margin of victory.

politician responds to the mandate at a constant rate. It approximates a proportional representation democracy. The solid line represents a "quasi-majority rule" under which the winning politician responds to the mandate at a decreasing rate. It approximates policy outcomes under direct election regimes such as the US Congress. Notice that winner take all majority rule can be viewed as the limiting case of a quasi-majority rule when G becomes a step function. The dashed line is an "unconventional rule" under which the winning politician responds to the mandate at an increasing rate. It is rarely observed empirically but is still admitted by the model.

Do politicians adopt a smooth policy rule? There are significantly little empirical evidence on this subject. Fowler (2005) shows that candidates in US Senate races respond to increases in previous election's Republican vote by adopting a more conservative position in the current race. Somer-Topcu (2009) finds that political parties respond to declining vote shares by changing the policies they support using a sample of 23 Democracies between 1945 and 1988 from the Comparative Manifesto Project. Peterson et al. (2003) proxy "mandate" to the winning party by analyzing newspaper coverage in US Presidential and off-year election and provide empirical evidence that Members of Congress deviate from their historical voting pattern in the direction of the mandate following a "mandate" election.

To provide more direct support for the presence of a mandate effect, we analyze the impact of margin of victory on the voting behavior of US Members of Congress for the 105th (1997–1998) through the 111th Congress (2009–2010). We identify the political position of a given Member of Congress in a given year by the first dimension of her DW-Nominate Score (Poole and Rosenthal, 2001).[11] This component of the DW-Nominate Score measures liberal–conservative leaning based on all roll-call votes during each Congressional Session. We construct the degree of partisanship for each Congressperson in each year as the distance between her DW-Nominate Score and the mid-point between the scores of the most liberal Republican and the most conservative Democrat (expressed in terms of absolute value).

We then regress the degree of partisanship on the margin of victory in the election in which the Congressperson won her seat, as well as on years in office. The first seven columns of Table 1 report the results for each of the seven included sessions. The eighth column pools the data for the seven congressional sessions into a single regression with the incorporation of congress-level fixed effects. In all cases the coefficient on margin of victory is highly statistically significant. To provide a more intuitive interpretation of the results, we also report the predicted marginal effect associated with a one standard deviation increase in the margin of victory − expressed in terms of standard deviations in partisanship. For example, in the 105th Congress, a one standard deviation increase in the margin of victory will be associated with a 0.2 standard deviation increase in degree of partisanship. The estimated marginal effects across all seven congresses range between 0.11 and 0.44, with a pooled estimate of 0.25.

Of course, these results do not necessarily imply a direct causal link. One alternative explanation is that more ideologically extreme candidates select into districts with larger partisan majorities. However, even in the presence of such a selection process, as long as election outcomes can be considered as the most direct measure of a given district's voter ideology, voters would still benefit by increasing their supported candidate's margin of victory as it would eventually lead to preferred policy outcomes through the recruitment of candidates with preferable policy positions.

---

[11] See Poole and Rosenthal (2001) for details on the construction of the DW-Nominate Score. A complete set of scores for the 1st through 111th Congress are available online at http://pooleandrosenthal.com/dwnominate.asp. We restrict our analysis to the 105th to 111th Congress because of the limited availability of digitized data in the earlier periods.

Nonetheless, we also use the panel structure of the data to investigate the direct, within representative, connection between margin of victory and partisan voting behavior. The final column of Table 1 reports the results of regressing the degree of partisanship on our explanatory variables with both Congress fixed effects and Congressperson fixed effects, using the pooled data from the seven Congresses. This allows us to evaluate how voting behavior of a given candidate evolves in response to changes in her margin of victory. Again the coefficient on the margin of victory is significant at the 1% level, with an estimated marginal effect of 0.12. Thus, even when we focus solely on the within candidate variation in ideological voting patterns, there is clear evidence supporting the assumption that mandate matters.

3. Model

We model an election with $N + 1$ citizens and two parties, A and B. Denote a generic citizen by i and a generic party by P. Each citizen has, independently, the same ex-ante probability $\alpha \in (0, 1)$ of being a supporter of party A and $1 - \alpha$ of supporting party B. Citizens choose simultaneously to vote for the candidate of party A, the candidate of party B or to abstain.[12] If citizen i votes (instead of abstaining), she bears a voting cost of c_i. The voting costs of citizens are identically and independently drawn from a non-degenerate interval $C \subseteq \mathbb{R}_+$ according to the cumulative distribution function F with a density function f, where $f(c) > 0$ for all $c \in C$. Let $\underline{c} = \inf C$. We assume that $\underline{c} < 1/2$ to eliminate uninteresting equilibria in which nobody votes simply because all voting costs are too high.[13] Each citizen's party alliance and voting cost are her private information, but their distributions are commonly known.

Let n_A and n_B be the numbers of votes cast for parties A and B respectively. Party P's mandate is equal to the proportion of the votes received by its candidate, except when no one votes, in which case each party receives a mandate of $1/2$. Thus,

$$m_P(n_P, n_Q) = \begin{cases} \frac{n_P}{n_P + n_Q} & \text{if } n_P + n_Q > 0 \\ \frac{1}{2} & \text{if } n_P + n_Q = 0 \end{cases}, \tag{1}$$

where $P, Q \in \{A, B\}$ and $P \neq Q$. To operationalize the mandate effect we assume that each member of party P receives a direct benefit as a result of her party's mandate. This benefit is denoted by $G(m_P)$ where $G : [0, 1] \to [0, 1]$ is a continuous, strictly increasing, twice differentiable function whose first derivative is bounded away from 0 and infinity.[14]

Citizens may be paternalistic. If a supporter of party P receives a gross benefit of b from the election, she will also enjoy a paternalism benefit equal to $\gamma_P b$ for each other citizen in the society,[15] where $\gamma_P \geq 0$. We label the case of $\gamma_P = 0$ for $P = A, B$ as "no paternalism", while $\gamma_A, \gamma_B > 0$ is labeled as paternalism. All payoff information is common knowledge.

Let i be a member of party P with voting cost c_i. Let v_P and v_Q be the number of votes cast for parties P and Q, both exclusive of i. Then i's utility is

$$[1 + \gamma_P N]G(m_P(v_P + 1, v_Q)) - c_i \quad \text{if } i \text{ votes for party } P;$$

$$[1 + \gamma_P N]G(m_P(v_P, v_Q + 1)) - c_i \quad \text{if } i \text{ votes for party } Q; \text{ and}$$

$$[1 + \gamma_P N]G(m_P(v_P, v_Q)) \quad \text{if } i \text{ abstains.}$$

As specified, the model defines a Bayesian game. A pure strategy for an individual is a pair of (measurable) voting rules, (s_i^A, s_i^B), where $s_i^P : C \to \{A, B, \text{Abstain}\}$ assigns a choice for each voting cost when i belongs to party P. Notice that for each party affiliation and voting cost realization, voting for the opposite party is always *strictly* worse than abstaining regardless of the number of votes cast for the own and opposite party. Thus at no voting cost would any best-responding citizen vote for the opposite party. In addition, conditional on party alliance and other citizens' voting rules, the difference between the expected payoffs from voting for own party and abstaining is strictly decreasing in individual i's cost of voting. Thus any best response of citizen i to any profile of other citizens' strategy must be monotone — i.e., for each party P, there exists a cut-off cost c_P such that when i is a P-supporter, she votes for P for all $c < c_P$ and abstains for all $c > c_P$.

In our analysis, we consider symmetric Bayesian–Nash equilibria in which all citizens use the same pair of voting rules, i.e., $(s_i^A, s_i^B) = (s^A, s^B)$ for all i. By our argument above, this equilibrium pair of voting rules involves two cut-off costs (c_A, c_B).[16] Define the turnout of party P as the probability that a randomly chosen supporter of P votes. In equilibrium, the turnout of party P is $F(c_P)$. Since there is a one-to-one relationship between c_P and party P's turnout, we will refer to an equilibrium solely by $(a, b) = (F(c_A), F(c_B))$.

---

[12] For ease of exposition, we use the terms "party P" and "candidate of party P" interchangeably.

[13] In the standard winner take all costly voting model, $\underline{c} < 1/2$ ensures that a voter has a strictly positive probability of voting when she knows she is pivotal with probability 1.

[14] One interpretation of the model's "preference for mandate" is that it captures the behavioral notion that voters simply enjoy winning with a larger mandate. However, we prefer to consider the model within the context of our previous discussion and interpret $G(.)$ as representing the process through which larger voting majorities are translated into increasingly partisan (and thus more preferred) policy outcomes.

[15] In order to obtain our limiting results (Theorem 2), the paternalism factor has to be asymptotically linear in N. However, given a fixed N (as in our computation section), all we need is the value of the paternalism factor at the chosen N to be strictly positive.

[16] Strictly speaking, an equilibrium should also specify the actions chosen at the two cut-off points. However, $c_i = c_A$ and $c_i = c_B$ are zero probability events. The actions at these two points would not affect ex-ante calculations. Hence we will only describe an equilibrium by the cut-offs.

4. Mandate and paternalism

When N is finite, an equilibrium exists by the Brouwer Fixed-Point Theorem[17]:

Proposition 1. *For finite N, a pure strategy symmetric monotone Bayesian–Nash equilibrium exists.*

Let (a_N, b_N) be an equilibrium turnout profile in an election with $N + 1$ citizens.[18] For each N, let $T_N = \alpha a_N + (1 - \alpha)b_N$ be the ex-ante equilibrium overall turnout rate (i.e., turnout unconditional on party alliance). Notice that $\{(a_N, b_N)\}_N \subset [0, 1]^2$ possesses a convergent subsequence. Passing to subsequence if necessary, let $\{(T_N)\}_N$ be the corresponding convergent subsequence of equilibrium overall turnout.

We will first consider the case of no paternalism (i.e., $\gamma_P = 0$ for both parties P). Define the *expected gross marginal benefit of voting* to be the difference between the expected benefit from voting and from abstaining, both without taking into account the voting cost. The following theorem gives the convergent rate of the voting benefit along any (not necessarily equilibrium) sequence of turnout profiles such that the limiting turnout is strictly positive.

Theorem 1. *Assume no paternalism. Let $\{(a_N, b_N)\}_N$ be a convergent sequence of turnout profiles (not necessarily in equilibrium) such that $\lim_{N \to \infty} T_N > 0$. The expected gross marginal benefit of voting to any voter along this sequence converges to 0 at the rate of $1/N$ as N approaches infinity.*

Corollary 1. *Assume no paternalism. Any convergent sequence of equilibrium overall turnout rates $\{(T_N)\}_N$ converges to 0.*

Proof. Suppose not. Then by Theorem 1 the expected gross marginal benefit of voting to any voter goes to 0 as N goes to infinity. But this means $(a_N, b_N) \to (0, 0)$. □

Corollary 1 can be viewed as an extension of the zero turnout result of Palfrey and Rosenthal (1985) to a general class of policy rules. However, Theorem 1 reveals a crucial difference between models with and without mandate concerns − had there been no mandate concern, the expected gross marginal benefit of voting would have converged to 0 exponentially (see the proof of Proposition 2). In other words, adding a mandate consideration slows down the vanishing rate of voting benefits (given a fixed turnout).

Theorem 1 is important for the paradox of voting despite it not being an equilibrium result. If we attempt to explain election turnouts by social preference (e.g., civic duty, other-regarding preference), the strength of such social preference required can be significantly smaller when there is a mandate consideration. Indeed, if $\gamma_A, \gamma_B > 0$, our model with mandate consideration predicts strictly positive limiting equilibrium turnout.

