Review of Economic Studies (2013) **80**, 277–312 doi:10.1093/restud/rds026

Advance access publication 20 June 2012

Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses

JOSEPH C. McMURRAY
Brigham Young University

First version received August 2010*; final version accepted February* 2012

This article analyzes participation and information aggregation in a common-value election with continuous private signals. In equilibrium, some citizens ignore their private information and abstain from voting, in deference to those with higher-quality signals. Even as the number of highly informed peers grows large, however, citizens with only moderate expertise continue voting, so that voter participation remains at realistic levels (*e.g.* 50 to 60 percent, for simple examples). The precise level of voter turnout, along with the margin of victory, are determined by the distribution of expertise. Improving a voter's information makes her more willing to vote, consistent with a growing body of empirical evidence, but makes her peers more willing to abstain, providing a new explanation for various empirical patterns of voting. Equilibrium participation is optimal, even though the marginal voter may have very little (*e.g.* below-average) expertise, and even though nonvoters' information is not utilized.

Key words: Information, Voting, Elections, Turnout, Roll-off, Swing Voter's Curse, Jury Theorem

JEL Codes: Codes: D72, D82

1. INTRODUCTION

One of the strongest endorsements of democracy ever made was in 1785 by the political philosopher and mathematician, Nicolas de Condorcet. His argument was a statistical one: if a large number of individuals independently attempt to assess the strengths of two policy alternatives (or candidates) and vote for the superior policy, then the policy that is in fact superior will almost surely receive a majority of votes (see Young 1988).[1] Condorcet's proof assumes that all citizens vote, however, and also assumes that voters' private opinions are equally reliable; in reality, many citizens abstain, and on any given issue, there are inevitably some voters with substantial expertise and others who know very little. An individual therefore faces a dilemma: if she has a tentative opinion about the merits of existing policy proposals, but also recognizes that she lacks expertise on an issue, should she vote? Or leave the election to others, who know more about the issue? On one hand, if she abstains, she deprives the electorate of her particular judgment and expertise; on the other hand, in a large electorate, there are likely to be many who know more than she does about the issue at hand.

1. Krishna and Morgan (2011) laud this *jury theorem* as the "first welfare theorem of political economy".

How voters answer these questions will determine collective outcomes, such as voter turnout, the margin of victory, and ultimately the quality of the collective decision. Many observers, for example, interpret low and declining voter turnout as indicative of poor decision quality, and take great effort to increase participation, whether by instilling a sense of civic duty, or by punishing nonvoters with stigma, or even fines. At the same time, others are troubled by voter ignorance, and make efforts to discourage or even prevent uninformed citizens from voting. From society's perspective, then, it is natural to ask, what level of voter participation aggregates information optimally?

To answer these questions, this article constructs a general model of information and voting. The basic structure is Condorcet's (1785) classic *common value* environment: members of an electorate pursue a common objective but disagree over which of two policies is most likely to achieve the desired outcome. Individual opinions are modeled as private signals of an unknown state of the world that determines which policy is superior. The precision of each individual's signal depends on her expertise, however, which is drawn from a known distribution. Voting is costless, but citizens may also abstain.

The distribution of expertise in this model is quite general. In particular, it may be continuous. In the context of the model, this implies that uninformative signals are realized with zero probability; in other words, *every* citizen tentatively favors one of the two policy alternatives. This may seem to suggest that everyone will vote, but to the contrary, many ignore their own signals in equilibrium, and abstain. This is because of the "swing voter's curse" identified by Feddersen and Pesendorfer (1996): since her peers are more likely to vote for the superior policy than the inferior policy, a citizen's own vote is most likely to be *pivotal* (*i.e.*, change the election outcome) when she mistakenly votes for the *inferior* policy.[2] In other words, a citizen relies on the judgment of her more expert peers—even if it contradicts her own—just as she might rely on expert professional advice.

