Centre for
Economic
and Financial
Research
at
New Economic
School

April 2012

Altruism and Voting: A Large-Turnout Result That Does not Rely on Civic Duty or Cooperative Behavior

Özgür Evren

Working Paper No 173

CEFIR / NES Working Paper series

Altruism and Voting: A Large-Turnout Result That Does not Rely on Civic Duty or Cooperative Behavior

Özgür Evren*

New Economic School
Nakhimovsky Prospekt 47, Suite 1721, Moscow 117418
E-mail: oevren@nes.ru

April 22, 2012

Abstract

I propose a game-theoretic model of costly voting that predicts significant turnout rates even when the electorate is arbitrarily large. The model has two key features that jointly drive the result: (i) some agents are altruistic (or ethical), (ii) among the agents who prefer any given candidate, the fraction of altruistic agents is uncertain. When deciding whether to vote or not, an altruistic agent compares her private voting cost with the expected contribution of her vote to the welfare of the society. Under suitable homogeneity assumptions, the asymptotic predictions of my model coincide with those of Feddersen and Sandroni (2006a) up to potential differences between the respective parameters that measure the importance of the election. I demonstrate with an example that these homogeneity assumptions are not necessary for qualitative predictions of my model. I also show that when the fractions of altruistic agents are known, turnout rates will typically be close to zero in a large election, despite the presence of altruism.

JEL classification: D64; D72
Keywords: Altruism; Utilitarianism; Voting; Turnout; Pivotal Voter

*I am grateful to Debraj Ray for his continuous guidance and support. I thank the editor Christian Hellwig and two anonymous referees for their thoughtful suggestions which improved the exposition immensely. I would also like to thank Daniel Barczyk, Gokhan Buturak, Timothy Feddersen, James Fowler, Guillaume Frechette, Andrew Gelman, Anna Gumen, Yilmaz Kocer, Alessandro Lizzeri, Efe Ok, Howard Rosenthal, Joel Sobel, Srinivasa Varadhan, Atilla Yilmaz, and the seminar participants at HSE and NES. When this project started, I was a Ph.D. student at NYU, Department of Economics. I owe special thanks to this institution. All remaining errors are mine.

1. Introduction

Why do we observe substantial turnout rates even among millions of voters? Since Downs (1957), this very basic question has been a major challenge for political economists. The difficulty is that voting is a time consuming, costly activity, but a single vote (among many others) is highly unlikely to influence the election outcome. Thus, according to the classical pivotal-voter approach, the turnout rate in a large election must be approximately zero (Palfrey and Rosenthal, 1985).[1]

In this paper, I propose a game-theoretic model that predicts significant turnout rates even when the electorate is arbitrarily large. The model has two key features that jointly derive the result: (i) altruism towards other voters, and (ii) uncertainty about aggregate voting behavior.

The objective of an altruistic agent is to maximize the expected value of an additive welfare function (determined by her ideological beliefs). This implies that when deciding whether to vote or not, a given altruistic agent compares her private voting cost with the expected contribution of her vote to the welfare of the society.[2] The latter term is asymptotically proportional to $n\mathcal{P}$, where n is the size of the electorate (excluding the agent) and $\mathcal{P}$ is the probability that the agent will be decisive (pivotal). Thus, compared with the classical model that assumes population invariant utility functions, the presence of altruism scales the perceived benefits to voting by n. In this framework, whether the agent would vote in a large election depends on the rate at which $\mathcal{P}$ converges to zero.

Uncertainty about aggregate voting behavior, the second key component of the model, ensures that the equilibrium value of $\mathcal{P}$ is asymptotically proportional to $1/n$. This, in turn, implies that the expected contribution of our agent's vote to the welfare of the society converges to a positive number, making it well possible for the agent to vote or abstain, depending on her voting cost.

To model such aggregate uncertainty, I assume that the agents' types are independently and identically distributed conditional on a parameter q, and that the true value of q is unknown. More specifically, q has two components q_ℓ and q_r, and q_i equals the (conditional) probability that a randomly chosen agent among the supporters of candidate i is altruistic. By the law of large numbers, in a large election, q_i can also be seen as the fraction of altruistic agents among the supporters of candidate i. Thus, effectively, I assume that the fraction of altruistic agents who support any given candidate is unknown.

[1] More precisely, Palfrey and Rosenthal show that the turnout rate must be approximately equal to the fraction of agents who perceive voting as a civic duty that is more important than the associated costs. In this paper, I abstract from the fact that citizens may perceive voting as a civic duty.

[2] It is also worth noting that I allow the welfare function in the mind of an altruistic agent to be biased towards her self-interest. The use of such biased welfare functions in the analysis of social choice problems dates back to Sen (1966).

Since selfish agents abstain, this uncertainty randomizes the equilibrium vote shares of the two candidates in such a way that the implied pivot probabilities are proportional to $1/n$. By contrast, when q_ℓ and q_r are known, pivot probabilities decline at an exponential rate, excluding the special cases in which the values of the parameters that characterize the electorate happen to imply tie in equilibrium. Hence, when q_ℓ and q_r are known, the conclusion of the impossibility theorem of Palfrey and Rosenthal (1985) typically survives (see Proposition 6 below).

I also show that my theory is compatible with several stylized phenomena related to political elections. Specifically, the model implies that the expected total turnout increases with the importance and expected closeness of the election, and that the expected turnout rate of the minority is larger than that of the majority.[3]

Why does an altruistic agent adopt a welfare maximizing behavior? There are two possible interpretations. The first, more traditional, interpretation is that what I refer to as a "welfare function" is, in fact, an altruistic von Neumann-Morgenstern utility index; and the agent is an expected utility maximizer in the standard sense. A disadvantage of this interpretation is that, *when given the chance*, such an altruistic agent would be willing to make huge sacrifices in order to influence the outcome of a large election. In turn, according to the second interpretation, the agent believes that the maximization of the welfare function in her mind corresponds to an ethical mode of behavior.[4] Moreover, the agent is "ethical" in the sense that she receives a payoff by adopting the ethical behavior. Thus, the agent votes if, and only if, (i) this is the ethical act, and (ii) the payoff of acting ethically exceeds her private voting cost. Here, the payoff associated with ethical behavior represents the intrinsic utility of such behavior. This alternative interpretation follows the "warm-glow" literature that focuses on intrinsic value of prosocial actions (e.g., Andreoni, 1990; Coate and Conlin, 2004; Feddersen and Sandroni 2006a, 2006b).

If, in the second interpretation, the intrinsic payoff of ethical behavior exceeds the maximum possible voting cost, the agents corresponding to two interpretations behave precisely in the same way. I focus on this particular case because allowing for a smaller intrinsic payoff does not lead to qualitatively different predictions within the context of my model (see Section 5 below). Furthermore, in real elections that motivate this paper, a single vote is not likely to influence the election outcome, implying that even altruistic agents in the traditional sense may not be willing to incur unreasonably large voting costs. Therefore, I remain agnostic about the two interpretations of the model.

Variety of empirical and experimental findings support the idea of other-regarding vot-

[3] These predictions are compatible with empirical and experimental findings (see Section 3).

[4] The ethical theory that promotes such behavior in a game-theoretic set-up is known as *act utilitarianism* (see Harsanyi, 1980). I compare this interpretation of my model with the earlier literature on ethical voters throughout the paper and in Online Appendix A (at https://files.nyu.edu/oe240/public/atsj1_o_app.pdf).

ers. Notably, there is considerable evidence that voting behavior is better explained by "sociotropic" concerns about the overall state of the macroeconomy rather than individual concerns (Kinder and Kiewiet, 1979; Markus, 1988). For instance a person, say, an economist, who might vote against a proponent of free trade policies may actually be concerned about the *number* of low skilled workers who may lose their job, rather than her personal financial situation. Moreover, recent experimental evidence points to a positive relation between subjects' participation in elections and (i) their level of altruism measured with their generosity in dictator games (Fowler, 2006; Fowler and Kam, 2007), and (ii) their moral concerns about the well-being of others (Feddersen, Gailmard and Sandroni, 2009).

Motivated by similar observations, recently, scholars have proposed several models of altruistic or ethical voters. In this literature, the closest model to mine is that of Feddersen and Sandroni (2006a, 2006b), which focuses on ethical voters who are concerned with the well-being of the society. Just as I do in the present paper, Feddersen and Sandroni assume that the fraction of ethical agents who support a given candidate is uncertain. The distinctive future of their model is the equilibrium concept that they utilize, which is not game-theoretic in the traditional sense. A key ingredient of their model is a group structure that divides the set of all types (i.e., agents' characteristics) into certain groups. A *rule* for a given group defines a type contingent behavior that the types in that group should follow. Ethical agents get an intrinsic payoff by acting as they should, and they compare this payoff with their private voting cost when deciding whether to vote or not (as in the second interpretation of my model). Thereby, a rule profile determines the actual behavior of all types. In equilibrium, this behavior is required to be *consistent* in the sense that it must be induced by a rule profile that achieves the best (expected) social outcome from the perspective of all types in any given group, taking as given the behavior of the types outside that group.

By definition, the consistency requirement above necessitates all types in a given group to agree about the optimality of a certain rule. This, in turn, implies that the permissible group structures depend on the level of homogeneity of agents' characteristics such as the intensity of their preferences towards the candidates.[5] By utilizing suitable homogeneity assumptions,[6] Feddersen and Sandroni (2006a, 2006b) focus on a particular case with two groups determined by the favored candidates of the agents. I also utilize analogous homogeneity assumptions in the main body of the present paper. In Online Appendix A, I show that, under these assumptions, the asymptotic predictions of my model coincides

[5] Specifically, higher levels of homogeneity facilitate larger group structures (in measure theoretic sense). Online Appendix A contains a more detailed discussion of the role of homogeneity assumptions in this set-up.

[6] "Heterogeneity within a group is possible, but must be restricted to differences among individuals' cost to vote." (Feddersen and Sandroni, 2006b, p.3)

precisely with those of Feddersen and Sandroni (2006a) for a suitable specification of social preferences in their model. Feddersen and Sandroni do not investigate if their homogeneity assumptions are salient for their findings. Fortunately, my game-theoretic approach is compatible with various forms of heterogeneity in agents' characteristics. I demonstrate this with an example in Appendix A, which shows that all my findings survive even under an extreme form of heterogeneity in agents' intensity of preferences. It is also worth noting that while Feddersen and Sandroni assume a continuum of agents, I study the asymptotic behavior of a finite model.

Another related paper is due to Edlin, Gelman and Kaplan (2007), who propose a model of altruistic voters with *exogenous* pivot probabilities. In line with my approach, Edlin et al. (2007) assume that the pivot probabilities are inversely proportional to the size of the electorate. By endogenizing pivot probabilities, in this paper I report more satisfactory comparative statics exercises. Specifically, my findings that relate turnout to expected closeness of the election and to the relative size of the supporters of the candidates are not within the scope of Edlin et al. (2007).

To the best of my knowledge, Jankowski (2007), and Faravelli and Walsh (2011) are the only other game-theoretic papers on altruistic voters. Unlike the present paper, Jankowski's model predicts tie in equilibrium, which is not compatible with substantial vote differentials that we observe in real elections. (More on this in Online Appendix C.) A key feature of Jankowski's model is that all agents have the same, deterministic voting cost. In equilibrium, all agents are indifferent between voting and abstaining, and they randomly select these actions in such a way that the expected fraction of votes for both candidates equals 1/2. In turn, my findings show that introducing cost uncertainty to Jankowski's model would typically lead to low turnout rates by eliminating such mixed strategy equilibria. (See, in particular, Footnote 17 below.)

Faravelli and Walsh (2011) is a concurrent, working paper that proposes a fundamentally different approach. They relax the usual winner-take-all assumption. Specifically, they assume that the winning candidate responds smoothly to her margin of victory. In their model, a single vote has always an effect on the policy outcome, but this effect becomes smaller in a large election. Assuming a form of altruism as in the present paper, they show that this alternative approach is also compatible with significant turnout rates, despite the fact that there is no uncertainty about aggregate voting behavior in their model.

From a technical point of view, a key finding of the present paper is a formula on the magnitude of pivot probabilities that requires q to be a continuous random variable (see Lemma E1 in Appendix E). Earlier, Good and Mayer (1975) have provided a related formula. The main novelty of my approach is that I allow for abstention. Put formally, Good and Mayer assume that a randomly chosen agent votes for a given candidate with an unknown probability P and for the other candidate with probability $1 - P$. By contrast, in my model, even the altruistic agents may abstain. Specifically, in the equilibrium of my model, a randomly chosen agent votes for candidate i with probability $\lambda_i q_i F(C_i^*)$, where λ_i is the fraction of agents who prefer candidate i, and $F(C_i^*)$ is the fraction of participants among altruistic agents with such preferences. Allowing for abstention in this way enables me to identify the interactions between individuals' participation decisions and turnout rates of the supporters of the two candidates. In turn, my comparative statics exercises build upon these observations.

In the next section I formally introduce my model. In Section 3, I present my main findings under uncertainty in q. Section 4 contains my negative result for the case of known q. In Section 5, I discuss several extensions of my basic model with uncertain q. In Section 6, I relate my model to Feddersen and Sandroni (2006a, 2006b). I conclude in Section 7. Appendix contains the proofs and some other supplementary material.

2. The Model

The society consists of $n+1$ agents with a generic member h. To simplify the exposition, I assume that all agents are eligible to vote and that n is a known positive integer.

Throughout the paper, I often use the same notation for a random variable and a possible value of that random variable.

Agent h has private knowledge of: (i) her **policy type** which can be ℓ (left) or r (right), (ii) her **personality type** which can be s (selfish) or a (altruistic), and (iii) her **voting cost** $C \in \mathbb{R}_+$. I denote by τ_h the three dimensional random vector that describes these characteristics. In what follows, I simply write "type" instead "policy type."

$\tau_1, ..., \tau_{n+1}$ are iid random variables conditional on a possibly random vector $q \equiv (q_\ell, q_r)$. Here, $q_i \in [0, 1]$ stands for the probability that a randomly chosen agent of type $i \in \{\ell, r\}$ is altruistic. I denote by G the joint distribution of (q_ℓ, q_r). I assume that q_ℓ and q_r are positive with probability 1. In turn, λ stands for the probability that a randomly chosen agent is of type ℓ. For simplicity, I assume that λ is known. The distribution function of a randomly chosen agent's voting cost is given by F.

When n is large, by the law of large numbers, λ can be seen as the fraction of type ℓ agents and q_i as the fraction of altruistic agents among type i agents. I set $\lambda \leq 1/2$ so that type ℓ agents is a *minority*.

I make the following assumption on the distribution of C.

(H1) The support of F is an interval of the form $[0, c] \subseteq \mathbb{R}_+$ for some $c > 0$. Moreover, F has a density f that is continuous and positive on $[0, c]$.

There are two candidates, also denoted by ℓ and r. Given any $i \in \{\ell, r\}$, 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 $u > 0$ 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. Accordingly, I assume that the behavior of an altruistic agent h' of type i maximizes the following function:

$$\mathbf{E}\left(\left(u\mathbf{1}_i - C_{h'}\mathbf{1}^{h'}\right) + \psi \sum_{h \neq h'} \left(u\mathbf{1}_i - C_h\mathbf{1}^h\right)\right). \tag{1}$$

Here, $\mathbf{E}$ is the expectation operator; $\psi \in (0, 1]$ is the weight that an altruistic agent's objective function places on others' payoffs; C_h is the voting cost of agent h, for $h = 1, ..., n+1$, and

$$\begin{aligned} \mathbf{1}^h &\equiv 1 \text{ if } h \text{ votes, } \mathbf{1}^h \equiv 0 \text{ otherwise;} \\ \mathbf{1}_i &\equiv 1 \text{ if candidate } i \text{ wins, } \mathbf{1}_i \equiv 0 \text{ otherwise.} \end{aligned}$$

As I discussed earlier, there are two interpretations of why the agent in question maximizes the function (1).

