Altruism, Turnout and Strategic Voting Behavior

Özgür Evren*

Department of Economics, New York University

19th West 4th Street, New York, NY 10012

E-mail: oe240@nyu.edu

July 28, 2009

Abstract

We consider the problem of costly voting in a game-theoretic framework where agents are altruistic. We show that if, as usual, agents' types are assumed to be independently and identically distributed, the classical theorem of Thomas R. Palfrey and Howard Rosenthal (1985) on the impossibility of large-scale turnout essentially survives, despite the introduction of altruism. We solve this problem by introducing additional uncertainty about the fraction of altruistic agents who support a given candidate. It turns out that under suitable homogeneity assumptions, this modified model is asymptotically equivalent to the rule utilitarian voter model of Timothy J. Feddersen and Alvaro Sandroni (2006, 2007), which is known to be consistent with large-scale turnout and strategic behavior. However, in contrast to rule utilitarian voter models, which typically require large-scale agreements among uncoordinated agents, our model is also compatible with several forms of heterogeneity in agents' characteristics. It is hoped that this paper will help to close the discrepancy between our understanding of costless or small elections, which is largely shaped by pivotal-voter models, and that of costly, large elections. (JEL D64, D72)

*I am grateful to Debraj Ray who supervised this research, which also benefited from various discussions with Daniel Barczyk, Gokhan Buturak, Guillaume Frechette, Anna Gumen, Melis Kartal, Yilmaz Kocer, Alessandro Lizzeri, Efe A. Ok, Pietro Ortoleva, Gil Riella and Florian Semani. I owe special thanks to Timothy J. Feddersen, Srinivasa Varadhan and Atilla Yılmaz. All remaining errors and deficiencies are mine.

Introduction

Why do we observe substantial turnout rates, even among millions of voters, given that voting is presumably costly and that in a large election the probability of being decisive (pivotal) for a single voter is presumably small? Since Anthony Downs (1957), this very basic question has been a major challenge for political economists. In a seminal game-theoretic work on costly voting, Thomas R. Palfrey and Howard Rosenthal (1985) showed that if there is a sufficient degree of uncertainty about the equilibrium number of votes for candidates, say, due to uncertain voting costs, the probability that a given agent will be pivotal indeed converges to zero as the size of the electorate becomes arbitrarily large. Assuming that the importance of an election (i.e., the utility that an agent receives by changing the winner) is independent of the size of the electorate, this led them to conclude that in a large election with positive net voting costs, total turnout rate must be approximately zero.

On the other hand, there is considerable evidence that voting behavior is better explained by "sociotropic" concerns about the overall state of the macroeconomy rather than individual concerns (Donald R. Kinder and D. Roderick Kiewiet, 1979; Gregory Markus, 1988; Peter Nannestad and Martin Paldam, 1994; Michael S. Lewis-Beck and Mary Stegmaier, 2000). For instance a person, say, an economist, who opposes free trade policies may be concerned with the *number* of low skilled workers who may lose their job, rather than her personal financial situation. Owing to such altruistic concerns, the possibility of choosing policies that may significantly affect the lives of millions of fellow citizens may be a big enough motivation to vote, despite the small likelihood of being decisive.[1]

Recent experimental evidence supports the view that altruism may be a factor in explaining voter turnout. The findings of David K. Levine and Palfrey (2007) indicate that a potential intensity in voters' preferences may lead to substantial turnout rates in large elections.[2] James H. Fowler (2006) and Fowler and Cindy D. Kam (2007) note that altruism may indeed cause such "preference intensity" and show that subjects' level of altruism measured with their generosity in dictator games is correlated with their participation in

[1] The reader may want to know why then only a relatively small number of people devote vast amounts of money (or time) to influence election outcomes. One reason may be the fact that if candidates' campaigns have already raised huge funds, the contribution of an additional dollar is not likely to influence the election outcome. Moreover, even if someone were offered a direct control of the election outcome in return for a large sum of money, she would most likely refuse due to the unfair nature of this offer that leaves all burden of a socially desirable outcome to a single agent. (For the importance of fairness motives in economic decisions, see, e.g., Ernst Fehr and Klaus M. Schmidt (1999).) I am grateful to Timothy J. Feddersen for calling my attention to this interesting question.

[2] Levine and Palfrey also show that allowing deviations from optimal behavior (instead of assuming perfect rationality) fits their data better.

elections.

Motivated by similar observations, Richard Jankowski (2002) and Aaron Edlin et al. (2007) present decision theoretic models that show that reasonable levels of altruism may explain high turnout rates in large, costly elections. However, since pivot probabilities in decision theoretic models are exogenously determined, such models do not offer a self-contained explanation of the strategic aspects of voting behavior.

In this paper, we present a model of costly voting with altruistic agents in a game-theoretic framework. By endogenizing pivot probabilities, we will be able to explain not only high turnout rates but also a variety of well known phenomena related to strategic voting behavior, such as a negative correlation between expected margin of victory and turnout, i.e., the *competition effect*, a positive correlation between the level of disagreement and turnout, i.e., the *polarization effect*, and a negative correlation between the fraction of supporters of a candidate and their turnout rate, i.e., the *underdog effect*.[3]

We also show that if, as it is frequently assumed in game-theoretic models on costly voting,[4] agents' types are independently and identically distributed (henceforth, *iid*), so that the equilibrium number of votes for a given candidate is a binomial random variable, (depending on parametric specifications) typically reasonable levels of altruism cannot generate significant turnout rates in large elections. Put differently, under the assumption that types are iid, the impossibility theorem of Palfrey and Rosenthal (1985) essentially survives, despite the introduction of altruism.

This finding is related to the knife-edge behavior of pivot probabilities in this binomial model noted by Gary Chamberlain and Michael Rothschild (1981): By the law of large numbers, vote shares of candidates can be predicted, almost precisely if the electorate is large enough. This, in turn, implies extremely small pivot probabilities unless the expected vote shares of candidates happen to be very close to each other.[5] Moreover, vote differentials in US presidential elections that we observe are big enough to confidently conclude that, ex-ante, the expected vote shares of candidates are so different that the implied pivot probabilities in this model are indeed negligible for reasonable levels of altruism (see Section II.A below).

Hints for a possible solution to this problem are provided by the work of I. J. Good and Lawrence S. Mayer (1975) who were first to note that if we instead assume that a randomly chosen agent votes for a given candidate with an unknown conditional probability, then the

[3]For evidence on competition effect see André Blais (2000) and Levine and Palfrey (2007), and for evidence on underdog effect see Levine and Palfrey (2007).

[4]See, e.g., Tilman Börgers (2004), Stefan Krasa and Mattias Polborn (2009) and Curtis R. Taylor and Huseyin Yildirim (2008).

[5]In stark contrast, when the expected vote shares are equal, pivot probabilities are larger than what the data suggests (Andrew Gelman et al., 2004).

implied unconditional pivot probabilities would be smoother. In fact, under the assumption that everyone votes, they have shown that, asymptotically, the implied pivot probabilities are inversely proportional to the size of the electorate.[6]

Inspired by this observation, in this paper we propose a modification of the binomial model where we randomize parameters that characterize the distribution of types, while keeping the assumption that types are iid conditional on these parameters. More specifically, we assume that there are two candidates and that among the supporters of a given candidate a randomly chosen agent is altruistic with an unknown conditional probability. In other words, we assume that the fraction of altruistic agents who prefer a given candidate is uncertain. Since selfish agents abstain in equilibrium, this additional uncertainty randomizes the equilibrium vote shares of the two candidates. Consistent with the finding of Good and Mayer (1975), it turns out that the implied pivot probabilities are inversely proportional to the size of the electorate. Thus, if the importance of the election grows linearly with the population, this modified model predicts significant turnout rates even in an arbitrarily large election.

The major conceptual difference between our formula on the magnitude of pivot probabilities and that of Good and Mayer (1975) is that we allow for abstention. This enables us to identify the interactions between the participation decisions of individuals and the turnout rates of the supporters of the two candidates which, in turn, form the base of our comparative statics exercises that reveal the aforementioned strategic aspects of voting behavior.

Assuming a linear link between the importance of an election (for a given agent) and the size of the population, as we do here, is equivalent to saying that a given agent weighs other agents' payoffs in a symmetric way and aggregates them additively, which seems to be a reasonable explanation of voter turnout in large elections. This is, however, not at all to believe that our approach provides a precise picture of the mental process through which altruism motivates voting or a complete list of the motivations of real voters.[7] Our purpose is simply to present a basic, reasonable extension of the standard pivotal-voter model, which is the most widely used tool in the analysis of costless or small elections, so that stylized facts about costly, large elections can be explained in a conceptually straightforward and solid way.

While this paper is inspired by the notion of altruism, a special version of our model can

[6] Chamberlain and Rothschild (1981) proved the same result independently.

[7] For instance, agents may vote because of a sense of civic duty (William Riker and Peter Ordeshook, 1968; Blais, 2000) or to express their moral concerns (Feddersen et al., 2008). But evidence on strategic behavior seems to be best explained by instrumentalist motivations. Such motivations can alternatively be modeled using utilitarian theories of ethics. We will shortly relate the present paper to the previous literature that follows this alternate route where the focus has been rule utilitarianism.

also be seen as a theory of *act utilitarian* voters. Act utilitarianism maintains that in a given situation the morally right action for a given agent is the one that maximizes per capita payoff, taking others' actions as given (John C. Harsanyi, 1977, 1980). Since maximization of per capita payoff is behaviorally indistinguishable from maximization of total payoff, an act utilitarian agent's voting behavior coincides with that of a purely altruistic agent who gives no special importance to her own payoff. More generally, an altruistic agent who places a (possibly) higher weight to her own payoff can be seen as a generalized act utilitarian agent whose moral judgements are (possibly) biased toward her self-interest.[8]

In a certain sense, act utilitarianism is the opposite of *rule utilitarianism* which, in its pure form, refers to the view that the ethically correct strategy is the one that maximizes per capita payoff if followed by *everyone* simultaneously (Harsanyi, 1977, 1980). Recently, there has been considerable interest in the notion of rule utilitarianism as a potential explanation of large-scale turnout (e.g., Stephen Coate and Michael Conlin, 2004; Feddersen and Alvaro Sandroni, 2006, 2007). The main idea of this literature is that rule utilitarian ethics may lead an agent to consider herself to belong to a group even if she has no direct contact with other agents in this group. In its most general form, as formalized by Feddersen and Sandroni (2006), the moral theory considered in this literature is rich enough to cover both act utilitarianism and pure rule utilitarianism as particular cases: An ethical agent of a given type considers herself as a member of a group of *similar* types and maintains that she should act according to a strategy that maximizes per capita payoff assuming other agents in her group will also do so, and taking as given the actions of other (dissimilar) agents.[9] According to this definition, act utilitarianism corresponds to the smallest group structure where each agent considers herself as a separate group that acts autonomously from the rest of the society, and pure rule utilitarianism corresponds to the largest group structure where the whole society forms a single group.[10]

However, it is sometimes argued in this literature that the formation of *large* groups (in agents' minds) is a prerequisite for large-scale turnout.[11] In contrast to this view, our findings show that there is no inherent tendency of autonomously acting utilitarian (or altruistic) agents to abstain. In fact, it turns out that (asymptotically) our basic model

[8] For evidence of a self-serving bias in moral judgements, see David Messick and Keith Sentis (1979) and Linda Babcock and George Loewenstein (1997).

[9] In Coate and Conlin (2004), agents care only about the welfare of their group, not the whole society.

[10] Let us add that, in fact, the largest group structure is ruled out by consistency conditions (Feddersen and Sandroni, 2006, Proposition 1), because agents with opposite policy preferences will have different judgements about which strategy should be followed. As we shall shortly see, one can think of many other forms of heterogeneity that may further restrict the size of admissible groups.

[11] "We conjecture that, as in Palfrey and Rosenthal (1985), agents who consider themselves to belong to a small group will tend to abstain because they will take as given the behavior of agents in other groups who are voting for their favorite candidate." (Feddersen and Sandroni, 2006, p.8)

predicts exactly the same turnout rates (among the supporters of both candidates) as the model of the type considered by Feddersen and Sandroni (2006, 2007) which is obtained by replacing act utilitarian agents with ethical agents who take as given the actions of only those agents with the opposite policy preferences.[12] (See Appendix C.) Thus, while preserving the conceptual simplicity of standard pivotal-voter models, our basic model performs as well as some canonical rule utilitarian (henceforth, RU) voter models.[13]

It is important to note that this equivalence result is derived under some homogeneity assumptions that facilitate the formation of large groups in RU-voter models without violating the consistency condition used in this literature which, roughly speaking, requires that agents in the same group must agree on the strategy that they should follow (given others' actions). While these homogeneity assumptions are of somewhat technical nature and do not seem to be very important for the purposes of political economics, if, as it is argued, the size of admissible groups in RU-voter models were a major determinant of the predicted turnout, whether or to what extent these assumptions hold would become matters of central importance. For instance, two agents who prefer the same candidate would presumably disagree on the welfare maximizing strategy provided that one of them feels more strongly towards this candidate. Thus, large-scale turnout could hardly occur if there is a significant uncertainty about the intensities of agents' preferences. Similarly, if agents' voting costs are correlated, two agents with different cost levels would presumably form different conjectures about the costs of others, and hence, would disagree on the welfare maximizing strategy. This, in turn, would imply that some imperfect correlation in agents' voting costs may dramatically reduce turnout rates.

Since we follow a standard game theoretic approach that is based on autonomy of individuals' actions, our explanation of turnout is conceptually independent from such homogeneity conditions that may facilitate large-scale, uncoordinated agreements among agents. As an example, in Appendix A we will present a brief extension of our basic model which shows that even an extreme form of heterogeneity in voters' estimation of the importance of the election is consistent with large-scale turnout and strategic behavior.[14]

[12] However, because of some informational issues, this does not mean that group sizes are entirely irrelevant. In our model, agents update their beliefs based on their knowledge of their types, while Feddersen and Sandroni (2006, 2007) work with prior beliefs. Modifying the information structure assumed by Feddersen and Sandroni accordingly leads to slightly higher turnout rates and closer elections (compared to our model and the original Feddersen-Sandroni model), but this is a consequence of some strategic considerations which are irrelevant for the present discussion. (See Remark C1 in Appendix C.)

[13] After showing that the structurally estimated version of their model outperforms reasonable alternatives based on simple expressive voting, Coate and Conlin (2004, p. 1496) add that "While there are good reasons to be skeptical about the pivotal-voter model's ability to explain turnout, it represents in many respects the simplest way of thinking about voting behavior."

[14] If our model is considered as a degenerate RU-voter model, it also follows from our findings that the explanatory power of the *general* RU-voter model is, in fact, also independent from homogeneity conditions.

The paper is organized as follows: In the next section we present our model under general assumptions about fractions of altruistic agents. The binomial version of the model is discussed in Section II. In Section III, we introduce uncertainty about the fractions of altruistic agents and relate this model to the RU-voter literature. In Section IV we conclude. Appendix contains the proofs and other supplementary material.

I. **The Model**

There are $n+1$ agents, who are eligible to vote, with the names $1, ..., n+1$, where n is a (strictly) positive integer. h stands for the name of a generic agent. τ_h denotes the privately known type of an agent h and consists of three components: h's policy type which can be ℓ (left) or r (right), h's personality type which can be a (altruistic) or s (selfish), and h's voting cost $C \in \mathbb{R}_+$.[15] We assume that conditional on a possibly random vector of parameters (λ, q_ℓ, q_r), types of agents (i.e., $\tau_1, ..., \tau_{n+1}$) are iid random variables. Here, $\lambda \in (0,1)$ is the conditional probability that a randomly chosen agent has the policy type ℓ and $q_i \in [0,1]$ is the conditional probability that a randomly chosen agent of policy type $i \in \{\ell, r\}$ is altruistic. The distribution function of a randomly chosen agent's voting cost is given by F, irrespective of the personality or policy type of this agent and irrespective of the values of parameters.

A *policy group* refers to the set of all agents with a given policy type.[16] If n is sufficiently large, by the law of large numbers, λ can be seen as the fraction of policy group ℓ and q_i as the fraction of altruistic agents in the policy group i. For simplicity we assume that λ is known, but to demonstrate the role of uncertainty about equilibrium vote shares, we allow q_ℓ and q_r to be random and denote their joint distribution function with G. To avoid trivialities, we suppose that q_ℓ and q_r are positive with probability 1. We assume $\lambda \leq 1/2$ so that policy group ℓ is a *minority*. We will often write λ_ℓ (resp. λ_r) instead of λ (resp. $1-\lambda$), and "type" instead of "policy type."

To ensure the existence of an equilibrium, a standard practice in game-theoretic models on costly voting is to assume that F is continuous. We will work with the following stronger assumption:

(H1) The support of F is an interval $[0, c] \subseteq \mathbb{R}_+$ with $c > 0$, over which F is continuously differentiable such that $f \equiv F' > 0$.

Thus, in the broadest sense, a message that follows from our analysis is that utilitarian theories of ethics can explain turnout under fairly general assumptions.

[15] Throughout the paper, we often use the same notation for a random variable and a possible value of this random variable. No confusion will arise.

[16] Of course, this is *not* meant to be a collectively acting group as in RU-voter models.

There are two candidates which we denote by a slight abuse of notation with ℓ and r. Given $i \in \{\ell, r\}$, j stands for the element of $\{\ell, r\}$ different from i. The election is decided by majority rule. In case of a tie, the winner is determined by tossing a fair coin.

An agent of type i believes that the victory of candidate i will bring a material policy payoff of $u > 0$ to every agent and the victory of candidate j is worth 0. Thus, agents of different types disagree about which candidate is good for the whole society. A selfish agent is assumed to care only about her own payoff. Since our focus is large elections, to avoid uninteresting details that will not affect our asymptotic results, we assume that selfish agents will abstain. In contrast, an altruistic agent is interested in the welfare of the whole society[17] and may choose to vote.

Each agent takes as given the actions of others. Hence, assuming that other agents' payoffs enter the utility of a given agent additively, their voting costs will not affect her participation decision. We thus assume that an altruistic agent places a total weight of $\Psi(n) > 0$ to the policy payoff of other n agents, and specify the utility of an altruistic agent of type i with voting cost C as:

$$\begin{aligned} &u\mathbf{1}_i\,(1+\Psi(n)) - C \quad \text{if she votes,} \\ &u\mathbf{1}_i\,(1+\Psi(n)) \quad \text{if she abstains,} \end{aligned}$$

where $\mathbf{1}_i \equiv 1$ if candidate i wins, and $\mathbf{1}_i \equiv 0$ otherwise.[18]

$\Psi(n)$ determines the level of altruism. We will assume that this function is linear:

(H2) (Linearity) There is a number $\psi \in (0, 1]$ such that $\Psi(n) = \psi n$ for every $n \in \mathbb{N}$.

(H2) corresponds to the idea that a given agent weighs other agents' payoffs in a symmetric way that is independent of the size of the electorate and aggregates them additively.[19] While this seems reasonable, as needed throughout the paper we will comment on consequences of relaxing the linearity assumption.

We focus on pure strategies. This causes no loss of generality, for as we shall see shortly, the continuity of distribution of costs ensure that in an equilibrium the event that a voter uses a mixed strategy is null. Therefore, a strategy for an agent h is a measurable map $Y_h : \{\ell, r\} \times \{a, s\} \times \mathbb{R}_+ \to \{-1, 0, 1\}$ such that $Y_h(i, s, C) \equiv 0$ for $i \in \{\ell, r\}$ and $C \in \mathbb{R}_+$.

[17] In this respect we follow Feddersen and Sandroni (2006b, 2007) rather than Coate and Conlin (2004).

[18] For the sake of clarity we choose to include an altruistic agent's own policy payoff in her utility, but this has no role in our asymptotic results.

[19] In our theoretical results we will ignore the fact that there may be agents who are not eligible to vote. The existence of ineligible agents would contribute to the importance of the election (without changing the size of the electorate) which can by modeled by setting $\Psi(n) \equiv \psi\eta n$, where $1 + \eta n$ is the size of the population and $1/\eta$ is the fraction of eligible agents. Increasing the importance of the election this way would contribute to turnout. (For an analogous comparative statics exercise, see Proposition 10 below.)

Here, -1, 1 and 0 stand for "vote for candidate ℓ," "vote for candidate r" and "abstain," respectively. Thus, as noted before, selfish agents are assumed to abstain. The action that an agent h will take is a random variable given by $X_h \equiv Y_h \circ \tau_h$. Since all agents are ex-ante symmetric, we assume that all agents use the same strategy, i.e., $Y_1 = Y_2 = \cdots = Y_{n+1}$. Since $\tau_1, ..., \tau_{n+1}$ are iid conditional on $q \equiv (q_\ell, q_r)$, it follows that so are $X_1, ..., X_{n+1}$.

We denote by $P_{i,q}$ the conditional probability that a randomly chosen agent casts a vote for candidate i, i.e., $P_{r,q} \equiv \Pr\{X_h = 1 \mid q\}$ and $P_{\ell,q} \equiv \Pr\{X_h = -1 \mid q\}$ for any agent h and any $q \in \mathfrak{I} \equiv [0,1]^2$. It should be noted that excluding a given agent, conditional on q, the number of votes for candidate i is a binomial random variable with the success probability $P_{i,q}$ and population size n, irrespective of the name of the agent excluded. Therefore, by suppressing the names of agents, we denote by S_i^- the number of votes for candidate i excluding a given agent.