In a small election, a citizen can reasonably expect to be among the best-informed members of the electorate. In a large election, however, she expects an arbitrarily large number of better-informed peers. This may seem to suggest the intuition that, no matter how accurate her own information is, a citizen should eventually abstain, when the electorate is sufficiently large. Again to the contrary, however, citizens with only moderate levels of expertise continue voting, even in the limit: although a pivotal vote does provide growing evidence of a mistake as the electorate grows large, this effect tapers off as an absolute difference of one or two votes declining in significance.[3] In simple examples, in fact, citizens of quite low (*e.g.* below-average) expertise vote, so that turnout approaches 50 to 60 percent of the electorate. Such levels are the same order of magnitude as turnout in actual public elections, which is an important result, because rational voting models are notorious for the inability to predict realistic levels of voter turnout.[4]

It may seem unlikely that citizens with such low expertise can improve the election decision of their (arbitrarily many) better-informed peers. The result that such citizens do not abstain on their own may therefore seem to vindicate efforts to dissuade or prevent them from voting. On the

2. The swing voter's curse takes its name from the "winner's curse" that arises in common-value auctions because a bid only wins (*i.e.* is "pivotal") when it exceeds all competing bids, suggesting that a bidder has overestimated the auction item's value (see Milgrom and Weber 1982).

3. The proof of this result utilizes the limiting ratios of pivot probabilities, derived by Myerson (2000), which is one technical contribution of this article.

4. Aldrich (1997, p. 373) calls voter turnout "the Achilles' heel of rational choice theory". For recent reviews, see Feddersen (2004) and Geys (2006b). In assuming that voting is costless, this article does not address the standard "paradox"of costly voting, but as Section 6 discusses, recognizes that costs do not deter voting, and that abstention occurs in costless environments as well.

other hand, the general rule that more information is always better may seem to endorse efforts to increase turnout, like compulsory voting laws. After all, voluntary elections fail to utilize nonvoters' information. Ultimately, neither of these intuitions is correct: the optimal level of voter participation arises precisely in equilibrium.[5] The first intuition ignores Condorcet's (1785) classic insight that a large number of poorly informed voters can be better informed (collectively) than any expert. The second intuition fails to recognize the informative content of a citizen's *decision* to vote, which is lost when voting is compulsory.

In addition to predicting realistic levels of turnout, this model highlights information as one of the central determinants of voter participation, consistent with a variety of available empirical evidence. The connection between information and voter turnout is nuanced, however, because improving one citizen's information makes that citizen more willing to vote, but makes her peers more willing to abstain. Thus, turnout is highest in homogeneous electorates, but may increase or decrease as an electorate becomes generally better informed. Evidence from laboratory experiments support these predictions, as do statistical analyses of public elections. Informational considerations can also explain why margins of victory are sometimes quite large, especially on collective decisions that are in some sense obvious, such as updating archaic government procedures or constitutional language. As Feddersen and Pesendorfer (1996) point out, strategic delegation also provides a rationale for abstention in costless voting environments, such as *roll off* (*i.e.* voting in some, but not all, races on a ballot).

In the context of public elections, of course, the assumption of common preferences is restrictive, since policies inevitably affect different citizens differently. Traditionally, political disputes are viewed primarily as conflicts of interests, with information playing a secondary role. The results of this article are robust to various kinds of heterogeneity, but only if a certain degree of homogeneity is maintained, so that one citizen values another's expertise. This seems quite plausible, as the broad goals of most policies—such as world peace, economic stability, and reducing crime, poverty, pollution, and corruption—hold essentially unanimous appeal. Furthermore, persuasive efforts such as debate, policy research, endorsements, and advertising, suggest that a citizen expects her peers (once informed) to adopt her own positions. Mueller (2003, ch. 3, 14) also argues that, as explained in more detail later, apparent conflicts of interest may be substantially mitigated by altruistic or insurance motivations.[6]

The remainder of this article is organized as follows. Section 2 begins by summarizing related literature, and Section 3 introduces the formal model. Sections 4.1 and 4.2 characterize equilibrium behavior in small and large electorates, respectively, and Section 4.3 presents illustrative examples. Sections 4.4 and 4.5 then analyze comparative statics and welfare, and Section 5 discusses the robustness of the central results to generalizing the structure of information and preferences. Section 6 discusses empirical applications, and Section 7 concludes. Mathematical proofs are presented in the appendix.