(I) Preference-Intensity Interpretation: The agent is an expected utility maximizer in the traditional sense, and the term inside the expectation operator in (1) is simply the von Neumann-Morgenstern utility index of the agent. Therefore, the agent *always* takes the action that maximizes (1).

(II) Warm-Glow Interpretation: (1) represents a welfare function (that is possibly biased towards the agent's self-interest). The agent believes that taking the action that maximizes this function is ethically right thing to do. Moreover, the agent is *ethical* in the sense that she receives an intrinsic payoff $D > 0$ by adopting the ethical mode of behavior. Thus, whenever she should vote from the ethical point of point of view, the agent compares D with her private voting cost $C_{h'}$. Specifically, the agent votes for a given candidate if this action maximizes (1) and if, at the same time, $D \geq C_{h'}$. In turn, the agent abstains if $D < C_{h'}$, or if this is the action that maximizes (1).

It is important to note that according to interpretation (II), the agent takes the action that maximizes (1) whenever $D \geq C_{h'}$. If $D \geq c$, this is a sure event, and hence, the two interpretations predict the same behavior. Henceforth, I assume $D \geq c$ within interpretation (II). (I discuss the case $D < c$ below, in Remark 1 and Section 5.)

When solving the optimization problem that I just described, each altruistic agent takes as given the strategies of others in a game-theoretic fashion. A (pure) strategy for 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 $C \in \mathbb{R}_+$. Here, -1, 1 and 0 stand for "vote for candidate ℓ," "vote for candidate r" and "abstain," respectively. Thus, selfish agents (who care only about their own payoff) are assumed to abstain. Since such agents would necessarily abstain as n tends to ∞, this assumption has no role in my asymptotic results; it simply serves to avoid technical details.

The action that agent h will take is a random variable given by $X_h \equiv Y_h \circ \tau_h$. Since the agents are ex-ante symmetric, I assume that all agents use the same strategy, i.e., $Y_1 = Y_2 = \cdots = Y_{n+1}$. Given that $\tau_1, ..., \tau_{n+1}$ are iid conditional on q, it follows that so are $X_1, ..., X_{n+1}$.

Since voting costs are positive, for any agent casting a vote against her favored candidate is strictly dominated by abstaining. Conditional on q, this makes the number of votes for candidate ℓ a Binomial random variable with "success probability" $\Pr\{X_h = -1 \mid q\}$. Similarly, $\Pr\{X_h = 1 \mid q\}$ gives the success probability for the conditional distribution of the number of votes for candidate r. I denote by $P_i(q)$ the implied conditional pivot probability for a given agent of type i. That is, $P_i(q)$ is the conditional probability of the event that the election is tied excluding the agent and candidate i loses the coin toss, or the agent's vote creates a tie and i wins the coin toss. The usual formula for $P_i(q)$ can be found in Appendix E.1.

G^i stands for the *posterior* distribution function of q from the perspective of an altruistic agent of type i. In turn, $\mathcal{P}_i$ denotes such an agent's assessment of the unconditional probability of being pivotal:

$$\mathcal{P}_i \equiv \int_{[0,1]^2} P_i(q)\, dG^i(q). \tag{2}$$

Consider an altruistic agent h' of type i whose voting cost is $C \equiv C_{h'}$. Holding fixed the strategies of all other agents, let $\mathcal{E}_i^+$ be the value of (1) that obtains if the agent votes for candidate i and if the expectation operator is applied with respect to G^i. Similarly, let $\mathcal{E}_i$ be the corresponding value of (1) that obtains if the agent abstains. It can easily be seen that $\mathcal{E}_i^+ - \mathcal{E}_i = u\,(1 + \psi n)\,\mathcal{P}_i - C$. Here, the term

$$\Pi_i \equiv u\,(1 + \psi n)\,\mathcal{P}_i \tag{3}$$

is the increase in (1) due to the potential contribution of the agent's vote to the welfare of the society through the election outcome. It follows that to maximize (1), an altruistic agent of type i should vote for candidate i if $\Pi_i > C$, and abstain if $C < \Pi_i$. These observations lead to the following notion of equilibrium.

Definition 1. An **equilibrium** consists of a pair of cutoff points $C_\ell^*, C_r^* \in \mathbb{R}_+$ such that $\Pi_i^* = C_i^*$ for $i \in \{\ell, r\}$, where Π_i^* is the value of (3) induced by the conditional voting probabilities $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell F(C_\ell^*)$ and $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r F(C_r^*)$.

Remark 1. Recall that according to interpretation (II), even if the voting cost C of a type i agent is less than Π_i, the agent would not vote whenever $C > D$. This point becomes immaterial when $D \geq c$, for then the event $C > D$ becomes null. On the other hand, when $D < c$, the notion of equilibrium above needs to be modified accordingly. In Section 5, I will show that the case $D < c$ does not lead to qualitatively different predictions (although the implied turnout rates would typically be lower). In fact, the modified notion of equilibrium can be transformed back to the equilibrium notion above upon a suitable adjustment of voting costs.

The following preliminary observation shows that an equilibrium exists, and the cutoff points are positive in any equilibrium.

Proposition 1. *An equilibrium exists. Moreover, in any equilibrium, $C_i^* > 0$ for every type i.*

As n tends to ∞, do the cutoff points converge to 0 or to positive numbers? By equation (3), the answer to this question depends on the rate at which $\mathcal{P}_\ell$ and $\mathcal{P}_r$ converge to 0. In turn, the rate of convergence of $\mathcal{P}_\ell$ and $\mathcal{P}_r$ depends on the distribution of q via the equation (2). In the remainder of the paper, I will examine the asymptotic behavior of equilibria under two alternative assumptions on the distribution of q.

3. Asymptotic Turnout when q Is Unknown

In this section, I assume that q is a continuous random variable. Specifically:

(H2) G has a continuous and positive density g on $[0,1]^2$.

I will also assume that:

(H3) $g(t+\varepsilon, t-\varepsilon)$ is nonincreasing in ε for $0 \leq \varepsilon \leq \min\{t, 1-t\}$ and every fixed $t \in [0,1]$.

(H4) $g(q_\ell, q_r) = g(q_r, q_\ell)$ for every $(q_\ell, q_r) \in [0,1]^2$.

(H2) is my most important assumption, for, by itself, it ensures that $\lim_n C_{i,n}^* > 0$ for any type i and any sequence of convergent equilibria $(C_{\ell,n}^*, C_{r,n}^*)$. Assumptions (H3) and (H4), on the other hand, enable me to show that there exists a unique such limit point for each type and to report clear-cut comparative statics exercises. Intuitively, (H3) implies that the prior density g does not place higher probabilities on asymmetric realizations of q_ℓ and q_r. More specifically, (H3) requires that a vertical movement away from the line

$q_\ell = q_r$ (caused by a fall in q_r) does not increase g. In turn, (H4) rules out the cases in which one of the candidates may have an ex-ante advantage in terms of the fractions of altruistic agents. In particular, (H4) implies that the expected values of q_ℓ and q_r with respect to g are equal to each other.

It is important to note that (H3) and (H4) allow a positive correlation between q_ℓ and q_r, as one may expect in reality. For example, it can easily be checked that both assumptions hold if the distribution of q is obtained by conditioning a bivariate normal distribution to the unit square $[0,1]^2$ provided that the marginals of the normal distribution are identical. A simple case which immediately implies (H3) and (H4) is when q_ℓ and q_r are independently and uniformly distributed. More generally, when q_ℓ and q_r are iid beta random variables with monotone or unimodal densities, then (H3) and (H4) hold.[7]

My final assumption in this section is that:

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

While (H5) simplifies the exposition of my comparative statics exercises, all my findings remain true for a more general class of cost distributions (see Section 5).

I start with a simple claim:

Claim 1. *The posterior distribution G^i admits a density g^i defined by*

$$g^i(q) \equiv \frac{q_i}{\overline{q}_i} g(q) \quad \textit{for every } q \in [0,1]^2, \tag{4}$$

where $\overline{q}_i$ is the expectation of q_i with respect to g.

In the next proposition, I show that the cutoff points of both types converge to positive numbers.[8] I also provide two equations that jointly determine these limit points. This is the main finding of the paper.

Proposition 2. *For each type i, there exists a unique number $C_i^\bullet$ such that $\lim_n C_{i,n}^* = C_i^\bullet$ for any sequence of equilibria $(C_{\ell,n}^*, C_{r,n}^*)$. Moreover, we have $0 < C_i^\bullet < \infty$ and*

$$C_i^\bullet = u\psi \frac{T_j^\bullet}{\overline{q}_i} \int_0^{\frac{1}{\max\{T_\ell^\bullet, T_r^\bullet\}}} \theta g(\theta T_r^\bullet, \theta T_\ell^\bullet)\, d\theta \quad \textit{for } i = \ell, r, \tag{5}$$

where $T_\ell^\bullet \equiv \lambda F(C_\ell^\bullet)$ and $T_r^\bullet \equiv (1-\lambda) F(C_r^\bullet)$.

In the proof of this proposition, I first provide a formula for the asymptotic magnitude

[7] In Appendix B, I provide explicit formulas for the compatible distributions that I mentioned in this paragraph.

[8] I do not have a result on the rate of this convergence, but I have performed some simulations. The results indicate that the cutoff points do not significantly differ from their limit points when $n > 1000$. (See Online Appendix B.)

of $\mathcal{P}_i$ under the assumption that the cutoff points of both types remain bounded away from 0. Then, I show that the cutoff points are indeed bounded away from 0. Thereby, I conclude that

$$n\mathcal{P}_i \to \int_0^{\frac{1}{\max\left\{T_\ell^\bullet, T_r^\bullet\right\}}} g^i(\theta T_r^\bullet, \theta T_\ell^\bullet)\, d\theta \tag{6}$$

along any subsequence of equilibria that converges to an arbitrary pair of numbers $(C_\ell^\bullet, C_r^\bullet)$. Then, I substitute (4) into expression (6) and invoke the definition of equilibrium to obtain the system of equations (5). The final step in the proof is to show that this system of equations has a unique solution $(C_\ell^\bullet, C_r^\bullet)$. Since it is somewhat involved, I will discuss this uniqueness issue after presenting my comparative statics exercises.

Now, I will examine expression (6) more closely to provide insight about the behavior of pivot probabilities. The first point to note is that the conditional pivot probability $P_i(q)$ exhibits a knife-edge behavior depending on the distance between the voting probabilities $\Pr\{X_h = -1 \mid q\}$ and $\Pr\{X_h = 1 \mid q\}$. If this distance remains bounded away from 0 as $n \to \infty$, then $P_i(q)$ declines at an exponential rate with n (see Corollary E1 in Appendix E). By contrast, $P_i(q)$ is asymptotically proportional to $1/\sqrt{n}$ when $\Pr\{X_h = -1 \mid q\} = \Pr\{X_h = 1 \mid q\}$. Thus, for large values of n, almost all contribution to unconditional pivot probability $\mathcal{P}_i$ comes from those q for which $\Pr\{X_h = -1 \mid q\}$ and $\Pr\{X_h = 1 \mid q\}$ are almost equal to each other. In turn, in equilibrium, the condition $\Pr\{X_h = -1 \mid q\} = \Pr\{X_h = 1 \mid q\}$ holds if and only if $\lambda q_\ell F(C_\ell^*) = (1-\lambda) q_r F(C_r^*)$. Therefore, $\mathcal{P}_i$ can be approximated by the posterior likelihood of those realizations of q that are very close to the following set:

$$\mathcal{CR} \equiv \left\{ (q_\ell, q_r) \in [0,1]^2 : \frac{q_\ell}{(1-\lambda)F(C_r^*)} = \frac{q_r}{\lambda F(C_\ell^*)} \right\}.$$

I refer to this set as the **critical ray,** and hence the notation $\mathcal{CR}$. Next note that if we denote by θ the common value of $\frac{q_\ell}{(1-\lambda)F(C_r^*)}$ and $\frac{q_r}{\lambda F(C_\ell^*)}$ along the critical ray, we obtain the following alternative expression:

$$\mathcal{CR} = \left\{ (\theta T_r^*, \theta T_\ell^*) : 0 \le \theta \le \frac{1}{\max\{T_\ell^*, T_r^*\}} \right\},$$

where $T_\ell^* \equiv \lambda F(C_\ell^*)$ and $T_r^* \equiv (1-\lambda)F(C_r^*)$. Expression (6) builds open this equality. Specifically, (6) tells us that, asymptotically, $\mathcal{P}_i$ is equal to $1/n$ multiplied by the integral of g^i (with respect to θ) over $\mathcal{CR}$.

Since $n\mathcal{P}_i$ is asymptotically proportional to Π_i, we can understand the asymptotic behavior of Π_i by examining the integral on the right side of expression (6). The following lemma uncovers two properties of this integral, which prove useful in the rest of my analysis.

Lemma 1. *Set* $\varphi^i(T_\ell, T_r) \equiv \int_0^{\frac{1}{\max\{T_\ell, T_r\}}} g^i(\theta T_r, \theta T_\ell)\, d\theta$ *for* $i = \ell, r$ *and* $(T_\ell, T_r) \in (0,1)^2$ *with* $T_\ell + T_r \leq 1$.

(i) (Level Effect) *The functions* φ^ℓ *and* φ^r *decrease with* $T_\ell + T_r$ *holding* T_r/T_ℓ *constant.*

(ii) (Ratio Effect) *Suppose* $T_r \geq T_\ell$. *Then, the functions* $\varphi^\ell + \varphi^r$ *and* φ^r *decrease with a further increase in* T_r/T_ℓ *holding* $T_\ell + T_r$ *constant.*

I present the proof of this lemma in Appendix E, which is a simple algebraic exercise. To gain insight, let T_ℓ and T_r stand for hypothetical values of the equilibrium objects T_ℓ^* and T_r^*, respectively. Suppose that excluding a given altruistic agent, all agents behave as implied by T_ℓ and T_r. Then, how would the changes in T_ℓ and T_r influence the motivation to vote for the particular agent that we excluded? The level and ratio effects answer this question. Specifically, level effect shows that an increase in $T_\ell + T_r$ decreases Π_ℓ and Π_r asymptotically. Thereby, the level effect formalizes the idea that a single agent would be less inclined to vote in elections with higher expected turnout. Similarly, the anticipation of a large margin of victory could decrease agents' motivation to vote. Ratio effect formalizes this point by showing that, asymptotically, $\Pi_\ell + \Pi_r$ decreases in response to an increase in $T_r/T_\ell \geq 1$. On the other hand, while Π_r behaves in the same direction, Π_ℓ may or may not decrease. The reason is the informational asymmetry between the two types of agents: for each q over the critical ray, we have $g^\ell(q)/g^r(q) = q_\ell/q_r = T_r/T_\ell$ implying that Π_ℓ/Π_r approximately equals T_r/T_ℓ. Thus, an increase in T_r/T_ℓ simply increases Π_ℓ/Π_r.

3.1 Comparative Statics

Let us take a large value of n and suppose that all agents behave as predicted by the equilibrium of the model. Then, by the law of large numbers, the observed fraction of votes for candidate i relative to the size of the electorate (including the abstainers) would approximately be equal to $q_i T_i^*$ (up to a small probabilistic error term). In turn, Proposition 2 implies that T_i^* must be close to the number $T_i^\bullet$. Thus, in my comparative statics exercises, I use the random variable $q_i T_i^\bullet$ to approximate the fraction of votes for candidate i relative to the size of the electorate. This motivates the following definition.