Theorem 2. *Suppose $\gamma_A, \gamma_B > 0$. Let $\{(a_N, b_N)\}_N$ be a convergent (sub)sequence of equilibrium turnout profiles and let (a^*, b^*) be its limit. Then $a^*, b^* > 0$.*

This is a consequence of Theorem 1, as the linearly increasing paternalism benefits counteracts the diminishing expected gross marginal benefit of voting (over abstaining), which vanishes exactly at the rate of $1/N$. Had the limiting turnout been zero, the large gross marginal benefit of voting would prompt citizens to vote, which cannot be an equilibrium.

Theorem 2 is not a mere consequence of the paternalism assumed. Indeed, if there is no preference for mandate, the limiting turnout is still zero even under our paternalism assumption.

Proposition 2 (No preference for mandate). *Suppose $\gamma_A, \gamma_B > 0$ and consider a winner take all election, i.e.,*

$$G(m_P) = \begin{cases} 0 & \text{if } m_P < 1/2 \\ 1/2 & \text{if } m_P = 1/2 \\ 1 & \text{if } m_P > 1/2. \end{cases}$$

If $\alpha \neq 1/2$,[19] then the limiting overall turnout is zero.

Proposition 2 reiterates the importance of Theorem 1 − that the rate of convergence matters. In both models with and without mandate concerns, the expected impact of a single vote on the outcome vanishes as the number of votes cast gets large, either because the pivotal probability is too small, or because the difference that a single vote makes on the mandate is too small. Yet as the expected benefit from voting under a preference for mandate decreases at the rate of $1/N$, a linearly

---

[17] The proofs of all theorems and propositions are presented in Appendices A and B.

[18] If there are multiple equilibria at N, just pick any one of them.

[19] The knife-edge case of $\alpha = 1/2$ is a well-known exception. In this case, the probability that each voter is pivotal vanishes at the rate of $1/\sqrt{N}$ in a symmetric equilibrium. With paternalism, the expected gross benefit goes to infinity.

increasing paternalism benefit is sufficient to balance it out. The same paternalism is insufficient for a positive limiting turnout in a winner take all election, where the expected benefit from voting vanishes exponentially.

Conversely, the limiting turnout rates need not be 1, which would be the case if the expected gross benefit from voting went to infinity with N. Indeed, the sequence of turnout profiles is interior in many settings.

Proposition 3. *If C, the support of voting cost, is unbounded from above, then $a_N, b_N < 1$ for all N. If C is bounded from above, then given α and G, there exist strictly positive γ_A and γ_B sufficiently small that $a_N, b_N < 1$ for all N sufficiently large.*

When the convergent sequence of equilibrium voting probabilities is interior, the limiting probabilities satisfy the following condition:

Proposition 4. *If $\{(a_N, b_N)\}_N \subset [0, 1)^2$ converges to (a^*, b^*), then (a^*, b^*) satisfies*

$$\gamma_A \frac{1 - m_A^*}{T^*} G'(m_A^*) = F^{-1}(a^*)$$

$$\gamma_B \frac{m_A^*}{T^*} G'(1 - m_A^*) = F^{-1}(b^*) \tag{2}$$

where $T^* = \alpha a^* + (1 - \alpha) b^*$ is the limiting total turnout (the probability that a citizen votes) and $m_A^* = \alpha a^* / T^*$ is the limiting mandate of party A.

Theorem 2 and Proposition 4 are statements about the limiting turnout only. They do not imply higher turnout for larger society. Nothing prohibits $\{a_N\}$ and $\{b_N\}$ from being decreasing sequences.

5. Numerical examples

Three questions remain: First, are the above limiting results good approximations for large but finite elections? Second, is our model giving quantitatively reasonable predictions? Third, is mandate making an economically significant difference when it comes to the paradox of voting? We answer these questions by performing a set of numerical exercises.

5.1. Parameterization

We consider three classes of G functions: the Winner take all Rule as described in Proposition 2, the Proportional Rule defined by $G(m) = m$, and the class of "Constant Absolute Risk Aversion" (CARA) G functions with parameter k, i.e.,

$$G(m) = \frac{\tilde{G}(m) - 1}{2(e^{k/2} - 1)},$$

where

$$\tilde{G}(m) = \begin{cases} e^{km} & \text{if } m \in [0, 1/2] \\ 2e^{k/2} - e^{k(1-m)} & \text{if } m \in [1/2, 1] \end{cases}.$$

The parameter $k = G'(m)/G(m)$ would be the coefficient of absolute risk aversion if G was a von Neumann–Morgenstern utility function. When $k \to 0$, G approaches the identity function (i.e., the proportional rule). When $k \to \infty$, G approaches the step function for winner take all election. One may therefore consider this class of functions as the policy rules "between" the winner take all rule and the proportional rule.

The size of the electorate (minus one), N, is set at 75 000. The limiting turnouts as $N \to \infty$ are also computed for comparison. Voting cost is assumed to be uniformly distributed on $[0, 1/2]$. We consider two values of the prior support for party A, α, 0.55 and 0.65. The former represents a close election, while the latter roughly matches the mean margin of victory in the empirical study in Section 2.[20] The two parties are assumed to have the same paternalism factor, that is, $\gamma_A = \gamma_B = \gamma$. We choose two (relatively low) values for γ: $\gamma = 0.05$ and $\gamma = 0.1$. All parameter values are summarized in Table 2. It should be noted, though, that our primary goal is to compare turnout rates under different policy rules at these parameter values, rather than to argue that these parameter values are necessarily "correct".

With $N = 75\,000$, solving for the exact equilibrium turnout is computationally heavy. The insight here is to note that the probability of a realization of a binomial random variable vanishes exponentially away from the mean. We can therefore truncate the distribution such that the expected benefit from the truncated tails is less than half of the approximation error allowed, which is set at 10^{-8}. The details of our computation strategy are provided in Appendix B.[21]

---

[20] Due to the underdog effect (see Section 6, Proposition 5), $\alpha = 0.55$ means that the margin of victory of party A is strictly smaller than 10%. In our numerical exercises, this margin is typically around 5%. Similarly, $\alpha = 0.65$ results in a margin of victory around 20%, which seems to be a typical victory margin from Fig. 2.

[21] The computation codes (written in Mathematica) are available from the authors upon request.
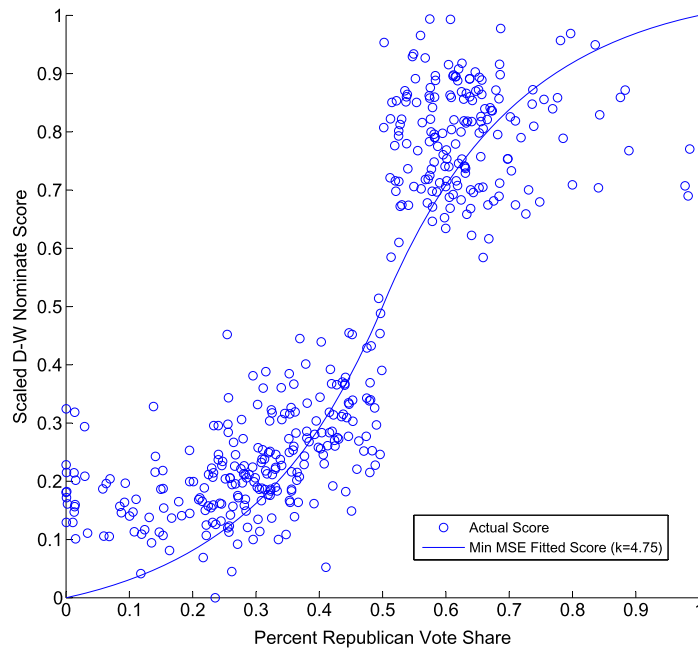
Fig. 2. Fitted CARA function for the US 2008 Congressional Election.

Table 2
Parameter values for numerical exercises.

Parameters	Notations	Value(s)
Population	N	75 000, $N \to \infty$ limit
Voting cost distribution	N/A	Uniform on $[0, 1/2]$
Prior support of A	α	0.55, 0.65
Paternalism factor	γ	0.05, 0.1
Absolute risk aversion coefficient	k	4.75, 10, 20
Approximation error allowed	ε	10^{-8}

5.2. Limiting approximation

We first check whether the $N \to \infty$ limiting case is a good approximation to large elections. To this end, we calculated the turnout rates when $N = 75\,000$ and $N \to \infty$ under the winner take all rule, the proportional rule and the CARA function with $k = 10$. Table 3 gives the results. In most cases, the difference is less than 0.01%, indicating that the limiting case is an excellent approximation to large elections of interests. The largest discrepancy occurs at the winner take all election with $\alpha = 0.55$, where the turnout rate at $N = 75\,000$ is about 4% while the limiting turnout is zero (Proposition 2). Nonetheless, a 4% turnout in 75 000-voter election is, politically and economically, not much different from a zero turnout.

However, the difference between the computation time required for the $N = 75\,000$ and the limiting cases is remarkable. Even for the proportional rule, which is computationally the lightest, the exact $N = 75\,000$ calculation is about 3000 times longer than the limiting approximation. For the CARA function, an $N = 75\,000$ exercise requires 80–110 hours while the limiting approximation takes at most 0.05 second. Thus, analyzing the $N \to \infty$ limiting case as an approximation for large elections saves a significant amount of analytical and computational costs without losing much precision.

5.3. A simple calibration

As the limiting approximation is an efficient tool for analyzing large elections, we would like to check whether our limiting model gives reasonable predictions. To this end we utilize the Congress roll-call votes record in the 111th Congress (2009–2010) as described in Section 2. We linearly transform the D-W Nominate scores so that they have a minimum of 0 and a maximum of 1. Then we fit a CARA function to the data by minimizing the mean squared error of the actual and predicted scores. This gives us $k = 4.75$, as plotted in Fig. 2. We then calculated the predicted limiting turnout rate under the CARA function with $k = 4.75$ for different combinations of values of α and γ. This gives overall turnout rates of about 30%–48% (see Column 3, Table 4).

If one thinks that Congresspersons' behavior in the 2009–2010 Congress will affect voters perception of their benefit function in the 2010 election (i.e., voters' perception of G is adaptive), we can compare our predicted turnouts to the actual

Table 3
Comparison between $N = 75\,000$ and limiting turnouts.