A vote for candidate i is decisive in the event that excluding this vote the election is tied and i loses the coin toss or this vote creates a tie and i wins the coin toss. Conditional on q, the probability of this event is equal to $\frac{1}{2} piv_i(P_{\ell,q}, P_{r,q}, n)$, where $piv_i(P_{\ell,q}, P_{r,q}, n) \equiv \Pr\{S_i^- - S_j^- \in \{0, -1\} \mid q\}$ is the conditional probability that, excluding a given agent, tie occurs or candidate i is 1 behind. Clearly, we have,

$$\begin{aligned} piv_i(P_{\ell,q}, P_{r,q}, n) &= \sum_{b=0}^{\lfloor n/2 \rfloor} \tfrac{n!}{(n-2b)!b!b!} (P_{\ell,q})^b (P_{r,q})^b (1 - P_{\ell,q} - P_{r,q})^{n-2b} \\ &+ \sum_{b=0}^{\lfloor (n-1)/2 \rfloor} \tfrac{n!}{(n-2b-1)!b!(b+1)!} (P_{i,q})^b (P_{j,q})^{b+1} (1 - P_{\ell,q} - P_{r,q})^{n-2b-1} . \end{aligned} \tag{1}$$

[20,21]

Since voting costs are assumed to be positive, for any agent casting a vote against her favorite candidate is strictly dominated by abstaining. On the other hand, excluding her voting cost, for an altruistic agent of type i casting a vote for candidate i, as opposed to abstaining, brings an additional expected utility of

$$\Pi_i \equiv u \left(1 + \Psi(n)\right) \frac{1}{2} PIV_i,$$

where PIV_i is this agent's estimation of the (unconditional) probability of the event that, excluding her vote, tie occurs or candidate i is 1 behind. This is given by

$$PIV_i \equiv \int_{\mathfrak{I}} piv_i(P_{\ell,q}, P_{r,q}, n) dG^i(q). \tag{2}$$

[20] Given a real number ω, we define $\lfloor \omega \rfloor \equiv \max\{b \in \mathbb{Z} : b \leq \omega\}$ and $\lceil \omega \rceil \equiv \min\{b \in \mathbb{Z} : b \geq \omega\}$, where $\mathbb{Z}$ is the set of all integers.

[21] For future use, we define $piv_i(P_{\ell,q}, P_{r,q}, 0) \equiv 1$.

Here, G^i is the *posterior* distribution function of q from the perspective of an altruistic agent of type i.

In *equilibrium*, for an altruistic agent of type i the choice between participation (i.e., voting for candidate i) and abstention is determined by the size of her voting cost, C, relative to Π_i: If $C > \Pi_i$ she must abstain and if $C < \Pi_i$ she must participate. Letting the choice in the null event $C = \Pi_i$ be arbitrary, we may therefore characterize an equilibrium by a pair of nonnegative cutoff points $C_\ell^*, C_r^* \in \mathbb{R}_+$ such that:

$$\begin{aligned} \Pi_i^* &= C_i^*, \quad \text{for every } i \in \{\ell, r\}, \quad \text{and,} \\ P_{i,q}^* &= \lambda_i q_i F(C_i^*), \quad \text{for every } i \in \{\ell, r\} \text{ and } q = (q_\ell, q_r) \in \mathfrak{I}. \end{aligned}$$

Here, Π_i^* and $P_{i,q}^*$ are the equilibrium values of Π_i and $P_{i,q}$, respectively, and $q_i F(C_i^*)$ is the conditional expected turnout rate of type i agents. It should be noted that, by our conditional independence assumption and the law of large numbers, in a large election the equilibrium turnout rate of the policy group i (which is a random variable) will be close to $q_i F(C_i^*)$ with a high probability.

Remark 1. As we mentioned earlier, act utilitarianism maintains that an agent should act in a way to maximize per capita payoff, taking as given the actions of other agents. It is clear that participation of an agent with voting cost C increases per capita voting cost by $\frac{1}{1+n}C$. Thus, if we modify our model in an obvious way to replace the notion of altruism with act utilitarianism, an act utilitarian agent of type i with voting cost C would vote if and only if $u\frac{1}{2}PIV_i \geq \frac{1}{1+n}C$. Clearly, when $\psi \equiv 1$, this behavior is indistinguishable from the behavior of an altruistic agent. More generally, the altruistic agent that we consider in this paper who places a weight of $\psi \leq 1$ to other agents' payoffs can be seen as a generalized act utilitarian agent who maximizes an additive welfare function that is possibly biased toward her self-interest.

We close this section with the next proposition which shows that an equilibrium exists. Moreover, in any equilibrium cutoff points are positive.

Proposition 1. *If* (H1) *holds, an equilibrium exists. Moreover, in any equilibrium,* $C_i^* > 0$ *for every type* i.

II. Turnout Rates when Fractions of Altruistic Agents Are Known

In this section, we investigate turnout rates under the assumption that q_ℓ and q_r are known so that the number of votes for a given candidate is a binomial random variable. To motivate the extension that we present in the following section, our task in this section will be to provide theoretical and empirical arguments against this minimal model, which extends a popular model of costly voting to incorporate altruism.

We start with the following impossibility result which shows that when q_ℓ and q_r are known, despite the introduction of altruism, as the size of the electorate becomes arbitrarily large turnout rates converge to 0 provided that λ is less than $1/2$ and other potential asymmetries do not offset the size advantage of the majority. (In the sequel, we write P_i instead of $P_{i,q}$ for the known value of q.)

Proposition 2. *Suppose that q_ℓ and q_r are known and equal, and that* (H1) *and* (H2) *hold. For every $n \in \mathbb{N}$ consider an equilibrium $\left(C^*_{\ell,n}, C^*_{r,n}\right)$. Then:*

(i) *If $\lambda < \frac{1}{2}$, we have $\lim_n C^*_{\ell,n} = \lim_n C^*_{r,n} = 0$.*

(ii) *If $\lambda = \frac{1}{2}$, we have $\lim_n C^*_{\ell,n} = \lim_n C^*_{r,n} = \infty$.*

The intuition for Proposition 2(i) consists of two parts. First, as is well known, by the law of large numbers, in this binomial model pivot probabilities decrease with the size of the electorate at exponential rate provided that P_ℓ and P_r are distant.[22] (For more on this, see Appendix D.) Since our linearity assumption rules out the (implausible) case where $\Psi(n)$ increases at exponential rate, it follows that in the limit, equilibria with high turnout rates can be sustained only if $P^*_\ell - P^*_r$ tends to 0 (as n tends to ∞). Thus, if, as hypothesized, P^*_ℓ and P^*_r are bounded away from 0, in fact, the ratio P^*_ℓ/P^*_r must be tending to 1.

It turns out, however, that when $\lambda < 1/2$, the ratio P^*_ℓ/P^*_r is bounded away from 1 (in fact, bounded below 1), unless P^*_ℓ and P^*_r are both readily close to 0.[23,24] To gain intuition, first we note that when $\lambda < 1/2$, the ratio P^*_ℓ/P^*_r can be close to 1 only if the expected turnout rate of the majority, $q_r F\left(C^*_r\right)$, is small enough relative to that of the minority, $q_\ell F\left(C^*_\ell\right)$, to offset the size advantage of the majority: $\frac{q_r F(C^*_r)}{q_\ell F\left(C^*_\ell\right)} \simeq \frac{\lambda}{1-\lambda} < 1.$[25] With $q_\ell = q_r$, unless both $F(C^*_\ell)$ and $F(C^*_r)$ tend to 0, in the limit this could be possible only if, for a subsequence, $F(C^*_\ell) - F(C^*_r)$ exceeds a positive number ξ. But the hypothesis that P^*_ℓ/P^*_r is close to 1 also implies that the probability of a type r agent being pivotal is close to that of a type ℓ agent being pivotal, for with some algebra it is verified that

[22] That pivot probabilities decline exponentially does not rigidly depend on the assumption that n is known. For instance, formulas provided by Roger B. Myerson (2000) show that pivot probabilities in large Poisson voting games also decline exponentially with the mean of n (unless the expected vote shares of the two candidates are equal).

[23] In particular, Proposition 2(i) can be generalized to include the case of voters who enjoy the act of voting, i.e., the case $F(0) > 0$. (For more on this, see Remark E1 below.)

[24] Under the selfishness hypothesis and slightly different technical conditions, Krasa and Polborn (2009) have shown that P^*_ℓ/P^*_r converges to 1. This is not at odds with our claim: All we argue is that P^*_ℓ/P^*_r can converge to 1 only if P^*_ℓ and P^*_r converge to 0, as in the model of Krasa and Polborn.

[25] Throughout the paper, when two eventually nonzero sequences (b_n) and (d_n) are asymptotically equal, i.e., when $\lim_n \frac{b_n}{d_n} = 1$, we write $b_n \simeq d_n$.

$\Pr\{S_r^- - S_\ell^- = -1\}/\Pr\{S_\ell^- - S_r^- = -1\} = P_\ell^*/P_r^*$.[26] Since $C_i^* = \Pi_i^*$, this conclusion on pivot probabilities implies that C_r^* is close to C_ℓ^*, which contradicts the hypothesis that $F(C_\ell^*) - F(C_r^*)$ exceeds ξ. Hence, the conclusion of Proposition 2(i) follows.[27]

Remark 2. In Proposition 2(i) the linearity assumption can be replaced with the following condition: $\lim_n \Psi(n)e^{-\varepsilon n} = 0$ for every $\varepsilon > 0$. (This generalization does not require a change in our proof.) It is clear that the said condition is a very weak requirement that would hold for any specification of Ψ that corresponds to reasonable levels of altruism. This condition holds, for instance, if $\Psi(n)$ is bounded from above by ηn^β for some positive numbers η and β. In fact, a natural approach would be to assume $\Psi(1) \leq 1$ so that an altruistic agent weighs her own policy payoff not less than that of a single other agent, and choose Ψ to be a concave function so that $\Psi(n) \leq n$ (assuming $\Psi(0) = 0$). (If ineligible agents are taken into account, a better bound may be ηn where $1/\eta$ is the fraction of eligible agents.) Thus, the negative message delivered by Proposition 2(i) is independent from our linearity assumption.

On the other hand, when a randomly chosen agent votes for the two candidates with equal probabilities, pivot probabilities are asymptotically proportional to $1/\sqrt{n}$ (see, e.g., Chamberlain and Rothschild, 1981).[28] Moreover, $\lambda = \frac{1}{2}$ implies $P_\ell^* = P_r^*$. In this case, therefore, our linearity assumption implies arbitrarily large turnout rates in large elections.

Proposition 2 shows that in the present model, despite the introduction of altruism, the impossibility theorem of Palfrey and Rosenthal (1985) essentially survives, excluding the special case in which the fractions of the two policy groups are equal. Though this observation undermines the value of the present model as a potential explanation of high turnout rates, the case $\lambda = \frac{1}{2}$ may have some theoretical appeal because of a potential Downsian platform convergence. Indeed, this special case has been the focus of some studies on elections with selfish voters (see, e.g., Börgers, 2004; Taylor and Yildirim, 2005).

[26] This equality seems to be first noted by John O. Ledyard (1984).

[27] While this argument relies on the assumption that $q_\ell = q_r$ and on other symmetry assumptions that are implicit in our model, the following generalization of Proposition 2(i) shows that our symmetry assumptions are dispensable. *Suppose that: (1) Voting cost of a type i agent is weakly less than that of a type j agent in the sense of first order stochastic dominance; (2) relative to a type j agent, changing the winner is not less important for a type i agent; (3)* $\lambda_i q_i > \lambda_j q_j$. *Then the conclusion of Proposition 2(i) still holds.* (This can be proved by modifying the proof of Proposition 2(i) slightly.) Thus, in the broadest sense, the message of Proposition 2(i) is that, asymptotically, the present model can support large scale turnout only under restrictive assumptions on parameters. Finally, let us note that in view of Footnote 22, a natural conjecture is that analogous results can be proved when n is a well behaved random variable.

[28] Chamberlain and Rothschild (and other formal treatments of the matter that we are aware of) assume $P_\ell = P_r = 1/2$, but using a central limit theorem (William Feller, 1966, p. 490), it is not difficult to see that, more generally, for $P \in (0, 1/2]$ we have $piv_i(P, P, n) \simeq \frac{1}{\sqrt{\pi n P}}$. At any rate, since $piv_i(P, P, n)$ is decreasing in $P \in [0, 1/2]$ for a fixed n (Börgers, 2004), focusing on the case $P = 1/2$ is sufficient for the present discussion.

The explanatory power of the present model, however, can be refuted from an empirical perspective in a definitive way. We next turn to this issue.

A. *Empirical Observations*

Some evidence that indicates a discrepancy between the data and the predictions of the present model is provided by Gelman et al. (2004) who focus on the critical case $P_\ell = P_r$. By a central limit theorem, they conclude that if, as in the present model, agents were acting independently and identically, standard deviations of the relative vote shares of candidates would fall very rapidly as the observed turnout increases. (The relative vote share of candidate i refers to S_i/M, where $M \equiv S_\ell + S_r$ is the total turnout and S_i is the number of votes for candidate i.) This implies in particular that when $P_\ell = P_r$ so that the expected relative vote shares of both candidates equal 1/2, in elections with higher turnouts we would observe much closer races (in percentage terms), compared to elections with smaller turnouts. However, looking at various data sets from US and Europe, Gelman et al. (2004) show that the observed values of $|S_i/M - 1/2|$ do not exhibit such a rapid decrease with respect to realizations of M.

In the same study, they also report that taking pivot probabilities to be proportional to $1/n$ fits the data better than assuming pivot probabilities that are proportional to $1/\sqrt{n}$, as implied by the present model when $P_\ell = P_r$.[29]

Another argument against the present model can be based on the law of large numbers which implies that, in a large election, the ratio $S_i/(n+1)$ will be close to P_i with a high probability. Using the fact that this convergence is exponentially fast, in Appendix D (Lemma D2) we provide lower and upper bounds on P_i as functions of the observed value of $S_i/(n+1)$ and the probability of "type 1" error that one is willing to allow. We can then use these confidence intervals to a find a lower bound for $|P_\ell - P_r|$ as a function of the observed value of $|S_\ell - S_r|/(n+1)$ and then decide whether the pivot probabilities predicted by the present model are compatible with substantial turnout rates. (We recall that even if $|P_\ell - P_r|$ is only slightly positive, the implied pivot probabilities will be extremely small.)

Using this procedure, in Appendix D we show that in 2004 US presidential elections, among the 15 states with highest voting age population in all but Ohio, the pivot probabilities predicted by the present model were less than $e^{-47}/2$ with a confidence level of (at least) $1 - 10^{-6}$. (This upper bound on pivot probabilities obtains whenever $\frac{|S_\ell - S_r|}{n+1} \geq 0.015$ and $n \geq 10^6$ so that $|P_\ell - P_r| \geq 0.0097$ with $1 - 10^{-6}$ confidence.) In any such state, even if we allow $\Psi(n)$ to be as large the world population, 6.75×10^9, to ensure that cutoff costs

[29] Casey B. Mulligan and Charles G. Hunter (2003) report similar findings in their empirical study of US House elections.

are larger than 10 cents one must assume that per capita value of winning the election, u, is at least 7.64 billion dollars, which seems to be an absurdity.

Similarly, in 2000 (resp. 1996), among the same states the same conclusion is valid for all but Florida (resp. Georgia and Virginia). It thus seems that the present model is short of explaining high turnout rates together with substantial vote differentials that we observe.

III. Turnout Rates when Fractions of Altruistic Agents are Random

The driving force behind the negative conclusions of the previous section is that when q_ℓ and q_r are known, our model reduces to one in which agents vote for a given candidate i with a known probability P_i, independently. In this framework, the induced pivot probabilities are extremely low even if P_ℓ is only slightly different than P_r, while the case $P_\ell = P_r$ predicts a very close race with a very high probability, which leaves observed vote differentials unexplained. The observation that the case $P_\ell \neq P_r$ predicts extremely small pivot probabilities led Chamberlain and Rothschild (1981) to suggest taking P_ℓ and P_r as random, rather than given. Earlier, motivated by philosophical difficulties associated with the assumption that P_ℓ and P_r can be predicted precisely, Good and Mayer (1975) made the same suggestion and showed that if $P_\ell + P_r = 1$, and if there is a sufficient degree of uncertainty about P_ℓ that can be modeled with a continuous density, then pivot probabilities are asymptotically proportional to $1/n$ provided that, conditional on P_ℓ, agents vote independently and identically.[30]

In our model, randomness of $q \equiv (q_\ell, q_r)$ creates an environment which resembles the one considered by Good and Mayer (1975): Given any possible q, in equilibrium, a randomly chosen agent votes for candidate i with probability $P^*_{i,q} \equiv \lambda_i q_i F(C^*_i)$. In this section, our first task will be to present an extension of Good-Mayer formula that corresponds to this scenario where the focus is the possibility of abstention which is ruled out by Good and Mayer. We will thereby conclude that if, as we assume in this paper, the importance of the election grows (at least) linearly with the population, large-scale turnout is possible even in the limit. Moreover, the strategic content of this new formula will allow us to clarify the interactions between the participation decisions of individuals and the turnout rates of the two policy groups which will, in turn, lead to a variety of conclusions on turnout and margin of victory that cannot be understood within the framework of Good and Mayer, where the turnout rates are fixed at maximum possible levels.[31]

[30] Following this finding, in their decision theoretic work on altruistic voters, Edlin et al. (2007) assume that pivot probabilities are proportional to $1/n$.

[31] The finding of Good and Mayer (1975) reads as $\int_0^1 piv_i\,(P, 1-P, n)\, v\,(P)\, dP \simeq v\,(1/2)\,/n$ for any

It must also be emphasized that *once we prove that for reasonable levels of altruism the randomness of q gives rise to cutoff points that differ from* 0 *significantly, we can immediately explain observed vote differentials by realizations of q such that $P^*_{\ell,q}$ is distant from $P^*_{r,q}$*. As we noted, however, the conditional pivot probabilities are extremely small unless $P^*_{\ell,q}$ is close to $P^*_{r,q}$. Thus, almost all contribution to the unconditional pivot probabilities come from those q such that $P^*_{\ell,q}$ is close to $P^*_{r,q}$. The critical set on which $P^*_{\ell,q} = P^*_{r,q}$ is a ray

$$\{(q_\ell, q_r) \in \mathfrak{I} : \frac{q_\ell}{T^*_r} = \frac{q_r}{T^*_\ell}\} = \left\{(\theta T^*_r, \theta T^*_\ell) : 0 \leq \theta \leq \frac{1}{\max\{T^*_\ell, T^*_r\}}\right\}, \tag{3}$$

where $T^*_i \equiv \lambda_i F(C^*_i)$ and (as usual) $\mathfrak{I} \equiv [0,1]^2$. (The uses of the transformation on the right side of (3), where $\theta \equiv \frac{q_\ell}{T^*_r} \equiv \frac{q_r}{T^*_\ell}$, will be apparent shortly.) The next assumption ensures that any neighborhood of this ray has a positive probability irrespective of the values of $F(C^*_\ell)$ and $F(C^*_r)$.

(H3) q has a continuous and positive density g on $\mathfrak{I}$.

Given the prior density g, the posterior of an altruistic agent h of type i satisfies, for every $(b_\ell, b_r) \in \mathfrak{I}$,

$$\begin{aligned} G^i(b_\ell, b_r) &\equiv \Pr\{q_\ell \leq b_\ell, q_r \leq b_r \mid h \text{ is of type } i \text{ and altruistic}\} \\ &= \frac{\Pr\{q_\ell \leq b_\ell, q_r \leq b_r,\ h \text{ is of type } i \text{ and altruistic}\}}{\Pr\{h \text{ is of type } i \text{ and altruistic}\}} \\ &= \frac{\int_0^{b_r}\int_0^{b_\ell} \Pr\{h \text{ is of type } i \text{ and altruistic} \mid q_\ell, q_r\} g(q_\ell, q_r) dq_\ell dq_r}{\int_{\mathfrak{I}} \Pr\{h \text{ is of type } i \text{ and altruistic} \mid q\} dG(q)} \\ &= \frac{\int_0^{b_r}\int_0^{b_\ell} \lambda_i q_i g(q_\ell, q_r) dq_\ell dq_r}{\int_{\mathfrak{I}} \lambda_i q_i dG(q)} = \int_0^{b_r}\int_0^{b_\ell} \frac{q_i}{\overline{q}_i} g(q_\ell, q_r) dq_\ell dq_r, \end{aligned}$$

where $\overline{q}_i$ denotes the mean of q_i. Thus, the density of G^i is defined by $g^i(q) \equiv \frac{q_i}{\overline{q}_i} g(q)$ for every $q \in \mathfrak{I}$.

density υ on $[0,1]$ that is continuous at $1/2$. For our purposes, a major limitation of this formula is that it is based on the assumption that P_ℓ and P_r sum to a constant, namely 1. Let alone performing comparative statics exercises, to be able to determine the equilibrium values of P_ℓ and P_r (as functions of q), we need to understand the behavior of pivot probabilities without any unwarranted assumptions on $P_\ell + P_r$ or P_ℓ/P_r. While the double integration exercise forced by our agenda is not straightforward (which we postpone to Appendix E), as we shall see shortly, the asymptotic behavior of pivot probabilities is easily understood. (An alternate way of reading Good-Mayer formula is to view n as the total turnout (excluding a given agent) and $P = P_\ell$ as the probability that a randomly chosen *voter* casts a vote for candidate ℓ conditional on n. With this interpretation, we readily have $P_\ell + P_r = 1$, but other complications arise: Because of the correlation between q and turnout, the density of P, i.e. υ, becomes a random variable that depends on the realization of n, which takes us to a discrete double summation exercise that does not seem to be easier. Our proof is based on the related idea of summing pivot probabilities conditional on $P_\ell + P_r$ (i.e., expected turnout rate), rather than directly conditioning to turnout.)