Definition 2. *Expected turnout rate* refers to $\overline{q}_\ell T_\ell^\bullet + \overline{q}_r T_r^\bullet$, while *expected margin of victory* is defined by

$$MV \equiv \int_{[0,1]^2} \left| \frac{q_r T_r^\bullet - q_\ell T_\ell^\bullet}{q_r T_r^\bullet + q_\ell T_\ell^\bullet} \right| g(q)\, dq.$$

In turn, the *winning probability of the majority* is the probability of the event $\left\{ \frac{q_\ell}{q_r} \leq \frac{T_r^\bullet}{T_\ell^\bullet} \right\}$ with respect to g.

By the logic of approximation that I just explained, we can also view the random variable $q_i F(C_i^\bullet)$ as the turnout rate of type i agents. Thus, in what follows, by the *expected turnout rate of the minority* (resp. *majority*) I mean the number $\overline{q}_\ell F(C_\ell^\bullet)$ (resp. $\overline{q}_r F(C_r^\bullet)$). Finally, I will refer to λ as the *level of disagreement*, for smaller values of λ correspond to those cases in which a large fraction of the society agrees that candidate r is better than candidate ℓ.

The next proposition compares the equilibrium behavior of the two types of agents.

Proposition 3. Suppose $\lambda < 1/2$. Then, $C_r^\bullet$ is smaller than $C_\ell^\bullet$, while the winning probability of the majority is larger than 1/2.[9]

That $C_r^\bullet$ is smaller than $C_\ell^\bullet$ means that an altruistic agent in the majority is less likely to vote than an altruistic agent in the minority. Since $\overline{q}_\ell = \overline{q}_r$, this also implies that the expected turnout rate of the majority is smaller than the expected turnout rate of the minority. This phenomenon is known as the *underdog effect* in the literature.[10] Proposition 3 also shows that the underdog effect is not strong enough to offset the size advantage of the majority in that they are more likely to win the election.

Analytically, Proposition 3 is a trivial consequence of the following observations that we obtain by dividing the first equation in (5) with the second one:

$$\frac{C_\ell^\bullet}{C_r^\bullet} = \frac{T_r^\bullet}{T_\ell^\bullet} \equiv \frac{(1-\lambda)\,F(C_r^\bullet)}{\lambda F(C_\ell^\bullet)}, \quad \text{i.e.,} \tag{7}$$

$$\frac{C_\ell^\bullet F(C_\ell^\bullet)}{C_r^\bullet F(C_r^\bullet)} = \frac{1-\lambda}{\lambda}. \tag{8}$$

On the other hand, conceptually, what drives Proposition 3 is the aforementioned informational asymmetry: In a large election with $T_i^* > T_j^*$, the outcome of the election can be close only if $q_j > q_i$. But for such q, we have $g^j(q) > g^i(q)$, implying that an altruistic agent of type j deems herself more likely to be pivotal than an altruistic agent of type i. Thus, in the limit, $T_i^\bullet > T_j^\bullet$ implies $C_j^\bullet > C_i^\bullet$. But with $\lambda < 1/2$, this can be true only if $C_\ell^\bullet > C_r^\bullet$ and $T_\ell^\bullet < T_r^\bullet$. Moreover, in view of the symmetry assumption (H4), that $T_\ell^\bullet < T_r^\bullet$ simply means that the winning probability of the majority is larger than 1/2.

The next proposition establishes a negative correlation between turnout and margin of victory. This phenomenon, which is known as the *competition effect*, has attracted considerable attention in the empirical literature (e.g., Shachar and Nalebuff, 1999; Blais, 2000).

Proposition 4. *An increase in the level of disagreement increases the expected turnout rate (unless the expected turnout rate is already at its maximum level, $\overline{q}_\ell\lambda + \overline{q}_r(1-\lambda)$).*

[9] When $\lambda = 1/2$, these strict relations turn into equalities.

[10] Levine and Palfrey (2007) provide experimental evidence for the underdog effect.

Moreover, an increase in the level of disagreement decreases the expected margin of victory.

The proof of the first part of Proposition 4 builds upon the level and ratio effects. To understand the basic arguments, assume that $C_\ell^\bullet$ and $C_r^\bullet$ are less than c. Then, under the uniformity assumption (H5), equations (7) and (8) imply that $T_r^\bullet/T_\ell^\bullet$ is decreasing with λ. Thus, an increase in λ causes an upward pressure on $C_r^\bullet$ and $C_\ell^\bullet + C_r^\bullet$ through the ratio effect. On the other hand, if the expected turnout rate were to fall, the level effect would push up $C_r^\bullet$ and $C_\ell^\bullet + C_r^\bullet$ further. But then, since $F(C)$ is linear in C, the expected turnout rate would also increase, which is a contradiction. In turn, if the expected turnout rate were to remain constant, we would obtain a similar contradiction. Hence, an increase in λ must actually increase the expected turnout rate. Finally, the second part of Proposition 4 follows from the fact that the expected margin of victory increases with $T_r^\bullet/T_\ell^\bullet$ (see Lemma E6 in Appendix E).

The last result of this section examines the effects of the parameters u, ψ and c.

Proposition 5. *An increase in $u\psi$ increases the expected turnout rate (unless the expected turnout rate is already at its maximum level). Moreover, the expected margin of victory is a nondecreasing function of $u\psi$. The consequences of a decrease in c are analogous to those of an increase in $u\psi$.*

A particular implication of Proposition 5 is that the expected turnout rate increases with u. The parameter u measures how strongly a given candidate would contribute to the welfare of the society according to her supporters' belief. Therefore, Proposition 5 is consistent with empirical findings which show that there is a positive relation between turnout and voters' perception of how important the election is (e.g., Teixeira, 1987, 1992).

The level effect is the main force behind Proposition 5. Indeed, if $C_\ell^\bullet$ and $C_r^\bullet$ are less than c, the ratio $T_r^\bullet/T_\ell^\bullet$ solely depends on λ. So, among such equilibria, the ratio effect is silent about the implications of a change in u. On the other hand, by the level effect, a fall in the expected turnout rate would push $C_\ell^\bullet$ and $C_r^\bullet$ upward. Moreover, an increase in u would only strengthen the upward pressure on $C_\ell^\bullet$ and $C_r^\bullet$, implying that the expected turnout rate should actually increase. Therefore, a fall in the expected turnout rate is not compatible with an increase in u. By the same logic, we can also rule out the case in which the expected turnout rate remains constant in response to an increase in u.

In turn, since the expected margin of victory is an increasing function of $T_r^\bullet/T_\ell^\bullet$, the second part of Proposition 5 is equivalent to saying that $T_r^\bullet/T_\ell^\bullet$ is a nondecreasing function of $u\psi$. The latter statement is true because: (i) $T_r^\bullet/T_\ell^\bullet$ is constant for smaller values of $u\psi$, (ii) as we increase $u\psi$, at some point $C_\ell^\bullet$ reaches c while $C_r^\bullet$ remains below c, and (iii) a further increase in $u\psi$ increases $T_r^\bullet$ while keeping $T_\ell^\bullet$ constant.

In passing, I will comment on the proof of the fact that the system of equations (5) has a unique solution. The first point to note is that if there were two solutions, say $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$, the sign of $C_\ell^\bullet - \overline{C}_\ell^\bullet$ would be the same as that of $C_r^\bullet - \overline{C}_r^\bullet$ by equation (8). Assume therefore that $\overline{C}_\ell^\bullet > C_\ell^\bullet$ and $\overline{C}_r^\bullet > C_r^\bullet$. Then, (H5) and (8) imply $\overline{C}_\ell^\bullet / \overline{C}_r^\bullet \geq C_\ell^\bullet / C_r^\bullet$. Thus, by (7), when moving from the smaller solution to the larger solution, the ratio effect creates a downward pressure on Π_r. However, the level effect also pushes Π_r downward, implying that $\overline{C}_r^\bullet \leq C_r^\bullet$, a contradiction as we seek.

4. Asymptotic Turnout when q Is Known

In this section, I investigate the asymptotic behavior of equilibria under the assumption that q_ℓ and q_r are known. Since λ is already taken as deterministic, this assumption rules out the only source of parametric uncertainty. Thereby, we obtain a binomial model with altruistic voters. In this set-up, the knife-edge behavior of pivot probabilities leads to the following negative result.

Proposition 6. *Suppose that q_ℓ and q_r are known and equal, and that* (H1) *holds. Then, for any sequence of equilibria $\left(C_{\ell,n}^*, C_{r,n}^*\right)$:*

(i) $\lambda < \frac{1}{2}$ *implies* $\lim_n C_{\ell,n}^* = \lim_n C_{r,n}^* = 0$.

(ii) $\lambda = \frac{1}{2}$ *implies* $\lim_n C_{\ell,n}^* = \lim_n C_{r,n}^* = \infty$.

The first part of Proposition 6 shows that when $\lambda < 1/2$, the cutoff points converge to 0 as n tends to ∞, provided that other potential asymmetries do not offset the size advantage of the majority. Hence, the impossibility theorem of Palfrey and Rosenthal (1985) essentially survives when q is known.

The proof of the first part of the proposition builds upon two main observations. First, as I noted in the previous section, when $\Pr\{X_h = -1 \mid q\}$ and $\Pr\{X_h = 1 \mid q\}$ are distant, the pivot probabilities $P_\ell(q)$ and $P_r(q)$ decrease at an exponential rate with n, while Π_i increases only linearly with n. Therefore, when q is known, equilibria with positive cutoff points can be sustained asymptotically only if the equilibrium values of $\Pr\{X_h = -1 \mid q\}$ and $\Pr\{X_h = 1 \mid q\}$ get arbitrarily close to each other as n tends to ∞.

The second point is that if, in equilibrium, $\Pr\{X_h = -1 \mid q\}$ and $\Pr\{X_h = 1 \mid q\}$ are close to each other, the implied pivot probabilities for the two types of agents must also be close, and hence, the cutoff points C_ℓ^* and C_r^* must be close as well. But, assuming $q_\ell = q_r$ and $\lambda < 1/2$, if C_ℓ^* and C_r^* are substantially large numbers that are close to each other, the continuity of F implies that $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell F(C_\ell^*)$ is significantly smaller than $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r F(C_r^*)$, which is a contradiction.

On the other hand, when $q_\ell = q_r$ and $\lambda = \frac{1}{2}$, in equilibrium we have $\Pr\{X_h = -1 \mid q\} = \Pr\{X_h = 1 \mid q\}$. Therefore, in this case, the implied pivot probabilities are asymptotically proportional to $1/\sqrt{n}$, and the cutoff points become arbitrarily large as n tends to ∞.

In Appendices C and D, I will provide some generalizations of Proposition 6 which show that the conclusion of this result is quite robust. It should also be noted that if we were to assume that the agents vote for the two candidates with equal probabilities in an iid fashion, the law of large numbers would virtually rule out significant vote differentials in a large election. (More on this in Online Appendix C.) By contrast, when q is uncertain in my model, not only is the majority more likely to win, but we might even observe extreme vote differentials at asymmetric realizations of q. Next, I will discuss some extensions of the basic model that I analyzed in Section 3.

5. Extensions of the Positive Results

As I discussed in Introduction, an important issue in earlier models on ethical voters is the role of homogeneity assumptions that imply the existence of large groups of agents with similar ethical judgments. (More on this in Online Appendix A.) So far, in this paper I have utilized similar homogeneity assumptions. However, my approach is conceptually independent from these assumptions, for in my model, each agent takes as given the behavior of others in a game-theoretic fashion. For example, one of my homogeneity assumptions (which might be especially questionable) is that each agent can be characterized by the same known level of u. This corresponds to a situation in which all agents agree about how strongly their favored candidates would contribute to the welfare of the society. In Appendix A, I will show that we can indeed dispense with this assumption. Specifically, if for each agent, u is a (privately known) random draw from a uniform distribution, then the conclusions of Propositions 1-4 and a suitable modification of Proposition 5 continue to hold.

When analyzing this example, I transform voting costs in a way that enables me to utilize the machinery that I have developed in Section 3. The transformed cost distribution is not uniform and does not have a bounded support. To cover such cases, in Appendix E, when proving my results I will replace (H1) and (H5) with more general assumptions. Specifically, I will show that except for the comparative statics exercise with respect to c, the conclusions of Propositions 1-6 remain true if:

(i) The support of F is a subinterval of $\mathbb{R}_+$ that contains 0;

(ii) F is continuously differentiable, strictly increasing and concave on its support;

(iii) $F(\gamma C)/F(C)$ is a nonincreasing function of $C \in \mathbb{R}_{++}$ for every fixed $\gamma \geq 1$.

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, (H5) implies properties (i)-(iii). It should also be noted that in the main body of the paper, instead of property (i) above I have chosen to use (H1) simply because this allows me to assume $D \geq c$ so that interpretations (I) and (II) predict the same behavior.

Within interpretation (II), all of my results continue to hold when $0 < D < c$ (although the implied turnout will be lower than that implied by interpretation (I)). To see the reason, let C be the voting cost of an altruistic agent of type i. Then, according to interpretation (II), the agent would vote for her favored candidate if and only if $C \leq \min\{\Pi_i, D\}$. Thus, the corresponding notion of equilibrium requires a pair of cutoff points C_ℓ^*, C_r^* such that $\Pi_i^* = C_i^*$, $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell F(\min\{C_\ell^*, D\})$ and $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r F(\min\{C_r^*, D\})$. In turn, if we set $\widetilde{F}(C) \equiv F(\min\{C, D\})$ for $C \in \mathbb{R}_+$, we can rewrite these equations as $\Pi_i^* = C_i^*$, $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell \widetilde{F}(C_\ell^*)$ and $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r \widetilde{F}(C_r^*)$. Notice that such a pair (C_ℓ^*, C_r^*) is simply the equilibrium of a model as in Section 2, the only difference being that $\widetilde{F}$ takes the role of F. Moreover, $\widetilde{F}$ possesses all relevant properties of a distribution function with support $[0, D]$ (although we have $\widetilde{F}(D) < 1$). In particular, $\widetilde{F}$ inherits the properties (i)-(iii) above whenever F satisfies these properties. Consequently, the implied comparative statics remain the same except that we should replace the parameter c in Proposition 5 with D.

By changing the objective function of an altruistic agent, we can obtain various modifications of my basic model. Such modifications would predict positive asymptotic turnout so long as the agent's objective function increases linearly with n. For example, an altruistic agent may care only about those agents who prefer the same candidate as her. A simple version of such a situation can be modeled by restricting the domain of the summation operator in (1) to the set $\{h : h \text{ is of type } i,\ h \neq h'\}$. This model can be analyzed by multiplying the right side of equations (5) with λ_i.[11]

The implications of introducing ineligible agents to my model would be akin to those of increasing the altruism parameter ψ. In fact, if we denote by η the fraction of eligible agents, the corresponding modification of the model can be analyzed by replacing $u\psi$ with $u\psi/\eta$.

Finally, it should be noted that my findings remain true when the size of the electorate is uncertain (see Appendix C).

[11] We can also envision non-altruistic voters who enjoy influencing others' payoffs. For example, a conservative voter might enjoy forcing policies upon liberals. (I am grateful to a referee for calling my attention to this point.)

6. Relations to Rule Utilitarian Voter Model of Feddersen and Sandroni

As I noted in Introduction, the ethical voter model of Feddersen and Sandroni (2006a, 2006b) is based on the notion of a rule, which refers to a type contingent behavior that applies to all types in a given group. In turn, their equilibrium notion (i.e., consistency requirement) formalizes the following scenario: A given agent contemplates the social consequences of various rules that her group could follow in principle (given the behavior of the types outside her group). Thereby, the agent identifies a socially optimal rule that maximizes her social preferences. She believes that this optimal rule is the ethical rule that she and other types in her group should follow.