2. RELATED LITERATURE

The basic logic of strategic abstention and the empirical application of abstention in costless voting environments are the focus of Feddersen and Pesendorfer (1996). As Section 4.3 explains,

5. As in Condorcet's (1785) original analysis, this implies that a large electorate almost surely identifies the better of two policies.

6. Ethical considerations such as altruism and civic duty are also frequently offered as explanations for costly voting (*e.g.* Riker and Ordeshook 1968; Edlin *et al.* 2007; Feddersen and Sandroni 2006; Evren 2010). Such motivations are most natural in a common-value environment, where voting provides a bona fide public good; in private-value settings, voting merely pulls policy outcomes in a voter's own favor, so there is no clear ethical basis for voting or encouraging others to vote.

that model (without partisans) can be viewed in the context of this article as the special case of a discrete distribution of expertise, for which signals are either perfectly informative or completely uninformative. Voting and abstention are less surprising in that case, since perfectly informed citizens have no reason to abstain, and those with no information have no reason to vote. More importantly, central results regarding the level of participation are not applicable, because the sizes of the two groups are specified exogenously. With continuous signals, every citizen possesses some private information, and none is infallible. Continuous signals also imply more realistic heterogeneity, avoid the need for complicated mixed strategies in equilibrium, and facilitate the more nuanced comparative statics analysis of Section 4.4, along with its applications.[7]

In a common-value election environment similar to this, Krishna and Morgan (2010) demonstrate the informational inefficiency of compulsory voting, and further show that this extends to costly elections, despite the fact that information is under-provided in that case, because only a tiny fraction of the electorate votes. While complementary, the two models differ in the source of inefficiency. Here, it arises from weighing heterogeneous signals equally. There, citizens are homogeneous, but signals of the two states of the world are not equally informative, so to compensate, citizens respond to the less informative signal by mixing between voting and abstaining—or by voting uninformatively, if abstention is prohibited.

Extensions of Feddersen and Pesendorfer (1996) focus primarily on the structure of preferences, rather than information. Feddersen and Pesendorfer (1999) extend both, but focus on the former. In their model, preferences have both a private- and a common-value dimension, and (in most specifications) signals are of heterogeneous quality, though not all are informative. Like this article, those authors analyze turnout in large elections. As Section 5.2 explains, however, abstention declines in that model as the electorate grows, because citizens increasingly vote on the basis of private values, trusting commonly valued information to be aggregated from their peers. Depending on the model specifications, then, abstention may all but vanish in the limit. In contrast, abstention in this model *increases* (consistent with evidence described in Section 6.2) as citizens in larger electorates increasingly defer to those with better information, but turnout remains substantial in the limit.[8]

In addition to studies of participation and information aggregation in democracy, this article relates to studies of participation and information aggregation in markets, and other settings. In such settings, expertise plays a similar role. In capital markets, for example, Grossman and Stiglitz (1976) discuss how uninformed traders use market prices to infer information from those who are informed. Ottaviani and Sorensen (2010) show that the least informed traders abstain from trading, to avoid losing wealth to those with superior information.[9] Their analysis assumes a small participation utility; as this decreases, only the best-informed traders continue to participate, until in the limit no one is willing to trade, as in Milgrom and Stokey (1982). Participation in elections does not similarly fall to zero because relationships are inherently cooperative, rather than adversarial.

3. THE MODEL

An electorate consists of N citizens where, following Myerson (1998, 2000), N is drawn from a Poisson distribution with mean n. Together, these citizens must choose a policy (or candidate)

7. Continuous signals are also quite standard in other areas of information economics, such as auctions and social learning models.

8. Feddersen and Pesendorfer (1999) also give an example in which turnout falls as information improves, but provide no general result comparable to those of Section 4.4.