	Winner-take-all		Proportional		CARA $k = 10$	
$\alpha = 0.55$, $\gamma = 0.05$						
N	75 000	∞*	75 000	∞	75 000	∞
a	4.1261%	0	21.2960%	21.2933%	42.1586%	42.1475%
b	4.4115%	0	23.5436%	23.5406%	46.6046%	46.5959%
T	4.2545%	0	22.3074%	22.3046%	44.1593%	44.1493%
Computation time[†]	2′2″	N/A	1′1″	0.02″	80h11′24″	0.03″
$\alpha = 0.55$, $\gamma = 0.1$						
N	75 000	∞*	75 000	∞	75 000	∞
a	4.4598%	0	30.1150%	30.1132%	59.6149%	59.6056%
b	4.7683%	0	33.2934%	33.2914%	65.9033%	65.8965%
T	4.5986%	0	31.5453%	31.5434%	62.4447%	62.4365%
Computation time[†]	2′18″	N/A	1′2″	0.02″	91h43′48″	0.03″
$\alpha = 0.65$, $\gamma = 0.05$						
N	75 000	∞*	75 000	∞	75 000	∞
a	0.5880%	0	19.3815%	19.3791%	29.6294%	29.6201%
b	0.7228%	0	26.4126%	26.4092%	40.3740%	40.3654%
T	0.6352%	0	21.8424%	21.8396%	33.3900%	33.3810%
Computation time[†]	1′50″	N/A	1′15″	0.02″	111h47′45″	0.05″
$\alpha = 0.65$, $\gamma = 0.1$						
N	75 000	∞*	75 000	∞	75 000	∞
a	0.6237%	0	27.4078%	27.4062%	41.8972%	41.8892%
b	0.7666%	0	37.3505%	37.3483%	57.0922%	57.0853%
T	0.6737%	0	30.8878%	30.8859%	47.2155%	47.2078%
Computation time[†]	1′46″	N/A	1′17″	0.02″	109h51′29″	0.05″

* Theoretical prediction due to Proposition 2.
[†] Of a Mathematica code on a desktop computer with a 3.06 GHz Duo Core Processor and 4 GB of memory.

Table 4
Turnout rates at $N = 75\,000$.

	Winner take all	Proportional	CARA $k = 4.75$	CARA $k = 10$	CARA $k = 20$
$\alpha = 0.55$, $\gamma = 0.05$					
a	4.1261%	21.2960%	32.4656%*	42.1586%	52.4090%*
b	4.4115%	23.5436%	35.8920%*	46.6046%	57.9403%*
T	4.2545%	22.3074%	34.0075%*	44.1593%	54.8981%*
$\alpha = 0.55$, $\gamma = 0.1$					
a	4.4598%	30.1150%	45.9132%*	59.6149%	74.1175%*
b	4.7683%	33.2934%	50.7590%*	65.9033%	81.9400%*
T	4.5986%	31.5453%	48.0938%*	62.4447%	77.6376%*
$\alpha = 0.65$, $\gamma = 0.05$					
a	0.5880%	19.3815%	26.1326%*	29.6294%	28.4409%*
b	0.7228%	26.4126%	35.6127%*	40.3740%	38.7585%*
T	0.6352%	21.8424%	29.4506%*	33.3900%	32.0521%*
$\alpha = 0.65$, $\gamma = 0.1$					
a	0.6237%	27.4078%	36.9571%*	41.8972%	40.2216%*
b	0.7666%	37.3505%	50.3640%*	57.0922%	54.8127%*
T	0.6737%	30.8878%	41.6495%*	47.2155%	45.3285%*

* Limiting approximation.

turnout rate in the 2010 US Congressional election, which is 42% among all registered voters and 38% among the voting age population.[22] Both numbers fit remarkably well with our predicted turnouts. If instead one believes that voters in the 2008 election correctly forecast the Congresspersons' behavior in the 2009–2010 Congress, we can compare our predicted turnout to the actual turnout in the 2008 US Congressional election, which is 64% among registered voters and 57% among the voting age population. While both numbers are higher than our predicted values, it should be noted that 2008 is a Presidential election year, possibly resulting a higher γ for voters (as they are likely to be more drawn into the election).

---

[22] All turnout data in this section are obtained from http://www.idea.int/vt/.
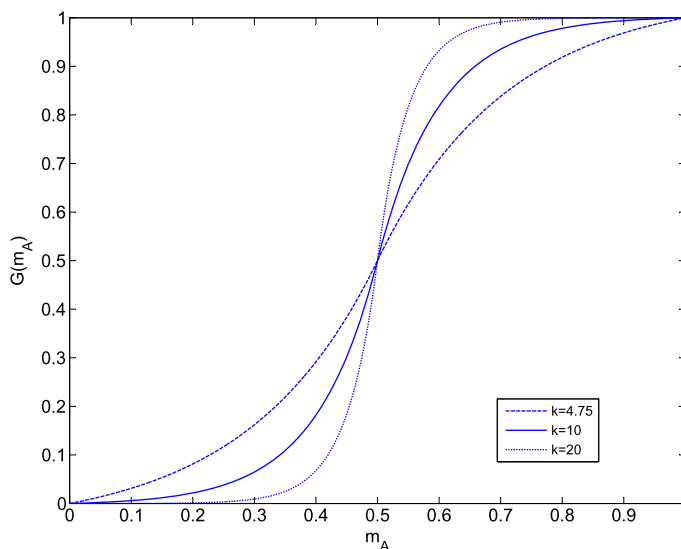
Fig. 3. Policy rules used in numerical exercises.

Indeed, with $\alpha = 0.55$ and $k = 4.75$, overall turnout rates of 59% and 68% can be generated by setting $\gamma = 0.15$ and $\gamma = 0.2$, respectively.

5.4. Effect of mandate consideration

More important, however, is the question of how mandate consideration affects turnout prediction quantitatively. To this end, we calculated the turnout rates under the winner take all rule, proportional rule, and CARA functions with $k = 4.75$, $k = 10$ and $k = 20$ for a comparison. Fig. 3 plots the CARA functions used.

The quantitative effect of mandate concerns on turnout is significant, as demonstrated by Table 4. For a close election with $\alpha = 0.55$, which favors the winner take all rule, the turnout rate under proportional rule is 5 to 7 times that under winner take all. The CARA functions give even higher turnout rates. At $\gamma = 0.05$, the CARA functions with parameter $k = 10$ and $k = 20$ result in the overall turnout of 44% and 55%, respectively, being 10 and 12 times of that under a winner take all rule. When $\gamma = 0.1$, the CARA functions with $k = 10$ and $k = 20$ yield the turnout rates of 62% and 78%, respectively. With the same α and γ, the winner take all rule generates a turnout of 4.6%, less than one-tenth of the turnouts under the CARA functions.

The effect of mandate concern is more visible at $\alpha = 0.65$, where the turnout under the winner take all rule converges to zero at a faster rate. For both values of γ used, the overall turnout under winner take all is less than 1%. Meanwhile, the turnout rates in models with considerable mandate concerns are generally around 30–45%, more than 45 times higher than that under winner take all.

If one finds our parameters reasonable, then our model yields turnout predictions matching empirically observed turnout rates in some national elections. For examples, when $\alpha = 0.55$, the turnout rates under the CARA functions with $k = 10$ and 20 are in the range of 44–78%, roughly matching typical national election turnouts in North America and Europe. On the lower end of empirical turnout rates, the 2011 Switzerland Parliamentary election records a turnout rate of 49% among registered voters and 40% among the voting age population. The registered voter and voting age population turnout rates in the UK 2010 Parliamentary election are 66% and 61%, respectively. On the higher end of empirical turnout rates, the Italy 2013 Parliamentary election turnout rates stand at 75% among registered voters and 68% among the voting age population. All these empirical turnout rates are compatible with the results of our numerical exercises. It should also be noted that γ does not have to be a lot higher for a higher turnout. In fact, when $\alpha = 0.55$, with a CARA $k = 10$ function, $\gamma = 0.3$ is sufficient to generate an overall turnout of 97.7%, with full turnout from the minority party.

If instead one finds our paternalism factors too high to be reasonable, then a drastically more unreasonable paternalism factor is required to generate the same turnout rate under the winner take all rule. For instance, when $\alpha = 0.55$, to generate an overall turnout of 22% (the predicted turnout under the proportional rule when $\gamma = 0.05$), a γ of 10 trillions (10^{13}) is required for the winner take all rule. If we would like to match the turnout rate under the CARA $k = 4.75$ function, the γ required would be even higher.

6. Equilibrium analysis

The numerical examples in Section 5 suggests that the limiting equilibrium in our model with low level of paternalism provides an excellent approximation to large elections. Given our parameterization, the numerical exercises also produce

turnout rates matching empirical observations. We now wish to demonstrate yet another advantage of our approximation — that its comparative statics match stylized facts about large elections.

The assumptions required for Proposition 4 are rather general. To allow for a meaningful equilibrium analysis, however, we must impose some symmetry across the two parties. Otherwise, the comparison across the two parties could arise simply out of the asymmetries and may have little to do with the election game per se.

To this end, we require that all citizens, conditional on their party alliance, are ex-ante identical in terms of their utilities. First, we assume that the two parties are identical in terms of their paternalism (i.e., $\gamma_A = \gamma_B = \gamma$). Second, we assume that the benefit from mandate is symmetric around $1/2$. That is, for any mandate $m \in [0, 1]$, $G(1 - m) = 1 - G(m)$.[23] Using the interpretation in Section 2, this is saying that, in case of victory, either candidate responds to mandate in the same magnitude. Under the above symmetry assumptions, Equation (2) can be rewritten as

$$\gamma \frac{1 - m_A^*}{T^*} G'(m_A^*) = F^{-1}(a^*) \tag{3}$$

$$\gamma \frac{m_A^*}{T^*} G'(1 - m_A^*) = F^{-1}(b^*). \tag{4}$$

The only asymmetry we allow is the (expected) size of the parties. Without loss of generality we assume that party A is the larger party. That is, $\alpha \geq 1/2$.

We begin by proving the existence of an *underdog effect*: that supporters of the smaller party vote more frequently, but the larger party still secures a greater mandate.

Proposition 5 *(Underdog effect). Suppose $\alpha \geq 1/2$. Then $a^* \leq b^*$ and $m_A^* \geq 1/2$. Moreover, both inequalities are strict whenever $\alpha > 1/2$.*

The term underdog effect was coined by Levine and Palfrey (2007). It is documented in large elections by Shachar and Nalebuff (1999), Blais (2000). Taylor and Yildirim (2010) formalize the notion in a general game theoretic framework for small elections. However, in their model the underdog effect disappears when $N \to \infty$, since turnout converges to zero. In contrast, our Proposition 5 offers a theoretical foundation for the underdog effect in a large election. Kartal (2014) obtains a similar finding. Unlike her result, our proposition does not require diminishing marginal mandate.

The underdog effect compares voting behaviors across parties given fixed party sizes. It is caused by the relatively higher incentive to free-ride experienced by the majority supporters, but it depends neither on the curvature of G (the smooth policy rule), nor on the distribution of voting costs. All it requires is that supporters across parties have the same preference parameters (so that none of them are more zealous about voting ex-ante) and that the preference for mandate be symmetric around $1/2$.

In contrast, the effect of a change in α on turnout is affected by both the curvature of G and the distribution of voting costs. The former governs the rate of returns to voting arising from changes in the expected number of votes for each party (and the associated change in mandate benefits). The latter affects responsiveness to changes in expected votes through changes in the gross benefits of voting. To make our analysis tractable, we impose two mild assumptions. First, we assume that the marginal benefit from mandate when a party is winning is decreasing. This is the usual "decreasing returns" assumption in economics. In elections, though, it has the additional interpretation that, when the votes are asymmetrically split, an extra vote for the winning party is worth less than an extra vote for the losing party. We think of this as a realistic feature of most elections. Mathematically this translates to G being concave on $[1/2, 1]$. That is, G has the same general shape as the quasi-majority rule described in Section 2. Second, we assume that the distribution F satisfies the following condition:

$$\frac{f(c)c}{F(c)} \quad \text{is weakly decreasing in } c.$$

In other words, the *elasticity* of the cumulative distribution function is decreasing in c. Interestingly, while in our work this assumption is sufficient to guarantee the presence of a competition effect, in Herrera et al. (2014a) a strict version of this assumption guarantees an increasing relation between the relative support ex-ante (i.e., in preferences) for the two parties and the relative support ex-post (i.e., in turnout) for the two parties. This assumption is also used by Che et al. (2013). We note that this is a fairly innocuous assumption, as it is satisfied by virtually all of the commonly used distributions with support in $\mathbb{R}_+$, including log-normal distributions, exponential distributions, Pareto distributions on $[\underline{c}, \infty)$ for $\underline{c} > 0$, uniform distributions on $[\underline{c}, \bar{c}]$ where $\underline{c} \geq 0$, Weibull distributions and single parameter Fréchet distributions.