This notion of ethical behavior is aptly called (group-based) *rule utilitarianism*. This is quite distinct from the notion of ethical behavior that I utilize in this paper, for interpretation (II) of my model builds upon the assumption that when maximizing her social preferences, each agent takes as given the behavior of everyone else (regardless of their types). Yet, qualitatively, my comparative statics exercises (Propositions 3-5) lead to the same conclusions as those of Feddersen and Sandroni (2006a). In Online Appendix A, I will prove that the relation between the two models is much deeper. In fact, the predictions of the two models are *cardinally* equivalent for a suitable specification of social preferences in Feddersen-Sandroni model.

A closer look into this rule utilitarian model is in order to understand the content of the equivalence result. In the general version of their model, Feddersen and Sandroni (2006b) assume that the social preferences of an ethical agent can be represented with a function of the following form:

$$wp - \vartheta(\phi). \tag{9}$$

Here, p is the probability that the agent's favored candidate wins the election, ϕ is the expected per capita cost of voting, ϑ is an increasing function, and w is a parameter that the authors refer to as "the importance of the election." In a special version of this model, Feddersen and Sandroni (2006a) focus on the case $\vartheta(\phi) \equiv \phi$. If, in addition, w corresponds to the agent's estimation of the per capita material benefit associated with the policies of her favored candidate, then (9) reduces to expected per capita payoff (from the perspective of the agent in question). This special form of (9) is simply the objective function of a purely altruistic agent in the sense of my model, who places the same weight to her own payoff as that of others. While this observation clarifies the conceptual connection between the two models, formally the equivalence result reads as follows:

If we set $\vartheta(\phi) \equiv \phi$ and $w \equiv u\psi$, the first order conditions of the rule utilitarian voter model coincides with equations (5). So, in this case, when the first order conditions in the *rule utilitarian voter model are sufficient, the two models predict exactly the same turnout rates.*[12]

In light of this equivalence result, Propositions 3-5 are simply extensions of the comparative statics exercises of Feddersen and Sandroni (2006a). My comparative statics exercises are more general because Feddersen and Sandroni (2006a) assume that the fractions of the two types of ethical agents are independent uniform random variables. The cost distributions that I discussed in Section 5 are also outside the scope of Feddersen and Sandroni (2006a).

More importantly, the driving forces behind the comparative statics of the two models are quite distinct. In my model, what drives the results is the behavior of individual pivot probabilities. In particular, two related phenomena that I have uncovered here, namely, the level and ratio effects, explain most of the comparative statics. Moreover, Bayesian updating of agents' beliefs underlie the ratio effect and the underdog effect. By contrast, in Feddersen and Sandroni (2006a), a single agent can never be pivotal,[13] and the agents do not update their beliefs according to their types.

A further implication of Bayesian updating is that, in the present model, the prior probability of winning for candidate i tends to be smaller than the posterior probability that an altruistic agent of type i places on this event. Hence, when there is no overwhelming majority, altruistic agents of *both* types may believe that their favored candidate is more likely to win the election.[14] This is consistent with 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.

7. Concluding Remarks

In the literature on voter participation, a prominent theory explains turnout in large elections with elites' ability to mobilize large groups of voters (e.g., Uhlaner, 1989; Morton, 1991; Shachar and Nalebuff, 1999; Herrera and Martinelli, 2006). In line with this approach, empirical evidence points to a positive relation between turnout and political leaders' efforts

[12] This is the content of Proposition O1 in Online Appendix A. The conclusion of this equivalence result is very robust. In particular, it remains valid if $0 < D < c$ and if we adopt interpretation (II) of my model. I have also verified that the same conclusion obtains when λ is a random variable that takes finitely many values and q is a continuous random variable, or vice versa. (The proofs are available upon request.)

[13] Instead, group-wise pivot probabilities play a comparable role in their model.

[14] For example, if q is uniform on $[0,1]^2$ 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.

(Shachar and Nalebuff, 1999). On the other hand, as Feddersen (2004) points out, if voting is costly, it is not so clear how elites persuade large groups of agents to vote. In particular, explaining turnout with large-scale, external reward/punishment mechanisms seems problematic in several respects.

Following a rule utilitarian approach, earlier models on ethical voters also focus on large groups of agents.[15] Unlike elite-based models, this literature does not require an external device to coordinate agents' behavior. Rather, agents' sense of group membership and ethical concerns give rise to equilibria in which the agents behave *as if* there are a few leaders each controlling a large fraction of agents.[16] However, such equilibria require a certain level of homogeneity in agents' characteristics such as their perception of the importance of the election, the strength of their ethical values, or their beliefs about the distribution of voting costs. (More on this in Online Appendix A.)

In this paper, I proposed a game-theoretic approach that separates the ethical-voter idea from group-based approaches. The basic model that I have studied in Section 3 utilizes the analogues of the homogeneity assumptions of Feddersen and Sandroni (2006a). The two models make the same predictions up to potential differences between the respective parameters that measure the importance of changing the winner (namely, $u\psi$ and w). Moreover, as I demonstrate with an example in Appendix A, my approach is conceptually independent from homogeneity of agents' characteristics.

It is important to note that suitable extensions of the present model may allow us to relate turnout to elites' efforts. For example, we can think of an extension in which candidates' campaign efforts cause horizontal shifts between the masses of the supporters of the two candidates and vertical shifts in the distribution of social preferences among the supporters of a given candidate (by changing supporters' perception of u). Thereby, it may be possible to obtain a model of mobilization in a voluntary-participation/costly-voting framework. I leave it as an open question to determine if the aforementioned empirical evidence can be explained along these lines.

In my model, uncertainty in the fractions of altruistic agents smoothens the behavior of pivot probabilities. The implied pivot probabilities are inversely proportional to the size of the electorate. Thus, if we were to identify the policies of the two candidates with two points on the real line, the expected effect of a single vote on the policy outcome would also be inversely proportional to the size of the electorate. Faravelli and Walsh (2011) have recently shown that, even with iid voters, the effect of a single vote can be of the same order if candidates respond smoothly to their margin of victory. More specifically, in their

[15]When setting up their model, Feddersen and Sandroni (2006b) allow for small groups, but they do not study the equilibria of their model for such group structures.

[16]Proposition 3 of Feddersen and Sandroni (2006b) formalizes this point.

model, a single vote has always some effect on the policy outcome, and the implied change in the outcome is inversely proportional to the size of the electorate. A comparison of the relative merits of the two models might be an interesting task for future research.

A few final remarks are in order. My model is not meant to provide a precise picture of the mental process through which altruism motivates voting or a complete list of the motivations of real voters. For instance, citizens may vote to *express* their ethical concerns (Feddersen et al., 2009), because of a sense of civic duty (Riker and Ordeshook, 1968; Blais, 2000), or because they fail to behave in a perfectly optimal/rational manner (Levine and Palfrey, 2007). As it is based on the notion of pivotality, my approach has relatively closer ties with the classical pivotal-voter model, which remains as a fundamental tool for political economists outside the realm of large, costly elections. Hence, I hope that this paper may help establishing a closer connection between our understanding of small or costless elections and that of large, costly elections.

Appendix

A. Heterogeneity in Agents' Valuations

In this appendix, I assume that for each agent, $u \geq 0$ is a random draw from a continuous distribution. Agents have private knowledge of their value of u. Moreover, the distribution of u is independent across agents (and from any other random variable in the model). This implies an extreme form of heterogeneity: any agent of a given type knows that almost surely there is no other agent with whom she can precisely agree about the value of u.

In this modified model, the value of u for a particular agent influences the rate of increase of her objective function as described in expression (3). Let us relable this expression as $\Pi_{u,i}$. Since the distribution of u is independent across agents, any two altruistic agents of type i face the same pivot probability $\mathcal{P}_i$. Hence, with

$$\Pi_i \equiv (1 + \psi n)\,\mathcal{P}_i, \tag{A-1}$$

we have $\Pi_{u,i} = u\Pi_i$.

Let us denote with $\Pi^*_{u,i}$ and Π^*_i the equilibrium values of $\Pi_{u,i}$ and Π_i, respectively. As in Section 2, in equilibrium, an altruistic agent of type i with a given u must vote with probability $F(\Pi^*_{u,i}) = F(u\Pi^*_i)$. This implies that $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell \mathbf{E} F(u\Pi^*_\ell)$ and $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r \mathbf{E} F(u\Pi^*_r)$ for a randomly chosen agent h. (Here, the expectation operator $\mathbf{E}$ is applied with respect to u, and as before, X_h specifies agent h's behavior as a function of her uncertain characteristics.) Thus, we seek a pair of cutoff points C^*_ℓ, C^*_r such that $C^*_i = \Pi^*_i$ for $i = \ell, r$, where Π^*_i is the value of expression (A-1) induced by the conditional voting probabilities $\Pr\{X_h = -1 \mid q\} = \lambda q_\ell \mathbf{E}F(uC_\ell^*)$ and $\Pr\{X_h = 1 \mid q\} = (1-\lambda) q_r \mathbf{E}F(uC_r^*)$.

Note that if we set $\widetilde{F}(C) \equiv \mathbf{E}F(uC)$ for $C \in \mathbb{R}_+$, these cutoff points C_ℓ^*, C_r^* coincide with the equilibrium of a *dual model* with $u \equiv 1$ and cost distribution $\widetilde{F}$. The two models are also equivalent in terms of the implied expected turnout and margin of victory.

Remark A1. The equilibrium of the model with random u that I described above is consistent with both interpretations (I) and (II) (provided that F has a bounded support and D exceeds the maximum possible voting cost). However, typically, the support of $\widetilde{F}$ will not be bounded. Thus, in the dual model, we should focus on interpretation (I) so that $[0, \Pi_i^*)$ defines the participation region for altruistic agents of type i. This causes no loss of generality, for the dual model is only a tool to solve the original model with random u.

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 $\omega \equiv \frac{c}{\mathfrak{u}}$,

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

Thus, for every $\gamma \ge 1$, we have

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

It is easily verified that the function $\widetilde{F}(\gamma C)/\widetilde{F}(C)$ is nonincreasing in $C \in \mathbb{R}_{++}$.

In turn, the density of $\widetilde{F}$ is as follows:

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

Note that $\widetilde{f}$ is continuous and nonincreasing. In particular, $\widetilde{F}$ is concave. To summarize, $\widetilde{F}$ satisfies the properties (i)-(iii) in Section 5.

From my analysis in Appendix E, it follows that, under the assumptions (H2)-(H4), the conclusions of Propositions 1-4 hold for the dual model without any modifications. Moreover, the expected turnout rate is increasing with ψ and $\frac{\mathfrak{u}}{c}$, while the expected margin of victory is nondecreasing with ψ and $\frac{\mathfrak{u}}{c}$. (See, in particular, Remark E1 in Appendix E.)

B. Examples of Compatible g

Let us denote by $\widetilde{g}$ the density of a bivariate normal random variable $\widetilde{q} = (\widetilde{q}_\ell, \widetilde{q}_r)$ with identical marginal distributions. That is, $\widetilde{g}(\widetilde{q}) \equiv \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} e^{-Q(\widetilde{q})/2}$ for $\widetilde{q} \in \mathbb{R}^2$, where $Q(\widetilde{q}) \equiv \frac{1}{1-\rho^2}\left(\left(\frac{\widetilde{q}_\ell-\mu}{\sigma}\right)^2 - 2\rho\left(\frac{\widetilde{q}_\ell-\mu}{\sigma}\right)\left(\frac{\widetilde{q}_r-\mu}{\sigma}\right) + \left(\frac{\widetilde{q}_r-\mu}{\sigma}\right)^2\right)$. Here, μ (resp. σ) is the common mean (resp. standard deviation) of the components of $\widetilde{q}$, and ρ is the correlation coefficient. Then, the density of $\widetilde{q}$ conditioned to $[0,1]^2$ satisfies the assumptions (H3) and (H4). This conditional density has the form $g(q) = \widetilde{g}(q)/K$ for $q \in [0,1]^2$, where $K > 0$ is the probability that $\widetilde{q}$ belongs to $[0,1]^2$.

As I noted in text, another important case is the class of beta distributions. If q_i has a beta distribution, its density has the form $K\,(q_i)^{\alpha-1}(1-q_i)^{\beta-1}$ for $0 \leq q_i \leq 1$, where α and β are nonnegative parameters, and $K > 0$ is a normalizing constant (that depends on α and β). A beta density is *unimodal* if $\alpha, \beta > 1$, and *monotone* if $\alpha \leq 1 \leq \beta$ or $\alpha \geq 1 \geq \beta$. When $\alpha = \beta = 1$, we obtain the uniform distribution. If q_ℓ and q_r have independent beta distributions, their joint density takes the form $g(q) = K^2\,(q_\ell q_r)^{\alpha-1}\,((1-q_\ell)(1-q_r))^{\beta-1}$. In this case, g satisfies (H3) and (H4) provided that $\alpha \geq 1$ and $\beta \geq 1$. In particular, unimodal densities and monotone densities with $\alpha = 1 \leq \beta$ or $\alpha \geq 1 = \beta$ are compatible with (H3) and (H4).

C. Population Uncertainty

This appendix demonstrates how the conclusions of Proposition 1-6 can be reproduced under population uncertainty. Suppose that n is a random, positive integer that is stochastically independent from every other random variable in the model. Also assume that for any realization of n, the distribution of agents' characteristics is as in Section 2.

In this framework, expression (3) depends on the realization of n, but its interpretation remains the same. Let us relable this expression as $\Pi_i(n)$. When the expected value of n is finite, $\Pi_i(n)$ would also have a finite expectation. Thus, in this case, we can modify our notion of equilibrium in an obvious way and show that such an equilibrium exists and is positive (as in Proposition 1).

Consider a sequence of probability distributions Q_m for n. Now, if $\lim_m Q_m\{n \leq b\} = 0$ for every positive integer b, large values of m would correspond to unambiguously large elections. In this case, it can be shown that the conclusions of Propositions 2-6 would also hold as $m \to \infty$.

Here, the key observation is that when n is stochastically independent from agents' characteristics, the cutoff points and the realization of n would determine $\mathcal{P}_i$ uniquely, irrespective of the distribution of n. Thus, we can apply Lemma E1 in Appendix E uniformly in m. Thereby, we can prove that equations (5) continue to hold. Similarly, $P_i(q)$ would exhibit a knife-edge behavior for large values of m, implying the conclusion of Proposition 6.[17] (The details are available upon request.)

D. More on the Case of Known q

The following result is a more general version of Proposition 6 that dispenses with the assumption that $q_\ell = q_r$. This result also allows u, ψ and F be type dependent. Moreover, we see that even if voting costs are bounded away from 0, low turnout rates are guaranteed. In other words, the assumption that 0 belongs to the support of voting costs, as demanded by (H1), serves only to conclude that the cutoff points are close to 0.

Proposition (General Impossibility Result). *Let u_ℓ and u_r be type dependent analogues of the parameter u, and similarly for ψ_ℓ, ψ_r, F_ℓ and F_r. Suppose that q_ℓ and q_r are known, F_ℓ and F_r are continuous distributions on $\mathbb{R}$ with $F_\ell(0) = F_r(0) = 0$, and there exists a type i such that:* (1) *F_i is weakly less than F_j in the sense of first order stochastic dominance;* (2) *$u_i\psi_i \geq u_j\psi_j$; and* (3) *$\lambda_i q_i > \lambda_j q_j$. Then, in the corresponding modification of the model in Section 2, we have $\lim_n F_\ell(C^*_{\ell,n}) = \lim_n F_r(C^*_{r,n}) = 0$ along any sequence of equilibria.*

For brevity, I omit the proof of this proposition, which is similar to the proof of Proposition 6 in Appendix E below.