9. Those authors treat the special case of betting markets, but the intuition applies more generally.

from the set $\{A,B\}$, by majority vote. The state of the world, $\omega \in \{\alpha,\beta\}$, is not directly observable when the vote is taken, but citizens agree that policy A is superior in state α and policy B is superior in state β. Specifically, every citizen's utility function is

$$u(A|\alpha) = u(B|\beta) = 1$$
$$u(A|\beta) = u(B|\alpha) = 0.$$

Expected utility is simply the probability of choosing the superior policy. Ex ante, both states of the world are equally likely.

Independently of ω, each citizen, i, is endowed with information quality $q_i \in \left[\frac{1}{2},1\right]$, representing her level of expertise on the issue at hand. This is drawn independently from a common distribution F, which has smooth density f that is strictly positive on $\left(\frac{1}{2},1\right)$.[10] Each citizen also observes a signal $s_i \in \{a,b\}$, representing her private opinion as to which policy is superior. The accuracy of a citizen's signal is determined by her level of expertise. Specifically, $\Pr(a|q_i,\alpha) = \Pr(b|q_i,\beta) = q_i$. The highest quality signal thus reveals ω perfectly, while the lowest quality signal reveals nothing. The distribution of expertise is common knowledge, but q_i and s_i are private information. Signal values are mutually independent (conditional on ω) and also independent of q_i and N.[11]

Conditional on observing a signal s of quality q, a citizen's updated belief $\theta(\omega|q,s)$ about the state of the world is given by Bayes' rule:

$$\theta(\alpha|q,a) = \theta(\beta|q,b) = \frac{\frac{1}{2}q}{\frac{1}{2}q + \frac{1}{2}(1-q)} = q. \quad (1)$$

Thus, q_i also measures a citizen's confidence that she has correctly identified the superior policy.

After observing her private information, an individual may vote (at no cost) for either policy, or may abstain. A strategy $\sigma : [1/2,1] \times \{a,b\} \to \{A,B,0\}$ specifies behavior for each realized quality type $q \in [1/2,1]$ and signal value $s \in \{a,b\}$, where a vote for policy 0 represents abstention.[12] The numbers N_A and N_B of votes for either policy depend on the total number of citizens in the electorate, and on each voter's strategy, in combination with her private information. The policy that receives more votes is implemented, breaking a tie if necessary by a fair coin toss. The strategy σ_i^* is a *best response* to opponent strategies if $\sigma_i^*(q,s)$ maximizes the probability of electing the superior policy, for every $(q,s) \in [1/2,1] \times \{a,b\}$. A profile of strategies is a *Bayesian Nash equilibrium* if it specifies a strategy for each player that is the best response to her opponents. In games of Poisson population uncertainty such as this, equilibria are necessarily *symmetric*, specifying identical behavior for each player.[13] Thus, an equilibrium profile can be represented by a single strategy σ^*.

10. As Section 5.1 discusses, these restrictions on F simplify the exposition, but are not necessary for the results below. Accordingly, not all of the distributions illustrated in Section 4.3 have full support.

11. A more standard formulation of private information posits a single, continuous signal. As Section 5.1 explains, the decomposition into q_i and s_i imposes symmetry on such a formulation, but is otherwise equivalent, and is useful for the comparative statics analysis of Section 4.4.

12. Mixed strategies play an important role in discrete models such as Feddersen and Pesendorfer (1996), and could be allowed here, but with continuous signals, mixed strategies would not be played in equilibrium (by Lemma 1).

13. This is because the finite set of citizens who actually play the game is a random draw from an infinite set of *potential* players, for whom strategies are defined (see Myerson, 1998). The distribution of opponent behavior is therefore the same for any two individuals (in contrast with a game between a finite set of players), implying that a best response for one citizen is a best response for all.