We next present the extension of Good-Mayer formula that applies to our framework. This shows that a type i agent will be pivotal with a probability that is asymptotically equal to $1/n$ multiplied by the integral of g^i over the critical ray induced by the transformation on the right side of (3).

Lemma 1. *Let ν be a continuous (but not necessarily positive) density on $\mathfrak{I}$. Fix a pair of positive numbers (T_ℓ, T_r) with $T_\ell + T_r \leq 1$. Then, for any type i,*

$$\lim_n n \int_0^1 \int_0^1 piv_i(q_\ell T_\ell, q_r T_r, n)\nu(q_\ell, q_r)dq_\ell dq_r = 2\int_0^{\frac{1}{\max\{T_\ell, T_r\}}} \nu(\theta T_r, \theta T_\ell)d\theta. \tag{4}$$

Moreover, the convergence is uniform on any set $\mathfrak{T}$ of such (T_ℓ, T_r) which is bounded from below by a (strictly) positive vector. In particular, given an increasing self-map $k \to n_k$ on $\mathbb{N}$, if for every n_k there is an equilibrium $(C^*_{\ell,n_k}, C^*_{r,n_k})$ such that the sequences $T^*_{\ell,n_k} \equiv \lambda F(C^*_{\ell,n_k})$ and $T^*_{r,n_k} \equiv (1-\lambda)F(C^*_{r,n_k})$ converge to positive numbers $T^\bullet_\ell$ and $T^\bullet_r$, respectively, then (H3) *implies that, for any type i,*

$$\lim_k \frac{n_k}{2} \int_0^1 \int_0^1 piv_i(q_\ell T^*_{\ell,n_k}, q_r T^*_{r,n_k}, n_k)g^i(q_\ell, q_r)dq_\ell dq_r = \varphi^i(T^\bullet_\ell, T^\bullet_r), \tag{5}$$

where, for any pair of positive numbers (T_ℓ, T_r) with $T_\ell + T_r \leq 1$,

$$\varphi^i(T_\ell, T_r) \equiv \int_0^{\frac{1}{\max\{T_\ell, T_r\}}} g^i(\theta T_r, \theta T_\ell)d\theta \equiv \int_0^{\frac{1}{\max\{T_\ell, T_r\}}} \frac{\theta T_j}{\overline{q}_i} g(\theta T_r, \theta T_\ell)d\theta. \tag{6}$$

A particular implication of (5) is that if we fix an arbitrary pair of cutoff points (C_ℓ, C_r), asymptotically, a type i agent will be pivotal with probability $\varphi^i(T_\ell, T_r)/n$, where $T_\ell \equiv \lambda F(C_\ell)$ and $T_r \equiv (1-\lambda)F(C_r)$. Therefore, in what follows, by a slight abuse of terminology, *type i (posterior) asymptotic pivot probability* refers to $\varphi^i(T_\ell, T_r)$. It should also be noted that, here, $q_i T_i \equiv P_{i,q}$ equals the conditional expected *vote share* of candidate i relative to the size of the electorate, and the (unconditional) expected vote share of candidate i is given by $\overline{q}_i T_i$.

Before we proceed, let us uncover two interesting properties of asymptotic pivot probabilities that will be useful in our subsequent analysis of equilibria.

First, holding T_r/T_ℓ constant, both of the asymptotic pivot probabilities fall with $T_\ell + T_r$. In fact, for $b > 1$, we have $\varphi^i(bT_\ell, bT_r) = \varphi^i(T_\ell, T_r)/b$.[32] We call this property the *level*

[32] The proof is a simple substitution exercise: With $\theta' \equiv \theta b$, we have $\int_0^{\frac{1}{\max\{bT_\ell, bT_r\}}} \frac{\theta b T_j}{\overline{q}_i} g(\theta b T_r, \theta b T_\ell)d\theta =$

effect which, assuming $\overline{q}_\ell = \overline{q}_r$, is equivalent to saying that asymptotic pivot probabilities fall with the *expected (total) turnout rate*, $\overline{q}_\ell T_\ell + \overline{q}_r T_r$. Contrary to what one may think at first sight, the level effect is not solely a consequence of the fact that the conditional pivot probability $piv_i\,(P, P, n)$ decreases with P. An equally important observation is that, keeping the critical ray constant (which depends only on T_r/T_ℓ), a fall in $T_\ell + T_r$ decreases the rate of change of $q_i T_i$ (with respect to q_i) and this contributes to the thickness of the set (of q's) around the critical ray over which conditional pivot probabilities are relatively large.[33]

Second, assuming $T_r \geq T_\ell$, a further increase in T_r/T_ℓ pushes the critical ray away from the 45 degree line (i.e., the line over which $q_\ell = q_r$) and makes it shorter. Consequently, holding $T_\ell + T_r$ constant and assuming away a possible adverse effect through the prior density (a formal assumption in this direction will be introduced shortly), the sum of asymptotic pivot probabilities falls with T_r/T_ℓ. (A formal proof of this fact can be found in Footnote 40 below.) Moreover, the fall in $\varphi^r(T_\ell, T_r)$ is even sharper because of the private information of altruistic agents of type r. Indeed, assuming $\overline{q}_\ell = \overline{q}_r$, we have

$$\frac{\varphi^r(T_\ell, T_r)}{\varphi^\ell(T_\ell, T_r)} = \frac{T_\ell}{T_r}$$

which is a consequence of the following equalities that hold along the critical ray:

$$\frac{piv_r(q_\ell T_\ell, q_r T_r, n) g^r(q)}{piv_\ell(q_\ell T_\ell, q_r T_r, n) g^\ell(q)} = \frac{g^r(q)}{g^\ell(q)} = \frac{q_r}{q_\ell} = \frac{T_\ell}{T_r}.$$

In the sequel, the term *ratio effect* refers to these two consequences of an increased asymmetry in the expected vote shares of the two policy groups: A fall in the sum of asymptotic pivot probabilities and a sharper fall in the asymptotic pivot probability of the policy group that is more likely to win.

A. *Asymptotic Properties of Equilibria*

In the remainder of the paper, an *asymptotic equilibrium* refers to a pair of nonnegative, extended real numbers $(C_\ell^\bullet, C_r^\bullet)$ that is the limit of a convergent subsequence of equilibria

---

$\frac{1}{b} \int_0^{\frac{1}{\max\{T_\ell, T_r\}}} \frac{\theta' T_j}{\overline{q}_i} g(\theta' T_r, \theta' T_\ell) d\theta'$.

[33] On its own, the former observation does not explain the level effect, since for smaller values of $P_\ell + P_r$ a given distance between P_ℓ and P_r is more meaningful in relative terms. Indeed, this precisely offsets the fact that $piv_i\,(P, P, n)$ is a decreasing function of P, in the sense that for any fixed $P \in (0, \frac{1}{2})$ we have $\int_0^P piv_i(P + y, P - y, n) dy \simeq \int_0^{1/2} piv_i(\frac{1}{2} + y, \frac{1}{2} - y, n) dy$, while, of course, for a large n almost all contribution to both of these integrals come from very small y. (The proof is available upon request.)

$(C^*_{\ell,n_k}, C^*_{r,n_k})$. It must be noted that since the cutoff points associated with an asymptotic equilibrium are allowed to be infinite, *there exists at least one asymptotic equilibrium*; in fact, any sequence equilibria has a subsequence that converges to an asymptotic equilibrium. Given an asymptotic equilibrium $(C^\bullet_\ell, C^\bullet_r)$, we define $T^\bullet_i \equiv \lambda_i F(C^\bullet_i)$ so that over the corresponding subsequence of equilibria, the vote share of candidate i converges to $q_i T^\bullet_i$, for any realization of q.

In view of equation (5), our linearity assumption implies that *if it is positive*, an asymptotic equilibrium is finite and solves the following equations:

$$C^\bullet_i = u\psi\varphi^i(T^\bullet_\ell, T^\bullet_r) \equiv u\psi\frac{T^\bullet_j}{\overline{q}_i}\int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} \theta g(\theta T^\bullet_r, \theta T^\bullet_\ell)d\theta \quad \text{for } i = \ell, r. \tag{7}$$

The next result shows that this is the case for all asymptotic equilibria.

Proposition 3. *Assume* (H1)-(H3). *Then any asymptotic equilibrium is positive, finite and solves* (7).[34]

Remark 3. All asymptotic equilibria would still be positive whenever $\liminf_n \Psi(n)/n \geq \psi$ for some $\psi > 0$. The additional strength of our linearity assumption (which could be replaced by the asymptotic linearity condition $\lim_n \Psi(n)/n = \psi$) is needed for the clear-cut comparative statics results that we report below.

It is clear that (7) has at least one solution, for, by Proposition 3, any asymptotic equilibrium solves (7). As we shall see shortly, mild conditions ensure the uniqueness of this solution which, in turn, implies that any sequence of equilibria converges to this unique solution, for otherwise, in contradiction with Proposition 3, we could find an asymptotic equilibrium that cannot solve (7). Put differently, under mild conditions, in a large election the equilibrium is close to the unique solution of (7) (and therefore, is also "essentially" unique). Thus, in the sequel, we analyze further properties of asymptotic equilibria using (7). (We will not offer a result on the speed of the convergence of equilibria. However, numerical computations that we performed for several parameters have shown that in a moderately large election with more than a couple of thousands of agents, the precise values of the cutoff points do not significantly differ from their asymptotic estimators and follow the qualitative patterns that we will report below (see Appendix B).)

[34] The converse, i.e., whether any solution to (7) is an asymptotic equilibrium seems to be a nontrivial question and will be left open. However, we will shortly see that this difficulty disappears when (7) has a unique solution.

The first point to note is that whenever $\overline{q}_\ell = \overline{q}_r$, (7) implies:

$$\frac{C_\ell^\bullet}{C_r^\bullet} = \frac{T_r^\bullet}{T_\ell^\bullet} \equiv \frac{(1-\lambda)\,F(C_r^\bullet)}{\lambda F(C_\ell^\bullet)}, \quad \text{i.e.,} \tag{8}$$

$$\frac{C_\ell^\bullet F(C_\ell^\bullet)}{C_r^\bullet F(C_r^\bullet)} = \frac{1-\lambda}{\lambda}. \tag{9}$$

In the remainder of this section, without further mention *we will assume that* $\overline{q}_\ell = \overline{q}_r$ *and that* (H1)-(H3) *hold*, so that (8) and (9) apply to any asymptotic equilibrium.

An immediate implication of (9) is the underdog effect:

Proposition 4. *In any asymptotic equilibrium, the cutoff point of the minority,* $C_\ell^\bullet$*, is weakly larger than that of the majority,* $C_r^\bullet$*. Moreover,* $C_\ell^\bullet = C_r^\bullet$ *if and only if* $\lambda = 1/2$.[35]

The intuition behind Proposition 4 is the informational asymmetry that we explained when discussing the ratio effect: In a large election, if T_i^* is larger than T_j^*, the outcome can be close only if the fraction of altruistic agents among type j agents is larger than that among type i agents to offset the difference between T_i^* and T_j^*. In other words, when $T_i^* > T_j^*$, along the critical ray we have $q_j > q_i$. But because of her private information, an altruistic agent of type j deems more likely any realization of q such that $q_j > q_i$ (relative to an altruistic agent of type i). (Put formally, $g^j(q) > g^i(q)$ iff $q_j > q_i$.) This implies that, according to her subjective information, an altruistic agent of type j is more likely to be pivotal, so that $C_j^* = \Pi_j^*$ exceeds $C_i^* = \Pi_i^*$. In the limit this reasoning becomes precise, and $T_i^\bullet > T_j^\bullet$ implies $C_j^\bullet > C_i^\bullet$. But with $\lambda < 1/2$, this can be true only if $C_\ell^\bullet > C_r^\bullet$, for otherwise we would have $C_r^\bullet \geq C_\ell^\bullet$ and $T_r^\bullet > T_\ell^\bullet$.

The converse of this intuition is also true: $C_\ell^\bullet$ exceeds $C_r^\bullet$ precisely because altruistic agents of type ℓ are more likely to be pivotal, according to their posteriors. This amounts to saying that along the critical ray q_ℓ exceeds q_r; i.e., that $T_r^\bullet$ is larger than $T_\ell^\bullet$. Thus, the underdog effect does not entirely offset the size advantage of the majority (as it is also clear from (8)):

Proposition 5. *In any asymptotic equilibrium, the expected vote share of the majority,* $\overline{q}_r T_r^\bullet$*, is weakly larger than that of the minority,* $\overline{q}_\ell T_\ell^\bullet$*. Moreover, the expected vote shares of the two policy groups are equal if and only if* $\lambda = 1/2$.

The (prior) probability of victory for candidate i is (asymptotically) equal to $\Pr\{\frac{q_j}{q_i} \leq \frac{T_i^\bullet}{T_j^\bullet}\}$. Thus, when g is symmetric (so that q_ℓ/q_r and q_r/q_ℓ are identically distributed),

[35] Jankowski (2002) and Edlin et al. (2007) assume that agents of both types face the same (exogenous and objective) pivot probability. Therefore, in their models there is no underdog effect.

Proposition 5 also implies that the majority is more likely to win.[36]

Now, for every $\lambda \in (0, 1/2]$, let us consider an asymptotic equilibrium $(C^\bullet_{\ell,\lambda}, C^\bullet_{r,\lambda})$ and denote by $(T^\bullet_{\ell,\lambda}, T^\bullet_{r,\lambda})$ the associated maximum possible vote shares. Following Feddersen and Sandroni (2007), we view λ as the *level of disagreement* in the society. Since $T^\bullet_{\ell,\lambda} \leq \lambda$, if we can show that $T^\bullet_{r,\lambda} \to 0$ as $\lambda \to 0$, we can immediately conclude that as the level of disagreement tends to 0, the expected turnout rate, $\overline{q}_\ell T^\bullet_{\ell,\lambda} + \overline{q}_r T^\bullet_{r,\lambda}$, also tends to 0. But if this were not the case, i.e., if $T^\bullet_{r,\lambda}$ were bounded away from 0 for some arbitrarily small λ, the ratio $T^\bullet_{r,\lambda}/T^\bullet_{\ell,\lambda}$ would become arbitrarily large, and along the critical ray q_r would be arbitrarily small. On the other hand, such q are not likely to occur, especially according to the posterior of altruistic agents of type r. In fact, (7) tells us that (in this hypothetical situation) this effect is strong enough to push $C^\bullet_{r,\lambda}$ down to 0, giving us the desired contradiction.[37] Thus, we proved our first result on the polarization effect:

Proposition 6. *As the level of disagreement tends to* 0*, the expected turnout rate also tends to* 0 *for any possible selection of asymptotic equilibria.*

In Proposition 6, the relation between the level of disagreement and the expected turnout rate is not monotonic. We find the following assumption useful to provide stronger comparative statics results and the uniqueness of asymptotic equilibrium:

(H4) $g(\vartheta + \varepsilon, \vartheta - \varepsilon)$ is nonincreasing in ε for $0 \leq \varepsilon \leq \min\{\vartheta, 1 - \vartheta\}$ and every fixed $\vartheta \in [0, 1]$.

(H4) requires that a vertical movement away from the 45 degree line (caused by a fall in q_r) does not increase the prior density. This allows a positive correlation between q_ℓ and q_r, as one may expect in reality. For example, it can easily be checked that (H4) holds if the distribution of q is obtained by conditioning a bivariate normal distribution to the unit square $\mathfrak{I}$ provided that the normal distribution in question is such that the correlation coefficient is nonnegative and marginals are identical.[38] The simplest case which trivially implies (H4) is when q_ℓ and q_r are independently and uniformly distributed. Much more generally, when q_ℓ and q_r are iid beta random variables with monotone or unimodal densities, then (H4) again holds.[39]

[36] In contrast, as we discussed in Footnote 24, in previous game-theoretic models on costly voting, the minority may be as likely to win a large election as the majority, even when the majority is overwhelming.

[37] Specifically, if $T^\bullet_{r,\lambda} \geq \varepsilon \geq T^\bullet_{\ell,\lambda}$, by (7), we have $C^\bullet_{r,\lambda} \leq \frac{T^\bullet_{\ell,\lambda}}{\overline{q}_r} u\psi \frac{\overline{g}}{2\varepsilon^2}$, where $\overline{g}$ is an upper bound to g.

[38] In this case, for $q \in \mathfrak{I}$ and with $Q(q) \equiv \frac{1}{1-\rho^2}\left(\left(\frac{q_\ell - \mu}{\sigma}\right)^2 - 2\rho\left(\frac{q_\ell - \mu}{\sigma}\right)\left(\frac{q_r - \mu}{\sigma}\right) + \left(\frac{q_r - \mu}{\sigma}\right)^2\right)$, we have $g(q) = \frac{K}{2\pi\sigma^2\sqrt{1-\rho^2}} e^{-Q(q)/2}$ where $K > 0$ is a constant, μ (resp. σ) is the common mean (resp. standard deviation) of the components of the unconditional bivariate normal random variable, and $\rho \geq 0$ is the correlation coefficient.

[39] For constants $K > 0$ and $\alpha, \beta \geq 0$, the beta density $K(q_i)^{\alpha-1}(1 - q_i)^{\beta-1}$ $(0 \leq q_i \leq 1)$ is unimodal if

The role of (H4) in our analysis is to ensure that the prior density does not place higher probabilities to asymmetric realizations of q_ℓ and q_r so that the ratio effect holds.[40] For simplicity, we also assume that:

(H5) C is uniformly distributed on $[0, c]$.[41]

Using the ratio and level effects, we can now easily prove the uniqueness of asymptotic equilibrium. Suppose by contradiction that there are two different solutions, say $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$, to (7). By (9), it is clear that the sign of $C_\ell^\bullet - \overline{C}_\ell^\bullet$ is the same as that of $C_r^\bullet - \overline{C}_r^\bullet$. But if, say, $\overline{C}_\ell^\bullet > C_\ell^\bullet$ and $\overline{C}_r^\bullet > C_r^\bullet$, from (9) and the uniformity assumption on C it follows that $\overline{C}_\ell^\bullet/\overline{C}_r^\bullet$ is greater than or equal to $C_\ell^\bullet/C_r^\bullet$.[42] Thus, by (8), when moving from the smaller solution to the larger solution, (if any) the ratio effect creates a downward pressure on φ^r. But this is a contradiction, for (if any) the level effect also pushes φ^r downward so that $\overline{C}_r^\bullet \leq C_r^\bullet$. Hence, we proved that (7) has a unique solution which, as we noted earlier, must be the unique asymptotic equilibrium around which equilibria of large elections accumulate:

Proposition 7. (H4) *and* (H5) *imply that* (7) *has a unique solution which is the unique asymptotic equilibrium.*[43]

Next, we note that, under the uniformity assumption on C, (8) and (9) imply that $T_{r,\lambda}^\bullet/T_{\ell,\lambda}^\bullet$ is decreasing with λ.[44] Thus, through the ratio effect, a fall in λ causes a downward

---

$\alpha, \beta > 1$, and monotone if $\alpha \leq 1 \leq \beta$ or $\alpha \geq 1 \geq \beta$. When $\alpha = \beta = 1$, we obtain the uniform density. If q_ℓ and q_r are iid beta random variables, we have $g(q) = K^2 (q_\ell q_r)^{\alpha-1} ((1-q_\ell)(1-q_r))^{\beta-1}$. Strictly speaking, (H4) requires a little more than monotonicity: We need $\alpha \geq 1$ and $\beta \geq 1$, particular cases being (unimodal densities and) monotone densities with $\alpha = 1 \leq \beta$ or $\alpha \geq 1 = \beta$.

[40] To formally prove the ratio effect, let $T_\ell + T_r = \widetilde{T}_\ell + \widetilde{T}_r$ and $T_\ell \leq T_r < \widetilde{T}_r$. Then, by (H4), $g(\theta T_r, \theta T_\ell) \geq g(\theta \widetilde{T}_r, \theta \widetilde{T}_\ell)$ for every $\theta \leq 1/\widetilde{T}_r$, and hence, $\varphi^\ell(T_\ell, T_r) + \varphi^r(T_\ell, T_r) = \frac{T_\ell + T_r}{\overline{q}_\ell} \int_0^{1/T_r} \theta g(\theta T_r, \theta T_\ell) d\theta > \frac{\widetilde{T}_\ell + \widetilde{T}_r}{\overline{q}_\ell} \int_0^{1/\widetilde{T}_r} \theta g(\theta \widetilde{T}_r, \theta \widetilde{T}_\ell) g d\theta = \varphi^\ell(\widetilde{T}_\ell, \widetilde{T}_r) + \varphi^r(\widetilde{T}_\ell, \widetilde{T}_r)$. Similarly, it is easily seen that $\varphi^r(T_\ell, T_r) > \varphi^r(\widetilde{T}_\ell, \widetilde{T}_r)$, which holds, in fact, even if $\overline{q}_\ell \neq \overline{q}_r$.