Given the above assumptions, we obtain the *competition effect*, which states that total turnout decreases as the election becomes more lopsided.

Proposition 6 *(Competition effect). When $\alpha \in [1/2, 1)$ increases,*

---

[23] This immediately implies $G'(1 - m) = G'(m)$ and $G''(1 - m) = -G''(m)$ for all $m \in [0, 1]$.

(1) *Party A's turnout, a^*, decreases;*
(2) *Expected proportion of votes for party B (out of all citizens), $(1-\alpha)b^*$, decreases;*
(3) *Mandate for party A, m_A^*, increases; and*
(4) *Total turnout, T^*, decreases.*

While the first three comparative statics in this proposition deal with each party's turnout and mandate, the last result is commonly known in the voting literature as the competition effect. Like the underdog effect, the term competition effect was coined by Levine and Palfrey (2007) and documented in large elections by Shachar and Nalebuff (1999), Blais (2000). The competition effect has been difficult to obtain in models of large elections (Krasa and Polborn, 2009). Indeed, Taylor and Yildirim (2010, p. 464) concede, "the widely held intuition that elections with a more evenly split electorate should generate a greater expected turnout appears to be a property of small elections". It should be pointed out, though, that the standard costly voting model fails to predict the competition effect in large election because it predicts zero limiting turnout. Thus, by adopting the assumptions that mandate matters and voters are paternalistic, we are able to overcome this obstacle and provide a theoretical prediction for the competition effect in large elections.

The intuition behind this result is fairly straightforward. Under our decreasing returns assumption, members of the majoritarian party have less incentives to vote when they expect to win by a landslide than in a close election. The decrease in their turnout offsets the increase in the size of the majority party under the assumption on the elasticity of the distribution function. Meanwhile, the proportion of votes for the smaller party decreases. The total turnout is therefore reduced.

7. Conclusion

This paper argues and provides support for the notion that mandate (magnitude of victory) has consequences in elections. We show that this new, realistic assumption makes a substantial difference on turnout in an otherwise standard two-candidate costly voting model. Together with the assumption of paternalism, we obtain strictly positive turnout ratios in large elections, with comparative statics that match stylized facts. Our numerical computation exercises indicate that the effect of introducing mandate concerns is quantitatively relevant.

In concluding, we remark on a few features of our approach. First, we are able to provide the rate of convergence of the voting benefit without relying on specific functional forms of the policy rule (or power-sharing rule as in Castanheira, 2003 and Herrera et al., 2014b). Second, although our theorems are stated for the limiting case, they are equally relevant for large, finite elections, as our numerical exercise has demonstrated. Third, we show that in a winner take all election a paternalism factor 200 trillion times larger is required to match the turnout rates predicted by our model. Finally, we devise an innovative computation strategy for calculating equilibrium turnout rates, which might prove helpful for future research.

Appendix A. Omitted proofs

A.1. Notations for the appendix

From the point of view of a voter, the number of votes for parties A and B, both exclusive of herself, are random variables. Given N, denote these variables as $\tilde{v}_A^N$ and $\tilde{v}_B^N$, respectively. Define also the random variable $\tilde{M}^N = \tilde{v}_A^N + \tilde{v}_B^N$. Let M and v be a generic realization of $\tilde{M}^N$ and $\tilde{v}_P^N$ (P may be A or B, depending on the context), respectively. Denote the "marginal mandate benefit from voting" (over abstaining) for a party P supporter as

$$u_P(M, v) = G\big(m_P(v+1, M-v)\big) - G\big(m_P(v, M-v)\big),$$

where m_P is defined in Equation (1).

Given a turnout profile (a, b), let $T(a, b) = \alpha a + (1-\alpha)b$ be the unconditional turnout rate. Let $p_P(a, b)$ be the ex-ante probability that a voter votes for P (i.e., $p_A(a, b) = \alpha a$ and $p_B(a, b) = (1-\alpha)b$). The dependence of T, and p on (a, b) will often be suppressed. We can write the expected marginal gross benefits of voting (over abstaining) for a member of party P as:

$$U_P^N(a, b, \gamma_P) = \big(1 + \gamma_P N\big) \sum_{M=0}^{N} \sum_{v=0}^{M} \binom{N}{M}\binom{M}{v}(1-T)^{N-M} p_P^v p_Q^{M-v} u_P(M, v),$$

where Q is the other party. This is an expectation under a multinomial distribution: each of the N other voters is either not voting (with probability $1-T$), voting for A (with probability $p_A = \alpha a$) and voting for B (with probability $p_B = (1-\alpha)b$). Thus we may also write

$$U_P^N(a, b, \gamma_P) = \mathrm{E}_{a,b}\left[\big(1 + \gamma_P N\big)u_P(\tilde{M}^N, \tilde{v}_P^N)\right].$$

A.2. Existence of symmetric Bayesian–Nash equilibrium (Proposition 1)

As argued in Section 3, any best response of citizen i is monotone in the voting cost. In equilibrium, the turnout of party P is the probability that its supporter has a voting cost below $U_P(a, b)$. Hence a turnout profile (a, b) can be supported as a symmetric Bayesian Nash equilibrium if and only if

$$a = F\left(U_A^N(a, b, \gamma_A)\right)$$
$$b = F\left(U_B^N(a, b, \gamma_B)\right). \tag{5}$$

Given N, γ_A and γ_B, define a function $\varphi : [0, 1]^2 \to [0, 1]^2$ by

$$\varphi(a, b) = \left(F\left(U_A^N(a, b, \gamma_A)\right), F\left(U_B^N(a, b, \gamma_B)\right)\right).$$

Since U_A^N, U_B^N are jointly continuous in (a, b), and F is a continuous function, φ is a continuous function mapping from a compact, convex Euclidean space into itself. By the Brouwer Fixed-Point Theorem, φ has a fixed-point. This fixed-point satisfies Equations (5) and is therefore a symmetric Bayesian Nash equilibrium turnout profile.

A.3. Convergent rate of voting benefits (Theorem 1)

We first give an application of Hoeffding's Inequality (Hoeffding, 1963), which is then used to prove a lemma on the bounds of the voting benefit. Theorem 1 follows as a corollary of the lemma.

As in Section A.1, let $\tilde{M}^N$ denote the random variable of the number of voters turning out to vote (exclusive of the decision-making voter) when there are $N + 1$ citizens. Given a turnout profile (a_N, b_N), write $T_N = \alpha a_N + (1 - \alpha) b_N$ as the overall turnout rate.

Fact 1. *For any $\varepsilon > 0$,*

$$\Pr\left[\left|\tilde{M}^N - T_N N\right| \geq \varepsilon N\right] \leq 2e^{-2\varepsilon^2 N}.$$

Proof. Since $\tilde{M}^N$ is the sum of N independent Bernoulli random variables with mean T_N, Hoeffding's Inequality requires

$$\Pr\left[\left|\tilde{M}^N - T_N N\right| \geq \varepsilon N\right] \leq 2e^{-\frac{2\varepsilon^2 N^2}{N}}$$
$$= 2e^{-2\varepsilon^2 N}. \qquad \square$$

Fixing an $\varepsilon > 0$, let $I_M^N(\varepsilon) = [(T_N - \varepsilon)N, (T_N + \varepsilon)N]$.

Lemma A.1. *Fix an $\varepsilon > 0$. For each N, each γ_P and each turnout profile (a, b),*

$$h_P^N(a, b, \gamma_P) \leq U_P^N(a, b, \gamma_P) \leq h_P^N(a, b, \gamma_P) + e^{-2\varepsilon^2 N}(1 + \gamma_P N),$$

where

$$h_P^N(a, b, \gamma_P) = \Pr\left[\tilde{M}^N \in I_M^N(\varepsilon)\right] E_{a,b}\left[(1 + \gamma_P N)u_P(\tilde{M}^N, \tilde{v}_P^N)\,\Big|\,\tilde{M}^N \in I_M^N(\varepsilon)\right].$$

Proof. We can write

$$u_P^N(a, b, \gamma_P) = \Pr\left[\tilde{M}^N \in I_M^N(\varepsilon)\right] E\left[(1 + \gamma_P N)u_P(\tilde{M}_N, \tilde{v}_N)\,\Big|\,\tilde{M}^N \in I_M^N(\varepsilon)\right]$$
$$+ \Pr\left[\tilde{M}^N \notin I_M^N(\varepsilon)\right] E\left[(1 + \gamma_P N)u_P(\tilde{M}_N, \tilde{v}_N)\,\Big|\,\tilde{M}^N \notin I_M^N(\varepsilon)\right].$$

The second term is weakly positive (since u_P is). Thus $h_P^N(a, b, \gamma_P) \leq u_P^N(a, b, \gamma_P)$.

Fact 1 implies $\Pr\left[\tilde{M}^N \notin I_M^N(\varepsilon)\right] \leq 2e^{-2\varepsilon^2 N}$. Meanwhile, u_P is bounded above by $1/2$. Thus the second term is weakly smaller than

$$(1 + \gamma_P N)2e^{-2\varepsilon^2 N}\frac{1}{2},$$

which gives the stated upper bound of $U_P^N(a, b, \gamma_P)$. $\square$

To prove Theorem 1, let $\{(a_N, b_N)\}_N$ be a convergent sequence of turnout profiles such that $T = \lim_N T_N > 0$. Due to Lemma A.1, it suffices to show that $h_P^N(a_N, b_N, 0)$ converges to 0 at the rate of $1/N$.

So take an $\varepsilon \in (0, T)$. For all sufficiently large N, each $M \in I_M^N(\varepsilon)$ is strictly positive. Hence,

$$
\begin{aligned}
u_P(M, v) &= G\left(\frac{v+1}{M+1}\right) - G\left(\frac{v}{M}\right) \\
&= G'\left(\frac{v}{M}\right)\Delta(M, v) + o\left(\Delta(M, v)\right),
\end{aligned}
\tag{6}
$$

where

$$
\Delta(M, v) = \frac{M - v}{M(M + 1)}.
$$

Each $M \in I_M^N(\varepsilon)$ can be written as tN for some $t \in [T_N - \varepsilon, T_N + \varepsilon]$, which is bounded away from zero and infinity for N large. Thus in this range,

$$
\Delta(M, v) = \left(1 - \frac{v}{M}\right)\frac{1}{tN + 1}
$$

for some finite $t > 0$. Thus,

$$
h_P^N(a_N, b_N, 0) = \Pr\left[\tilde{M}^N \in I_M^N(\varepsilon)\right] \mathrm{E}_{a_N, b_N}\left[G'\left(\frac{\tilde{v}_P^N}{\tilde{M}^N}\right)\left(1 - \frac{\tilde{v}_P^N}{\tilde{M}^N}\right)\frac{1}{\tilde{t}N + 1}\,\bigg|\,\tilde{M}^N \in I_M^N(\varepsilon)\right]
$$

where $\tilde{t} = \tilde{M}^N/N$. By Fact 1, $\Pr\left[\tilde{M}^N \in I_M^N(\varepsilon)\right]$ converges to 1. $G'(m)$ is bounded away from zero and infinity for all $m \in [0, 1]$. As $1/(tN + 1)$ converges to 0 at the rate of $1/N$, so does $h_P^N(a_N, b_N, 0)$.