4. ANALYSIS

4.1. *Equilibrium*

If a citizen follows the strategy σ then, in state $\omega \in \{\alpha, \beta\}$, she takes action $x \in \{A, B, 0\}$ with probability $v_x(\omega)$, given by the following,

$$v_x(\omega) = \int_{1/2}^{1} \sum_{s=a,b} I_{[\sigma(q,s)=x]} \Pr(s|q,\omega)\, dF(q), \tag{2}$$

where I is an indicator function. Since the total number of citizens N has Poisson distribution with parameter n, and since fractions $v_A(\omega)$ and $v_B(\omega)$ of the electorate vote for policies A and B, respectively, the numbers N_A and N_B of A and B votes are independent Poisson random variables, with parameters $nv_A(\omega)$ and $nv_B(\omega)$, by the decomposition property of Poisson random variables (see Myerson, 1998). Accordingly, the probability of any voting outcome is merely the product of Poisson probabilities. For example, let $\pi_m(\alpha)$ denote the probability in state α that the superior policy (*i.e.* policy A) wins the election by a margin of exactly m votes (or loses by $|m|$, if m is negative). This is the probability of observing $k+m$ votes for policy A and k votes for policy B, summed over all possible values of k,

$$\pi_m(\alpha) = \sum_{k=\min\{0,m\}}^{\infty} \frac{e^{-nv_A(\alpha)}[nv_A(\alpha)]^{k+m}}{(k+m)!} \frac{e^{-nv_B(\alpha)}[nv_B(\alpha)]^{k}}{k!}.$$

Similarly, let $\pi_m(\beta)$ denote the probability in state β with which the superior policy (*i.e.*, policy B) wins by exactly m votes,

$$\pi_m(\beta) = \sum_{k=\min\{0,m\}}^{\infty} \frac{e^{-nv_A(\beta)}[nv_A(\beta)]^{k}}{k!} \frac{e^{-nv_B(\beta)}[nv_B(\beta)]^{k+m}}{(k+m)!}.$$

By the environmental equivalence property of Poisson games (see Myerson 1998), an individual from within the game reinterprets N_A and N_B as the number of A and B votes cast by her peers; by voting herself, she can add one to either total. Of particular interest, therefore, are the probabilities $\pi_0(\omega)$, $\pi_1(\omega)$, and $\pi_{-1}(\omega)$ of a tie, a one-vote win, and a one-vote loss—events in which a single additional vote would be *pivotal* in the election.[14] Specifically, an additional vote for the superior policy is pivotal if that policy either ties the election and loses the tie-breaking coin toss, or wins the coin toss but loses the election by exactly one vote. Let $P(\omega)$ denote the combined probability of such an event.

$$P(\omega) = \frac{1}{2}\pi_0(\omega) + \frac{1}{2}\pi_{-1}(\omega).$$

14. A common objection to strategic voting models is that voters in the real world do not seem cognizant of pivot probabilities, which are also presumably quite miniscule, and therefore difficult to estimate accurately. In favor of strategic voting, Abramson *et al.* (1992), Alvarez and Nagler (2000), and Fujiwara (2011) present evidence that voters support second-choice candidates when first-choice candidates seem unlikely to win, consistent with the logic of Duverger's (1954) law. Voters also behave strategically in the laboratory setting of Battaglini *et al.* (2008, 2009) and Morton and Tyran (2011), although they fail to do so even in the very simple setting of Esponda and Vespa (2011). In this model, Theorem 5 ultimately proves an equivalence between strategic behavior and socially optimal behavior. Thus, even without thinking about pivot probabilities, a voter could behave *as if* she were strategic, simply by determining which voter types should be voting, from a social planner's perspective, and then following her own recommendation.