E. Proofs

For convenience, the order proofs and the format of the assumptions will be different than the results and the assumptions appear in text. I will state an assumption right before proving a result which demands that particular assumption. After introducing an assumption, without further mention, throughout the remainder of the appendix I will assume that the property in question holds.

E.1 Preliminaries

Definition E1. $a_n \simeq b_n$ means $\lim_n \frac{a_n}{b_n} = 1$ for sequences of positive real numbers (a_n), (b_n).

[17] In line with this argument, Myerson (2000) has shown that in Poisson voting games, unconditional pivot probabilities (found by averaging over n) also decline exponentially with the expected size of the electorate, whenever the expected vote shares of the two candidates are different. In particular, it follows that in the Poisson voting game of Jankowski (2007), relaxing the assumption of deterministic costs renders large-scale turnout impossible unless the fractions of the two types of agents are the same.

Set $\mathfrak{p}_\ell(q) \equiv \Pr\{X_h = -1 \mid q\}$ and $\mathfrak{p}_r(q) \equiv \Pr\{X_h = 1 \mid q\}$ for a randomly selected agent h and any $q \in [0,1]^2$. The following expression gives the conditional probability of the event that candidate i is one behind or tie occurs (excluding any given agent):

$$\begin{aligned} piv_i(q) \equiv & \sum_{b=0}^{\lfloor n/2 \rfloor} \tfrac{n!}{(n-2b)!b!b!} \mathfrak{p}_\ell(q)^b \mathfrak{p}_r(q)^b (1 - \mathfrak{p}_\ell(q) - \mathfrak{p}_r(q))^{n-2b} \\ & + \sum_{b=0}^{\lfloor (n-1)/2 \rfloor} \tfrac{n!}{(n-2b-1)!b!(b+1)!} \mathfrak{p}_i(q)^b \mathfrak{p}_j(q)^{b+1} \left(1 - \mathfrak{p}_\ell(q) - \mathfrak{p}_r(q)\right)^{n-2b-1}. \end{aligned} \tag{E-1}$$

Here, $\lfloor \omega \rfloor$ stands for the largest integer less than or equal to a number ω. Similarly, I will denote by $\lceil \omega \rceil$ the smallest integer greater than or equal to ω. On occasion, I will consider specific functional forms for $\mathfrak{p}_\ell(q)$ and $\mathfrak{p}_r(q)$. In such cases, instead of $piv_i(q)$ I will write $piv_i(\mathfrak{a}, \mathfrak{b}, n)$, where $\mathfrak{a}$ and $\mathfrak{b}$ stand for the functional forms of $\mathfrak{p}_\ell(q)$ and $\mathfrak{p}_r(q)$, respectively.

The tie breaking rule implies that

$$P_i(q) = \frac{1}{2} piv_i(q) \quad \text{for any } q \in [0,1]^2. \tag{E-2}$$

Therefore, in what follows, I examine the asymptotic behavior of $piv_i(q)$.

The first point to note is that, by a central limit theorem, $\mathfrak{p}_\ell(q) = \mathfrak{p}_r(q) > 0$ implies $piv_i(q) \simeq \frac{1}{\sqrt{\pi n \mathfrak{p}}}$, where $\mathfrak{p}$ stands for the common value of $\mathfrak{p}_\ell(q)$ and $\mathfrak{p}_r(q)$ (see Feller, 1966, p. 90). On the other hand, as I will show momentarily, when $\mathfrak{p}_\ell(q)$ and $\mathfrak{p}_r(q)$ are distinct, $piv_i(q)$ converges to 0 at an exponential rate. In fact, this observation is a simple consequence of the following classical theorem, which shows that the rate of convergence in the law of large numbers is exponential.

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 first part of the above result is a straightforward modification of the statement of Theorem 2 of Hoeffding (1963). In turn, the second part can easily be derived from the first part.

The implied bounds on conditional pivot probabilities read as follows.

Corollary E1. *Suppose that for some $q \in [0,1]^2$ and $n \in \mathbb{N}$, we have $n \left| \mathfrak{p}_\ell(q) - \mathfrak{p}_r(q) \right| \geq 1$.*

Set $\delta \equiv |\mathfrak{p}_\ell(q) - \mathfrak{p}_r(q)|$ *and let* i *be such that* $\mathfrak{p}_i(q) > \mathfrak{p}_j(q)$*. Then*

$$piv_i\,(q) \leq e^{-n\delta^2/2} \quad \text{and} \quad piv_j\,(q) \leq e^{-(n\delta-1)^2/2n}.$$

Proof. Set $S^- \equiv \sum_{h=1}^{n} X_h$, so that $piv_\ell(q) = \Pr\{S^- \in \{0,1\}\}$ and $piv_r(q) = \Pr\{S^- \in \{0,-1\}\}$, where the probability operator refers to conditional probabilities at the given q. Assume first $\mathfrak{p}_r(q) > \mathfrak{p}_\ell(q)$. Note that in this case, $\mathbf{E}S^- = n\delta$. Thus,

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

Here, the last inequalities in both lines follow from part (ii) of Hoeffding inequality with $b_h \equiv -1$ and $d_h \equiv 1$ $(h = 1, ..., n)$. Similarly, when $\mathfrak{p}_r(q) < \mathfrak{p}_\ell(q)$, we have $\mathbf{E}S^- = -n\delta$. In this case, the desired conclusion follows from part (i) of Hoeffding inequality. $\square$

The next corollary is another routine application of Hoeffding inequality, which provides bounds for binomial tail probabilities. (I omit the proof.)

Corollary E2. *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* $\beta \leq \varrho$*,* $\Pr\{\mathbb{S} \leq \beta n\} \leq e^{-2(\beta-\varrho)^2 n}$*.*

E.2 Proofs of Propositions 1 and 6

Proposition 1 only demands the following property.

Assumption 1. F is a continuous distribution on $\mathbb{R}_+$ with $F(0) = 0$.

Proof of Proposition 1. Set $U \equiv u\,(1 + \psi n)\,/2$ and $\Omega \equiv [0,1]^2 \times [0, U]^2$. Let us denote a generic element of Ω by $(q, \mathbf{C}) \equiv (q_\ell, q_r, C_\ell, C_r)$. Define a function $\mathfrak{P} : \Omega \to \mathbb{R}$ as $\mathfrak{P}(q, \mathbf{C}) \equiv piv_i(\lambda q_\ell F(C_\ell), (1-\lambda) q_r F(C_r), n)$. Since F is continuous, the function $\mathfrak{P}$ is continuous on Ω. In fact, since Ω is compact, $\mathfrak{P}$ must be uniformly continuous. This, in turn, implies that for any i, the map $\mathbf{C} \to \Pi_i(\mathbf{C}) \equiv U \int_{[0,1]^2} \mathfrak{P}(q, \mathbf{C})\, dG^i(q)$ 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. Since $F(0) = 0$, this implies that type i agents abstain with probability 1 (excluding a given agent). Therefore, the election would be tied if all agents are of type i. The probability of this event equals $(\lambda_i)^n$. It follows that $\mathcal{P}_i \geq \frac{1}{2}(\lambda_i)^n > 0$. Thus, we must have $\Pi_i^* > 0$, a contradiction. $\square$

Proposition 6 demands the following additional property.

Assumption 2. 0 belongs to the support of F.

Proof of Proposition 6. For the given value of q, let us relabel expression (E-1) as $piv_{i,n}$. If we denote by S_i^- the number of votes for candidate i excluding a given agent, from definitions it follows that

$$\frac{piv_{r,n_k}}{piv_{\ell,n_k}} \equiv \frac{\Pr\left\{S_\ell^- - S_r^- = 0\right\} + \Pr\left\{S_r^- - S_\ell^- = -1\right\}}{\Pr\left\{S_\ell^- - S_r^- = 0\right\} + \Pr\left\{S_\ell^- - S_r^- = -1\right\}}. \tag{E-3}$$

Set $\mathfrak{p}_{\ell,n}^* \equiv \lambda q_\ell F(C_{\ell,n}^*)$ and $\mathfrak{p}_{r,n}^* \equiv (1-\lambda) q_r F(C_{r,n}^*)$. It is easy to see that (in equilibrium) we have

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

Therefore,

$$\Pr\left\{S_r^- - S_\ell^- = -1\right\} = \frac{\mathfrak{p}_{\ell,n}^*}{\mathfrak{p}_{r,n}^*} \Pr\left\{S_\ell^- - S_r^- = -1\right\}, \tag{E-4}$$

whenever $\mathfrak{p}_{r,n}^*$ is positive.

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

$$\limsup_k \frac{\mathfrak{p}_{\ell,n_k}^*}{\mathfrak{p}_{r,n_k}^*} < 1, \tag{E-5}$$

for any subsequence $\mathfrak{p}_{r,n_k}^*$ that is bounded away from 0. Assume by contradiction that $\lim_k \frac{\mathfrak{p}_{\ell,n_k}^*}{\mathfrak{p}_{r,n_k}^*} \geq 1$ for a subsequence $\mathfrak{p}_{r,n_k}^*$ that is bounded away from 0. Then, (E-3), (E-4) and the definition of equilibrium imply that

$$\liminf_k \frac{C_{r,n_k}^*}{C_{\ell,n_k}^*} = \liminf_k \frac{piv_{r,n_k}}{piv_{\ell,n_k}} \geq 1. \tag{E-6}$$

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-6) 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 $F(\infty) \equiv 1$. But then, $\lambda < 1/2$ and $q_\ell = q_r$ imply $\lim_k \mathfrak{p}_{r,n_k}^* = (1-\lambda) q_r \lim_k F(C_{r,n_k}^*) > \lambda q_\ell \lim_k F(C_{\ell,n_k}^*) = \lim_k \mathfrak{p}_{\ell,n_k}^*$, for $\mathfrak{p}_{r,n_k}^*$ is bounded away from 0 so that $\lim_k F(C_{r,n_k}^*) > 0$. This contradicts the supposition that $\lim_k \frac{\mathfrak{p}_{\ell,n_k}^*}{\mathfrak{p}_{r,n_k}^*} \geq 1$ and proves (E-5).

To complete the proof of part (i), suppose that for some i, there is a subsequence C_{i,n_k}^*

that is bounded away from 0. Then, $\mathfrak{p}^*_{i,n_k}$ is also bounded away from 0 since 0 belongs to the support of F. Moreover, given that C^*_{i,n_k} is bounded away from 0, Corollary E1 implies $\lim_k(\mathfrak{p}^*_{\ell,n_k} - \mathfrak{p}^*_{r,n_k}) = 0$, for Π_i increases only linearly with n. It follows that $\mathfrak{p}^*_{\ell,n_k}$ and $\mathfrak{p}^*_{r,n_k}$ are both bounded away from 0, and that $\lim_k \frac{\mathfrak{p}^*_{\ell,n_k}}{\mathfrak{p}^*_{r,n_k}} = 1$. This contradicts (E-5), as we sought.

To prove part (ii), I will first show that $\mathfrak{p}^*_{\ell,n} = \mathfrak{p}^*_{r,n}$ for every $n \in \mathbb{N}$. Suppose by contradiction $\mathfrak{p}^*_{i,n} > \mathfrak{p}^*_{j,n}$ for some i and n. Then, $\Pr\{S^-_i - S^-_j = -1\} \leq \Pr\{S^-_j - S^-_i = -1\}$, and hence, $piv_{i,n} \leq piv_{j,n}$. This, in turn, implies that $C^*_{i,n} \leq C^*_{j,n}$. Since $q_\ell = q_r$ and $\lambda = \frac{1}{2}$, it then follows that $\mathfrak{p}^*_{i,n} \leq \mathfrak{p}^*_{j,n}$, a contradiction.

Börgers (2004, Remark 1) shows that, for a fixed n, if a randomly chosen agent votes for the two candidates with the same probability $\mathfrak{p}$, then pivot probabilities decrease with $\mathfrak{p}$. Since $\mathfrak{p}^*_{\ell,n} = \mathfrak{p}^*_{r,n}$, it follows that $C^*_{i,n} \geq u\,(1+\psi n)\,\frac{1}{2}piv_n$ for every n and i, where piv_n stands for the value of expression (E-1) at $\mathfrak{p}_\ell(q) = \mathfrak{p}_r(q) = 1/2$. Moreover, as I noted earlier, we have $piv_n \simeq 1/\sqrt{\pi n \frac{1}{2}}$. Thereby, we obtain the desired the conclusion: $\lim_n C^*_{i,n} = \infty$ for $i = \ell, r$. $\square$

E.3 On the Magnitude of $\mathcal{P}_i$ when q Is Unknown

Good and Mayer (1975) have shown that $\lim_n n \int_0^1 piv_i\,(\mathfrak{p}, 1-\mathfrak{p}, n)\,\upsilon\,(\mathfrak{p})\;d\mathfrak{p} = \upsilon\,(1/2)$ for any density υ on $[0,1]$ that is continuous at $1/2$. Later, Chamberlain and Rothschild (1981) proved the same result independently.

In my model, randomness of q creates an analogous environment: given any possible q, in equilibrium, a randomly chosen agent votes for candidates ℓ and r with probabilities $\lambda q_\ell F(C^*_\ell)$ and $(1-\lambda)q_r F(C^*_r)$, respectively. The next lemma is an extension of Good-Mayer formula that corresponds to this scenario.

Lemma E1. *Let ν be a continuous (but not necessarily positive) density on $[0,1]^2$. Fix a pair of positive numbers (T_ℓ, T_r) with $T_\ell + T_r \leq 1$. Then, for any type i,*

$$\lim_n n \int_0^1 \int_0^1 piv_i(q_\ell T_\ell, q_r T_r, n)\nu(q_\ell, q_r)\, dq_\ell dq_r = 2\int_0^{\frac{1}{\max\{T_\ell, T_r\}}} \nu(\theta T_r, \theta T_\ell)\, d\theta. \tag{E-7}$$

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.

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

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

onto $(0,1)^2$. (The inverse of W is defined by $W^{-1}(q_\ell, q_r) \equiv (T_\ell q_\ell + T_r q_r, \frac{T_\ell q_\ell}{T_\ell q_\ell + T_r q_r}) = (t, \mathfrak{p})$.) Moreover, W is continuously differentiable, and $J \equiv \begin{bmatrix} \mathfrak{p}/T_\ell & t/T_\ell \\ (1-\mathfrak{p})/T_r & -t/T_r \end{bmatrix}$ 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 $t \in (0, T_\ell + T_r)$,

$$\Upsilon_{t,n} \equiv n \frac{t}{T_\ell T_r} \int_{I_t} piv_i\,(t\mathfrak{p}, t(1-\mathfrak{p}), n)\ \nu\left(\frac{t\mathfrak{p}}{T_\ell}, \frac{t(1-\mathfrak{p})}{T_r}\right) d\mathfrak{p}, \tag{E-8}$$

and $I_t \equiv \left(\max\left\{0, 1 - \frac{T_r}{t}\right\}, \min\left\{1, \frac{T_\ell}{t}\right\}\right)$ (see Billingsley, 1995, Theorem 17.2, p. 225). Note that the interval I_t is nondegenerate, because $0 < t < T_\ell + T_r$ implies $\frac{T_\ell}{t} > 1 - \frac{T_r}{t}$.