A.4. Limiting equilibrium under mandate and paternalism (Theorem 2, Propositions 3 and 4)

We will first prove, using two lemmas, that the limiting overall turnout rate $T = \alpha a + (1 - \alpha)b$, must be strictly positive. This allows us to derive an expression for the limiting payoff in Lemma A.4. Theorem 2, Propositions 3 and 4 follow from this lemma.

Lemma A.2. *If $\gamma_P > 0$ for at least one party P, then $T_N N \to \infty$ as $N \to \infty$.*

Proof. Suppose by contradiction that $\lim_N T_N N = r$ for some finite r. Then the distribution of $\tilde{M}_N$ converges to a Poisson distribution with mean r. Hence,

$$
\lim_N \Pr\left[\tilde{M}_N = 0\right] = e^{-r}.
$$

Since u_P is positive,

$$
U_P^N(a_N, b_N, \gamma_P) \geq (1 + \gamma_P N)\Pr\left[\tilde{M}_N = 0\right] u_P(0, 0).
$$

Taking limits on both sides,

$$
\lim_N U_P^N(a_N, b_N, \gamma_P) \geq \lim_N (1 + \gamma_P N)\frac{e^{-r}}{2}.
$$

As the right hand side goes to infinity, so must $U_P^N(a_N, b_N, \gamma_P)$. Yet this means members of party P turnout to vote with strictly positive probability at the limit, contradicting $\lim_N T_N N$ being finite. $\square$

Lemma A.3. *If $\gamma_A, \gamma_B > 0$, then $\lim_N T_N > 0$.*

Proof. By Lemma A.2, for all but a finite set of N, at least one of a_N and b_N is strictly positive. Name the parties so that $\lim_N(b_N/a_N) \neq 0$.[24] Suppose by contradiction that $b_N \to 0$.

Consider a party A supporter. For each N,

$$
U_A^N(a_N, b_N, \gamma_A) \geq \Pr\left[\tilde{M}^N \geq T_N N\right] \mathrm{E}_{a_N, b_N}\left[(1 + \gamma_A N)u_A(\tilde{M}^N, \tilde{v}_A^N)\,\bigg|\,\tilde{M}^N \geq T_N N\right].
$$

---

[24] As a convention, if $a_N = 0$ for all N sufficiently large, define $\lim_N(b_N/a_N) = \infty \neq 0$.

By Lemma A.2, any $M \geq T_N N$ is strictly positive for N large. The Taylor approximation in Equation (6) is therefore legitimate. Define the random variable $\tilde{x}_N = \tilde{M}^N - T_N N$ and write

$$U_A^N(a_N, b_N, \gamma_A) \geq \Pr\left[\tilde{M}^N \geq T_N N\right] \mathrm{E}\left[G'\left(\frac{\tilde{v}_A^N}{\tilde{M}^N}\right) \frac{\tilde{v}_B^N}{\tilde{M}^N} \frac{\frac{1}{N} + \gamma_A}{T_N + \frac{\tilde{x}_N + 1}{N}} + o\left(\Delta\right) \middle| \tilde{M}^N \geq T_N N \right]. \tag{7}$$

Recall that the median of $\tilde{M}_N$ must lie between $\lfloor T_N N \rfloor$ and $\lceil T_N N \rceil$. Hence

$$\Pr\left[\tilde{M}_N \geq T_N N\right] = \frac{1}{2} > 0 \quad \text{for all } N.$$

Moreover, we claim that, conditional on $\tilde{M}^N \geq T_N N$, $\tilde{v}_B > 0$ with positive probability. Notice that for any $M > 0$,

$$\begin{aligned}
\Pr\left[\tilde{v}_A^N = M \mid \tilde{M}^N = M\right] &= \Pr\left[\tilde{v}_A^N \geq M \mid \tilde{M}^N = M\right] \\
&\leq \frac{\mathrm{E}\left[\tilde{v}_A^N \mid M\right]}{M} \quad \text{(Markov's Inequality)} \\
&= \frac{(\alpha a_N / T_N) M}{M} \\
&= \frac{\alpha a_N}{\alpha a_N + (1 - \alpha) b_N}.
\end{aligned}$$

Thus for each N and each $M > 0$,

$$\begin{aligned}
\Pr\left[\tilde{v}_B^N > 0 \mid \tilde{M}^N = M\right] &= \Pr\left[\tilde{v}_A^N < \tilde{M}_N \mid \tilde{M}^N = M\right] \\
&\geq \frac{(1 - \alpha) b_N}{\alpha a_N + (1 - \alpha) b_N}.
\end{aligned}$$

The limit of the last expression (as $N \to \infty$) is non-zero since $\lim_N(b_N/a_N) \neq 0$. This is true for each $M > 0$, meaning that $\lim_N \Pr\left[\tilde{v}_B^N > 0 \mid \tilde{M}^N = M\right] > 0$.

The above argument implies that the limit of Equation (7) is governed by the order of

$$\frac{\frac{1}{N} + \gamma_A}{T_N + \frac{\tilde{x}_N + 1}{N}},$$

which goes to infinity since $T_N \to 0$. But this contradicts $a_N \to 0$. □

By now we have shown that the limiting overall turnout rate must be strictly positive. This allows us to rewrite the expression for the limiting gross marginal benefit of voting. To do so, define the following random variables:

$$\begin{aligned}
\tilde{a}^N &= \frac{\tilde{v}_A^N}{\alpha N} \\
\tilde{b}^N &= \frac{\tilde{v}_B^N}{(1 - \alpha) N} \\
\tilde{T}^N &= \frac{\tilde{M}^N}{N} = \tilde{a}^N + \tilde{b}^N.
\end{aligned}$$

Fact 2. *Suppose $\{(a_N, b_N)\} \to (a^*, b^*)$ such that $T = \alpha a^* + (1 - \alpha) b^* > 0$. Then*

(1) $\operatorname{plim}_N \tilde{a}^N = a^*$;
(2) $\operatorname{plim}_N \tilde{b}^N = b^*$; *and*
(3) $\operatorname{plim}_N \tilde{T}^N = T$.

Proof. We will only prove the statement for $\tilde{a}^N$. That for $\tilde{b}^N$ is similar, and that for $\tilde{T}^N$ will then follow from the continuous mapping theorem.

Take any $\varepsilon > 0$.

$$\Pr\left[\left|\tilde{a}^N - a^*\right| \geq \varepsilon\right] \leq \Pr\left[\left|\tilde{a}^N - a_N\right| + \left|a_N - a^*\right| \geq \varepsilon\right]$$

$$= \Pr\left[\left|\tilde{a}^N - a_N\right| \geq \varepsilon - \left|a_N - a^*\right|\right]$$

$$= \Pr\left[\left|\tilde{v}_A^N - \alpha a_N N\right| \geq \left(\varepsilon - \left|a_N - a^*\right|\right)\alpha N\right]$$

$$\leq 2 e^{-2\left(\varepsilon - \left|a_N - a^*\right|\right)^2 \alpha^2 N},$$

where the last inequality follows from Hoeffding's Inequality. Taking limit on both sides,

$$\lim_N \Pr\left[\left|\tilde{a}^N - a^*\right| \geq \varepsilon\right] \leq \lim_N 2 e^{-2\left(\varepsilon - \left|a_N - a^*\right|\right)^2 \alpha^2 N} = 0. \qquad \square$$

We can now derive an expression for the limiting gross benefit of voting as follows:

Lemma A.4. *Suppose* $\{(a_N, b_N)\} \to (a^*, b^*)$. *If* $T = \alpha a^* + (1 - \alpha b^*) = \lim_N T_N > 0$, *then*

$$\lim_N U_A(a_N, b_N, \gamma_A) = \gamma_A G'\left(\frac{\alpha a^*}{T}\right)\frac{(1 - \alpha)b^*}{T^2}, \text{ and}$$

$$\lim_N U_B(a_N, b_N, \gamma_B) = \gamma_B G'\left(\frac{(1 - \alpha)b^*}{T}\right)\frac{\alpha a^*}{T^2}.$$

Proof. We will prove this for party A. That for party B is analogous.

Recall the expression h_A^N in Lemma A.1. Take an $\varepsilon < T$. Notice that $\tilde{M}^N \in I_M^N(\varepsilon)$ implies $\tilde{T}^N \in [T_N - \varepsilon, T_N + \varepsilon]$. As $T_N \to T > 0$, for $\varepsilon < T$, $T_N - \varepsilon > 0$ for all N sufficiently large.

$$\lim_N U_A(a_N, b_N, \gamma_A) = \lim_N h_A^N(a_N, b_N, \gamma_A) \quad \text{(Lemma A.1)}$$

$$= \lim_N \mathrm{E}\left[(1 + \gamma_A N)u_P(\tilde{M}^N, \tilde{v}_A^N)\,\Big|\,\tilde{M}^N \in I_M^N(\varepsilon)\right] \quad \text{(Fact 1)}$$

$$= \lim_N \mathrm{E}\left[(1 + \gamma_A N)\left[G'\left(\frac{\tilde{v}_A^N}{\tilde{M}^N}\right)\Delta(\tilde{M}^N, \tilde{v}_A^N) + o\left(\Delta\right)\right]\,\Big|\,\tilde{M}^N \in I_M^N(\varepsilon)\right] \quad \text{(Equation (6))}$$

$$= \lim_N \mathrm{E}\left[G'\left(\frac{\alpha \tilde{a}^N}{\tilde{T}^N}\right)\left(\frac{(1 - \alpha)\tilde{b}^N}{\tilde{T}^N}\right)\frac{1 + \gamma_A N}{\tilde{T}^N N + 1}\,\Big|\,\tilde{T}^N \in [T_N - \varepsilon, T_N + \varepsilon]\right]$$

$$= \lim_N \gamma_A \mathrm{E}\left[G'\left(\frac{\alpha \tilde{a}^N}{\tilde{T}^N}\right)\frac{(1 - \alpha)\tilde{b}^N}{(\tilde{T}^N)^2}\,\Big|\,\tilde{T}^N \in [T - \varepsilon, T + \varepsilon]\right].$$

Define the function

$$\pi_A(a, b) = \begin{cases} G'\left(\frac{\alpha a}{\alpha a + (1 - \alpha)b}\right)\frac{(1 - \alpha)b}{(\alpha a + (1 - \alpha)b)^2} & \text{if } a + b > 0 \\ 1/2 & \text{otherwise} \end{cases}.$$

By Fact 2, $\mathrm{plim}\,\tilde{T}^N = T$, which means

$$\lim_N \mathrm{E}\left[\pi_A(\tilde{a}^N, \tilde{b}^N)\,\Big|\,\tilde{T}^N \in [T - \varepsilon, T + \varepsilon]\right] = \lim_N \mathrm{E}\left[\pi_A(\tilde{a}^N, \tilde{b}^N)\right].$$

Also, $\mathrm{plim}\,\tilde{T}^N = T$ implies π_A is continuous except on a set of (a, b) that occurs with zero probability at the limit. By the continuous mapping theorem and Fact 2,

$$\mathrm{plim}_N \pi_A(\tilde{a}^N, \tilde{b}^N) = \pi_A(\mathrm{plim}\,\tilde{a}^N, \mathrm{plim}\,\tilde{b}^N) = \pi_A(a^*, b^*).$$

It follows, then

$$\lim_N U_P^N(a_N, b_N, \gamma_A) = \lim_N \gamma_A \mathrm{E}\left[\pi_A(\tilde{a}^N, \tilde{b}^N)\,\Big|\,\tilde{T}^N \in [T - \varepsilon, T + \varepsilon]\right]$$

$$= \gamma_A \pi_A(a^*, b^*)$$

$$= \gamma_A G'\left(\frac{\alpha a^*}{T}\right)\frac{(1 - \alpha)b^*}{T^2}. \qquad \square$$

We can now prove Theorem 2, Proposition 3 and Proposition 4.