[41] In Section III.C, we will discuss a more general class of cost distributions for which all of our results remain valid.

[42] This observation is trivially true among interior solutions. To deal with the general case, we first note that $F(\gamma C)/F(C)$ is a nonincreasing function of $C \in \mathbb{R}_{++}$ for every fixed $\gamma \geq 1$. Thus, $\overline{\gamma} \equiv \overline{C}_\ell^\bullet/\overline{C}_r^\bullet < \gamma \equiv C_\ell^\bullet/C_r^\bullet$ implies, by (9), $\gamma F(\gamma C_r^\bullet)/F(C_r^\bullet) = \overline{\gamma} F(\overline{\gamma}\overline{C}_r^\bullet)/F(\overline{C}_r^\bullet) < \gamma F(\gamma \overline{C}_r^\bullet)/F(\overline{C}_r^\bullet)$, a contradiction to the said property of F. Hence, $\overline{\gamma} \geq \gamma$, as needed. (To avoid case by case eliminations and for future use, in our formal arguments we do not rely on the exact shape of F.)

[43] Since the ratio effect is applied only to φ^r, Proposition 7 does not require the (implicit) assumption that $\overline{q}_\ell = \overline{q}_r$. (See Footnote 40.)

[44] As in Footnote 42, only boundary solutions require caution here. By contradiction, if we assume $\lambda < \overline{\lambda} \leq 1/2$ and $\gamma \equiv T_{r,\lambda}^\bullet/T_{\ell,\lambda}^\bullet \leq \overline{\gamma} \equiv T_{r,\overline{\lambda}}^\bullet/T_{\ell,\overline{\lambda}}^\bullet$, (8) and (9) imply $\overline{\gamma} F(\overline{\gamma} C_{r,\overline{\lambda}}^\bullet)/F(C_{r,\overline{\lambda}}^\bullet) < \gamma F(\gamma C_{r,\lambda}^\bullet)/F(C_{r,\lambda}^\bullet) \leq \overline{\gamma} F(\overline{\gamma} C_{r,\lambda}^\bullet)/F(C_{r,\lambda}^\bullet)$. Since $F(\overline{\gamma} C)/F(C)$ is nonincreasing in C, we must thus have $C_{r,\lambda}^\bullet < C_{r,\overline{\lambda}}^\bullet$ so that $C_{\ell,\lambda}^\bullet = \gamma C_{r,\lambda}^\bullet < \overline{\gamma} C_{r,\overline{\lambda}}^\bullet = C_{\ell,\overline{\lambda}}^\bullet$ and $T_{\ell,\lambda}^\bullet < T_{\ell,\overline{\lambda}}^\bullet$. Hence, $T_{\ell,\lambda}^\bullet + T_{r,\lambda}^\bullet = T_{\ell,\lambda}^\bullet(1+\gamma) < T_{\ell,\overline{\lambda}}^\bullet(1+\overline{\gamma}) = T_{\ell,\overline{\lambda}}^\bullet + T_{r,\overline{\lambda}}^\bullet$. Thus, by the level effect (and the ratio effect if any) $C_{r,\lambda}^\bullet > C_{r,\overline{\lambda}}^\bullet$, a contradiction. Conclusion: $\lambda < \overline{\lambda} \leq 1/2$

pressure on the cutoff point of the majority and the sum of the two cutoff points. Assuming the case of full participation away, this must met by a fall in the expected turnout rate, for otherwise the level effect (if any) would further decrease $C_r^\bullet$ and $C_\ell^\bullet + C_r^\bullet$, while (in contradiction with the supposition that the expected turnout rate did not fall) concavity of F on $\mathbb{R}_+$ implies that the loss in the expected turnout rate caused by a decrease in the (smaller) cutoff of the majority cannot be compensated by a smaller increase in the (larger) cutoff of the minority.[45] Thus, we have the following further result on the polarization effect which establishes a monotone relation between the level of disagreement and the expected turnout rate:

Proposition 8. *Suppose* (H4) *and* (H5) *hold. Then an increase in the level of disagreement increases the expected turnout rate provided that there is a room for such an increase. Otherwise, the expected turnout rate remains constant at its maximum possible level as a response to an increase in the level of disagreement.*

The winning probability of candidate r is an increasing function of $\frac{T_r^\bullet}{T_\ell^\bullet}$. When g is a symmetric function that satisfies (H4), with some algebra (that we carry out in Appendix E) it can be shown that the same is true for the *expected margin of victory*, which is defined as $MV \equiv \mathbf{E}\left|\frac{T_r^\bullet q_r - T_\ell^\bullet q_\ell}{T_r^\bullet q_r + T_\ell^\bullet q_\ell}\right|$. Thus:

Proposition 9. *Assuming that* (H4) *and* (H5) *hold and that g is a symmetric function, the expected margin of victory and the winning probability of the majority are decreasing functions of the level of disagreement.*

Together, Propositions 8 and 9 imply a negative correlation between the expected turnout rate and the expected margin of victory, which is known as the competition effect.

Let us now investigate comparative statics with respect to the per capita value of the election, u. The first point to note is that, among interior solutions, $T_r^\bullet/T_\ell^\bullet$ depends only on λ. Hence, assuming the case of full participation away, the expected turnout rate is an increasing function of u, for an increase in u with a falling or constant expected turnout rate would increase both cutoff points due to a nonnegative level effect, which is a contradiction. (If boundary solutions are taken into account, the ratio effect may work in the diversion, i.e., $T_r^\bullet/T_\ell^\bullet$ may increase with u, but it is not difficult to see that this does not change

implies $\gamma > \overline{\gamma}$.

[45] Formally, suppose by contradiction that $\lambda < \overline{\lambda} \le 1/2$ and $T_{\ell,\lambda}^\bullet + T_{r,\lambda}^\bullet \ge T_{\ell,\overline{\lambda}}^\bullet + T_{r,\overline{\lambda}}^\bullet$. Then we have $C_{r,\lambda}^\bullet < C_{r,\overline{\lambda}}^\bullet$ and $C_{\ell,\lambda}^\bullet + C_{r,\lambda}^\bullet < C_{\ell,\overline{\lambda}}^\bullet + C_{r,\overline{\lambda}}^\bullet$, for the the level effect is not neutralizing the ratio effect. This implies $\lambda F(C_{\ell,\lambda}^\bullet) + (1-\lambda)F(C_{r,\lambda}^\bullet) \le \overline{\lambda} F(C_{\ell,\lambda}^\bullet) + (1-\overline{\lambda})F(C_{r,\lambda}^\bullet) \le \overline{\lambda} F(C_{\ell,\overline{\lambda}}^\bullet) + (1-\overline{\lambda})F(C_{r,\overline{\lambda}}^\bullet)$. Here, the second inequality is obtained by some algebra using concavity of F, and, in fact, both inequalities are strict unless $C_{r,\lambda}^\bullet \ge c$, which gives us the desired contradiction: $T_{\ell,\lambda}^\bullet + T_{r,\lambda}^\bullet < T_{\ell,\overline{\lambda}}^\bullet + T_{r,\overline{\lambda}}^\bullet$.

the conclusion that the expected turnout rate increases with u.[46]) Moreover, $T_r^\bullet/T_\ell^\bullet$ is nondecreasing with u, for when $C_\ell^\bullet$ reaches c, $C_r^\bullet$ is still below c, and a further increase in u increases $C_r^\bullet$ while keeping $T_\ell^\bullet$ constant up to the point where $C_r^\bullet$ reaches c, after which $T_r^\bullet/T_\ell^\bullet$ remains constant at its maximum level, $\frac{1-\lambda}{\lambda}$.[47] Finally, we note that in these arguments we can replace u with $u\psi$ and that the comparative statics with respect to c^{-1} are analogous.[48] Thus, the following observation is true:

Proposition 10. *Suppose* (H4) *and* (H5) *hold. Then the expected turnout rate increases with $u\psi$ provided that there is a room for such an increase. Otherwise, the expected turnout rate remains constant at its maximum possible level as a response to an increase in $u\psi$. When g is symmetric, the expected margin of victory is also a nondecreasing function of $u\psi$. Moreover, the consequences of a fall in c are analogous to those of a rise in $u\psi$.*[49]

Remark 3. In Appendix A, by means of an example, we will formally demonstrate that our explanation of large-scale turnout and strategic behavior are conceptually independent from the homogeneity assumptions that we employed in this section. More specifically, we will show that when for each agent, u is a random draw from a uniform distribution, Propositions 3-9 and a suitable modification of Proposition 10 continue to hold.

B. *Relations to the RU-voter Literature*

In their RU-voter model, where q_i is interpreted as the fraction of utilitarian agents in the policy group i who take as given the actions of only type j agents, Feddersen and Sandroni (2007) prove analogues of Propositions 4-10 under the additional assumption that q_ℓ and q_r are independently and uniformly distributed. In Appendix C, we will show that

[46] Indeed, if $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$ solve (7) for u and $\overline{u} > u$, respectively, by an obvious modification of the arguments in Footnote 44, the assumption $\gamma \equiv C_\ell^\bullet/C_r^\bullet < \overline{\gamma} \equiv \overline{C}_\ell^\bullet/\overline{C}_r^\bullet$ yields $C_r^\bullet < \overline{C}_r^\bullet$ and $T_\ell^\bullet + T_r^\bullet < \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. So, the expected turnout rate that corresponds to $\overline{u}$ is higher, as we seek.

[47] The formal proof of this argument is tedious and therefore omitted. However, that $T_r^\bullet/T_\ell^\bullet$ is nondecreasing with u is easily proved: Using the notation of Footnote 46, $\gamma > \overline{\gamma}$ would imply $C_r^\bullet > \overline{C}_r^\bullet$ and $T_\ell^\bullet + T_r^\bullet > \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. But this is a contradiction, for the ratio and level effects together with the fact that $u < \overline{u}$ would then imply $C_r^\bullet < \overline{C}_r^\bullet$.

[48] In fact, $(F(C_\ell^\bullet), F(C_r^\bullet))$ depends only on (λ and) $u\psi/c$, but not on particular values of u, ψ and c. To prove this point, the critical observation is that, under (H5), given two positive numbers c and $\overline{c}$, we have $F_c(C) = F_{\overline{c}}(\frac{\overline{c}}{c}C)$ for every $C \in \mathbb{R}_+$, where F_c (resp. $F_{\overline{c}}$) is the distribution that corresponds to c (resp. $\overline{c}$). Thus, (C_ℓ, C_r) solves (7) iff $(\frac{\overline{c}}{c}C_\ell, \frac{\overline{c}}{c}C_r)$ solves the modified version of (7) that is obtained by replacing $u\psi$ and c with $\frac{\overline{c}}{c}u\psi$ and $\overline{c}$, respectively.

[49] Neither Jankowski (2002) nor Edlin et al. (2007) comment on determinants of the expected closeness of an election, which can be measured in our model by $-MV$. Both of these former papers do provide intuitive explanations of the relation between turnout and the importance of an election. While the discussion of Edlin et al. (2007) is closer in spirit to the present model, it takes as given the expected closeness of the election.

the relation between the two models is beyond this similarity of comparative statics. In fact, assuming $\psi \equiv 1$, first order conditions of the model of Feddersen and Sandroni (2007) turn out to be the same as (7). Hence, when the first order conditions in the RU-voter model are sufficient, the two models predict exactly the same turnout rates.

In light of this equivalence result, from a formal point of view, our comparative statics exercises are merely generalizations of those of Feddersen and Sandroni (2007). While these generalizations seem material,[50] perhaps a more important merit of our discussion has been the identification of level and ratio effects as the driving forces behind the comparative statics of our model.

Moreover, there are some important differences between the interpretations of the comparative statics of the two models: In our model, the underdog effect and the ratio effect are driven by Bayesian updating of agents' beliefs, while in the RU-voter model of Feddersen and Sandroni (2007) the analogues of these properties are consequences of an analytical relation (equation (C-4) below) between the expected turnout rate of a policy group and their winning probability, which comes about despite the fact that there is no informational asymmetry between agents.

In passing, let us emphasize that the information structure that we use in this paper have interesting, peculiar implications. Notably, the prior probability of victory for candidate i tends to be smaller than the posterior probability that an altruistic agent of type i places to this event. In particular, when there is no overwhelming majority, altruistic agents in *both* policy groups may believe that their favorite candidate is more likely to win the election.[51] This is consistent with A. J. Fischer's (1999) observations on Australian voters which show that in 1994, a large majority of the supporters of the Australian Labor Party thought that they are going to win the next election, while a large majority of the supporters of the Liberal National Coalition held the opposite belief. Of course, such difference in beliefs cannot be understood from a prior point of view.

C. *Extensions*

Using Lemma 1, it is a straightforward exercise to extend equations (7) to the case λ is a simple random variable that is independent from q. Following our discussion in Appendix C,

[50] In particular, we allowed q_ℓ and q_r to be stochastically dependent. In the extended version of their paper, Feddersen and Sandroni (2002) keep the independence assumption, replace (H4) with the assumption that g is nonincreasing and relax (H5) at the cost of the conclusions of Proposition 8 and (partially) Proposition 10.

[51] For example, if q is uniform on $\mathfrak{I}$ and $\lambda \equiv 1/2$, according to the posterior of altruistic agents of type ℓ, the probability of their victory is $\int_0^1 \left(\int_0^{q_\ell} \frac{q_\ell}{1/2} dq_r \right) dq_\ell = \int_0^1 2(q_\ell)^2 dq_\ell = \frac{2}{3}$; and similarly, for altruistic agents of type r. (Taking λ as also random may create a similar pattern even in selfish agents' views.)

it is also easy to see that these extended equations are nothing but the first order conditions of the corresponding model of the type considered by Feddersen and Sandroni (2006, 2007). We also worked through the case where λ is a continuous random variable and q is simple, and obtained analogous conclusions. (The details are available upon request.) It remains as an open question to determine whether in these observations the simple random variable can be replaced with an arbitrary random variable.

Finally, let us add that in Propositions 7-10 the assumption that C is uniformly distributed can be replaced with the following conditions: (i) F is continuously differentiable and strictly increasing on its support which is a closed subinterval of $\mathbb{R}_+$ that contains 0; (ii) F is concave on $\mathbb{R}_+$; and (iii) $F(\gamma C)/F(C)$ is a nonincreasing function of $C \in \mathbb{R}_{++}$ for every fixed $\gamma \geq 1$. (The uses of the latter property are demonstrated in Footnotes 42, 44, 46 and 47. The extension that we present in Appendix A also benefits from this property.) Examples of such distribution functions include the exponential distribution $F(C) \equiv 1-e^{-\beta C}$ $(C \in \mathbb{R}_+, \beta > 0)$ (or the exponential distribution conditioned to an interval $[0, c]$) and functions of the form $F(C) \equiv c^{-\beta} C^{\beta}$ for $C \in [0, c]$ and some fixed $\beta \in (0, 1]$. In particular, excluding the comparative statics exercise with respect to c, the assumption that F has a bounded support is dispensable in all of our results.

IV. Concluding Remarks

A long tradition in the literature gives a central importance to collectively acting groups in explaining voter turnout in costly, large elections. In the elite-driven mobilization models of Carole J. Uhlaner (1989), Rebecca B. Morton (1991) and Ron Shachar and Bany Nalebuff (1999), turnout is explained by elites' ability to mobilize large groups of voters while the exact nature of this mobilization process is left unmodelled/unexplained.[52] Under some homogeneity assumptions, RU-voter models provide a solution to this coordination problem in which agents' sense of group membership takes the role of leaders, but, unfortunately, they do not explain why, or to what extent, a potential heterogeneity in voters' characteristics can be ignored.[53] (For a formal discussion of this matter, see Appendix C.)

On the other hand, while it is free of such conceptual difficulties, the standard pivotal-voter model has difficulty in explaining turnout in costly, large elections. In light of the mounting evidence on the role of altruistic/ethical concerns in shaping political/economic decisions, we introduced a game-theoretic, pivotal-voter model that relates the importance

[52] Feddersen (2004) presents a detailed discussion of conceptual issues related to mobilization models.

[53] In fact, Feddersen and Sandroni (2007) are aware of this matter and argue that by providing credible information, say, about the importance of the election, elites may be contributing to turnout. Thus, future research in this direction may focus on modelling how elites communicate with individuals and how this helps the formation of collectively acting large groups.

of an election to the size of the electorate. Under homogeneity assumptions, our model performs as well as some canonical RU-voter models, and has the further advantage of being compatible even with extreme forms of heterogeneity in agents' characteristics. It is thus hoped that this paper will help to close the discrepancy between our understanding of costless or small elections, which is largely shaped by pivotal-voter models, and that of costly, large elections.

Appendix

A. *Heterogeneity in Agents' Valuations*

Here, we consider an extension of the model that we discussed in Section III where for each agent per capita value of winning the election, $u \geq 0$, is a random draw from a continuous distribution that is independent from any other random variable in the model. We also assume that u is independently distributed across agents. This framework corresponds to an extreme scenario where any agent of a given type knows that almost surely there is no other agent with whom she can precisely agree about the per capita payoff of changing the winner.

In this modified model, the maximum cost level at which a given altruistic agent of type i is willing to vote (for her favorite candidate) also depends on the realization of u for this particular agent, which we denote by $C^*_{u,i}$. In fact, because of independence assumptions, any two altruistic agents of type i face the same pivot probability, and hence, we have $C^*_{u,i} = uC^*_i$ for every $u \geq 0$, where $C^*_i \equiv C^*_{1,i}$. Thus, conditional on q, a randomly chosen agent votes for candidate i with probability $\lambda_i q_i \mathbf{E}F(uC^*_i)$. It is clear that (C^*_ℓ, C^*_r) is nothing but the equilibrium of a *dual model* where u is known to be 1 and costs are distributed according to $\widetilde{F}(C) \equiv \mathbf{E}F(uC)$ $(C \in \mathbb{R}_+)$, and that this dual model is equivalent to the model with random u for all practical purposes.

As a concrete example, let us assume that u and C are uniformly distributed on $[0, \mathfrak{u}]$ and $[0, c]$, respectively. We then find that, with $\omega \equiv \frac{c}{\mathfrak{u}}$,

$$\widetilde{F}(C) = \begin{cases} \frac{1}{2\omega}C & \text{if } 0 \leq C \leq \omega, \\ 1 - \frac{\omega}{2}C^{-1} & \text{if } C > \omega. \end{cases}$$

It is easily verified that, for every $\gamma \geq 1$, $\widetilde{F}(\gamma C)/\widetilde{F}(C)$ is a nonincreasing function of $C \in \mathbb{R}_{++}$ which is in fact given by

$$\frac{\widetilde{F}(\gamma C)}{\widetilde{F}(C)} = \begin{cases} \gamma & \text{if } 0 < C \leq \frac{\omega}{\gamma}, \\ \omega\left(2C^{-1} - \frac{\omega}{\gamma}C^{-2}\right) & \text{if } \frac{\omega}{\gamma} < C \leq \omega, \\ \gamma^{-1}\left(\frac{\gamma-1}{1-\omega/(2C)} + 1\right) & \text{if } C > \omega. \end{cases}$$

Moreover, the density of C is found as

$$\widetilde{f}(C) = \begin{cases} \frac{1}{2\omega} & \text{if } 0 \leq C \leq \omega, \\ \frac{\omega}{2}C^{-2} & \text{if } C > \omega, \end{cases}$$

which is continuous and nonincreasing. (In particular, $\widetilde{F}$ is concave.)

As we discussed in Section III.C, it thus follows that Propositions 3-9 apply to the dual model (provided that g is a symmetric function that satisfies (H4)). Moreover, it is not difficult to see that the expected turnout rate is decreasing with ω/ψ, and the expected margin of victory is nonincreasing with ω/ψ. (The key observation here is that $\widetilde{F} \equiv \widetilde{F}_\omega$ depends only on ω, and given two positive numbers ω and $\overline{\omega}$, we have $\widetilde{F}_\omega(C) = \widetilde{F}_{\overline{\omega}}(\frac{\overline{\omega}}{\omega}C)$ for every $C \geq 0$. Thus, following the argument in Footnote 48, $(\widetilde{F}(C_\ell^\bullet), \widetilde{F}(C_r^\bullet))$ depends only on λ and ω/ψ. Moreover, the discussion in Section III about the comparative statics with respect to ψ applies as is (see Footnote 46). Hence the conclusion on the comparative statics with respect to ω/ψ.)