Pick any $\beta > 0$. First, I 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).[18] Let S_ℓ^-, S_r^- be as in the proof of Proposition 6, so that total turnout equals $S_\ell^- + S_r^-$ (excluding a given agent). Denote by m a possible value of $S_\ell^- + S_r^-$. Let $\mathcal{B}(\cdot; n, t)$ be the binomial probability distribution with population size n and success probability t. For a fixed $(t, \mathfrak{p}) \in V$, let us suppose that a randomly chosen agent votes for candidates ℓ and r with probabilities $t\mathfrak{p}$ and $t(1-\mathfrak{p})$, respectively. Then, we would have $\Pr\{S_\ell^- + S_r^- = m\} = \mathcal{B}(m; n, t)$ for every nonnegative integer m and positive integer n. Moreover, among those who participate, a randomly chosen agent would vote for candidates ℓ and r with probabilities $\mathfrak{p}$ and $1-\mathfrak{p}$, respectively. Thus, $piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)$ would give us the probability of the event that the election is tied or candidate i is 1 behind conditional on the event $S_\ell^- + S_r^- = m$ for any $m = 0, 1, \ldots$, where $piv_i(\cdot, \cdot, 0) \equiv 1$. Hence, for every $n \in \mathbb{N}$ and $(t, \mathfrak{p}) \in V$, we have

$$piv_i(t\mathfrak{p}, t(1-\mathfrak{p}), n) = \sum_{m=0}^{n} \mathcal{B}(m; n, t) piv_i(\mathfrak{p}, 1-\mathfrak{p}, m). \tag{E-9}$$

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

[18] Over a region of integration, if the integrand is not explicitly defined, I assume that it equals zero.

$\int_0^1 piv_r\left(\mathfrak{p},1-\mathfrak{p},m\right)d\mathfrak{p}=\frac{1}{m+1}$ and $\int_0^1 \mathcal{B}(m;n,t)dt=\frac{1}{n+1}$. Hence,

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

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 follows that there is a positive number $\varepsilon<1/2$ such that, for all $(t,\mathfrak{p})\in V$ with $|\mathfrak{p}-1/2|\leq\varepsilon$,

$$\left|\nu\left(\frac{t\mathfrak{p}}{T_\ell},\frac{t(1-\mathfrak{p})}{T_r}\right)-\nu^t\right|\leq\varepsilon',\quad\text{where }\nu^t\equiv\nu\left(\frac{t}{2T_\ell},\frac{t}{2T_r}\right).\tag{E-10}$$

(Notice that $\left|\frac{t\mathfrak{p}}{T_\ell}-\frac{t}{2T_\ell}\right|=\left|\mathfrak{p}-\frac{1}{2}\right|\frac{t}{T_\ell}$ and t/T_ℓ is bounded from above on $\mathfrak{T}$; and similarly for $\left|\frac{t(1-\mathfrak{p})}{T_r}-\frac{t}{2T_r}\right|$. 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 $2\varepsilon n^{1-\beta}\geq 1$. When $t\geq n^{-\beta}$ and $|\mathfrak{p}-1/2|>\varepsilon$, i.e., $|\mathfrak{p}-(1-\mathfrak{p})|>2\varepsilon$, we then have $n\delta_{t,\mathfrak{p}}\equiv nt\,|\mathfrak{p}-(1-\mathfrak{p})|\geq 2\varepsilon n^{1-\beta}\geq 1$. Thus, in this case, Corollary E1 implies $piv_i(t\mathfrak{p},t(1-\mathfrak{p}),n)\leq e^{-(n\delta_{t,\mathfrak{p}}-1)^2/2n}\leq e^{-\left(2\varepsilon n^{1-\beta}-1\right)^2/2n}$. That is, the integrand in (E-8) 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 $ne^{-\left(2\varepsilon n^{1-\beta}-1\right)^2/2n}\to 0$. Hence, 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

$$\Phi_{t,n}\equiv n\frac{t}{T_\ell T_r}\int_{\Xi_t}piv_i(t\mathfrak{p},t(1-\mathfrak{p}),n)\,\nu\left(\frac{t\mathfrak{p}}{T_\ell},\frac{t(1-\mathfrak{p})}{T_r}\right)\,d\mathfrak{p}$$

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 $t>\min\left\{\frac{2T_\ell}{1-2\varepsilon},\frac{2T_r}{1-2\varepsilon}\right\}$, because for such t we have $T_\ell/t<1/2-\varepsilon$ or $1-T_r/t>1/2+\varepsilon$ so that $\Xi_t=\varnothing$. We therefore conclude that, with $t_\varepsilon\equiv\min\left\{\frac{2T_\ell}{1-2\varepsilon},\frac{2T_r}{1-2\varepsilon},T_\ell+T_r\right\}$,

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

(uniformly on $\mathfrak{T}$).

Now notice that, by (E-9),

$$\int_{n^{-\beta}}^{t_\varepsilon} \Phi_{t,n}\, dt = \int_{n^{-\beta}}^{t_\varepsilon} \frac{tn}{T_\ell T_r} \sum_{m=0}^{n} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i\left(\mathfrak{p}, 1-\mathfrak{p}, m\right) \nu \left(\frac{t\mathfrak{p}}{T_\ell}, \frac{t(1-\mathfrak{p})}{T_r} \right) d\mathfrak{p} \right) dt. \quad \text{(E-11)}$$

Moreover, by Corollary E2, whenever $t \geq n^{-\beta}$, we have $n\sum_{m>t(1+\varepsilon)n} \mathcal{B}(m;n,t) \leq ne^{-2t^2\varepsilon^2 n} \leq ne^{-2\varepsilon^2 n^{1-2\beta}}$. Since $\beta < 1/2$, clearly, $ne^{-2\varepsilon^2 n^{1-2\beta}} \to 0$. Thus, the sequence $n\sum_{m>t(1+\varepsilon)n} \mathcal{B}(m;n,t)$ converges to 0 uniformly on $t \geq n^{-\beta}$. Similarly, for the sequence $n\sum_{m<t(1-\varepsilon)n} \mathcal{B}(m;n,t)$. Moreover, in (E-11) the integral inside the parenthesis and $\frac{t}{T_\ell T_r}$ are bounded from above (for relevant values of t and $(T_\ell, T_r) \in \mathfrak{T}$). It follows that we can focus on nonnegative integers m such that $t(1-\varepsilon)n \leq m \leq t(1+\varepsilon)n$. Combining this observation with (E-10), we conclude that for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$):

$$-\varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t - \varepsilon'\right) \frac{t}{T_\ell T_r} n \sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p} \right) dt \quad \text{(E-12)}$$

$$\leq \Upsilon_n$$

$$\leq \varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) \frac{t}{T_\ell T_r} n \sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m;n,t) \left(\int_{\Xi_t} piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p} \right) dt. \quad \text{(E-13)}$$

In (E-13), for each $m \geq t(1-\varepsilon)n$, the integral in parenthesis is at most $(m+1)^{-1} \leq (t(1-\varepsilon)n)^{-1}$. Thus, we see that for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$):

$$\Upsilon_n \leq \varepsilon + \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) \frac{t}{T_\ell T_r} n \left(\frac{1}{t(1-\varepsilon)n} \right) dt = \varepsilon + \frac{1}{T_\ell T_r (1-\varepsilon)} \int_{n^{-\beta}}^{t_\varepsilon} \left(\nu^t + \varepsilon'\right) dt. \quad \text{(E-14)}$$

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 $\int_0^1 piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p} = \frac{1}{m+1}$, it clearly follows that $\int_{\Xi_t} piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p} \geq m^{-1}(1-\varepsilon)$ 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-12) is at least $-\varepsilon + \int_{n^{-\beta}}^{t'_\varepsilon} \left(\nu^t - \varepsilon'\right) \frac{t}{T_\ell T_r} n \left(\sum_{m=\lceil t(1-\varepsilon)n\rceil}^{\lfloor t(1+\varepsilon)n\rfloor} \mathcal{B}(m;n,t) \right) \frac{1-\varepsilon}{t(1+\varepsilon)n}\, dt$ for all sufficiently large n (and every $(T_\ell, T_r) \in \mathfrak{T}$). As I 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 $(T_\ell, T_r) \in \mathfrak{T}$): $\Upsilon_n \geq -\varepsilon + \int_{n^{-\beta}}^{t'_\varepsilon} \left(\nu^t - \varepsilon'\right) \frac{(1-\varepsilon)^2}{T_\ell T_r (1+\varepsilon)}\, dt$. Since we can choose ε and ε' arbitrarily small, by the definitions of t_ε and t'_ε, this observation along with (E-14) 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 $\frac{1}{T_\ell T_r}\int_0^{2\min\{T_\ell,T_r\}} \nu^t\, dt = 2\int_0^{\frac{1}{\max\{T_\ell,T_r\}}} \nu(T_r\theta, T_\ell\theta)\, d\theta$. This completes the proof. $\square$

The following property will be crucial in what follows.

Assumption 3. G has a continuous density g on $[0,1]^2$.

Proof of Claim 1. Recall that $G^i(b_\ell, b_r) \equiv \Pr\{q_\ell \le b_\ell, q_r \le b_r \mid h \text{ is of type } i \text{ and altruistic}\}$ for every $(b_\ell, b_r) \in [0,1]^2$, that is,

$$\begin{aligned}
G^i(b_\ell, b_r) &= \frac{\Pr\{q_\ell \le b_\ell, q_r \le b_r,\ h \text{ is of type } i \text{ and altruistic}\}}{\Pr\{h \text{ is of type } i \text{ and altruistic}\}} \\
&= \frac{\int_0^{b_r}\int_0^{b_\ell} \Pr\{h \text{ is of type } i \text{ and altruistic} \mid q_\ell, q_r\}\, g(q_\ell, q_r)\, dq_\ell dq_r}{\int_{[0,1]^2} \Pr\{h \text{ is of type } i \text{ and altruistic} \mid q\} dG(q)} \\
&= \frac{\int_0^{b_r}\int_0^{b_\ell} \lambda_i q_i g(q_\ell, q_r)\, dq_\ell dq_r}{\int_{[0,1]^2} \lambda_i q_i\, dG(q)} = \int_0^{b_r}\int_0^{b_\ell} \frac{q_i}{\overline{q}_i} g(q_\ell, q_r)\, dq_\ell dq_r,
\end{aligned}$$

where $\overline{q}_i$ denotes the mean of q_i, $\lambda_\ell \equiv \lambda$ and $\lambda_r \equiv 1-\lambda$. Thus, the function $g^i(q) \equiv \frac{q_i}{\overline{q}_i} g(q)$ is a density for G^i on $[0,1]^2$. $\square$

Throughout the remainder of the appendix, g^i denotes the density of G^i as defined in the above proof. In the next lemma, I derive a formula for the equilibrium value of $\mathcal{P}_i$, assuming that the cutoff points are bounded away from 0.

Lemma E2. *Let $k \to n_k$ be an increasing self-map on $\mathbb{N}$. Assume that for every k, the voting game with n_k agents admits 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. For every k, let $\mathcal{P}_{i,n_k}$ denote the corresponding value of expression* (2). *Then, for any type i,*

$$\lim_k n_k \mathcal{P}_{i,n_k} = \int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} g^i(\theta T^\bullet_r, \theta T^\bullet_\ell)\, d\theta. \tag{E-15}$$

Proof. By equation (E-2), we have $\mathcal{P}_{i,n_k} = \frac{1}{2}\int_{[0,1]^2} piv_i(q_\ell T^*_{\ell,n_k}, q_r T^*_{r,n_k}, n_k)\, dG^i(q)$, that is,

$$\mathcal{P}_{i,n_k} = \frac{1}{2}\int_{[0,1]^2} piv_i(q_\ell T^*_{\ell,n_k}, q_r T^*_{r,n_k}, n_k) g^i(q)\, dq. \tag{E-16}$$

Notice that g^i is continuous by continuity of g. Moreover, since $\lim_k T^*_{i,n_k} = T^\bullet_i > 0$, the sequences T^*_{ℓ,n_k} and T^*_{r,n_k} are bounded away from 0. So, Lemma E1 applies to the right side of (E-16). That is, for each fixed k, we have

$$\lim_m \frac{m}{2}\int_{[0,1]^2} piv_i(q_\ell T^*_{\ell,n_k}, q_r T^*_{r,n_k}, m) g^i(q)\, dq = \int_0^{\frac{1}{\max\left\{T^*_{\ell,n_k}, T^*_{r,n_k}\right\}}} g^i(\theta T^*_{r,n_k}, \theta T^*_{\ell,n_k})\, d\theta.$$

Moreover, this convergence is uniform in k. Also note that the right side of the above equality is simply the function φ^i (as defined in Lemma 1) evaluated at $(T^*_{\ell,n_k}, T^*_{r,n_k})$. Since this function is continuous, it follows that

$$\lim_k \frac{n_k}{2} \int_{[0,1]^2} piv_i(q_\ell T^*_{\ell,n_k}, q_r T^*_{r,n_k}, n_k) g^i(q)\, dq = \int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} g^i(\theta T^\bullet_r, \theta T^\bullet_\ell)\, d\theta.$$

The desired conclusion follows from (E-16): $\lim_k n_k \mathcal{P}_{i,n_k} = \int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} g^i(\theta T^\bullet_r, \theta T^\bullet_\ell)\, d\theta.$ $\square$

E4. Proofs of Proposition 2 and Lemma 1

In what follows, an **asymptotic equilibrium** refers to a pair of nonnegative, extended real numbers $(C^\bullet_\ell, C^\bullet_r)$ that is the limit of a convergent subsequence of equilibria $(C^*_{\ell,n_k}, C^*_{r,n_k})$. Note that any sequence of equilibria has a subsequence that converges to an asymptotic equilibrium. Given $(C^\bullet_\ell, C^\bullet_r)$, as usual, let us set $T^\bullet_\ell \equiv \lambda F(C^\bullet_\ell)$ and $T^\bullet_r \equiv (1-\lambda) F(C^\bullet_r)$, where $F(\infty) \equiv 1$.

Consider an asymptotic equilibrium $(C^\bullet_\ell, C^\bullet_r)$, and let $\mathcal{P}_{i,n_k}$ be defined as in Lemma E2 for the subsequence of equilibria $(C^*_{\ell,n_k}, C^*_{r,n_k})$ that converges to $(C^\bullet_\ell, C^\bullet_r)$. Then, by continuity of F, $T^*_{\ell,n_k} \equiv \lambda F(C^*_{\ell,n_k})$ and $T^*_{r,n_k} \equiv (1-\lambda) F(C^*_{r,n_k})$ converge to $T^\bullet_\ell$ and $T^\bullet_r$, respectively. Hence, if $T^\bullet_\ell$ and $T^\bullet_r$ are positive numbers, equations (E-15) must hold. Since $C^*_{i,n_k} \equiv u\,(1+\psi n_k)\,\mathcal{P}_{i,n_k}$, we can then conclude that:

$$C^\bullet_i = u\psi \int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} g^i(\theta T^\bullet_r, \theta T^\bullet_\ell)\, d\theta \quad \text{for } i = \ell, r. \tag{E-17}$$

Moreover, since $g^i(\theta T^\bullet_r, \theta T^\bullet_\ell) \equiv \frac{\theta T^\bullet_j}{\overline{q}_i} g(\theta T^\bullet_r, \theta T^\bullet_\ell)$, the equations above are equivalent to:

$$C^\bullet_i = u\psi \frac{T^\bullet_j}{\overline{q}_i} \int_0^{\frac{1}{\max\left\{T^\bullet_\ell, T^\bullet_r\right\}}} \theta g(\theta T^\bullet_r, \theta T^\bullet_\ell)\, d\theta \quad \text{for } i = \ell, r. \tag{E-18}$$

It is also clear that if these equations hold, then $C^\bullet_i < \infty$ for $i = \ell, r$.