Similarly, a vote for the inferior policy is pivotal with probability $\tilde{P}(\omega)$.

$$\tilde{P}(\omega)=\frac{1}{2}\pi_0(\omega)+\frac{1}{2}\pi_1(\omega).$$

If the state of the world turns out to be β and a citizen's vote for policy B reverses the election outcome, she receives a utility benefit of $u(B|\beta)-u(A|\beta)=1$. On the other hand, if the state of the world is α and her B vote changes the election outcome, she suffers a utility penalty of $u(B|\alpha)-u(A|\alpha)=-1$. From the voter's perspective, these events occur with probabilities $\theta(\beta|q,s)P(\beta)$ and $\theta(\alpha|q,s)\tilde{P}(\alpha)$, respectively. Therefore, the expected benefit $\Delta_{0B}(q,s)$ of voting for policy B instead of abstaining is given by

$$\Delta_B(q,s)=\theta(\beta|q,s)P(\beta)-\theta(\alpha|q,s)\tilde{P}(\alpha).$$

Substituting $\theta(\alpha|q,s)=1-\theta(\beta|q,s)$, this is positive if and only if $\theta(\beta|q,s)$ exceeds the threshold $\hat{\theta}_{0B}$,

$$\hat{\theta}_{0B}=\frac{\tilde{P}(\alpha)}{\tilde{P}(\alpha)+P(\beta)}. \tag{3}$$

In other words, a citizen prefers voting for policy B to abstaining if she is sufficiently confident that the state is β. Similarly, the expected benefit

$$\Delta_A(q,s)=\theta(\alpha|q,s)P(\alpha)-\theta(\beta|q,s)\tilde{P}(\beta)$$

of voting A is positive if and only if $\theta(\beta|q,s)$ is less than

$$\hat{\theta}_{A0}=\frac{P(\alpha)}{P(\alpha)+\tilde{P}(\beta)}, \tag{4}$$

reflecting a strong belief that the state is α. The benefit of voting B exceeds the benefit of voting A if and only if $\Delta_B(q,s)>\Delta_A(q,s)$, or equivalently, if and only if $\theta(\beta|q,s)$ exceeds the threshold

$$\hat{\theta}_{AB}=\frac{P(\alpha)+\tilde{P}(\alpha)}{P(\alpha)+\tilde{P}(\alpha)+P(\beta)+\tilde{P}(\beta)}. \tag{5}$$

Definition 1 defines a *belief threshold strategy*, which simply prescribes that a citizen votes for policy A if state β is sufficiently *un*likely, and votes for policy B if state β is sufficiently likely. If belief thresholds T_A and T_B do not coincide, a citizen with moderate beliefs does not vote for either policy.

Definition 1. *The strategy σ_{T_A,T_B} is a belief threshold strategy, with belief thresholds $0\le T_A\le T_B\le 1$, if* $\sigma_{T_A,T_B}(q,s)=\begin{cases} A & \text{if } \theta(\beta|q,s)\le T_A \\ B & \text{if } \theta(\beta|q,s)\ge T_B \\ 0 & \text{otherwise}\end{cases}$ *(where inequalities may be strict instead of weak), for every* $(q,s)\in\left[\frac{1}{2},1\right]\times\{a,b\}$.

The discussion above makes clear that a voter's best response to σ is to vote B when her posterior $\theta(\beta|q,s)$ exceeds both $\hat{\theta}_{AB}$ and $\hat{\theta}_{0B}$, and to vote A when $\theta(\beta|q,s)$ falls below both $\hat{\theta}_{AB}$ and $\hat{\theta}_{A0}$. In other words, the best response to any opponent strategy is a belief threshold strategy, as Proposition 1 now states.

Proposition 1. *σ_i^* is a best response to σ only if σ_i^* is a belief threshold strategy.*

Given the symmetry of this model, it is natural to focus attention on a belief threshold strategy $\sigma_{1-T,T}$ that is symmetric around 1/2, which can be rewritten with a single subscript, as a *quality threshold strategy* σ_T, defined below. With such a strategy, a citizen who receives an a signal never votes for policy B, and a citizen who receives a b signal never votes for policy A. In either case, she votes if and only if her expertise exceeds a *quality threshold* T, and abstains otherwise.[15]

Definition 2. *σ_T is a quality threshold strategy, with quality threshold $1/2 \leq T \leq 1$, if $\sigma_T(q,s) = \begin{cases} A & \text{if } s=a \text{ and } q \geq T \\ B & \text{if } s=b \text{ and } q \geq T \\ 0 & \text{otherwise} \end{cases}$ (where inequalities may be strict instead of weak), for every $(q,s) \in \left[\frac{1}{2},1\right] \times \{a,b\}$.*