B. *Simulations*

In this part of the Appendix, we will present the results of simulations that we performed to test the accuracy of the asymptotic approximation that we used in Section III. In these simulations, we focused on the case where q and C are uniformly distributed on $[0,1]^2$ and $[0,c]$, respectively. In this case, the closed form solutions of (7) can easily be found. As functions of parameters, the unique asymptotic equilibrium, $(C_\ell^\bullet, C_r^\bullet)$, and the asymptotic turnout rates of altruistic agents, $(F(C_\ell^\bullet), F(C_r^\bullet))$, are as follows:[54]

$$(C_\ell^\bullet, C_r^\bullet) = \begin{cases} \left(\frac{\sqrt{u\psi c}}{\sqrt[4]{\lambda(1-\lambda)}}, \sqrt{u\psi c}\sqrt[4]{\frac{\lambda}{(1-\lambda)^3}}\right) & \text{if } \frac{u\psi}{\sqrt{\lambda(1-\lambda)}} \leq c, \\ \left(\frac{(u\psi)^{2/3}c^{1/3}}{\sqrt[3]{\lambda(1-\lambda)}}, \sqrt[3]{\frac{u\psi\lambda c^2}{(1-\lambda)^2}}\right) & \text{if } \frac{u\psi\lambda}{(1-\lambda)^2} \leq c < \frac{u\psi}{\sqrt{(1-\lambda)\lambda}}, \\ \left(\frac{u\psi}{(1-\lambda)}, \frac{u\psi\lambda}{(1-\lambda)^2}\right) & \text{if } \frac{u\psi\lambda}{(1-\lambda)^2} > c; \end{cases}$$

$$(F(C_\ell^\bullet), F(C_r^\bullet)) = \begin{cases} \left(\frac{\sqrt{u\psi/c}}{\sqrt[4]{\lambda(1-\lambda)}}, \sqrt{u\psi/c}\sqrt[4]{\frac{\lambda}{(1-\lambda)^3}}\right) & \text{if } \frac{u\psi}{\sqrt{\lambda(1-\lambda)}} \leq c, \\ \left(1, \sqrt[3]{\frac{u\psi}{c}\frac{\lambda}{(1-\lambda)^2}}\right) & \text{if } \frac{u\psi\lambda}{(1-\lambda)^2} \leq c < \frac{u\psi}{\sqrt{(1-\lambda)\lambda}}, \\ (1,1) & \text{if } \frac{u\psi\lambda}{(1-\lambda)^2} > c. \end{cases}$$

The values of n that we considered in simulations range between 249 and 4999. By slightly deviating from our theoretical model, in simulations we ignored an altruistic agent's own material payoff. In other words, we set $\Pi_i = u\psi n\frac{1}{2}PIV_i$. This minor deviation ensures that $(F(C_\ell^*), F(C_r^*))$ depends only on (λ and) $u\psi/c$, but not on the particular values of u, ψ and c.[55]

To test the performance of our asymptotic approximation, we set $u\psi/c = 1/4$ and considered three different values of λ, namely, 1/3, 0.4 and 0.5. The assumption $u\psi/c = 1/4$ holds, for instance, in the reasonable scenario where $u = 250\$$, $c = 100\$$ and $\psi = 1/10$. We observed that for smaller values of λ (namely, for 1/3 and 0.4), the asymptotic turnout rates of both policy groups agree almost precisely with their finite counterparts whenever $n \geq 999$ (Table B1 below). For larger values of λ the convergence seems to be slower. However, even with $\lambda = 1/2$, as n reaches 4999 the errors involved in approximations

[54]The reader will note that $F(C_\ell^\bullet)$ and $F(C_r^\bullet)$ coincide with the corresponding predictions of the RU-voter model of Feddersen and Sandroni (2007) where a parameter, w, that measures the importance of the election takes the role of $u\psi$. In the next section, we will show that this equivalence is a general phenomenon.

[55]When computing Π_i and when finding the equilibria we used the standard double integration and fixed point algorithms of MATLAB (namely, dblquad and fsolve), respectively.

fall below 1.5%. It thus seems that our asymptotic results are likely to be useful in the analysis of most political elections of interest.

Table B1–Convergence

	λ=1/3				λ=0.4				λ=0.5	
	type ℓ		type r		type ℓ		type r		$i=\ell,r$	
n	$F(C_\ell^*)$	*error*	$F(C_r^*)$	*error*	$F(C_\ell^*)$	*error*	$F(C_r^*)$	*error*	$F(C_i^*)$	*error*
249	.7231	.72%	.5153	−.07%	.7052	1.30%	.5790	.74%	.6628	6.68%
499	.7258	.34%	.5154	−.09%	.7112	.44%	.5827	.11%	.6762	4.58%
749	.7268	.21%	.5153	−.07%	.7126	.25%	.5833	.01%	.6820	3.69%
999	.7271	.17%	.5152	−.05%	.7133	.15%	.5834	−.02%	.6854	3.17%
1500	.7271	.17%	.5150	−.01%	.7137	.10%	.5834	−.02%	.6887	2.68%
2500	.7277	.08%	.5150	−.01%	.7138	.08%	.5833	.01%	.6935	1.96%
4999	.7280	.03%	.5150	−.01%	.7141	.04%	.5833	.00%	.6977	1.35%
∞	.7283		.5150		.7144		.5833		.7071	

Notes: The entries correspond to the case $u\psi/c = 1/4$ and are subject to rounding errors up to .01%. The last row gives the asymptotic figures as found in text. *error* is defined as $(F(C_i^\bullet) - F(C_i^*))/F(C_i^*)$.

Table B1 also shows that even for relatively small values of n, the results of our simulations were consistent with the underdog effect. The same is true for the competition effect: We observed that the expected turnout rate, $\frac{1}{2}(T_\ell^* + T_r^*)$, increases with λ while the expected margin of victory, $\mathbf{E}\left|\frac{q_r T_r^* - q_\ell T_\ell^*}{q_r T_r^* + q_\ell T_\ell^*}\right|$, falls with λ (Table B2 below). (We recall that the expected margin of victory is an increasing function of T_r^*/T_ℓ^*.) Moreover, in line with Proposition 5, when λ equals 1/3 or 0.4 we observed that T_r^* is larger than T_ℓ^*, i.e., the majority is more likely to win.

Table B2–The Competition Effect

	expected turnout rate (in %)			T_r^*/T_ℓ^*		
n	λ=1/3	λ=0.4	λ=0.5	λ=1/3	λ=0.4	λ=0.5
249	29.23	31.47	33.14	1.4254	1.2317	1
499	29.28	31.70	33.81	1.4202	1.2289	1
749	29.29	31.75	34.10	1.4181	1.2277	1
999	29.29	31.77	34.27	1.4172	1.2269	1
1500	29.28	31.78	34.43	1.4167	1.2262	1
2500	29.30	31.77	34.68	1.4154	1.2257	1
4999	29.30	31.78	34.88	1.4148	1.2252	1
∞	29.30	31.79	35.36	1.4142	1.2248	1

Notes: See the notes in Table B1.

Finally, for the case λ=0.4, we verified that even for the smallest value of n in our sample (i.e., 249), the expected turnout rate increases with $u\psi/c$. Moreover, among interior solutions, the expected margin of victory is roughly independent from $u\psi/c$.

Table B3–The Effects of $u\psi/c$

	expected turnout rate (in %)				T_r^*/T_ℓ^*			
n	$\frac{u\psi}{c}=\frac{1}{8}$	$\frac{u\psi}{c}=\frac{2}{8}$	$\frac{u\psi}{c}=\frac{3}{8}$	$\frac{u\psi}{c}=0.4899^\dagger$	$\frac{u\psi}{c}=\frac{1}{8}$	$\frac{u\psi}{c}=\frac{2}{8}$	$\frac{u\psi}{c}=\frac{3}{8}$	$\frac{u\psi}{c}=0.4899^\dagger$
249	22.12	31.47	38.65	44.24	1.2337	1.2317	1.2307	1.2301
∞	22.48	31.79	38.93	44.5	1.2248	1.2248	1.2248	1.2248

Notes: The entries correspond to the case $\lambda = 0.4$ and are subject to rounding errors up to .01%. The last row gives the asymptotic figures as found in text.

(†) The largest value of $u\psi/c$ that ensures interiority of the asymptotic equilibrium for $\lambda = 0.4$.

Remark B1. The experimental findings of Levine and Palfrey (2007) indicate a negative correlation between the turnout rate and the size of the electorate. While this is not consistent with the results of our simulations, our model can be made compatible with falling turnout rates by considering a strictly concave function $\Psi(n)$. (Such an extension would require a more careful analysis of pivot probabilities for a given n, which is beyond the scope of the present paper.) Let us note, however, that whether turnout rates in real elections fall with the size of the electorate is a debatable issue (see, e.g., Edlin et al. (2007)).

C. *On RU-voter Models*

Here we will consider a RU-voter model which is obtained by replacing the altruistic agents in Section III with rule utilitarian agents. Our focus will be the limiting case of infinitely many agents. Excluding some differences in presentation, the model that we will discuss is simply the one introduced by Feddersen and Sandroni (2006, 2007). We will first show that the predictions of this RU-voter model coincide with those of Section III. We will then discuss the role of homogeneity assumptions in this framework.

The set of agents is $\{1, 2, ...\}$. As in Feddersen and Sandroni (2006, 2007), we could equivalently assume that there is a continuum of agents and (instead of assuming conditional independence,) directly interpret λ as the fraction of agents of type ℓ, and similarly for q_ℓ and q_r. The reason for our choice will be apparent shortly. Throughout the discussion, (H1) and (H3) will be implicitly assumed.

Each agent h acts according to a type contingent (measurable) action profile $A^h : \Theta \to \{-1, 0, 1\}$ (to be determined endogenously), which takes the role of a strategy in the previous model. Here, $\Theta \equiv \{\ell, r\} \times \{a, s\} \times [0, c]$ is the set of types, and an agent of personality type a is now understood as a rule utilitarian agent. We assume $A^1 = A^2 = ...$ so that $A^1 \circ \tau_1, A^2 \circ \tau_2, ...$ are iid random variables conditional on q. For every realization of q, the winner is determined by the excess of the vote share of candidate r over that of candidate ℓ, $\mathcal{X} \equiv \lim_n \frac{\sum_{h=1}^n A^h \circ \tau_h}{n}$, which is well defined (and equals the conditional mean of $A^h \circ \tau_h$) by the law of large numbers. If $\mathcal{X} > 0$ (resp. $\mathcal{X} < 0$) candidate r wins (resp. loses), and in case of a tie, i.e. when $\mathcal{X} = 0$, the winner is determined by tossing a fair coin.

Each agent of a given type $\tau \in \Theta$ considers herself to belong to an exogenously given *(type) group* $\Theta_\tau \subseteq \Theta$ with $\tau \in \Theta_\tau$, subject to a consistency condition to be discussed shortly. A *rule* R_τ for type τ is a map from Θ_τ into the set of actions $\{-1, 0, 1\}$. The rule that type τ chooses is understood as the ethical rule which, according to type τ, *should* be followed by all types in Θ_τ. That is, if $\widetilde{\tau} \in \Theta_\tau$ and R_τ is the rule accepted by type τ, then type τ understands that $R_\tau(\widetilde{\tau})$ is the action which should be taken by all agents of type $\widetilde{\tau}$.

Only rule utilitarian agents, however, get a payoff $\overline{\mathcal{D}}$ from acting as they should. In other words, for a given agent, the value of acting ethically, $\mathcal{D}$, is a binary random variable which equals $\overline{\mathcal{D}}$ if the agent under consideration is rule utilitarian and 0 otherwise. Moreover, an agent votes for a given candidate if and only if (i) she should do so (according to the rule that she deems ethical); and (ii) $\mathcal{D}$ is above her voting cost. Following Feddersen and Sandroni (2007), we assume $\overline{\mathcal{D}} \geq c$ so that an agent who should vote for candidate i does so if and only if she is rule utilitarian and abstains otherwise. Put formally, the choice of a rule R_τ determines an action profile $A_{R_\tau} : \Theta_\tau \to \{-1, 0, 1\}$ defined by $A_{R_\tau}(i, s, C) \equiv 0$ and $A_{R_\tau}(i, a, C) \equiv R_\tau(i, a, C)$. Thereby, given an action profile $A_{-\Theta_\tau}$ for types outside Θ_τ, the choice of a rule R_τ also determines the winning probability of candidate i and the expected per capita cost, which we denote by $\mathbb{P}_i(A_{R_\tau}, A_{-\Theta_\tau})$ and $\mathbb{C}(A_{R_\tau}, A_{-\Theta_\tau})$, respectively.[56] Following Feddersen and Sandroni (2006, 2007), we assume that when computing $\mathbb{P}_i$ each agent uses the prior density g.

The ethical rule for an agent of type τ with policy component i maximizes the expected per capita payoff $u\mathbb{P}_i(A_{R_\tau}, A_{-\Theta_\tau}) - \mathbb{C}(A_{R_\tau}, A_{-\Theta_\tau})$ (from her own perspective) while taking as given the actions of types outside her group, Θ_τ. Each type τ solves this optimization problem under the assumption that a type $\widetilde{\tau} \in \Theta_\tau$ should follow R_τ, and therefore, will take the action $A_{R_\tau}(\widetilde{\tau})$. However, in fact, a type $\widetilde{\tau} \in \Theta_\tau$ takes the action $A_{R_{\widetilde{\tau}}}(\widetilde{\tau})$, i.e., the action induced by the rule that she herself deems ethical. Thus, as a basic *consistency requirement*, we seek a rule profile $(R^*_\tau)_{\tau\in\Theta}$ such that, for every $\tau, \widetilde{\tau} \in \Theta$ with $\widetilde{\tau} \in \Theta_\tau$,

$$A_{R^*_{\widetilde{\tau}}}(\widetilde{\tau}) = A_{R^*_\tau}(\widetilde{\tau}). \tag{C-1}$$

In words, each type who considers herself in the same group with $\widetilde{\tau}$ must correctly anticipate the action that $\widetilde{\tau}$ will take.[57] Moreover, optimality requires that for every type τ with policy component i and for every possible rule R_τ for the group Θ_τ, we have

$$u\mathbb{P}_i(A_{R^*_\tau}, A_{R^*_{-\Theta_\tau}}) - \mathbb{C}(A_{R^*_\tau}, A_{R^*_{-\Theta_\tau}}) \geq u\mathbb{P}_i(A_{R_\tau}, A_{R^*_{-\Theta_\tau}}) - \mathbb{C}(A_{R_\tau}, A_{R^*_{-\Theta_\tau}}), \tag{C-2}$$

where $A_{R^*_{-\Theta_\tau}}$ is the action profile on $\Theta\backslash\Theta_\tau$ defined by $A_{R^*_{-\Theta_\tau}}(\widetilde{\tau}) \equiv A_{R^*_{\widetilde{\tau}}}(\widetilde{\tau})$ for $\widetilde{\tau} \in \Theta\backslash\Theta_\tau$.

Following Feddersen and Sandroni (2006, 2007), we focus on the case where type groups coincide with policy groups. Hence, for $i = \ell, r$, we define $\Theta_\tau \equiv \Theta_i$ for every $\tau \in \Theta_i \equiv \{i\} \times \{a, s\} \times [0, c]$, and denote by R^*_i a common rule accepted by all types in Θ_i. Thus, $R^*_\tau \equiv R^*_i$ for every $\tau \in \Theta_i$ so that the consistency conditions (C-1) trivially hold. Feddersen and Sandroni (2006, Proposition 2) show that if the rule profile (R^*_ℓ, R^*_r) also satisfies the optimality conditions (C-2), then it must have a cutoff structure. That is, there must be a pair of numbers $(C^+_\ell, C^+_r) \in [0, c]^2$ such that for (almost every) $C \in [0, c]$, $R^*_\ell(\ell, a, C)$ equals -1 if $C \leq C^+_\ell$ and 0 otherwise; and $R^*_r(r, a, C)$ equals 1 if $C \leq C^+_r$ and 0 otherwise. (What a selfish agent should do has no consequences, and therefore, cannot be determined.) Moreover, both cutoff points must be positive, for when $C^+_j = 0$, the optimization problem of group i has no solution (since then, for any positive value of C^+_i candidate i wins (with probability 1), while $C^+_i = 0$ implies a tie). It is also worth noting that since cutoff points are positive, the event of a tie, $\left\{\frac{q_\ell}{q_r} = \frac{(1-\lambda)F(C^+_r)}{\lambda F(C^+_\ell)}\right\}$, is null.

The following is the promised equivalence result:

Proposition C1. *Suppose that* (H1)-(H5) *hold and that g is nonincreasing (i.e., that G is concave). Also assume $\psi \equiv 1$. Then, in terms of the predicted turnout rates among both policy groups, the unique asymptotic equilibrium of the altruistic-voter model is the same as the unique solution of the RU-voter model.*

Proof of Proposition C1 is simple enough to present here. First, we note that if the two groups were to choose a pair of cutoff points $(C_\ell, C_r) \in (0, c]^2$, the induced expected per capita cost would be

$$\mathbb{C} = \lambda_\ell \bar{q}_\ell \int_0^{C_\ell} f(C)CdC + \lambda_r \bar{q}_r \int_0^{C_r} f(C)CdC, \tag{C-3}$$

and candidate i would win with probability

$$\mathbb{P}_i = \Pr\left\{\frac{q_j}{q_i} \leq \frac{T_i}{T_j}\right\},$$

[56] It is important to note that $\mathbb{P}_i$ depends (among other things) on the distribution of $\mathcal{D}$ which is, in the present case where $\overline{\mathcal{D}}$ is fixed, entirely determined by the distribution of q. When discussing heterogeneity issues, we will comment on consequences of taking $\overline{\mathcal{D}}$ as random.

[57] It can easily be seen that (C-1) is equivalent to Definition 1.2 of Feddersen and Sandroni (2006b).

where $T_i \equiv \lambda_i F(C_i)$. Thus, we have $\frac{\partial \mathbb{P}_i}{\partial T_i} = g_{q_j/q_i}(\frac{T_i}{T_j})\frac{1}{T_j}$, where g_{q_j/q_i} is the prior density of q_j/q_i. Using a well known formula, it easily follows that $g_{q_\ell/q_r}(w) = \int_0^{\min\{1,1/w\}} g(yw, y)ydy$ and $g_{q_r/q_\ell}(w) = \int_0^{\min\{1,1/w\}} g(y, yw)ydy$ for every $w > 0$ (see, e.g., Vijay K. Rohatgi, p. 141). Substituting $y = \theta T_j$ in the integral representation of $g_{q_j/q_i}(\frac{T_i}{T_j})$, we therefore find that, for $i = \ell, r$,

$$\frac{\partial \mathbb{P}_i}{\partial T_i} = \int_0^{\frac{1}{\max\{T_\ell, T_r\}}} g(\theta T_r, \theta T_\ell)\theta T_j dy = \overline{q}_i \varphi^i(T_\ell, T_r). \tag{C-4}$$

This tells us that $\frac{\partial \mathbb{P}_i}{\partial T_i}$ is proportional to the type i *posterior* asymptotic pivot probability. It should be noted that this equality comes about despite the fact that agents use the *prior* density g when computing $\mathbb{P}_i$. We will discuss this matter after completing the proof.

Next, we note that $\frac{\partial \mathbb{C}}{\partial C_i} = \lambda_i \overline{q}_i f(C_i) C_i$ and $\frac{\partial \mathbb{P}_i}{\partial C_i} = \frac{\partial \mathbb{P}_i}{\partial T_i}\frac{\partial T_i}{\partial C_i} = \lambda_i f(C_i)\overline{q}_i \varphi^i(T_\ell, T_r)$ by (C-3) and (C-4). Thus, the first order conditions for the optimality requirements (C-2) can be written as,

$$\text{for } i = \ell, r, \quad u\varphi^i(T_\ell^+, T_r^+) - C_i^+ \geq 0 \quad \text{with equality if } C_i^+ < c, \tag{C-5}$$

where $T_i^+ \equiv \lambda_i F(C_i^+)$. One moment's thought will convince the reader to the fact that, with $\psi \equiv 1$, if $(C_\ell^\bullet, C_r^\bullet)$ solves (7), then $(\min\{C_\ell^\bullet, c\}, \min\{C_r^\bullet, c\})$ solves (C-5); and conversely, if (C_ℓ^+, C_r^+) solves (C-5), then $(u\varphi^\ell(T_\ell^+, T_r^+), u\varphi^r(T_\ell^+, T_r^+))$ solves (7). Moreover, these transformations leave the turnout rates of the two groups unchanged; that is, we have $T_\ell^\bullet = T_\ell^+$ and $T_r^\bullet = T_r^+$. Finally, if g is nonincreasing, as can easily be seen from the right side of (6), φ^i is nonincreasing in T_i (and C_i). This implies that the first order conditions (C-5) are sufficient for optimality, *which proves Proposition C1.*

To gain a better intuition, we now wish to take a closer look at equation (C-4). Let us consider the finite case with $n+1$ agents. Suppose that conditional on q, all agents but one vote for candidates ℓ and r with probabilities $\overline{P}_\ell$ and $\overline{P}_r$, while the agent h that we excluded votes for candidates ℓ and r with probabilities $P_{\ell,h}$ and $P_{r,h}$, respectively. Then if we denote by $\mathbb{P}_{i,q}(n)$ the winning probability of candidate i conditional on q, we clearly have

$$\begin{aligned}\mathbb{P}_{i,q}(n) \quad = \quad & (1 - P_{\ell,h} - P_{r,h})\left(\Pr\left\{S_i^- \geq S_j^- + 1; \overline{P}_\ell, \overline{P}_r\right\} + \frac{1}{2}\Pr\left\{S_i^- = S_j^-; \overline{P}_\ell, \overline{P}_r\right\}\right) \\ & + P_{i,h}\left(\Pr\left\{S_i^- \geq S_j^-; \overline{P}_\ell, \overline{P}_r\right\} + \frac{1}{2}\Pr\left\{S_i^- = S_j^- - 1; \overline{P}_\ell, \overline{P}_r\right\}\right) \\ & + P_{j,h}\frac{1}{2}\Pr\left\{S_i^- = S_j^- + 1; \overline{P}_\ell, \overline{P}_r\right\}.\end{aligned}$$

After some simple algebra, it follows that

$$\frac{\partial \mathbb{P}_{i,q}(n)}{\partial P_{i,h}} = \frac{1}{2} piv_i(\overline{P}_\ell, \overline{P}_r, n).$$

Moreover, since $\mathbb{P}_{i,q}(n)$ is a symmetric function of $P_{i,1}, ..., P_{i,n+1}$, if we set $P_i \equiv P_{i,h}$ for every $h = 1, ..., n+1$, differentiating $\mathbb{P}_{i,q}(n)$ with respect to P_i and evaluating at $P_i = \overline{P}_i$ gives

$$\frac{\partial \mathbb{P}_{i,q}(n)}{\partial P_i} = \sum_{h=1}^{n+1} \frac{\partial \mathbb{P}_{i,q}(n)}{\partial P_{i,h}} = \frac{n+1}{2} piv_i(\overline{P}_\ell, \overline{P}_r, n). \tag{C-6}$$

With $P_\ell \equiv q_\ell T_\ell$ and $P_r \equiv q_r T_r$, we therefore find that

$$\frac{\partial \mathbb{P}_{i,q}(n)}{\partial T_i} = \frac{\partial \mathbb{P}_{i,q}(n)}{\partial P_i}\frac{\partial P_i}{\partial T_i} = \frac{n+1}{2} q_i piv_i(q_\ell T_\ell, q_r T_r, n). \tag{C-7}$$

It follows that the unconditional probability of the victory for candidate i, $\mathbb{P}_i(n)$, satisfies

$$\frac{\partial \mathbb{P}_i(n)}{\partial T_i} = \int_{\mathfrak{I}} \frac{\partial \mathbb{P}_{i,q}(n)}{\partial T_i} g(q)dq = \frac{n+1}{2} \int_{\mathfrak{I}} q_i piv_i(q_\ell T_\ell, q_r T_r, n) g(q)dq. \tag{C-8}$$

We can view (C-4) as the limiting case of (C-8), for by Lemma 1, the right side of (C-8) converges to the right side of (C-4). The algebra that leads to (C-8) contains two critical observations. First, from the perspective of a planner whose policy preferences coincide with those of a type i agent, the rate of increase of the conditional per capita gross payoff with respect to the probability that a randomly chosen agent votes for candidate i equals, assuming $u = 1$, the sum of individual conditional pivot probabilities (equation (C-6)), which is the same as the (additional) conditional gross benefit that an act utilitarian (i.e., a purely altruistic) agent of type i gets by voting. This establishes the conceptual link between RU-voter models and our model. Second, the rate of increase of the collective conditional voting probability, $q_i T_i$, with respect to T_i equals q_i. Combined with the first observation, this implies that for a given q, the rate of increase of $\mathbb{P}_{i,q}(n)$ with respect to T_i is equal to q_i multiplied by the aggregate conditional type i pivot probability (equation (C-7)); and, this leads to (C-4).