In view of these arguments, to complete the proof Proposition 2 it suffices to show that: **(a)** $T^\bullet_\ell$ and $T^\bullet_r$ are positive at any asymptotic equilibrium; **(b)** equations (E-18) have a unique solution $(C^\bullet_\ell, C^\bullet_r)$.

The following technical observation will be useful in the proof of the point (a). (In what follows, $\mathbb{Z}_+$ denotes the set of all nonnegative integers.)

Lemma E3. *For every $\mathfrak{p} \in [0,1]$, $m \in \mathbb{Z}_+$ and $i \in \{\ell, r\}$, define*

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

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

Proof. Let us write $piv_i(m)$ and $\mu_i(m)$ instead of $piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)$ and $\mu_i(\mathfrak{p}, m)$, respectively. First notice that, with $\mathfrak{p}_\ell \equiv \mathfrak{p}$ and $\mathfrak{p}_r \equiv 1-\mathfrak{p}$, if m is odd, we have $piv_i(m) = \frac{m!}{\frac{m-1}{2}!\frac{m+1}{2}!}(\mathfrak{p}_i)^{\frac{m-1}{2}}(\mathfrak{p}_j)^{\frac{m+1}{2}}$, and if m is even, we have $piv_i(m) = \frac{m!}{\frac{m}{2}!\frac{m}{2}!}(\mathfrak{p}_\ell)^{\frac{m}{2}}(\mathfrak{p}_r)^{\frac{m}{2}}$. Therefore, for every $i \in \{\ell, r\}$ and $m \in \mathbb{Z}_+$,

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

Thus, when $\mathfrak{p}_i > 0$ and m is odd, we have $\frac{1}{2\mathfrak{p}_i}\mu_i(m) \equiv \frac{1}{2\mathfrak{p}_i} piv_i(m+1) = piv_i(m)$, so that $\left|\frac{1}{2\mathfrak{p}_i} - 1\right| \leq \varepsilon$ 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 $piv_i(m+2) \leq 4\mathfrak{p}_\ell\mathfrak{p}_r piv_i(m) \leq piv_i(m)$ for every m. 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. $\square$

Next, I will prove the point (a) with the help of the following two assumptions.

Assumption 4. The support of F is a subinterval of $\mathbb{R}_+$. Moreover, F has a density f that is continuous and positive on its support.

Assumption 5. $g(q) > 0$ for every $q \in [0,1]^2$.

Lemma E4. *$C_\ell^\bullet > 0$ and $C_r^\bullet > 0$ at any asymptotic equilibrium $(C_\ell^\bullet, C_r^\bullet)$.*

Proof. Consider a subsequence of equilibria $(C^*_{\ell,n_k}, C^*_{r,n_k})$ that converges to an asymptotic equilibrium $(C_\ell^\bullet, C_r^\bullet)$. 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 $C_\ell^\bullet$ or $C_r^\bullet$ equals 0.

Fix a number $\varepsilon' \in (0,1)$ and choose an $\varepsilon \in (0, \frac{1}{2})$ such that for every $\mathfrak{p} \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right]$ we have $\left|\frac{1}{2\mathfrak{p}} - 1\right| \leq \varepsilon'$ and $\left|\frac{1}{2(1-\mathfrak{p})} - 1\right| \leq \varepsilon'$. By Lemma E3, for every such $\mathfrak{p}$,

$$(1-\varepsilon')\mu_i(\mathfrak{p}, m) \leq piv_i(\mathfrak{p}, 1-\mathfrak{p}, m) \quad \text{for every } m \in \mathbb{Z}_+ \text{ and } i \in \{\ell, r\}. \tag{E-19}$$

Fix an $i \in \{\ell, r\}$. As in the proof of Lemma E1, for every $k \in \mathbb{N}$ and $0 < t < T^*_{\ell,n_k} + T^*_{r,n_k}$, set $t'_{\varepsilon,k} \equiv \min\left\{\frac{2T^*_{\ell,n_k}}{1+2\varepsilon}, \frac{2T^*_{r,n_k}}{1+2\varepsilon}\right\}$,

$$\begin{aligned}
\Xi_{t,k} &\equiv \left(\max\left\{0, 1 - \frac{T^*_{r,n_k}}{t}\right\}, \min\left\{1, \frac{T^*_{\ell,n_k}}{t}\right\}\right) \cap [1/2-\varepsilon, 1/2+\varepsilon], \\
\Phi_{t,k} &\equiv n_k \frac{t}{T^*_{\ell,n_k} T^*_{r,n_k}} \int_{\Xi_{t,k}} piv_i(t\mathfrak{p}, t(1-\mathfrak{p}), n_k) g^i\left(\frac{t\mathfrak{p}}{T^*_{\ell,n_k}}, \frac{t(1-\mathfrak{p})}{T^*_{r,n_k}}\right) d\mathfrak{p}.
\end{aligned}$$

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 E1, $t \in (0, t'_{\varepsilon,k})$ implies $\Xi_{t,k} = [1/2-\varepsilon, 1/2+\varepsilon]$. Thus, from equation (E-9) and the definition of g^i it easily follows that, for every $k \in \mathbb{N}$ and $t \in (0, t'_{\varepsilon,k})$,

$$\Phi_{t,k} \geq \frac{g_0}{\overline{q}_i}\left(\frac{1}{2} - \varepsilon\right) \frac{n_k t^2}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \sum_{m=0}^{n_k} \mathcal{B}(m; n_k, t) \left(\int_{\Xi_{t,k}} piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p}\right), \qquad \text{(E-20)}$$

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

Clearly, there is an $m_0 \in \mathbb{N}$ such that $\int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(\mathfrak{p}, m) d\mathfrak{p} \geq m^{-1}(1-\varepsilon')$ for every integer $m \geq m_0$. Moreover, since μ_i is nonincreasing in m, for every nonnegative integer $m < m_0$, we have $\int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(\mathfrak{p}, m) d\mathfrak{p} \geq \int_{1/2-\varepsilon}^{1/2+\varepsilon} \mu_i(\mathfrak{p}, m_0) d\mathfrak{p} \geq m_0^{-1}(1-\varepsilon')$. Combining these observations with (E-19), we conclude that $\int_{\Xi_{t,k}} piv_i(\mathfrak{p}, 1-\mathfrak{p}, m)\, d\mathfrak{p} \geq \frac{(1-\varepsilon')^2}{\max\{m, m_0\}}$, for every $k \in \mathbb{N}$, $m \in \mathbb{Z}_+$ and $t \in (0, t'_{\varepsilon,k})$. In view of (E-20), it follows that 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 E2(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 $k \in \mathbb{N}$,

$$\int_0^{t'_{\varepsilon,k}} \Phi_{t,k}\, dt \geq \phi \frac{n_k}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \frac{b_k}{\max\{\varepsilon n_k, m_0\}} \int_0^{t'_{\varepsilon,k}} t^2\, dt = \phi \frac{n_k}{T^*_{j,n_k}\left(T^*_{i,n_k}\right)^2} \frac{b_k}{\max\{\varepsilon n_k, m_0\}} \frac{(t'_{\varepsilon,k})^3}{3}. \qquad \text{(E-21)}$$

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-21) 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.

It remains to show 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 note that, as in (E-4), for every $q \in [0,1]^2$, 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:

$$\begin{aligned}\int_{[0,1]^2} \Pr\left\{S^-_r - S^-_\ell = -1 \mid q\right\} g^r(q)\, dq = \\ \tfrac{\bar{q}_\ell T^*_\ell}{\bar{q}_r T^*_r} \int_{[0,1]^2} \Pr\left\{S^-_\ell - S^-_r = -1 \mid q\right\} g^\ell(q)\, dq.\end{aligned} \tag{E-22}$$

Moreover, the conditional probability of tie at $(q_\ell, q_r) \in [0,1]^2$ is

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

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

$$\int_{q_r > q_\ell} g^r(q) tie(q)\, dq < b \int_{q_r < q_\ell} g^\ell(q) tie(q)\, dq, \tag{E-23}$$

where $b > 0$ is the maximum value of $\frac{\bar{q}_\ell}{\bar{q}_r} \frac{g(q_r, q_\ell)}{g(q_\ell, q_r)}$ for $(q_\ell, q_r) \in [0,1]^2$. Moreover,

$$\begin{aligned}\int_{q_r < q_\ell} g^r(q) tie(q)\, dq &= \int_{q_r < q_\ell} \tfrac{q_r}{\bar{q}_r} g(q) tie(q)\, dq \\ &< \tfrac{\bar{q}_\ell}{\bar{q}_r} \int_{q_r < q_\ell} \tfrac{q_\ell}{\bar{q}_\ell} g(q) tie(q)\, dq = \tfrac{\bar{q}_\ell}{\bar{q}_r} \int_{q_r < q_\ell} g^\ell(q) tie(q)\, dq.\end{aligned} \tag{E-24}$$

Combining (E-22)-(E-24), 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 $F(C^*_{r,n_k})/F(C^*_{\ell,n_k}) = \int_0^{C^*_{r,n_k}} f(C)\, dC / \int_0^{C^*_{\ell,n_k}} f(C)\, dC \simeq f(0) C^*_{r,n_k}/f(0) C^*_{\ell,n_k} = C^*_{r,n_k}/C^*_{\ell,n_k}$; that is, $C^*_{r,n_k}/C^*_{\ell,n_k}$ is asymptotically equal to $F(C^*_{r,n_k})/F(C^*_{\ell,n_k})$, a contradiction. Similarly, $T^*_{\ell,n_k}/T^*_{r,n_k}$ is also bounded from above. $\square$

The uniqueness result requires (H3) and (H4) as well as a further assumption on F.

Assumption 6. (H3) and (H4) hold.

Assumption 7. $\frac{F(\gamma C)}{F(C)}$ is a nonincreasing function of $C \in \mathbb{R}_{++}$ for every fixed $\gamma \geq 1$.

I will now prove Lemma 1, and then proceed to the proof of uniqueness.

Proof of Lemma 1. The level effect amounts to saying that $\varphi^i(bT_\ell, bT_r) < \varphi^i(T_\ell, T_r)$ for $b > 1$. Notice that the substitution $\theta' \equiv \theta b$ implies $\int_0^{\frac{1}{\max\{bT_\ell, bT_r\}}} \frac{\theta b T_j}{\overline{q}_i} g(\theta b T_r, \theta b T_\ell)\, d\theta = \frac{1}{b}\int_0^{\frac{1}{\max\{T_\ell, T_r\}}} \frac{\theta' T_j}{\overline{q}_i} g(\theta' T_r, \theta' T_\ell)\, d\theta'$. Thus, $\varphi^i(bT_\ell, bT_r) = \varphi^i(T_\ell, T_r)/b$ for $b > 1$. This proves the level effect.

For the ratio effect, let $T_\ell + T_r = \widetilde{T}_\ell + \widetilde{T}_r$ and $T_\ell \leq T_r < \widetilde{T}_r$. By (H4), we have $\overline{q}_\ell = \overline{q}_r$. Hence, $\varphi^\ell(T_\ell, T_r) + \varphi^r(T_\ell, T_r) = \frac{T_\ell + T_r}{\overline{q}_\ell}\int_0^{1/T_r} \theta g(\theta T_r, \theta T_\ell)\, d\theta$ and $\varphi^\ell(\widetilde{T}_\ell, \widetilde{T}_r) + \varphi^r(\widetilde{T}_\ell, \widetilde{T}_r) = \frac{\widetilde{T}_\ell + \widetilde{T}_r}{\overline{q}_\ell}\int_0^{1/\widetilde{T}_r} \theta g(\theta \widetilde{T}_r, \theta \widetilde{T}_\ell)\, d\theta$. Moreover, (H3) implies that $g(\theta T_r, \theta T_\ell) \geq g(\theta \widetilde{T}_r, \theta \widetilde{T}_\ell)$ for every $\theta \leq 1/\widetilde{T}_r$. It immediately follows that $\varphi^\ell(T_\ell, T_r) + \varphi^r(T_\ell, T_r) > \varphi^\ell(\widetilde{T}_\ell, \widetilde{T}_r) + \varphi^r(\widetilde{T}_\ell, \widetilde{T}_r)$. Similarly, we have $\varphi^r(T_\ell, T_r) > \varphi^r(\widetilde{T}_\ell, \widetilde{T}_r)$. $\square$

Lemma E5. *Equations* (E-18) *have a unique solution.*

Proof. Suppose by contradiction that equations (E-18) have two different solutions, $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$. Then, both solutions must satisfy (8). Hence, the sign of $C_\ell^\bullet - \overline{C}_\ell^\bullet$ is the same as that of $C_r^\bullet - \overline{C}_r^\bullet$. Without loss of generality, let us assume $\overline{C}_\ell^\bullet > C_\ell^\bullet$ and $\overline{C}_r^\bullet > C_r^\bullet$.

Set $\overline{\gamma} \equiv \overline{C}_\ell^\bullet/\overline{C}_r^\bullet$ and $\gamma \equiv C_\ell^\bullet/C_r^\bullet$. By (8), we have $\gamma F(\gamma C_r^\bullet)/F(C_r^\bullet) = \overline{\gamma} F(\overline{\gamma}\overline{C}_r^\bullet)/F(\overline{C}_r^\bullet)$. Thus, $\overline{\gamma} < \gamma$ implies $\gamma F(\gamma C_r^\bullet)/F(C_r^\bullet) < \gamma F(\gamma \overline{C}_r^\bullet)/F(\overline{C}_r^\bullet)$, that is, $F(\gamma C_r^\bullet)/F(C_r^\bullet) < F(\gamma \overline{C}_r^\bullet)/F(\overline{C}_r^\bullet)$, which contradicts Assumption 7. Hence, we must have $\overline{\gamma} \geq \gamma$. By (7), this also implies that $\overline{T}_r^\bullet/\overline{T}_\ell^\bullet \geq T_r^\bullet/T_\ell^\bullet \geq 1$. But then, by applying the ratio and level effects successively, we see that $\varphi^r(\overline{T}_\ell^\bullet, \overline{T}_r^\bullet) < \varphi^r(T_\ell^\bullet, T_r^\bullet)$. From (E-17), it then follows that $\overline{C}_r^\bullet < C_r^\bullet$, a contradiction. $\square$

E.5 Proofs of Proposition 3-5

Proof of Proposition 3. If $\lambda < 1/2$, equation (8) immediately implies $C_\ell^\bullet > C_r^\bullet$. Then, by invoking (7), we see that $T_r^\bullet > T_\ell^\bullet$. Moreover, since g is symmetric, the distributions of q_r/q_ℓ and q_ℓ/q_r are identical. Hence, $\Pr\{q_\ell/q_r \leq T_r^\bullet/T_\ell^\bullet\} = \Pr\{q_r/q_\ell \leq T_r^\bullet/T_\ell^\bullet\} > \Pr\{q_r/q_\ell \leq T_\ell^\bullet/T_r^\bullet\} = 1 - \Pr\{q_\ell/q_r \leq T_r^\bullet/T_\ell^\bullet\}$. Thus, $\Pr\{q_\ell/q_r \leq T_r^\bullet/T_\ell^\bullet\} > 1/2$. $\square$

Lemma E6. *MV is an increasing function of $T_r^\bullet/T_\ell^\bullet$.*

Proof. First note that

$$MV = \int_{\frac{q_\ell}{q_r} \leq \frac{T_r^\bullet}{T_\ell^\bullet}} \frac{T_r^\bullet q_r - T_\ell^\bullet q_\ell}{T_r^\bullet q_r + T_\ell^\bullet q_\ell} g(q)dq + \int_{\frac{q_r}{q_\ell} \leq \frac{T_\ell^\bullet}{T_r^\bullet}} \frac{T_\ell^\bullet q_\ell - T_r^\bullet q_r}{T_r^\bullet q_r + T_\ell^\bullet q_\ell} g(q)dq$$