Proof of Theorem 2. We have already shown that if $\{(a_N, b_N)\}$ is a sequence of equilibrium turnout profiles converging to (a^*, b^*), then $a^* + b^* > 0$. It remains to prove that if $a^* > 0$, so must b^*. Suppose by contradiction that $a^* > 0$ but $b^* = 0$. By Lemma A.4,

$$\lim_N U_A^N(a_N, b_N, \gamma_A) = \gamma_A G'(1) \frac{0}{(\alpha a^*)^2} = 0.$$

But this contradicts $a_N \to a^* > 0$. $\square$

Proof of Proposition 3. First suppose C, the support of voting cost, is unbounded from above. For each finite N, $U_P^N(a_N, b_N, \gamma_P)$ is finite for each party P and each $(a_N, b_N) \in [0, 1]^2$. Thus $a_N = F(U_A^N(a_N, b_N, \gamma_A))$ and $b_N = F(U_B^N(a_N, b_N, \gamma_B))$ are both strictly less than 1.

Now suppose C is bounded above by $\bar{c}$. When $\gamma_A, \gamma_B > 0$, by Lemma A.3 the $T > 0$. Thus the expression

$$G'\left(\frac{\alpha a^*}{T}\right) \frac{(1 - \alpha)b^*}{T^2}$$

is finite. Lemma A.4 implies $\lim_N U_A^N(a_N, b_N, \gamma_A) < \bar{c}$ for γ_A sufficiently small. In other words, there exists an open interval below $\bar{c}$ such that an A-supporter with voting cost within this interval will abstain for N sufficiently large. Hence $a_N < 1$ for all N sufficiently large. The case for party B is similar and is therefore omitted. $\square$

Proof of Proposition 4. By Theorem 2, if $\{(a_N, b_N)\}_N \subset [0, 1)^2$ is convergent, then for all but finitely many N, $(a_N, b_N) \in (0, 1)^2$. For any $(a_N, b_N) \in (0, 1)^2$, the equilibrium condition requires

$$U_A^N(a_N, b_N, \gamma_A) = F^{-1}(a_N)$$
$$U_B^N(a_N, b_N, \gamma_B) = F^{-1}(b_N).$$

Taking limits and using Lemma A.4, we obtain the limiting expressions in Proposition 4. $\square$

A.5. Winner take all election (Proposition 2)

In a winner take all election, voting yields a gross marginal benefit of $1/2$ if the vote cast is pivotal and 0 otherwise. Write

$$\Pr\left[\text{piv}_A^N\right] = \Pr\left[\tilde{v}_A^N = \tilde{v}_B^N \text{ or } \tilde{v}_A^N + 1 = \tilde{v}_B^N\right]$$
$$\Pr\left[\text{piv}_B^N\right] = \Pr\left[\tilde{v}_A^N = \tilde{v}_B^N \text{ or } \tilde{v}_B^N + 1 = \tilde{v}_A^N\right]$$

as the probability that a vote for party A and B is pivotal when there are N citizens, respectively. Given a turnout profile (a, b), the expected gross marginal benefits of voting of the two parties are

$$U_A^N(a, b, \gamma_A) = \frac{1}{2} \mathrm{E}\left[(1 + \gamma_A N) \Pr\left[\text{piv}_A^N\right]\right] \tag{8}$$

$$U_B^N(a, b, \gamma_B) = \frac{1}{2} \mathrm{E}\left[(1 + \gamma_B N) \Pr\left[\text{piv}_B^N\right]\right]. \tag{9}$$

Suppose by contradiction that there is a sequence of equilibria $\{(a_N, b_N)\}_N \to (a^*, b^*)$ where $a^* + b^* > 0$. Given N, the ex-ante probability that any citizen would vote is $T_N = \alpha a_N + (1 - \alpha)b_N$. Conditional on voting, a citizen votes for A with probability $m_A^N = \alpha a_N / T_N$. Using Poisson approximations (see Myerson, 2000, Eq. (5.5)),[25]

$$\Pr\left[\text{piv}_A^N\right] \approx \frac{e^{-2T_N N\left(\sqrt{m_A^N} - \sqrt{1 - m_A^N}\right)^2}}{4\sqrt{\pi T_N N}\sqrt{m_A^N(1 - m_A^N)}} \left(\frac{\sqrt{m_A^N} + \sqrt{1 - m_A^N}}{\sqrt{m_A^N}}\right)$$

$$\Pr\left[\text{piv}_B^N\right] \approx \frac{e^{-2T_N N\left(\sqrt{m_A^N} - \sqrt{1 - m_A^N}\right)^2}}{4\sqrt{\pi T_N N}\sqrt{m_A^N(1 - m_A^N)}} \left(\frac{\sqrt{m_A^N} + \sqrt{1 - m_A^N}}{\sqrt{1 - m_A^N}}\right).$$

---

[25] Two functions of N are approximately equal if their ratio converges to 1 as N goes to infinity. This equivalent relation is denoted by the symbol $\approx$.

(Note that Myerson's approximation is valid as long as $a_N + b_N > 0$.) Plugging back to Equations (8) and (9):

$$U_A^N \approx \frac{1}{2} \mathrm{E} \left[(1 + \gamma_A N) \frac{e^{-2 T_N N \left(\sqrt{m_A^N} - \sqrt{1 - m_A^N} \right)^2}}{4 \sqrt{\pi T_N N} \sqrt{m_A^N (1 - m_A^N)}} \left(\frac{\sqrt{m_A^N} + \sqrt{1 - m_A^N}}{\sqrt{m_A^N}} \right) \right]$$

$$U_B^N \approx \frac{1}{2} \mathrm{E} \left[(1 + \gamma_B N) \frac{e^{-2 T_N N \left(\sqrt{m_A^N} - \sqrt{1 - m_A^N} \right)^2}}{4 \sqrt{\pi T_N N} \sqrt{m_A^N (1 - m_A^N)}} \left(\frac{\sqrt{m_A^N} + \sqrt{1 - m_A^N}}{\sqrt{1 - m_A^N}} \right) \right].$$

Fact 3. *If $\alpha \neq 1/2$, then $\lim_N N(\sqrt{m_A^N} - \sqrt{1 - m_A^N})^2 > 0$.*

Proof. Suppose $\lim_N N(\sqrt{m_A^N} - \sqrt{1 - m_A^N})^2 = 0$. Then

$$\lim_N U_A^N = \lim_N \mathrm{E} \left[\frac{1 + \gamma_A N}{4 \sqrt{\pi T_N N(1/2)}} \right] = \infty$$

$$\lim_N U_B^N = \lim_N \mathrm{E} \left[\frac{1 + \gamma_B N}{4 \sqrt{\pi T_N N(1/2)}} \right] = \infty.$$

Thus $(a^*, b^*) = (1, 1)$. But this means

$$\frac{1}{2} = m_A^* = \lim_N \frac{\alpha a_N}{\alpha a_N + (1 - \alpha) b_N} = \frac{\alpha}{\alpha + (1 - \alpha)} = \alpha. \qquad \square$$

If $\alpha \neq 1/2$, U_P^N is of the same order as

$$e^{-2 T_N N (\sqrt{m_P^N} - \sqrt{1 - m_P^N})^2} \sqrt{N} \to 0$$

as $N \to \infty$. This contradicts $a^* + b^* > 0$.

A.6. Underdog effect (Proposition 5)

Suppose by contradiction that $\alpha \geq 1/2$ and $a^* > b^*$. Dividing Equation (3) by (4) and using the symmetry of G we get

$$\frac{(1 - \alpha) b^*}{\alpha a^*} = \frac{F^{-1}(a^*)}{F^{-1}(b^*)}. \qquad (10)$$

Since $\alpha \geq 1/2$ and $a^* > b^*$, the left hand side is strictly less than 1. Meanwhile, the right hand side is strictly greater than 1 since F^{-1} is a strictly increasing function. Contradiction. Moreover, if $\alpha > 1/2$ and $a^* = b^*$, then the left hand side is still strictly less than 1 while the right hand side is exactly 1, which is also a contradiction.

Now if $a^* \leq b^*$, Equation (10) implies

$$(1 - \alpha) b^* \leq \alpha a^*$$

$$1 - m_A^* \leq m_A^*,$$

with strictly inequalities if $a^* < b^*$. Therefore $m_A^* \geq 1/2$, with a strict inequality when $\alpha > 1/2$.

A.7. Competition effect (Proposition 6)

This proof is mainly about elasticities. Thus for any variable x to be determined, write

$$\varepsilon_x = \frac{dx}{d\alpha} \frac{\alpha}{x}$$

as the elasticity of x with respect to α (evaluated at the original value of α). For example, ε_a would be the elasticity of a^* with respect to α. Since we will only consider party A's mandate, m_A^*, we will write ε_m as the elasticity of m_A^* with respect to α.

For each party P, let $c_P^* = F^{-1}(p^*)$ be the limiting cut-off voting cost below which a P-supporter would vote. Write also:

$$\sigma_G^* = \frac{-G''(m_A^*)m_A^*}{G'(m_A^*)},$$

$$\sigma_F^A = \frac{F(c_A^*)}{f(c_A^*)c_A^*} \quad \text{and} \quad \sigma_F^B = \frac{F(c_B^*)}{f(c_B^*)c_B^*}.$$

By Proposition 5, $b^* \geq a^*$. Hence $c_B^* \geq c_A^*$. Given the assumption that $f(c)c/F(c)$ is decreasing in c, $\sigma_F^B \geq \sigma_F^A$.
Differentiate Equations (3) and (4) with respect to α. Completing elasticities we obtain:

$$-\varepsilon_T - \left(\frac{m_A^*}{1-m_A^*} + \sigma_G^*\right)\varepsilon_m = \sigma_F^A \varepsilon_a \tag{11}$$

$$-\varepsilon_T + \left(1 - \sigma_G^*\right)\varepsilon_m = \sigma_F^B \varepsilon_b. \tag{12}$$

Using the expressions for T^* and m_A^*,

$$\varepsilon_T = m_A^*(1 + \varepsilon_a) + (1 - m_A^*)\left(-\frac{\alpha}{1-\alpha} + \varepsilon_b\right)$$

$$\varepsilon_m = 1 + \varepsilon_a - \varepsilon_T$$

$$= (1 - m_A^*)\left[(1 + \varepsilon_a) - \left(-\frac{\alpha}{1-\alpha} + \varepsilon_b\right)\right].$$