Remark C1. In the RU-voter model under consideration, the comparative statics of turnout rates are robust with respect to the use of posterior or prior beliefs. Assuming q and C are uniformly distributed, our analysis of closed form solutions (which we omit here) has shown that the RU-voter model based on posterior beliefs leads to slightly closer election outcomes and slightly higher turnout rates. The intuition is that since along the critical ray $q_\ell > q_r$, updating beliefs creates an extra margin between the benefits that the two groups get by marginally increasing their maximum vote shares. This, in turn, pushes T_r^+/T_ℓ^+ downward.

i. *Voter Heterogeneity*

Assuming, for simplicity, a partitional group structure and ignoring the selfish agents, the consistency conditions (C-1) amount to saying that any two types in the same group must agree on the optimal/ethical rule that they should follow. This restricts the level of heterogeneity a group can involve. That is why, for instance, agents with different policy preferences must be classified in different groups (Feddersen and Sandroni, 2006, Proposition 1).

Of course, one can think of many other forms of heterogeneity that may severely restrict the size of admissible groups. For example, a significant uncertainty in u would make it impossible to form large groups without violating the consistency condition, since (as is clear from (C-2)) any two agents who disagree on the per capita payoff of changing the winner would also disagree on the optimal rule (that should be followed by a hypothetical group that contains these agents). Similarly, since winning probabilities, $\mathbb{P}_\ell$ and $\mathbb{P}_r$, induced by a rule accepted by a group depend on the distributions of C and $\mathcal{D}$, any two agents who disagree on these distributions would also disagree on the rule which should be followed. Since agents' private information about their costs and sense of ethic duty might affect their conjectures about those of others (say, because of a correlation across agents), by considering possible values of C or $\mathcal{D}$, in principle, we may obtain a continuum of different types who may have to be classified in different groups.

In fact, it can be shown that if, as in Appendix A, we assume that u is a continuous random variable and redefine Θ accordingly, the consistency conditions (C-1) and the optimality conditions (C-2) would imply that each group Θ_τ is of measure 0. Because of obvious reasons, a finite model with $n + 1$ agents is more suitable to analyze such extreme cases. In a finite model where the distribution of u is as described in Appendix A, consistency would require that a rule utilitarian agent of a given type must consider herself to belong to a group of measure $1/(n+1)$, since with probability 1 all other agents would have different judgements about the importance of the election. This takes us back to the case of an act utilitarian agent whose behavior can be analyzed as in Appendix A.

D. *On Pivot Probabilities in the Binomial Model*

As Chamberlain and Rothschild (1981) demonstrate assuming $P_\ell + P_r = 1$, the pivot probabilities in the binomial model fall very rapidly around the 45 degree line. While in the general case finding the exact orders of magnitudes is not a trivial task (that we do not pursue here), using known facts about the speed of convergence in the law of large numbers (see Hoeffding inequality in Appendix E), it can easily be shown that when $P_\ell \neq P_r$ the following exponential bounds on pivot probabilities apply:

Lemma D1. *Set* $\delta \equiv |P_\ell - P_r|$ *and take any* $n \in \mathbb{N}$ *with* $n\delta \geq 1$. *If* $P_i \geq P_j$ *for some* $i \in \{\ell, r\}$, *then*

$$piv_i\left(P_\ell, P_r, n\right) \leq e^{-n\delta^2/2} \quad \text{and} \quad piv_j\left(P_\ell, P_r, n\right) \leq e^{-(n\delta-1)^2/2n}. \tag{D-1}$$ [58]

Another simple consequence of Hoeffding inequality is the following lemma that provides upper and lower bounds on P_i as a function of the observed value of $\frac{S_i}{n+1}$ and the probability of type 1 error that one is willing to allow, α. (For the proof, see Appendix E.)

Lemma D2. *Let* $i \in \{\ell, r\}$, $\rho_i \in [0,1]$ *and* $\alpha \in (0,1]$. *Then, for any* $n \in \mathbb{N}$:

(i) $P_i < \rho_i - \sqrt{\frac{-\ln\alpha}{2(n+1)}}$ *implies* $\Pr\{\frac{S_i}{n+1} \geq \rho_i\} < \alpha$. *Thus, if the observed value of* $\frac{S_i}{n+1}$ *is greater than or equal to* ρ_i, *with a confidence level of (at least)* $1-\alpha$ *we have* $P_i \geq \rho_i - \sqrt{\frac{-\ln\alpha}{2(n+1)}}$.

(ii) $P_i > \rho_i + \sqrt{\frac{-\ln\alpha}{2(n+1)}}$ *implies* $\Pr\{\frac{S_i}{n+1} \leq \rho_i\} < \alpha$. *Thus, if the observed value of* $\frac{S_i}{n+1}$ *is less than or equal to* ρ_i, *with a confidence level of (at least)* $1-\alpha$ *we have* $P_i \leq \rho_i + \sqrt{\frac{-\ln\alpha}{2(n+1)}}$.

In light of Lemmas D1 and D2, we now take a look at the 2004 US presidential elections (Table D1 below). If we denote by S_ℓ and S_r the number of democrat and republican votes in a given state, respectively, we see that among the 15 states with highest voting age population (henceforth, VAP) in all but Ohio, $|S_\ell - S_r|$ was at least 1.5 percent of VAP.[59] Thus, if we estimate the number of eligible agents, $n+1$, with VAP, we see that except Ohio in all these states we had $\frac{|S_\ell - S_r|}{n+1} \geq 0.015$. For any such state and any $\alpha \in (0,1]$, it easily follows from Lemma D2 that $\delta \equiv |P_\ell - P_r| \geq 0.015 - 2\sqrt{\frac{-\ln\alpha}{2(n+1)}}$ with $1-\alpha$

[58] For our purposes, Lemma D1 is sufficient as is. However, in private communications with Srinivasa Varadhan and Atilla Yılmaz, I learned a method that leads to the following improvement of the first inequality in (D-1): $piv_i\left(P_\ell, P_r, n\right) \leq \left(1 - \left(\sqrt{P_\ell} - \sqrt{P_r}\right)^2\right)^n \leq \left(1-\delta^2\right)^{n/2}$. (The proof is available upon request.) However, obtaining a similar, elegant improvement upon the second inequality in (D-1) does not seem to be easy to the present author. It is also worth noting that the second bound in (D-1) is slightly larger than the first one, in fact, $e^{-(n\delta-1)^2/2n} \simeq e^{\delta}e^{-n\delta^2/2}$.

[59] VAP is the number of US residents above 18 years old.

confidence.[60] Since in all states under focus we have $n \geq 10^6$, even the extremely small choice $\alpha \equiv 10^{-6}$ implies therefore that $\delta \geq 0.0097$.[61] Using these bounds for δ and n, Lemma D1 then shows that pivot probabilities are less than $e^{-47}/2$. Similarly, in 2000 (resp. 1996) the same conclusion is valid for all states in Table D1 except Florida (resp. Georgia and Virginia).[62]

Table D1—US Presidential Elections

Top 15 states	VAP (in 1000's)			Plurality (in 1000's)			Plurality/VAP (in %)		
by 2004 VAP	*1996*	*2000*	*2004*	*1996*	*2000*	*2004*	*1996*	*2000*	*2004*
California	22,871	24,749	26,085	1,291	1,294	1,236	5.65	5.23	4.74
Texas	13,431	14,533	15,813	276	1,366	1,694	2.06	9.40	10.71
New York	13,408	13,725	14,492	1,823	1,704	1,352	13.59	12.42	9.33
Florida	10,886	11,633	13,133	302	1	381	2.78	0.00*	2.90
Pennsylvania	8,996	8,950	9,356	415	205	144	4.61	2.29	1.54
Illinois	8,598	8,859	9,303	755	570	546	8.78	6.43	5.86
Ohio	8,192	8,301	8,469	288	167	119	3.52	2.01	1.40*
Michigan	7,018	7,231	7,452	508	217	165	7.24	3.00	2.22
New Jersey	5,944	6,109	6,413	549	505	241	9.24	8.26	3.76
Georgia	5,303	5,775	6,338	27	303	548	0.51*	5.26	8.65
North Carolina	5,364	5,629	6,250	118	373	435	2.20	6.63	6.97
Virginia	4,955	5,177	5,364	47	220	262	0.95*	4.25	4.89
Massachusetts	4,560	4,614	4,840	854	738	733	18.72	15.99	15.14
Washington	4,059	4,314	4,596	283	139	205	6.96	3.22	4.47
Indiana	4,238	4,380	4,536	119	344	510	2.81	7.85	11.25

Notes: VAP is defined in Footnote 59. Plurality is the absolute difference between the numbers of democrat and republican votes. Because of rounding errors, 8th column may differ from 5th column/2nd column up to 0.01%, and similarly for 9th and 10th columns.

Sources: VAP—U.S. Bureau of the Census. Plurality—CQ Electronic Library, Voting and Elections Collection.

(*) Less than 1.5%.

E. *Proofs*

Proof of Proposition 1. Set $U \equiv u\left(1+\Psi(n)\right)/2$ and $\Omega \equiv \mathfrak{I}\times[0,U]^2$. For every $(q,\mathbf{C}) \equiv (q_\ell, q_r, C_\ell, C_r) \in \Omega$ and $i \in \{\ell, r\}$, define $P_i(q,\mathbf{C}) \equiv \lambda_i q_i F(C_i)$. Since F is continuous on $\mathbb{R}_+$, the function $(q,\mathbf{C}) \to (P_\ell(q,\mathbf{C}), P_r(q,\mathbf{C}))$ is continuous on Ω. Therefore, if we replace $P_{\ell,q}$ and $P_{r,q}$ in (1) with $P_\ell(q,\mathbf{C})$ and $P_r(q,\mathbf{C})$, respectively, the obtained function $(q,\mathbf{C}) \to piv_i(P_\ell(q,\mathbf{C}), P_r(q,\mathbf{C}), n)$ is also continuous on Ω. Since Ω is compact, this latter function must in fact be uniformly continuous. This, in turn, implies that

[60] This inequality is derived as follows: Assume $S_i \geq S_j$; set $\rho_\ell \equiv S_\ell/(n+1)$, $\rho_r \equiv S_r/(n+1)$; using Lemma D2(i) (resp. Lemma D2(ii)) conclude: $P_i \geq \rho_i - \sqrt{\frac{-\ln\alpha}{2(n+1)}}$ (resp. $P_j \leq \rho_j + \sqrt{\frac{-\ln\alpha}{2(n+1)}}$) with $1-\alpha$ confidence, so that $P_i - P_j \geq \rho_i - \rho_j - 2\sqrt{\frac{-\ln\alpha}{2(n+1)}}$.

[61] Though much of the literature on voter turnout uses VAP, because of the high number of ineligible US residents, VAP tends to overestimate the number of eligible agents (Michael P. McDonald and Samuel L. Popkin, 2001). To be on the safe side, we focus here on the largest 15 states so that 1 million is a clear lower bound for the number of eligible agents. (The use of VAP when computing $\frac{|S_\ell - S_r|}{n+1}$ is readily a conservative choice.)

[62] In the elections under focus, minor party candidates were not strong contenders. Hence, in our computations we implicitly consider votes for minor party candidates as abstentions.

for any i, the map $\mathbf{C} \to \Pi_i(\mathbf{C}) \equiv U \int_{\mathfrak{I}} piv_i(P_\ell(q, \mathbf{C}), P_r(q, \mathbf{C}), n) dG^i(q)$ is continuous on $[0, U]^2$. Moreover, clearly, $(\Pi_\ell(\cdot), \Pi_r(\cdot))$ is a self map on $[0, U]^2$. Hence, by Brouwer fixed point theorem this map has a fixed point which proves the existence of an equilibrium.

To establish positivity of cutoff points, suppose by contradiction that in an equilibrium we have $C_i^* = 0$ for a type i. Then, type i agents abstain with probability 1. It follows that $PIV_i \geq (\lambda_i)^n > 0$, for whenever all agents (excluding a given agent) are of type i the election is tied, which happens with probability $(\lambda_i)^n$. Thus, we must have $\Pi_i^* > 0$, a contradiction. □

The next theorem provides exponential bounds on the tail probabilities of the sum of a sequence of bounded, independent random variables. Its first part is a straightforward modification of the statement of Theorem 2 of Wassily Hoeffding (1963). The second part follows from the first part in an obvious way.

Hoeffding inequality. *Let $Z_1, ..., Z_n$ be independent random variables such that, for every $h = 1, ..., n$, we have $b_h \leq Z_h \leq d_h$ for a pair of real numbers b_h, d_h. Put $\mathbb{S} \equiv \sum_{h=1}^n Z_h$. Then:*

(i) *For any $\xi \geq 0$, $\Pr\{\mathbb{S} - \mathbf{E}\mathbb{S} \geq \xi\} \leq e^{-2\xi^2 / \sum_{h=1}^n (d_h - b_h)^2}$.*

(ii) *For any $\xi \leq 0$, $\Pr\{\mathbb{S} - \mathbf{E}\mathbb{S} \leq \xi\} \leq e^{-2\xi^2 / \sum_{h=1}^n (d_h - b_h)^2}$.*

The following observation provides bounds for binomial tail probabilities. Its proof is a routine application of Hoeffding inequality, and therefore, omitted.

Corollary E1. *Let $Z_1, ..., Z_n$ be independent Bernoulli random variables each with success probability ϱ, and set $\mathbb{S} \equiv \sum_{h=1}^n Z_h$. Then:*

(i) *For any number $\beta \geq \varrho$, $\Pr\{\mathbb{S} \geq \beta n\} \leq e^{-2(\beta - \varrho)^2 n}$.*

(ii) *For any number $\beta \leq \varrho$, $\Pr\{\mathbb{S} \leq \beta n\} \leq e^{-2(\beta - \varrho)^2 n}$.*

We now prove Lemmas D1 and D2.

Proof of Lemma D1. Set $S^- \equiv \sum_{h=1}^n X_h$. Notice that $piv_\ell(P_\ell, P_r, n) = \Pr\{S^- \in \{0, 1\}\}$ and $piv_r(P_\ell, P_r, n) = \Pr\{S^- \in \{0, -1\}\}$. Assume first $P_r \geq P_\ell$. Then, we have

$$\begin{aligned} piv_r(P_\ell, P_r, n) &\leq \Pr\left\{S^- \leq 0\right\} = \Pr\left\{S^- - n\delta \leq -n\delta\right\} \leq e^{-2(n\delta)^2/4n} = e^{-n\delta^2/2}, \\ piv_\ell(P_\ell, P_r, n) &\leq \Pr\left\{S^- \leq 1\right\} = \Pr\left\{S^- - n\delta \leq 1 - n\delta\right\} \leq e^{-2(n\delta - 1)^2/4n} = e^{-(n\delta - 1)^2/2n}. \end{aligned}$$

Here, since $\mathbf{E}S^- = n\delta$ and $1 - n\delta \leq 0$, the last inequalities in both lines follow from part (ii) of Hoeffding inequality with $b_h \equiv -1$ and $d_h \equiv 1$ $(h = 1, ..., n)$. The case $P_r \leq P_\ell$ can be handled similarly using part (i) of Hoeffding inequality and the fact that $\mathbf{E}S^- = -n\delta$. □

Proof of Lemma D2. Assume $\rho_i - P_i > \sqrt{\frac{-\ln \alpha}{2(n+1)}}$. Then, Corollary E1(i) implies $\Pr\left\{S_i \geq \rho_i(n+1)\right\} \leq e^{-2(\rho_i - P_i)^2(n+1)}$. Since $2\left(\rho_i - P_i\right)^2(n+1) > -\ln \alpha$, we conclude: $\Pr\left\{S_i \geq \rho_i(n+1)\right\} < e^{\ln \alpha} = \alpha$. This proves part (i), and part (ii) is similarly proved using Corollary E1(ii). □

Proof of Proposition 2. For every $n \in \mathbb{N}$ and $i \in \{\ell, r\}$, set $P_{i,n}^* \equiv \lambda_i q_i F(C_{i,n}^*)$. First notice that for every n and i,

$$\Pr\left\{S_i^- - S_j^- = -1\right\} = \sum_{\substack{b=0,1,\ldots,n \\ n-b \text{ is odd}}} \frac{n!}{b! \frac{n-b-1}{2}! \frac{n-b+1}{2}!} \left(1 - P_{\ell,n}^* - P_{r,n}^*\right)^b \left(P_{\ell,n}^* P_{r,n}^*\right)^{\frac{n-b-1}{2}} P_{j,n}^*,$$

and therefore,

$$\Pr\left\{S_r^- - S_\ell^- = -1\right\} = \frac{P_{\ell,n}^*}{P_{r,n}^*} \Pr\left\{S_\ell^- - S_r^- = -1\right\}, \tag{E-1}$$

whenever $P_{r,n}^*$ is positive.

To prove part (i), assume $\lambda < 1/2$. We will now show that

$$\limsup_k \frac{P_{\ell,n_k}^*}{P_{r,n_k}^*} < 1, \tag{E-2}$$

for any subsequence P_{r,n_k}^* that is bounded away from 0. Assume by contradiction that $\lim_k \frac{P_{\ell,n_k}^*}{P_{r,n_k}^*} \geq 1$ for a subsequence P_{r,n_k}^* that is bounded away from 0. Then, since

$$\frac{piv_r\left(P_{\ell,n_k}^*, P_{r,n_k}^*, n_k\right)}{piv_\ell\left(P_{\ell,n_k}^*, P_{r,n_k}^*, n_k\right)} \equiv \frac{\Pr\left\{S_\ell^- - S_r^- = 0\right\} + \Pr\left\{S_r^- - S_\ell^- = -1\right\}}{\Pr\left\{S_\ell^- - S_r^- = 0\right\} + \Pr\left\{S_\ell^- - S_r^- = -1\right\}},$$

by (E-1) it follows that

$$\liminf_k \frac{C_{r,n_k}^*}{C_{\ell,n_k}^*} = \liminf_k \frac{piv_r\left(P_{\ell,n_k}^*, P_{r,n_k}^*, n_k\right)}{piv_\ell\left(P_{\ell,n_k}^*, P_{r,n_k}^*, n_k\right)} \geq 1, \tag{E-3}$$

where the first equality is a consequence of definitions of C_{ℓ,n_k}^* and C_{r,n_k}^*.