Let g_{q_r/q_ℓ} and g_{q_ℓ/q_r} stand for the densities of q_r/q_ℓ and q_ℓ/q_r, respectively. If we express both of the above integrands in terms of $x \equiv \frac{T_r^\bullet}{T_\ell^\bullet}$ and $\frac{q_\ell}{q_r}$, by changing variables in an obvious way, we find that MV is equal to:

$$\int_0^x \left(\frac{1}{1+x^{-1}z} - \frac{1}{xz^{-1}+1}\right) g_{q_\ell/q_r}(z)\,dz + \int_0^{x^{-1}} \left(\frac{1}{xz+1} - \frac{1}{1+x^{-1}z^{-1}}\right) g_{q_r/q_\ell}(z)\,dz.$$

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 $z = x$ the first integrand equals 0 and at $z = x^{-1}$ 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 $\frac{dMV}{dx}$ is equal to:

$$\int_0^x \left(\frac{x^{-2}z}{(1+x^{-1}z)^2} + \frac{z^{-1}}{(xz^{-1}+1)^2}\right) \mathfrak{g}(z)\,dz - \int_0^{x^{-1}} \left(\frac{z}{(xz+1)^2} + \frac{x^{-2}z^{-1}}{(1+x^{-1}z^{-1})^2}\right) \mathfrak{g}(z)\,dz.$$

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

$$\frac{dMV}{dx} = \int_0^x \left(\frac{x^{-2}z}{(1+x^{-1}z)^2} + \frac{z^{-1}}{(xz^{-1}+1)^2}\right) \left(\mathfrak{g}(z) - \mathfrak{g}(x^{-2}z)x^{-2}\right) dz.$$

Recall that $x \geq 1$. At $x = 1$ the expression above equals 0. Assume now $x > 1$. It suffices to show that, for every $z \in (0, x)$,

$$\mathfrak{g}(z) > \mathfrak{g}(x^{-2}z)x^{-2}. \tag{E-25}$$

First, fix a $z \in (0, 1]$. Recall that $\mathfrak{g}(z) = \int_0^1 g(y, yz)y\,dy$ (see, e.g., Rohatgi, 1976, p. 141). Hence, the substitution $y = \frac{t}{1+z}$ gives $\mathfrak{g}(z) = \int_0^{1+z} g(\frac{t}{1+z}, \frac{tz}{1+z})\frac{t}{(1+z)^2}\,dt$. Since $x > 1$, similarly, we find that $\mathfrak{g}(x^{-2}z)x^{-2} = \int_0^{1+x^{-2}z} g(\frac{t}{1+x^{-2}z}, \frac{tx^{-2}z}{1+x^{-2}z})\frac{tx^{-2}}{(1+x^{-2}z)^2}\,dt$. Now, $\frac{t}{1+z} + \frac{tz}{1+z} = t = \frac{t}{1+x^{-2}z} + \frac{tx^{-2}z}{1+x^{-2}z}$ and $\frac{t}{1+x^{-2}z} \geq \frac{t}{1+z} \geq \frac{tz}{1+z}$ imply, by (H3), that $g(\frac{t}{1+z}, \frac{tz}{1+z}) \geq g(\frac{t}{1+x^{-2}z}, \frac{tx^{-2}z}{1+x^{-2}z})$ for every $t \in [0, 1+x^{-2}z]$. Moreover, it is easily seen that, for the given values of x and z, we have $\frac{tx^{-2}}{(1+x^{-2}z)^2} = \frac{t}{(x+x^{-1}z)^2} < \frac{t}{(1+z)^2}$ whenever $t > 0$. This proves (E-25) for the case $z \in (0, 1]$.

Now let $1 < z < x$. Then, applying (E-25) to $\widetilde{z} \equiv z^{-1} < 1$ and $\widetilde{x} \equiv xz^{-1} > 1$ gives $\mathfrak{g}(\widetilde{z}) > \mathfrak{g}(\widetilde{x}^{-2}\widetilde{z})\widetilde{x}^{-2}$, that is, $z^{-2}\mathfrak{g}(z^{-1}) > \mathfrak{g}(x^{-2}z)x^{-2}$. But since $\Pr\{\frac{q_r}{q_\ell} \leq z\} = 1 - \Pr\{\frac{q_\ell}{q_r} \leq z^{-1}\}$ for $z > 0$, and since $g_{q_r/q_\ell} = g_{q_\ell/q_r}$, we have $\mathfrak{g}(z) = z^{-2}\mathfrak{g}(z^{-1})$. This completes the proof. $\square$

Assumption 8. F is a concave function on its support.

Proof of Proposition 4. Let $\lambda < \overline{\lambda} \leq 1/2$, and denote by $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$ the asymptotic equilibria that correspond to λ and $\overline{\lambda}$, respectively. Using the notation in the proof of Lemma E5, I will first show that $\gamma > \overline{\gamma}$.

By contradiction, suppose $\gamma \leq \overline{\gamma}$. Then, (8) implies $\overline{\gamma}F(\overline{\gamma}\overline{C}_r^\bullet)/F(\overline{C}_r^\bullet) < \gamma F(\gamma C_r^\bullet)/F(C_r^\bullet) \leq$

$\overline{\gamma}F(\overline{\gamma}C_r^\bullet)/F(C_r^\bullet)$. By Assumption 7, we must thus have $C_r^\bullet < \overline{C}_r^\bullet$ so that $C_\ell^\bullet = \gamma C_r^\bullet < \overline{\gamma}\overline{C}_r^\bullet = \overline{C}_\ell^\bullet$ and $T_\ell^\bullet < \overline{T}_\ell^\bullet$. Hence, by (7), $T_\ell^\bullet + T_r^\bullet = T_\ell^\bullet(1+\gamma) < \overline{T}_\ell^\bullet(1+\overline{\gamma}) = \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Moreover, $T_r^\bullet/T_\ell^\bullet \leq \overline{T}_r^\bullet/\overline{T}_\ell^\bullet$. Thus, by the level and ratio effects, we must have $C_r^\bullet > \overline{C}_r^\bullet$, a contradiction.

It follows that $\gamma > \overline{\gamma}$, as we sought. This implies $T_r^\bullet/T_\ell^\bullet > \overline{T}_r^\bullet/\overline{T}_\ell^\bullet$. Thus, by Lemma E6, expected margin of victory at $\overline{\lambda}$ is smaller than that at λ.

Let c be the supremum of the support of F. It remains to show that $C_r^\bullet < c$ implies $\overline{q}_\ell T_\ell^\bullet + \overline{q}_r T_r^\bullet < \overline{q}_\ell \overline{T}_\ell^\bullet + \overline{q}_r \overline{T}_r^\bullet$. Since $\overline{q}_\ell = \overline{q}_r$, the latter inequality can be rewritten as $T_\ell^\bullet + T_r^\bullet < \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Suppose by contradiction that $T_\ell^\bullet + T_r^\bullet \geq \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Then, by the level and ratio effects, we have

$$C_r^\bullet < \overline{C}_r^\bullet \quad \text{and} \quad C_\ell^\bullet + C_r^\bullet < \overline{C}_\ell^\bullet + \overline{C}_r^\bullet.$$

Notice that if F is concave on its support, it is also concave on $\mathbb{R}_+$. Therefore, from the inequalities above it easily follows that

$$\overline{\lambda}F(C_\ell^\bullet) + (1-\overline{\lambda})F(C_r^\bullet) \leq \overline{\lambda}F(\overline{C}_\ell^\bullet) + (1-\overline{\lambda})F(\overline{C}_r^\bullet). \tag{E-26}$$

Moreover, since f is positive on the support of F, $C_r^\bullet < c$ implies $F(C_r^\bullet) < F(C_\ell^\bullet)$. Whence,

$$\lambda F(C_\ell^\bullet) + (1-\lambda)F(C_r^\bullet) < \overline{\lambda}F(C_\ell^\bullet) + (1-\overline{\lambda})F(C_r^\bullet). \tag{E-27}$$

By combining (E-26) and (E-27), we see that $T_\ell^\bullet + T_r^\bullet < \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$, a contradiction. □

Proof of Proposition 5. Let $(C_\ell^\bullet, C_r^\bullet)$ and $(\overline{C}_\ell^\bullet, \overline{C}_r^\bullet)$ stand for the solutions of equations (E-18) for $u\psi$ and $\overline{u\psi} > u\psi$, respectively.

Using the usual notation, I will first show that $\gamma \leq \overline{\gamma}$. Suppose by contradiction that $\gamma > \overline{\gamma}$. Then, as in the proof of Proposition 4, Assumption 7 implies $C_r^\bullet > \overline{C}_r^\bullet$ and $T_\ell^\bullet + T_r^\bullet > \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Moreover, $T_r^\bullet/T_\ell^\bullet > \overline{T}_r^\bullet/\overline{T}_\ell^\bullet$. Hence, by the level and ratio effects, $\varphi^r(T_\ell^\bullet, T_r^\bullet) < \varphi^r(\overline{T}_\ell^\bullet, \overline{T}_r^\bullet)$ so that $u\psi\varphi^r(T_\ell^\bullet, T_r^\bullet) < \overline{u\psi}\varphi^r(\overline{T}_\ell^\bullet, \overline{T}_r^\bullet)$. But then, equation (E-17) implies $C_r^\bullet < \overline{C}_r^\bullet$, a contradiction. Hence, $\gamma \leq \overline{\gamma}$ and $T_r^\bullet/T_\ell^\bullet \leq \overline{T}_r^\bullet/\overline{T}_\ell^\bullet$. By Lemma E6, this immediately implies that the expected margin of victory at $\overline{u\psi}$ is greater than or equal to that at $u\psi$.

Now, let c be the supremum of the support of F and suppose $C_r^\bullet < c$. First assume that $\gamma < \overline{\gamma}$. Then, as I noted above, $C_r^\bullet < \overline{C}_r^\bullet$ and $T_\ell^\bullet + T_r^\bullet < \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Thus, in this case, the expected turnout rate that corresponds to $\overline{u\psi}$ is higher, as we seek. Suppose now $\gamma = \overline{\gamma}$ so that $T_r^\bullet/T_\ell^\bullet = \overline{T}_r^\bullet/\overline{T}_\ell^\bullet$. Assume by contradiction that $T_\ell^\bullet + T_r^\bullet \geq \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$. Then, just as in the previous paragraph, $u\psi < \overline{u\psi}$ and equations (E-17) imply $C_\ell^\bullet < \overline{C}_\ell^\bullet$ and $C_r^\bullet < \overline{C}_r^\bullet$. Since F is strictly increasing on its support, these observations yield a contradiction: $T_\ell^\bullet + T_r^\bullet < \overline{T}_\ell^\bullet + \overline{T}_r^\bullet$.

Finally, consider an alternative cost distribution $\overline{F}$ and let $\bar{c}$ be the supremum of the support of $\overline{F}$. Suppose $c < \bar{c} < \infty$. As in the case of uniform distributions, also assume that $F(C) = \overline{F}(\frac{\bar{c}}{c}C)$ for every $C \in \mathbb{R}_+$. Then, clearly, $(\frac{\bar{c}}{c}C_\ell^\bullet, \frac{\bar{c}}{c}C_r^\bullet)$ solves the modified version of (E-18) that is obtained by replacing F and $u\psi$ with $\overline{F}$ and $\frac{\bar{c}}{c}u\psi$, respectively. Since $\overline{F}(\frac{\bar{c}}{c}C_i^\bullet) = F(C_i^\bullet)$ for $i = \ell, r$, it obviously follows that the implications of replacing F with $\overline{F}$ (while holding $u\psi$ fixed) are the same as the implications of replacing $u\psi$ with $\frac{\bar{c}}{c}u\psi$ in the model with $\overline{F}$. $\square$

Remark E1. In the dual model of Appendix A, the distribution function $\widetilde{F}$ depends only on $\frac{c}{\mathrm{u}}$. Moreover, if we consider two different values of this parameter, say ω and $\overline{\omega}$, the associated distributions satisfy $\widetilde{F}_\omega(C) = \widetilde{F}_{\overline{\omega}}(\frac{\overline{\omega}}{\omega}C)$ for every $C \geq 0$. Thus, following the argument above, the implications of increasing $\frac{c}{\mathrm{u}}$ are the same as those of decreasing ψ.

References

J. Andreoni, Impure altruism and donations to public goods: a theory of warm-glow giving, Econ. J. 100 (1990) 464–477.

P. Billingsley, Probability and Measure, Wiley, New York, 1995.

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

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

G. Chamberlain, M. Rothschild, A note on the probability of casting a decisive vote, J. Econ. Theory 25 (1981) 152-162.

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

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

A. Edlin, A. Gelman, N. Kaplan, Voting as a rational choice: Why and how people vote to improve the well-being of others, Ration. Society 19 (2007) 293-314.

M. Faravelli, R. Walsh, Smooth politicians and paternalistic voters: A theory of large elections, NBER Working paper no. 17397 (2001).

T.J. Feddersen, Rational choice theory and the paradox of not voting, J. Econ. Perspect. 18 (2004) 99-112.

T.J. Feddersen, S. Gailmard, A. Sandroni, Moral bias in large elections: Theory and experimental evidence, Amer. Polit. Sci. Rev. 103 (2009) 175-192.

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

T.J. Feddersen, A. Sandroni, The calculus of ethical voting, Int. J. Game Theory 35 (2006b) 1-25.

W. Feller, An Introduction to Probability Theory and Its Applications: Volume II, Wiley, New York, 1966.

A.J. Fischer, The probability of being decisive, Public Choice 101 (1999) 267–283.

J.H. Fowler, Altruism and turnout, J. Polit. 68 (2006) 674–683.

J.H. Fowler, C.D. Kam, Beyond the self: Altruism, social identity, and political participation, J. Polit. 69 (2007) 811–825.

I.J. Good, L.S. Mayer, Estimating the efficacy of a vote, Behav. Sci. 20 (1975) 25 -33.

J.C. Harsanyi, Rule utilitarianism, rights, obligations and the theory of rational behavior, Theory Dec. 12 (1980) 115-133.

H. Herrera, C. Martinelli, Group formation and voter participation, Theoretical Econ. 1 (2006) 461-487.

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

R. Jankowski, Altruism and the decision to vote: Explaining and testing higher voter turnout, Ration. Society 19 (2007) 5–34.

D.R. Kinder, D.R. Kiewiet, Economic discontent and political behavior: The role of personal grievances and collective economic judgements in congressional voting, Amer. J. Polit. Sci. 23 (1979) 495-527.

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

G. Markus, The impact of personal and national economic conditions on the presidential vote: A pooled cross-sectional analysis, Amer. J. Polit. Sci. 32 (1988) 137-154.

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

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

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

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

V.J. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley, New York, 1976.

A. Sen, Labour allocation in a cooperative enterprise, Rev. Econ. Stud. 33 (1966) 361-371.

R. Shachar, B. Nalebuff, Follow the leader: theory and evidence on political participation, Amer. Econ. Rev. 89 (1999) 525-47.

R.A. Teixeira, Why Americans Don't Vote: Turnout Decline in the United States: 1960-1984, Greenwood Press, New York, 1987.

R.A. Teixeira, The Disappearing American Voter, The Brookings Institution, Washington, 1992.

C.J. Uhlaner, Rational turnout: The neglected role of groups, Amer. J. Polit. Sci. 33 (1989) 390-422.