Adopt the following shorthands:

$$x = \frac{d\alpha a^*}{d\alpha}\frac{\alpha}{\alpha a^*} = 1 + \varepsilon_a; \quad y = \frac{d(1-\alpha)b^*}{d\alpha}\frac{\alpha}{(1-\alpha)b^*} = -\frac{\alpha}{1-\alpha} + \varepsilon_b;$$

$$z = 2m_A^* - 1 + \sigma_G^*(1 - m_A^*).$$

Notice that since $1/2 \leq m_A^* < 1$ and $\sigma_G^* \geq 0$, $z \geq 0$.
Using the expressions above, we can rewrite Equations (11) and (12) as the following system of linear equations:

$$\begin{pmatrix} -z - (1 + \sigma_F^A) & z \\ -z & z - (1 + \sigma_F^B) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\sigma_F^A \\ \frac{\alpha}{1-\alpha}\sigma_F^B \end{pmatrix}.$$

The determinant of this system is

$$D = \left[-z - (1 + \sigma_F^A)\right]\left[z - (1 + \sigma_F^B)\right] + z^2$$

$$= z(\sigma_F^B - \sigma_F^A) + (1 + \sigma_F^A)(1 + \sigma_F^B)$$

$$> 0.$$

Using Cramer's rule:

$$x = \frac{1}{D}\left[-\sigma_F^A(z - (1 + \sigma_F^B)) - \frac{\alpha}{1-\alpha}\sigma_F^B z\right]$$

$$= \frac{1}{D}\left[-z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) + \sigma_F^A(1 + \sigma_F^B)\right]$$

$$\leq \frac{(1 + \sigma_F^A)(1 + \sigma_F^B)}{z(\sigma_F^B - \sigma_F^A) + (1 + \sigma_F^A)(1 + \sigma_F^B)}$$

$$\leq 1.$$

Therefore $x = 1 + \varepsilon_a \leq 1$, meaning that $\varepsilon_a \leq 0$. In other words, a^* falls when α increases. However, it is not possible to determine the sign of x without further assumption.
Similarly,

$$y = \frac{1}{D}\left[-(z + (1 + \sigma_F^A))\frac{\alpha}{1-\alpha}\sigma_F^B - z\sigma_F^A\right]$$

$$= \frac{1}{D}\left[-z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) - \frac{\alpha}{1-\alpha}(1 + \sigma_F^A)\sigma_F^B\right]$$

$$< 0.$$

Hence $(1-\alpha)b^*$ falls as α increases. However, it is not possible to tell whether y is greater than or smaller than $-\alpha/(1-\alpha)$. The sign of ε_b is unknown.

Recall that (the relation $\overset{\text{sgn}}{=}$ means the two expressions are of the same sign):

$$
\begin{aligned}
\varepsilon_m &\overset{\text{sgn}}{=} x - y \\
&\overset{\text{sgn}}{=} -z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) + \sigma_F^A(1+\sigma_F^B) \\
&\quad + z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) + \frac{\alpha}{1-\alpha}(1+\sigma_F^A)\sigma_F^B \\
&> 0.
\end{aligned}
$$

Therefore m_A^* increases as α increases.

Finally we would like to consider the turnout.

$$
\begin{aligned}
\varepsilon_T &\overset{\text{sgn}}{=} m_A^*\left[-z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) + \sigma_F^A(1+\sigma_F^B)\right] \\
&\quad + (1-m_A^*)\left[-z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) - \frac{\alpha}{1-\alpha}(1+\sigma_F^A)\sigma_F^B\right] \\
&= -z\left(\sigma_F^A + \frac{\alpha}{1-\alpha}\sigma_F^B\right) \\
&\quad + m_A^*\sigma_F^A(1+\sigma_F^B) - (1-m_A^*)\frac{\alpha}{1-\alpha}(1+\sigma_F^A)\sigma_F^B.
\end{aligned}
\tag{13}
$$

The first term of expression (13) is weakly negative. Meanwhile,

$$
m_A^* = \frac{\alpha a^*}{T^*} \le \frac{\alpha b^*}{T^*} = \frac{\alpha}{1-\alpha}(1-m_A^*),
$$

and

$$
\sigma_F^A(1+\sigma_F^B) = \sigma_F^A + \sigma_F^A\sigma_F^B \le \sigma_F^B + \sigma_F^A\sigma_F^B = \sigma_F^B(1+\sigma_F^A).
$$

Thus the last two terms in expression (13) together are also negative. Total turnout falls as α increases.

Appendix B. Numerical computation strategy

As in Appendix A, let $u_P(M, v)$ be a party P supporter's marginal mandate benefit from voting (over abstaining) when there are M other voters who votes and among them v votes for P. Write $U_P(a, b)$ as the expected marginal gross benefits of voting for a member of party P, that is

$$
U_P(a, b) = (1 + \gamma N)\mathrm{E}_{a,b}\left[u_P(\tilde{M}, \tilde{v})\right].
$$

Under the uniform distribution assumption adopted in the numerical exercise, a pair of equilibrium turnout rates (a, b) satisfies

$$
U_A(a, b) = \max\left\{\min\left\{\frac{a}{2}, \frac{1}{2}\right\}, 0\right\}
\tag{14}
$$

$$
U_B(a, b) = \max\left\{\min\left\{\frac{b}{2}, \frac{1}{2}\right\}, 0\right\}.
\tag{15}
$$

For the limiting case, we solve for the solution to the system in Equations (3) and (4), using the uniform distribution.

B.1. Solution regarding large factorials

Since $\tilde{M}$ and $\tilde{v}$ are distributed according to a binomial distribution, to evaluate the expected marginal gross benefit functions, one would need to evaluate many binomial coefficients. This is computationally heavy when N is at the range we would like to consider.

To get around this problem, we realize that the probability of a binomial random variable vanishes exponentially away from the mean. Now suppose we set the precision of the numerical exercise to ε (i.e., the expected benefits and the inverse of the c.d.f. will be considered "the same" if they are less than ε apart). Since u_A and u_B are bounded above by $(1+\gamma N)/2$, ignoring a subset of the support with probability smaller than $\varepsilon/(1+\gamma N)$ will at most result in an approximation error of $\varepsilon/2$ on the benefit function side.

By Hoeffding's Inequality,

$$\Pr\left[\left|\tilde{M} - NT\right| \geq \delta\right] \leq 2e^{\frac{-2\delta^2}{N}}.$$

Solving for δ such that the right hand side of the above inequality is equal to $\varepsilon/(1 + \gamma N)$ we get:

$$2e^{\frac{-2\delta^2}{N}} = \frac{\varepsilon}{1 + \gamma N}$$

$$\delta = \sqrt{N \ln \sqrt{\frac{2(1 + \gamma N)}{\varepsilon}}}.$$

Therefore, instead of summing the relevant terms for all M from 0 to N, we sum only those M falling in the range $[NT - \delta, NT + \delta]$. Since δ increases only at the rate of $\sqrt{N}$, this reduces the computation burden.

B.2. Search

The stability of the system to be solved (Equations (14) and (15)) varies as different G functions are being used. We adopt a method of search that applies to all G functions considered to ensure that our search will stop at the solution.

Due to the underdog effect, any solution (a, b) satisfies $a \leq b$ and $\alpha a \geq (1 - \alpha)b$. Thus the solution must lie in the triangle bounded by $a = b$, $\alpha a = (1 - \alpha)b$ and $b = 1$ in the $[0, 1]^2$ box. Next, it can be verified that the limiting approximations[26] of U_A and U_B under each of the G specifications satisfy the following conditions in the triangular area considered:

(0) U_A and U_B are continuous in (a, b).
(1) $2U_A - a$ and U_B are strictly decreasing in a.
(2) Along the $a = b$ ray,
    (a) $2U_A - a$ is strictly decreasing in a;
    (b) $2U_B - b$ is strictly decreasing in b; and
    (c) $U_A < U_B$ at each point on the ray.
(3) Along the $\alpha a = (1 - \alpha)b$ ray,
    (a) $2U_A - a$ is strictly decreasing in a;
    (b) $2U_B - b$ is strictly decreasing in b; and
    (c) $U_A = U_B$ at each point on the ray.

Given these conditions, we search the triangular area in the following manner:

Step 1. Check whether $2U_A(1, 1) \geq 1$ and $2U_B(1, 1) \geq 1$. If yes, then $(a, b) = (1, 1)$ is an equilibrium. Halt and output $(1, 1)$.

Step 2. If $2U_B(1, 1) \geq 1$ and $2U_A(1, 1) < 1$, by (1), $2U_B(a, 1) \geq 1$ for all $a \in [\frac{1-\alpha}{\alpha}, 1]$. Also, since $(\frac{1-\alpha}{\alpha}, 1)$ is on the $\alpha a = (1 - \alpha)b$ ray, by (3c) $2U_A(\frac{1-\alpha}{\alpha}, 1) = 2U_B(\frac{1-\alpha}{\alpha}, 1) \geq 1 > \frac{1-\alpha}{\alpha}$. Thus we can use the method of bisection on the $b = 1$ line to search for an $\bar{a} \in \left[\frac{1-\alpha}{\alpha}, 1\right]$ such that $2U_A(\bar{a}, 1) = \bar{a}$. Then $(a, b) = (\bar{a}, 1)$ is an equilibrium. Halt and output $(\bar{a}, 1)$.

Step 3. If $2U_B(1, 1) < 1$, by (2c) $2U_A(1, 1) < 1$. There are two cases possible:

Case I: $2U_B(\frac{1-\alpha}{\alpha}, 1) > 1$.
    By (3c), $2U_A(\frac{1-\alpha}{\alpha}, 1) = 2U_B(\frac{1-\alpha}{\alpha}, 1) \geq 1 > \frac{1-\alpha}{\alpha}$. Use the method of bisection on the $b = 1$ line to search for an $\bar{a} \in \left[\frac{1-\alpha}{\alpha}, 1\right]$ such that $2U_A(\bar{a}, 1) = \bar{a}$. If $2U_B(\bar{a}, 1) \geq 1$, then $(a, b) = (\bar{a}, 1)$ is an equilibrium and we can stop. Otherwise, there exists an $\hat{a} \in \left[\frac{1-\alpha}{\alpha}, \bar{a}\right]$ such that $2U_B(\hat{a}, 1) = 1$.
    Meanwhile, since $2U_B(0, 0) > 0$, there exists a $\underline{b} \in [0, 1]$ such that $2U_B(\underline{b}, \underline{b}) = \underline{b}$. By (2c), $2U_A(\underline{b}, \underline{b}) < \underline{b}$. Yet as $2U_A(\frac{1-\alpha}{\alpha}, 1) > \frac{1-\alpha}{\alpha}$, by (3a) $2U_A(\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}) > \frac{1-\alpha}{\alpha}\underline{b}$. Thus there exists an $\underline{a} \in \left[\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}\right]$ such that $2U_A(\underline{a}, \underline{b})$.
    Conditions (2a), (2b), (3a) and (3b) imply that each of the $2U_A(a, b) = a$ and $2U_B(a, b) = b$ loci can cut each of the two boundary rays only once. The above argument therefore means that there exists an (a, b) within the convex hull of $\left(\frac{1-\alpha}{\alpha}, 1\right)$, $(1, 1)$, $(\underline{b}, \underline{b})$ and $\left(\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}\right)$ such that $2U_A(a, b) = a$ and $2U_B(a, b) = b$. See Fig. 4a for an illustration.