By passing to a further subsequence of n_k if necessary, assume C_{ℓ,n_k}^* and C_{r,n_k}^* converge, possibly to ∞, and let the corresponding limits be $C_\ell^\bullet$ and $C_r^\bullet$, respectively. Then, (E-3) implies $C_r^\bullet \geq C_\ell^\bullet$. Since F is continuous, it follows that $\lim_k F(C_{r,n_k}^*) = F(C_r^\bullet) \geq F(C_\ell^\bullet) = \lim_k F(C_{\ell,n_k}^*)$, where we set $F(\infty) \equiv 1$. But then $\lambda < 1/2$ and $q_\ell = q_r$ imply $\lim_k P_{r,n_k}^* = (1-\lambda) q_r \lim_k F(C_{r,n_k}^*) > \lambda q_\ell \lim_k F(C_{\ell,n_k}^*) = \lim_k P_{\ell,n_k}^*$, for P_{r,n_k}^* is bounded away from 0 so that $\lim_k F(C_{r,n_k}^*) > 0$. This contradicts the supposition that $\lim_k \frac{P_{\ell,n_k}^*}{P_{r,n_k}^*} \geq 1$ and proves (E-2).

To complete the proof of (i), notice that if for a type i there is a subsequence C_{i,n_k}^* that is bounded away from 0, then P_{i,n_k}^* is also bounded away from 0. But by (H2) and Lemma D1 both of these sequences can be bounded away from 0 only if $\lim_k \frac{P_{\ell,n_k}^*}{P_{r,n_k}^*} = 1$, which contradicts (E-2).

To prove (ii), we will first show that $P_{\ell,n}^* = P_{r,n}^*$ for every $n \in \mathbb{N}$. Suppose by contradiction $P_{i,n}^* > P_{j,n}^*$ for some i and n. Then, $\Pr\{S_i^- - S_j^- = -1\} \leq \Pr\{S_j^- - S_i^- = -1\}$, and hence, $piv_i(P_{\ell,n}^*, P_{r,n}^*, n) \leq piv_j(P_{\ell,n}^*, P_{r,n}^*, n)$ so that $C_{i,n}^* \leq C_{j,n}^*$. Since $q_\ell = q_r$ and $\lambda = \frac{1}{2}$, this implies $P_{i,n}^* \leq P_{j,n}^*$, a contradiction.

Let then $P_n^* \equiv P_{\ell,n}^* = P_{r,n}^*$ for every n. Börgers (2004, Remark 1) shows that the function $piv(P,n) \equiv piv_\ell(P,P,n) = piv_r(P,P,n)$ is decreasing in $P \in [0, 1/2]$, for every n. Thus, $C_{i,n}^* \geq u\,(1+\psi n)\,\frac{1}{2} piv\left(\frac{1}{2}, n\right)$ for every n and i. Finally, since $piv\left(\frac{1}{2}, n\right) \simeq 1/\sqrt{\pi n \frac{1}{2}}$ (see, e.g., Chamberlain and Rothschild, 1981), we conclude that $\lim_n C_{i,n}^* = \infty$, as we seek. □

Remark E1. From the proof of Proposition 2(i) it is clear that when $\lambda < 1/2$, for an arbitrary continuous distribution function F on $\mathbb{R}$, we can still conclude that $\lim_n F(C_{i,n}^*) = F(0)$ for any type i, since otherwise we can find a subsequence n_k such that $F(C_{i,n_k}^*)$ and C_{i,n_k}^* are both bounded away from 0. It may also be useful to note once again that we may in fact have $\lim_n \frac{P_{\ell,n}^*}{P_{r,n}^*} = 1$ (Karasa and Polborn, 2009; Taylor and Yildirim, 2008), for we derived (E-2) under the hypothesis that P_{r,n_k}^* is bounded away from 0, which is incorrect unless $F(0) > 0$.

Proof of Lemma 1. Set $\Upsilon_n \equiv n \int_0^1 \int_0^1 piv_i(q_\ell T_\ell, q_r T_r, n) \nu(q_\ell, q_r) dq_\ell dq_r$ for every $n \in \mathbb{N}$ and some

$i \in \{\ell, r\}$. To evaluate Υ_n, consider the substitution $(q_\ell, q_r) = W(t,P) \equiv (\frac{tP}{T_\ell}, \frac{t(1-P)}{T_r})$. It is a routine task to verify that W is a bijection from the set

$$V \equiv \{(t,P) : 0 < t < T_\ell + T_r, \max\{0, 1 - T_r/t\} < P < \min\{1, T_\ell/t\}\}$$

onto $(0,1)^2$. (The inverse of W is defined by $W^{-1}(q_\ell, q_r) \equiv (T_\ell q_\ell + T_r q_r, \frac{T_\ell q_\ell}{T_\ell q_\ell + T_r q_r}) = (t,P)$.) Moreover, W is continuously differentiable, and $J \equiv \begin{bmatrix} P/T_\ell & t/T_\ell \\ (1-P)/T_r & -t/T_r \end{bmatrix}$ is its Jacobian matrix. Since $|\det J| = \frac{t}{T_\ell T_r}$, from the change of variables formula it follows that for every $n \in \mathbb{N}$,

$$\Upsilon_n = \int_0^{T_\ell + T_r} \Upsilon_{t,n} dt$$

where, for every $t \in (0, T_\ell + T_r)$,

$$\Upsilon_{t,n} \equiv n \frac{t}{T_\ell T_r} \int_{I_t} piv_i\left(tP, t(1-P), n\right) \nu\left(\frac{tP}{T_\ell}, \frac{t(1-P)}{T_r}\right) dP, \tag{E-4}$$

and $I_t \equiv \left(\max\left\{0, 1 - \frac{T_r}{t}\right\}, \min\left\{1, \frac{T_\ell}{t}\right\}\right)$ (see Patrick Billingsley, 1995, Theorem 17.2, p. 225). It may also be useful to note that if $0 < t < T_\ell + T_r$, we have $\frac{T_\ell}{t} > 1 - \frac{T_r}{t}$ so that the interval I_t is nondegenerate.

Pick any $\beta > 0$. First we will show that $\int_0^{n^{-\beta}} \Upsilon_{t,n} dt$ converges to 0 as $n \to \infty$ (uniformly on a set $\mathfrak{T}$ of the given form).[63] Denote by m a possible value of the turnout $S_\ell^- + S_r^-$ excluding a given agent. Let $\mathcal{B}(\cdot; n, t)$ be the binomial probability distribution with population size n and success probability t. Notice that for a fixed $(t,P) \in V$, if we consider tP (resp. $t(1-P)$) as the conditional probability that a type ℓ (resp. r) agent votes, since $t = tP + t(1-P)$, we have $\Pr\{S_\ell^- + S_r^- = m \mid (t,P)\} = \mathcal{B}(m; n, t)$ for every nonnegative integer m and positive integer n. Moreover, in this case, among those who participate, a randomly chosen agent votes for candidate ℓ with probability $P = tP/t$, which implies that conditional on the event $S_\ell^- + S_r^- = m$ (and conditional on (t,P)) the probability of the event that the election is tied or candidate i is 1 behind equals $piv_i(P, 1-P, m)$. Hence, for every $n \in \mathbb{N}$ and $(t,P) \in V$, we have

$$piv_i(tP, t(1-P), n) = \sum_{m=0}^{n} \mathcal{B}(m; n, t) piv_i(P, 1-P, m). \tag{E-5}$$

Thus, $\int_0^{n^{-\beta}} \Upsilon_{t,n} dt \le n \frac{n^{-\beta}}{T_\ell T_r} \overline{\nu} \int_0^1 \sum_{m=0}^n \mathcal{B}(m;n,t) \left(\int_0^1 piv_i(P, 1-P, m) dP\right) dt$, where $\overline{\nu}$ is an upper bound for ν. Note that $\int_0^1 piv_\ell(P, 1-P, m) dP \equiv \int_0^1 \binom{m}{\lfloor m/2 \rfloor} P^{\lfloor m/2 \rfloor} (1-P)^{m - \lfloor m/2 \rfloor} dP \equiv \int_0^1 \mathcal{B}(\lfloor m/2 \rfloor; m, P) dP = \frac{1}{m+1}$ (see Chamberlain and Rothschild, 1981). Similarly, $\int_0^1 piv_r(P, 1-P, m) dP = \frac{1}{m+1}$ and $\int_0^1 \mathcal{B}(m; n, t) dt = \frac{1}{n+1}$. Hence,

$$\int_0^{n^{-\beta}} \Upsilon_{t,n} dt \le n \frac{n^{-\beta}}{T_\ell T_r} \overline{\nu} \sum_{m=0}^{n} \frac{1}{m+1} \int_0^1 \mathcal{B}(m; n, t) dt = n \frac{n^{-\beta}}{T_\ell T_r} \overline{\nu} \sum_{m=0}^{n} \frac{1}{m+1} \frac{1}{n+1} \le \frac{n^{-\beta}}{T_\ell T_r} \overline{\nu} \sum_{m=0}^{n} \frac{1}{m+1}.$$

Since the Harmonic series diverges at logarithmic rate, $n^{-\beta} \sum_{m=0}^n \frac{1}{m+1}$ tends to 0. It thus follows that, for any fixed $\beta > 0$,

$$\int_0^{n^{-\beta}} \Upsilon_{t,n} dt = \Upsilon_n - \int_{n^{-\beta}}^{T_\ell + T_r} \Upsilon_{t,n} dt \to 0$$

(uniformly on $\mathfrak{T}$ where $\frac{1}{T_\ell T_r}$ is bounded from above).

Fix $\varepsilon' > 0$. Since ν is continuous on the compact set $[0,1]^2$, it must be uniformly continuous. It thus

[63] Throughout the proof, over a region of integration where the integrand is not explicitly defined, it is assumed to be zero.

follows that there is a positive number $\varepsilon < 1/2$ such that, for all $(t,P) \in V$ with $|P - 1/2| \le \varepsilon$,

$$\left|\nu\left(\frac{tP}{T_\ell}, \frac{t(1-P)}{T_r}\right) - \nu^t\right| \le \varepsilon', \quad \text{where } \nu^t \equiv \nu\left(\frac{t}{2T_\ell}, \frac{t}{2T_r}\right). \tag{E-6}$$

(Notice that $\left|\frac{tP}{T_\ell} - \frac{t}{2T_\ell}\right| = \left|P - \frac{1}{2}\right| \frac{t}{T_\ell}$ and t/T_ℓ is bounded from above on $\mathfrak{T}$; and similarly for $\left|\frac{t(1-P)}{T_r} - \frac{t}{2T_r}\right|$. Thus, such a number ε can be chosen uniformly on $\mathfrak{T}$.)

Now fix a positive number $\beta < 1/2$ and consider any n such that $2\varepsilon n^{1-\beta} \ge 1$. When $t \ge n^{-\beta}$ and $|P - 1/2| > \varepsilon$, i.e., $|P - (1-P)| > 2\varepsilon$, we then have $n\delta_{t,P} \equiv nt\,|P - (1-P)| \ge 2\varepsilon n^{1-\beta} \ge 1$; and Lemma D1 implies $piv_i(tP, t(1-P), n) \le e^{-(n\delta_{t,P}-1)^2/2n} \le e^{-\left(2\varepsilon n^{1-\beta}-1\right)^2/2n}$; i.e., the integrand in (E-4) is less than $\overline{\nu} e^{-\left(2\varepsilon n^{1-\beta}-1\right)^2/2n}$. Since $\beta < 1/2$, it is easily verified that $ne^{-\left(2\varepsilon n^{1-\beta}-1\right)^2/2n} \to 0$. Thus, it follows that $\int_{n^{-\beta}}^{T_\ell+T_r} \Upsilon_{t,n} dt - \int_{n^{-\beta}}^{T_\ell+T_r} \Phi_{t,n} dt$ tends to 0 (uniformly on $\mathfrak{T}$) where

$$\Phi_{t,n} \equiv n\frac{t}{T_\ell T_r} \int_{\Xi_t} piv_i(tP, t(1-P), n)\nu\left(\frac{tP}{T_\ell}, \frac{t(1-P)}{T_r}\right) dP$$

and $\Xi_t \equiv I_t \cap [1/2 - \varepsilon, 1/2 + \varepsilon]$, for $0 < t < T_\ell + T_r$ and $n \in \mathbb{N}$. In particular, we can ignore any t such that $t > \min\left\{\frac{2T_\ell}{1-2\varepsilon}, \frac{2T_r}{1-2\varepsilon}\right\}$ which implies $T_\ell/t < 1/2 - \varepsilon$ or $1 - T_r/t > 1/2 + \varepsilon$ so that $\Xi_t = \varnothing$. We therefore conclude that

$$\Upsilon_n - \int_{n^{-\beta}}^{t_\varepsilon} \Phi_{t,n} dt \to 0$$

(uniformly on $\mathfrak{T}$) where $t_\varepsilon \equiv \min\left\{\frac{2T_\ell}{1-2\varepsilon}, \frac{2T_r}{1-2\varepsilon}, T_\ell + T_r\right\}$.

Now notice that, by (E-5),

$$\int_{n^{-\beta}}^{t_\varepsilon} \Phi_{t,n} dt = \int_{n^{-\beta}}^{t_\varepsilon} \frac{t}{T_\ell T_r} n \sum_{m=0}^{n} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i\left(P, 1-P, m\right) \nu\left(\frac{tP}{T_\ell}, \frac{t(1-P)}{T_r}\right) dP\right) dt. \tag{E-7}$$

Moreover, by Corollary E1, whenever $t \ge n^{-\beta}$, we have $n\sum_{m>t(1+\varepsilon)n} \mathcal{B}(m;n,t) \le ne^{-2t^2\varepsilon^2 n} \le ne^{-2\varepsilon^2 n^{1-2\beta}}$. Since $\beta < 1/2$, clearly, $ne^{-2\varepsilon^2 n^{1-2\beta}} \to 0$. It thus follows that $n\sum_{m>t(1+\varepsilon)n} \mathcal{B}(m;n,t) \to 0$ uniformly on $t \ge n^{-\theta}$; and similarly for $n\sum_{m<t(1-\varepsilon)n} \mathcal{B}(m;n,t)$. Since in (E-7) the integral inside the parenthesis and $\frac{t}{T_\ell T_r}$ are bounded from above (for relevant values of t and $(T_\ell, T_r) \in \mathfrak{T}$), we can therefore focus on nonnegative integers m such that $t(1-\varepsilon)n \le m \le t(1+\varepsilon)n$. Combining this observation with (E-6) it follows that, for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$):

$$-\varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t - \varepsilon'\right) \frac{t}{T_\ell T_r} n \sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i(P, 1-P, m) dP\right) dt \tag{E-8}$$

$$\le \Upsilon_n$$

$$\le \varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) \frac{t}{T_\ell T_r} n \sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i(P, 1-P, m) dP\right) dt. \tag{E-9}$$

Since in (E-9) for each $m \ge t(1-\varepsilon)n$ the integral in parenthesis is at most $(m+1)^{-1} \le (t(1-\varepsilon)n)^{-1}$, we see that for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$):

$$\Upsilon_n \le \varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) \frac{t}{T_\ell T_r} n \left(\frac{1}{t(1-\varepsilon)n}\right) dt = \varepsilon + \frac{1}{T_\ell T_r (1-\varepsilon)} \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) dt. \tag{E-10}$$

Next notice that for $t < t'_\varepsilon \equiv \min\left\{\frac{2T_\ell}{1+2\varepsilon}, \frac{2T_r}{1+2\varepsilon}\right\}$ we have $T_\ell/t > 1/2+\varepsilon$ and $1 - T_r/t < 1/2-\varepsilon$, and thus, $\Xi_t = [1/2-\varepsilon, 1/2+\varepsilon]$. Since $\int_0^1 piv_i(P, 1-P, m)dP = \frac{1}{m+1}$, it clearly follows that $\int_{\Xi_t} piv_i(P, 1-P, m)dP \geq m^{-1}(1-\varepsilon)$ for all sufficiently large m, say, for $m \geq \overline{m}$, and every $t < t'_\varepsilon$. Since $n^{-\beta}(1-\varepsilon)n$ eventually exceeds $\overline{m}$, and since $t'_\varepsilon < t_\varepsilon$, we conclude that the expression in (E-8) is at least $-\varepsilon + \int_{n^{-\beta}}^{t'_\varepsilon} (\nu^t - \varepsilon') \frac{t}{T_\ell T_r} n \left(\sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m; n, t)\right) \frac{1-\varepsilon}{t(1+\varepsilon)n} dt$ for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$). As noted before, here, the term inside the parenthesis converges to 1 uniformly for $t \geq n^{-\beta}$. It thus follows that, for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$):

$$\Upsilon_n \geq -\varepsilon + \int_{n^{-\beta}}^{t'_\varepsilon} (\nu^t - \varepsilon') \frac{(1-\varepsilon)^2}{T_\ell T_r (1+\varepsilon)} dt.$$

Since we can choose ε and ε' arbitrarily small, by the definitions of t_ε and t'_ε, this observation along with (E-10) imply that $\Upsilon_n \to \frac{1}{T_\ell T_r}\int_0^{2\min\{T_\ell, T_r\}} \nu^t dt$ (uniformly on $\mathfrak{T}$). Finally, the substitution $t = 2T_\ell T_r \theta$ gives $\frac{1}{T_\ell T_r}\int_0^{2\min\{T_\ell,T_r\}} \nu^t dt = 2\int_0^{\frac{1}{\max\{T_\ell,T_r\}}} \nu(T_r\theta, T_\ell\theta)d\theta$, as we seek.

Proving that (4) along with the fact that the convergence is uniform imply (5) is a routine exercise, and therefore, omitted. □

Conditional pivotal probabilities do not decrease monotonically with n. However, as the next lemma shows, in the special case where everyone votes, they can be approximated by a nonincreasing function provided that voting probabilities are close to 1/2. This will be useful when proving Proposition 3. (In the sequel, $\mathbb{Z}_+$ denotes the set of all nonnegative integers.)

Lemma E1. *For every $P \in [0,1]$, $m \in \mathbb{Z}_+$ and $i \in \{\ell, r\}$, define*

$$\mu_i(P, m) \equiv \begin{cases} piv_i(P, 1-P, m+1) & \text{if } m \text{ is odd,} \\ piv_i(P, 1-P, m) & \text{if } m \text{ is even.} \end{cases}$$

Then, for every $P \in [0,1]$ and $i \in \{\ell, r\}$, the function $\mu_i(P, \cdot)$ is nonincreasing on $\mathbb{Z}_+$. Moreover, for any $\varepsilon > 0$ and $P \in (0,1)$ such that $\left|\frac{1}{2P} - 1\right| \leq \varepsilon$ and $\left|\frac{1}{2(1-P)} - 1\right| \leq \varepsilon$, we have

$$(1-\varepsilon)\mu_i(P,m) \leq piv_i(P, 1-P, m) \leq (1+\varepsilon)\mu_i(P,m) \quad \text{for every } m \in \mathbb{Z}_+ \text{ and } i \in \{\ell, r\}.$$

Proof. To simplify our notation we will write $piv_i(m)$ and $\mu_i(m)$ instead of $piv_i(P, 1-P, m)$ and $\mu_i(P,m)$, respectively. First notice that, with $P_\ell \equiv P$ and $P_r \equiv 1-P$, if m is odd, we have $piv_i(m) = \frac{m!}{\frac{m-1}{2}!\frac{m+1}{2}!}(P_i)^{\frac{m-1}{2}}(P_j)^{\frac{m+1}{2}}$, and if m is even, we have $piv_i(m) = \frac{m!}{\frac{m}{2}!\frac{m}{2}!}(P_\ell)^{\frac{m}{2}}(P_r)^{\frac{m}{2}}$. Therefore, for every $i \in \{\ell, r\}$ and $m \in \mathbb{Z}_+$,

$$piv_i(m+1) = \begin{cases} 2P_i piv_i(m) & \text{if } m \text{ is odd,} \\ \frac{m+1}{m+2} 2P_j piv_i(m) & \text{if } m \text{ is even.} \end{cases}$$

Thus, when $P_i > 0$ and m is odd, we have $\frac{1}{2P_i}\mu_i(m) \equiv \frac{1}{2P_i} piv_i(m+1) = piv_i(m)$, so that $\left|\frac{1}{2P_i} - 1\right| \leq \varepsilon$ implies $(1-\varepsilon)\mu_i(m) \leq piv_i(m) \leq (1+\varepsilon)\mu_i(m)$. Since $\mu_i(m) \equiv piv_i(m)$ for every even m, the desired inequalities between $\mu_i(m)$ and $piv_i(m)$ are proved.

To show that $\mu_i(m)$ is nonincreasing, note that $piv_i(m+2) \leq 4P_\ell P_r piv_i(m) \leq piv_i(m)$ for every m.[64] It follows that $\mu_i(m) \geq \mu_i(m+1)$ for every even m, and $\mu_i(m) = \mu_i(m+1)$ for every odd m. □

[64] Taylor and Yildirim (2008, Lemma 1(iv)) show that even if $P_\ell + P_r < 1$ the conclusion $piv_i(m+2) \leq piv_i(m)$ is still true.

Proof of Proposition 3. Consider a subsequence of equilibria $(C^*_{\ell,n_k}, C^*_{r,n_k})$ that converges to an asymptotic equilibrium $(C^\bullet_\ell, C^\bullet_r)$. For every $k \in \mathbb{N}$, let $(T^*_{\ell,n_k}, T^*_{r,n_k})$ be the associated maximum vote shares. It suffices to show that for $i = \ell, r$, we have $C^\bullet_i > 0$, since then $T^\bullet_\ell$ and $T^\bullet_r$ are positive and equations (5) hold, and consequently, equations (7) also hold. (This implies, in particular, that $C^\bullet_i < \infty$ for $i = \ell, r$.)