---

[26] At the large N we have considered, the actual functions are close to the limiting approximations. Since we make only qualitative restrictions below, we check only the limiting functions, rather than the actual ones.
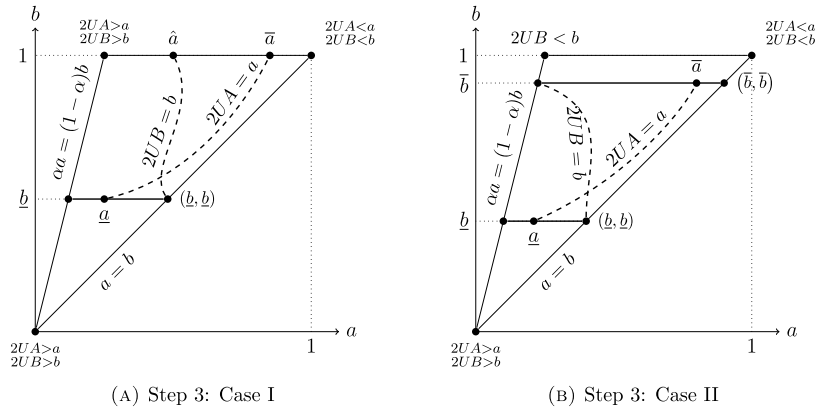
(A) Step 3: Case I   (B) Step 3: Case II

Fig. 4. Step 3: two possible cases.

(A)   (B)   (C)
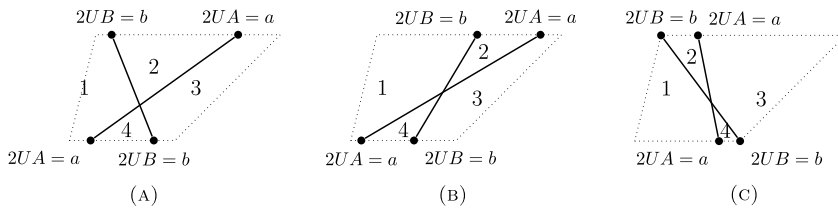
Fig. 5. Four areas for the search in Step 3.

Case II: $2U_B(\frac{1-\alpha}{\alpha}, 1) < 1$.

Since $2U_B(0,0) > 0$, there exists a $\bar{b} \in [0,1]$ such that $2U_B(\frac{1-\alpha}{\alpha}\bar{b}, \bar{b}) = \bar{b}$. By (3c), $2U_A(\frac{1-\alpha}{\alpha}\bar{b}, \bar{b}) > \frac{1-\alpha}{\alpha}\bar{b}$. Also, since U_B is decreasing in a, $2U_B(\bar{b}, \bar{b}) < \bar{b}$. By (2c), then, $2U_A(\bar{b}, \bar{b}) < \bar{b}$. Continuity of U_A implies the existence of an $\bar{a} \in \left[\frac{1-\alpha}{\alpha}\bar{b}, \bar{b}\right]$ such that $2U_A(\bar{a}, \bar{b}) = \bar{a}$.

Meanwhile, by (2b), there exists a $\underline{b} \in [0, \bar{b}]$ such that $2U_B(\underline{b}, \underline{b}) = \underline{b}$. By (2c), $2U_A(\underline{b}, \underline{b}) < \underline{b}$. Yet as $2U_A(\frac{1-\alpha}{\alpha}\bar{b}, \bar{b}) > \frac{1-\alpha}{\alpha}\bar{b}$, (3a) implies $2U_A(\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}) > \frac{1-\alpha}{\alpha}\underline{b}$. Hence there is an $\underline{a} \in \left[\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}\right]$ such that $2U_A(\underline{a}, \underline{b}) = \underline{a}$.

Conditions (2a), (2b), (3a) and (3b) imply that each of the $2U_A(a,b) = a$ and $2U_B(a,b) = b$ loci can cut each of the two boundary rays only once. The above argument therefore means that there exists an (a,b) within the convex hull of $\left(\frac{1-\alpha}{\alpha}\bar{b}, \bar{b}\right)$, $(\bar{b}, \bar{b})$, $(\underline{b}, \underline{b})$ and $\left(\frac{1-\alpha}{\alpha}\underline{b}, \underline{b}\right)$ such that $2U_A(a,b) = a$ and $2U_B(a,b) = b$. See Fig. 4b for an illustration.

In either case, the locus of $2U_A(a,b) = a$ cuts that of $2U_B(a,b) = b$ from above at least once. Around this intersection, the two loci separate four areas (see Fig. 5):

Area 1: $2U_A(a,b) > a$ and $2U_B(a,b) > b$.
Area 2: $2U_A(a,b) > a$ and $2U_B(a,b) < b$.
Area 3: $2U_A(a,b) < a$ and $2U_B(a,b) < b$.
Area 4: $2U_A(a,b) < a$ and $2U_B(a,b) > b$.

(The signs of the inequalities are determined by Condition (1) above.) We can then find the equilibrium, which is the intersection of the loci, using the following search algorithm:

State Variables: b_h, b_l (the high and low bounds of b)
Initialization: $b_h = 1$ (Case I) or $\bar{b}$ (Case II), $b_l = \underline{b}$
Iteration: Given b_h and b_l, set $b = (b_h + b_l)/2$. Let a_h and a_l be the high and low bounds of a given b. Set $a_h = b$, $a_l = \frac{1-\alpha}{\alpha}b$ and $a = (a_h + a_l)/2$. Evaluate $2U_A(a,b)$ and $2U_B(a,b)$ to check the area to which (a,b) belongs. Transit according to the following rules:
- If (a,b) is in Area 1, set $a_l = a$. Take the new $a = (a_h + a_l)/2$ and keep b. Re-evaluate the area of the new point and transit.
- If (a,b) is in Area 2, set $b_h = b$ and repeat the iteration.
- If (a,b) is in Area 3, set $a_h = a$. Take the new $a = (a_h + a_l)/2$ and keep b. Re-evaluate the area of the new point and transit.
- If (a,b) is in Area 4, set $b_l = b$ and repeat the iteration.

Halting: Stop either the distance of the current a_h and a_l is smaller than the required precision, or when the distance of the current b_h and b_l is smaller than the required precision. In the former case, output the equilibrium a to be the mid-point between the final a_h and a_l, and the equilibrium b to be the b at which these bounds of a are evaluated. In the latter case, output the equilibrium b to be the mid-point between the final b_h and b_l, and the a to be the mid-point between the last two a's when b_h and b_l are reset.

References

Blais, A., 2000. To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory. University of Pittsburgh Press, Pittsburgh.

Börgers, T., 2004. Costly voting. Amer. Econ. Rev. 94 (1), 57–66.

Campbell, C., 1999. Large electorates and decisive minorities. J. Polit. Economy 107 (6), 1199–1217.

Castanheira, M., 2003. Victory margins and the paradox of voting. Europ. J. Polit. Economy 19 (4), 817–841.

Che, Y.K., Dessein, W., Kartik, N., 2013. Pandering to persuade. Amer. Econ. Rev. 103 (1), 47–79.

Coate, S., Conlin, M., 2004. A group rule-utilitarian approach to voter turnout: theory and evidence. Amer. Econ. Rev. 94 (5), 1476–1504.

Coate, S., Conlin, M., Moro, A., 2008. The performance of pivotal-voter models in small-scale elections: evidence from Texas liquor referenda. J. Public Econ. 92, 582–596.

Downs, A., 1957. An Economic Theory of Democracy. Harper and Row, New York.

Evren, Ö., 2012. Altruism and voting: a large-turnout result that does not rely on civic duty or cooperative behavior. J. Econ. Theory 147 (6), 2124–2157.

Faravelli, M., Sanchez-Pages, S., 2014. (Don't) make my vote count. J. Theor. Polit. http://dx.doi.org/10.1177/0951629814556174.

Feddersen, T.J., Sandroni, A., 2006. A theory of participation in elections. Amer. Econ. Rev. 96 (4), 1271–1282.

Fowler, J.H., 2005. Turnout in a small world. In: Zuckerman, A. (Ed.), Social Logic of Politics. Temple University Press, Philadelphia, pp. 269–287.

Goeree, J., Grosser, J., 2007. Welfare reducing polls. Econ. Theory 31 (1), 51–68.

Harsanyi, J.C., 1977. Morality and the theory of rational behavior. Soc. Res. 44 (4), 623–656.

Harsanyi, J.C., 1992. Game and decision theoretic models in ethics. In: Aumann, R., Hart, S. (Eds.), Handbook of Game Theory with Economic Applications, vol. I. Elsevier, Amsterdam, pp. 669–707.

Herrera, H., Morelli, M., Nunnari, S., 2014a. Turnout across democracies. NBER working paper No. 20451.

Herrera, H., Morelli, M., Palfrey, T.R., 2014b. Turnout and power sharing. Econ. J. 124, F131–F162.

Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statistical Assoc. 58 (301), 13–30.

Kartal, M., 2014. A comparative welfare analysis of electoral systems with endogenous turnout. Econ. J. http://dx.doi.org/10.1111/ecoj.12150.

Krasa, S., Polborn, M.K., 2009. Is mandatory voting better than voluntary voting? Games Econ. Behav. 66 (1), 275–291.

Krishna, V., Morgan, J., 2012. Voluntary voting: costs and benefits. J. Econ. Theory 147 (6), 2083–2123.

Ledyard, J., 1981. The paradox of voting and candidate competition: a general equilibrium analysis. In: Horwich, G., Quirk, J.P. (Eds.), Essays in Contemporary Fields of Economics. Purdue University Press, West Lafayette.

Ledyard, J., 1984. The pure theory of two candidate elections. Public Choice 44 (1), 7–41.

Levine, D.K., Palfrey, T.R., 2007. The paradox of voter participation? A laboratory study. Amer. Polit. Sci. Rev. 101 (1), 143–158.

McMurray, J.C., 2013. Aggregating information by voting: the Wisdom of the experts versus the Wisdom of the masses. Rev. Econ. Stud. 80, 277–312.

Morton, R.B., 1991. Groups in rational turnout models. Amer. J. Polit. Sci. 35 (3), 758–776.

Myatt, D.P., 2012. A rational choice theory of voter turnout. Working paper. London Business School.

Myerson, R.B., 2000. Large Poisson games. J. Econ. Theory 94 (1), 7–45.

Palfrey, T.R., Rosenthal, H., 1983. A strategic calculus of voting. Public Choice 41 (1), 7–53.

Palfrey, T.R., Rosenthal, H., 1985. Voter participation and strategic uncertainty. Amer. Polit. Sci. Rev. 79 (1), 62–78.

Peterson, D.A.M., Grossback, L.J., Stimson, J.A., Gangl, A., 2003. Congressional response to mandate elections. Amer. J. Polit. Sci. 47 (3), 411–426.

Poole, K.T., Rosenthal, H., 2001. D-NOMINATE after 10 years: a comparative update to congress: a political economic history of roll-call voting. Legis. Stud. Q. 26 (1), 5–29.

Razin, R., 2003. Signaling and election motivations in a voting model with common values and responsive candidates. Econometrica 71 (4), 1083–1119.

Riker, W.H., Ordeshook, P.C., 1968. A theory of the calculus of voting. Amer. Polit. Sci. Rev. 62 (1), 25–42.

Shachar, R., Nalebuff, B., 1999. Follow the leader: theory and evidence on political participation. Amer. Econ. Rev. 89 (3), 525–547.

Smirnov, O., Fowler, J.H., 2007. Mandates, Parties, and Voters: How Elections Shape the Future. Temple University Press, Philadelphia.

Somer-Topcu, Z., 2009. Timely decisions: the effects of past national elections on party policy change. J. Polit. 71 (1), 238–248.

Taylor, C., Yildirim, H., 2010. A unified analysis of rational voting with private values and group-specific costs. Games Econ. Behav. 70 (2), 457–471.