To prove that both asymptotic cutoff points are positive, let us first assume that the ratio $T^*_{\ell,n_k}/T^*_{r,n_k}$ remains bounded away from 0 and ∞. Suppose by contradiction that $C^\bullet_\ell$ or $C^\bullet_r$ equals 0.

Fix a number $\varepsilon' \in (0,1)$ and choose an $\varepsilon \in (0, 1/2)$ such that for every $P \in [1/2-\varepsilon, 1/2+\varepsilon]$ we have $\left|\frac{1}{2P} - 1\right| \leq \varepsilon'$ and $\left|\frac{1}{2(1-P)} - 1\right| \leq \varepsilon'$. By Lemma E1, for every such P we have

$$(1-\varepsilon')\,\mu_i(P,m) \leq piv_i(P, 1-P, m) \quad \text{for every } m \in \mathbb{Z}_+ \text{ and } i \in \{\ell, r\}\,. \tag{E-11}$$

Fix an $i \in \{\ell, r\}$. For every $k \in \mathbb{N}$ and $0 < t < T^*_{\ell,n_k} + T^*_{r,n_k}$, as in the proof of Lemma 1, set $t'_{\varepsilon,k} \equiv \min\left\{\frac{2T^*_{\ell,n_k}}{1+2\varepsilon}, \frac{2T^*_{r,n_k}}{1+2\varepsilon}\right\}$, $\Xi_{t,k} \equiv \left(\max\left\{0, 1-\frac{T^*_{r,n_k}}{t}\right\}, \min\left\{1, \frac{T^*_{\ell,n_k}}{t}\right\}\right) \cap [1/2-\varepsilon, 1/2+\varepsilon]$ and

$$\Phi_{t,k} \equiv n_k \frac{t}{T^*_{\ell,n_k} T^*_{r,n_k}} \int_{\Xi_{t,k}} piv_i(tP, t(1-P), n_k) g^i\left(\frac{tP}{T^*_{\ell,n_k}}, \frac{t(1-P)}{T^*_{r,n_k}}\right) dP.$$

Notice that $t'_{\varepsilon,k} \to 0$ as $k \to \infty$ and that $t'_{\varepsilon,k} < T^*_{\ell,n_k} + T^*_{r,n_k}$ for every $k \in \mathbb{N}$. As we have seen in the proof of Lemma 1, $t \in (0, t'_{\varepsilon,k})$ implies $\Xi_{t,k} = [1/2-\varepsilon, 1/2+\varepsilon]$. Thus, from equation (E-5) and the definition of g^i it easily follows that, for every $k \in \mathbb{N}$ and $t \in (0, t'_{\varepsilon,k})$,

$$\Phi_{t,k} \geq \frac{g_0}{\overline{q}_i}\left(\frac{1}{2} - \varepsilon\right) \frac{n_k t^2}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \sum_{m=0}^{n_k} \mathcal{B}(m; n_k, t) \left(\int_{\Xi_{t,k}} piv_i(P, 1-P, m) dP\right), \tag{E-12}$$

where $g_0 > 0$ is a lower bound for g.

Clearly, there is an $m_0 \in \mathbb{N}$ such that for every integer $m \geq m_0$, we have $\int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(P,m)dP \geq m^{-1}(1-\varepsilon')$. Moreover, since μ_i is nonincreasing in m, for every nonnegative integer $m < m_0$, we have $\int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(P,m)dP \geq \int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(P,m_0)dP \geq m_0^{-1}(1-\varepsilon')$. Combining these observations with (E-11), we conclude that for every $k \in \mathbb{N}$, $m \in \mathbb{Z}_+$ and $t \in (0, t'_{\varepsilon,k})$,

$$\int_{\Xi_{t,k}} piv_i(P, 1-P, m)dP \geq \frac{(1-\varepsilon')^2}{\max\{m, m_0\}}.$$

By combining this with (E-12), we obtain: For every $k \in \mathbb{N}$ and $t \in (0, t'_{\varepsilon,k})$,

$$\Phi_{t,k} \geq \phi \frac{n_k t^2}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \sum_{m=0}^{n_k} \mathcal{B}(m; n_k, t) \frac{1}{\max\{m, m_0\}},$$

where $\phi \equiv \frac{g_0}{\overline{q}_i}\left(\frac{1}{2} - \varepsilon\right)(1-\varepsilon')^2 > 0$. Notice that since $t'_{\varepsilon,k} \to 0$, by Corollary E1(i), there is a sequence of numbers $b_k \to 1$ such that for every $k \in \mathbb{N}$ and every $t \in (0, t'_{\varepsilon,k})$ we have $\sum_{m \leq \varepsilon n_k} \mathcal{B}(m; n_k, t) \geq b_k$. Since the function $\frac{1}{\max\{\cdot, m_0\}}$ is nonincreasing on $\mathbb{Z}_+$, it follows that, for every $k \in \mathbb{N}$,

$$\int_0^{t'_{\varepsilon,k}} \Phi_{t,k} dt \geq \phi \frac{n_k}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \frac{b_k}{\max\{\varepsilon n_k, m_0\}} \int_0^{t'_{\varepsilon,k}} t^2 dt = \phi \frac{n_k}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \frac{b_k}{\max\{\varepsilon n_k, m_0\}} \frac{(t'_{\varepsilon,k})^3}{3}. \tag{E-13}$$

Since $T^*_{\ell,n_k}/T^*_{r,n_k}$ is bounded away from 0 and ∞, obviously so is $\frac{(t'_{\varepsilon,k})^3}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2}$. Hence, for large k, the right side of (E-13) is proportional to $\phi\varepsilon^{-1}$. Since we can choose ε and ε' arbitrarily small, we therefore conclude that, for any type i, $C^*_{i,k} \to \infty$ as $k \to \infty$, a contradiction.

What remains to show is that $T^*_{\ell,n_k}/T^*_{r,n_k}$ is bounded away from 0 and ∞. By suppressing the dependence on k, assume $T^*_r > T^*_\ell$. First, we note that, as in (E-1), for every $q \in \mathfrak{I}$, we have

$$\Pr\left\{S^-_r - S^-_\ell = -1 \mid q\right\} q_r T^*_r = \Pr\left\{S^-_\ell - S^-_r = -1 \mid q\right\} q_\ell T^*_\ell .$$

Thus:

$$\int_{\mathfrak{I}} \Pr\left\{S^-_r - S^-_\ell = -1 \mid q\right\} g^r(q)dq = \frac{\overline{q}_\ell T^*_\ell}{\overline{q}_r T^*_r} \int_{\mathfrak{I}} \Pr\left\{S^-_\ell - S^-_r = -1 \mid q\right\} g^\ell(q)dq \tag{E-14}$$

Moreover, the conditional probability of tie at $(q_\ell, q_r) \in \mathfrak{I}$ is

$$tie(q_\ell, q_r) \equiv \sum_{m=0}^{\left\lfloor \frac{n}{2} \right\rfloor} \frac{n!}{m!m!(n-2m)!}(1 - T^*_\ell q_\ell - T^*_r q_r)^{n-2m}(T^*_\ell q_\ell)^m (T^*_r q_r)^m .$$

Since $T^*_r > T^*_\ell$, whenever $q_r > q_\ell$ we have $T^*_\ell q_\ell + T^*_r q_r > T^*_\ell q_r + T^*_r q_\ell$, which obviously implies that $tie(q_\ell, q_r) < tie(q_r, q_\ell)$. Hence:

$$\int_{q_r > q_\ell} \frac{q_r}{\overline{q}_r} g(q_\ell, q_r) tie(q_\ell, q_r) d(q_\ell, q_r) < \int_{q_r > q_\ell} \frac{q_r}{\overline{q}_r} g(q_\ell, q_r) tie(q_r, q_\ell) d(q_\ell, q_r).$$

Clearly, here, the latter integral can be rewritten as $\int_{q_r < q_\ell} \frac{q_\ell}{\overline{q}_r} g(q_r, q_\ell) tie(q_\ell, q_r) d(q_\ell, q_r)$. We therefore see that

$$\int_{q_r > q_\ell} g^r(q) tie(q) dq < b \int_{q_r < q_\ell} g^\ell(q) tie(q) dq, \tag{E-15}$$

where $b > 0$ is the maximum value of $\frac{\overline{q}_\ell}{\overline{q}_r} \frac{g(q_r, q_\ell)}{g(q_\ell, q_r)}$ for $(q_\ell, q_r) \in \mathfrak{I}$. Moreover,

$$\int_{q_r < q_\ell} g^r(q) tie(q) dq = \int_{q_r < q_\ell} \frac{q_r}{\overline{q}_r} g(q) tie(q) dq < \frac{\overline{q}_\ell}{\overline{q}_r} \int_{q_r < q_\ell} \frac{q_\ell}{\overline{q}_\ell} g(q) tie(q) dq = \frac{\overline{q}_\ell}{\overline{q}_r} \int_{q_r < q_\ell} g^\ell(q) tie(q) dq. \tag{E-16}$$

Combining (E-14)-(E-16), we see that for $T^*_{r,n_k}/T^*_{\ell,n_k} > 1$, the ratio $C^*_{r,n_k}/C^*_{\ell,n_k}$ is bounded from above. In particular, $T^*_{r,n_k}/T^*_{\ell,n_k}$ can be arbitrarily large only if both C^*_{ℓ,n_k} and C^*_{r,n_k} are arbitrarily close to 0. Since $f(0) > 0$, this implies, by L'Hôpital's rule, that $F(C^*_{r,n_k})/F(C^*_{\ell,n_k})$ is approximately equal to $C^*_{r,n_k}/C^*_{\ell,n_k}$, a contradiction. Similarly, $T^*_{\ell,n_k}/T^*_{r,n_k}$ is also bounded away from ∞. □

Proofs of Propositions 4-8 and C1 can be found in Section III and Appendix C, respectively. We proceed to:

Proofs of Propositions 9 and 10. The arguments in text are complete except for the proof of the fact that MV is an increasing function of $T^\bullet_r/T^\bullet_\ell \geq 1$. To prove this point, we first note that

$$MV = \int_{\frac{q_\ell}{q_r} \leq \frac{T^\bullet_r}{T^\bullet_\ell}} \frac{T^\bullet_r q_r - T^\bullet_\ell q_\ell}{T^\bullet_r q_r + T^\bullet_\ell q_\ell} g(q) dq + \int_{\frac{q_r}{q_\ell} \leq \frac{T^\bullet_\ell}{T^\bullet_r}} \frac{T^\bullet_\ell q_\ell - T^\bullet_r q_r}{T^\bullet_r q_r + T^\bullet_\ell q_\ell} g(q) dq$$

If we express both integrands in terms of $x \equiv \frac{T^\bullet_r}{T^\bullet_\ell}$ and $\frac{q_\ell}{q_r}$, and then change variables in an obvious way, we find that

$$MV = \int_0^x \left(\frac{1}{1 + x^{-1} w} - \frac{1}{x w^{-1} + 1} \right) g_{q_\ell/q_r}(w) dw + \int_0^{x^{-1}} \left(\frac{1}{x w + 1} - \frac{1}{1 + x^{-1} w^{-1}} \right) g_{q_r/q_\ell}(w) dw.$$

By symmetry of g, we have $\mathfrak{g} \equiv g_{q_r/q_\ell} = g_{q_\ell/q_r}$. Moreover, in the above line, at $w = x$ the first integrand equals 0 and at $w = x^{-1}$ the second integrand equals 0. Thus, a marginal change in x does not affect MV through the integration bounds. From Leibniz rule it therefore follows that

$$\frac{dMV}{dx} = \int_0^x \left(\frac{x^{-2}w}{(1+x^{-1}w)^2} + \frac{w^{-1}}{(xw^{-1}+1)^2} \right) \mathfrak{g}(w)dw - \int_0^{x^{-1}} \left(\frac{w}{(xw+1)^2} + \frac{x^{-2}w^{-1}}{(1+x^{-1}w^{-1})^2} \right) \mathfrak{g}(w)dw.$$

In the second integral, if we substitute $w = x^{-2}\widetilde{w}$ and then write w instead of $\widetilde{w}$, we see that

$$\frac{dMV}{dx} = \int_0^x \left(\frac{x^{-2}w}{(1+x^{-1}w)^2} + \frac{w^{-1}}{(xw^{-1}+1)^2} \right) \left(\mathfrak{g}(w) - \mathfrak{g}(x^{-2}w)x^{-2} \right) dw.$$

At $x = 1$ this expression equals 0. Assume now $x > 1$. It suffices to show that, for every $w \in (0, x)$,

$$\mathfrak{g}(w) > \mathfrak{g}(x^{-2}w)x^{-2}. \tag{E-17}$$

First fix a $w \in (0, 1]$. We recall that $\mathfrak{g}(w) = \int_0^1 g(y, yw)ydy$. The substitution $y = \frac{\vartheta}{1+w}$ gives $\mathfrak{g}(w) = \int_0^{1+w} g(\frac{\vartheta}{1+w}, \frac{\vartheta w}{1+w})\frac{\vartheta}{(1+w)^2}d\vartheta$. Since $x > 1$, similarly, $\mathfrak{g}(x^{-2}w)x^{-2} = \int_0^{1+x^{-2}w} g(\frac{\vartheta}{1+x^{-2}w}, \frac{\vartheta x^{-2}w}{1+x^{-2}w})\frac{\vartheta x^{-2}}{(1+x^{-2}w)^2}d\vartheta$. Now, $\frac{\vartheta}{1+w} + \frac{\vartheta w}{1+w} = \vartheta = \frac{\vartheta}{1+x^{-2}w} + \frac{\vartheta x^{-2}w}{1+x^{-2}w}$ and $\frac{\vartheta}{1+x^{-2}w} \geq \frac{\vartheta}{1+w} \geq \frac{\vartheta w}{1+w}$ imply, by (H4), that $g(\frac{\vartheta}{1+w}, \frac{\vartheta w}{1+w}) \geq g(\frac{\vartheta}{1+x^{-2}w}, \frac{\vartheta x^{-2}w}{1+x^{-2}w})$ for every $\vartheta \in [0, 1+x^{-2}w]$. Moreover, it is easily seen that, for the given values of x and w, we have $\frac{\vartheta x^{-2}}{(1+x^{-2}w)^2} = \frac{\vartheta}{(x+x^{-1}w)^2} < \frac{\vartheta}{(1+w)^2}$ whenever $\vartheta > 0$. This proves (E-17) for the case $w \in (0, 1]$.

Now let $1 < w < x$. Then, applying (E-17) to $\widetilde{w} \equiv w^{-1} < 1$ and $\widetilde{x} \equiv xw^{-1} > 1$ gives $\mathfrak{g}(\widetilde{w}) > \mathfrak{g}(\widetilde{x}^{-2}\widetilde{w})\widetilde{x}^{-2}$, i.e., $w^{-2}\mathfrak{g}(w^{-1}) > \mathfrak{g}(x^{-2}w)x^{-2}$. But since $\Pr\{\frac{q_r}{q_\ell} \leq w\} = 1 - \Pr\{\frac{q_\ell}{q_r} \leq w^{-1}\}$ for $w > 0$, and since $g_{q_r/q_\ell} = g_{q_\ell/q_r}$, we have $\mathfrak{g}(w) = w^{-2}\mathfrak{g}(w^{-1})$. This completes the proof. □

REFERENCES

Babcock, Linda and Loewenstein, George. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives,* 1997, 11(1), pp. 109-126.

Billingsley, Patrick. *Probability and measure.* New York: Wiley, 1995.

Blais, André. *To vote or not to vote: The merits and limits of rational choice theory.* Pittsburgh: University of Pittsburgh Press, 2000.

Börgers, Tilman. "Costly Voting," *American Economic Review,* 2004, 94(1), pp. 57-66.

Chamberlain, Gary and Rothschild, Michael. "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory,* 1981, 25(1), pp. 152-162.

Coate, Stephen and Conlin, Michael. "A Group Rule Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review,* 2004, 94(5), pp. 1476-1504.

Downs, Anthony. *An economic theory of democracy.* New York: Harper and Row, 1957.

Edlin, Aaron; Gelman, Andrew and Kaplan, Noah. "Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others," *Rationality and Society,* 2007, 19(3), pp. 293-314.

Feddersen, Timothy J. "Rational Choice Theory and the Paradox of Not Voting," *Journal of Economic Perspectives,* 2004, 18(1), pp. 99-112.

Feddersen, Timothy J.; Gailmard, Sean and Sandroni, Alvaro. "Moral Bias in Large Elections: Theory and Experimental Evidence," Northwestern University, Department of Managerial Economics and Decision Sciences, 2008, mimeo.

Feddersen, Timothy J. and Sandroni, Alvaro. "A Theory of Participation in Elections," Northwestern University, Department of Managerial Economics and Decision Sciences, 2002, mimeo.

_____. "The Calculus of Ethical Voting," *International Journal of Game Theory,* 2006, 35(1), pp. 1-25.

_____. "A Theory of Participation in Elections," *American Economic Review,* 2007, 96(4), pp. 1271-1282.

Fehr, Ernst and Schmidt, Klaus M. "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics,* 1999, 114(3), pp. 817-868.

Feller, William. *An introduction to probability theory and its applications. Volume II.* New York: Wiley, 1966.

Fischer, A.J. "The Probability of Being Decisive," *Public Choice,* 1999, 101(3-4), pp. 267–283.

Fowler, James H. "Altruism and Turnout," *Journal of Politics,* 2006, 68(3), pp. 674–683.

Fowler, James H., and Kam, Cindy D. "Beyond the Self: Altruism, Social Identity, and Political Participation," *Journal of Politics,* 2007, 69(3), pp. 811–825.

Gelman, Andrew; Katz, Jonathan N. and Bafumi, Joseph. "Standard Voting Power Indexes Do Not Work: An Empirical Analysis," *British Journal of Political Science,* 2004, 34(4), pp. 657–674.

Good, I.J. and Mayer, Lawrence S. "Estimating the Efficacy of a Vote," *Behavioral Science,* 1975, 20(1), pp. 25 -33.

Harsanyi, John C. "Morality and the Theory of Rational Behavior," *Social Research,* 1977, 44(4), pp. 623-656.

_____. "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior," *Theory and Decision,* 1980, 12(2), pp. 115-133.

Hoeffding, Wassily. "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association,* 1963, 58(301), pp. 13-30.

Jankowski, Richard. "Buying a Lottery Ticket to Help the Poor: Altruism, Civic Duty, and Self-Interest in the Decision to Vote," *Rationality and Society,* 2002, 14(1), pp. 55–77.

Kinder, Donald R. and Kiewiet, D. Roderick. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgements in Congresional Voting," *American Journal of Political Science,* 1979, 23(3), 495-527.

Krasa, Stefan and Polborn, Mattias. "Is Mandatory Voting Better than Voluntary Voting," *Games and Economic Behavior,* 2009, forthcoming.

Ledyard, John O. "The Pure Theory of Large Two-Candidate Elections," *Public Choice,* 1984, 44(1), pp. 7-41.

Levine, David K. and Palfrey, Thomas R. "The Paradox of Voter Participation? A Laboratory Study," *American Political Science Review,* 2007, 101(1), pp.143-158

Lewis-Beck, Michael S. and Stegmaier, Mary. "Economic Determinants of Electoral Outcomes," *Annual Review of Political Science,* 2000, 3, pp. 183-219.

Markus, Gregory. "The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis," *American Journal of Political Science,* 1988, 32(1), pp. 137-154.

McDonald, Michael P. and Popkin, Samuel L. "The Myth of the Vanishing Voter," *American Political Science Review,* 2001, 95(4), pp. 963-974.

Messick, David and Sentis, Keith. "Fairness and Preference," *Journal of Experimental Social Psychology,* 1979, 15(4), pp. 418-434.

Morton, Rebecca B. "Groups in Rational Turnout Models," *American Journal of Political Science,* 1991, 35(3), pp. 758-776.

Mulligan, Casey B. and Hunter, Charles G. "The Empirical Frequency of a Pivotal Vote," *Public Choice,* 2003, 116(1-2), pp. 31-54.

Myerson, Roger B. "Large Passion Games," *Journal of Economic Theory*, 2000, 94(1), pp. 7-45.

Nannestad, Peter and Paldam, Martin. "The VP-function: A Survey of the Literature on Vote and Popularity Functions after 25 Years," *Public Choice,* 1994, 79(3-4), pp. 213-45.

Palfrey, Thomas R. and Rosenthal, Howard. "Voter Participation and Strategic Uncertainty," *American Political Science Review,* 1985, 79(1), pp. 62-78.

Riker, William and Ordeshook, Peter. "A Theory of the Calculus of Voting," *American Political Science Review,* 1968, 62(1), pp. 25-42.

Rohatgi, Vijay K. *An introduction to probability theory and mathematical statistics.* New York: Wiley, 1976.

Shachar, Ron and Nalebuff, Bany. "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review,* 1999, 89(3), pp. 525-47.

Taylor, Curtis R. and Yildirim, Huseyin. "Re-examining Voter Turnout in Large Elections," Duke University, Department of Economics, 2005, mimeo.

____. "A Unified Analysis of Rational Voting with Private Values and Cost Uncertainty," Duke University, Department of Economics, 2008, mimeo.

Uhlaner, Carole J. "Rational Turnout: The Neglected Role of Groups," *American Journal of Political Science,* 1989, 33(2), pp. 390-422.