Since a quality threshold strategy prescribes symmetric behavior in response to a and b signals, and since signals arise symmetrically in the two states of the world, behavior is also symmetric with respect to the state. In either state, for example, a citizen votes for the superior policy (*i.e.* for A in state α, and for B in state β) with probability

$$v_+ = v_A(\alpha) = v_B(\beta) = \int_T^1 q dF(q), \tag{6}$$

votes for the inferior policy with probability

$$v_- = v_B(\alpha) = v_A(\beta) = \int_T^1 (1-q) dF(q), \tag{7}$$

and abstains with probability

$$v_0 = v_0(\alpha) = v_0(\beta) = \int_{1/2}^T dF(q) = F(T). \tag{8}$$

Symmetric voting behavior produces symmetric electoral outcomes. In either state, for example, the numbers N_+ and N_- of votes for the superior policy and the inferior policy have Poisson distributions, with parameters np_+ and np_-, respectively. The probability π_m with which the superior candidate wins by exactly m votes is also invariant across states.

$$\pi_m = \pi_m(\alpha) = \pi_m(\beta) = \sum_{k=\min\{0,m\}}^{\infty} \frac{e^{-nv_+}(nv_+)^{k+m}}{(k+m)!} \frac{e^{-nv_-}(nv_-)^k}{k!}. \tag{9}$$

So are the probabilities P and $\tilde{P}$ with which votes for the superior policy and votes for the inferior policy, respectively, are pivotal.

$$P = P(\alpha) = P(\beta) = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_{-1} \tag{10}$$

$$\tilde{P} = \tilde{P}(\alpha) = \tilde{P}(\beta) = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1. \tag{11}$$

15. That is, $q \geq T$ simultaneously implies that $\theta(\beta|q,a) = 1-q \leq 1-T$ and that $\theta(\beta|q,b) = q \geq T$.

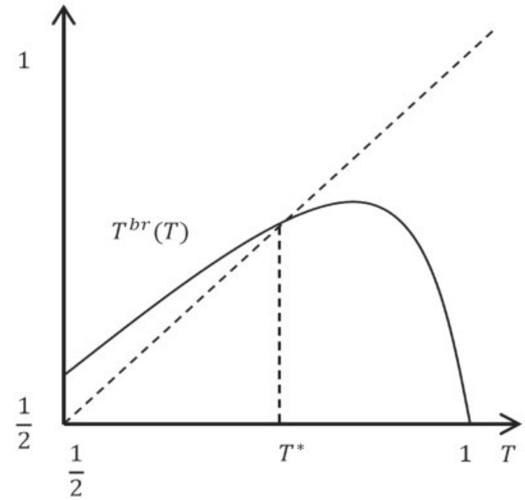

FIGURE 1
Equilibrium threshold T^*

With symmetric pivot probabilities, the belief threshold in (5) reduces to $\hat{\theta}_{AB} = 1/2$, and the thresholds in (3) and (4) reduce symmetrically around $1/2$:

$$\hat{\theta}_{0B} = 1 - \hat{\theta}_{A0} = \frac{\tilde{P}}{P + \tilde{P}} \equiv T^{br}. \tag{12}$$

Thus, the best response to a quality threshold strategy σ_T is another quality threshold strategy, with a quality threshold T^{br} given by (12). Proposition 2 states this result formally, and also points out that $T^{br} > 1/2$ as long as $T < 1$, implying that a positive fraction of the electorate abstains (*i.e.* $v_0 > 0$) in response to σ_T.

Proposition 2 (Swing voter's curse) *σ_i^* is a best response to the quality threshold strategy σ_T only if σ_i^* is a quality threshold strategy $\sigma_{T^{br}}$, with quality threshold $T^{br} = \frac{\tilde{P}}{P + \tilde{P}}$. Furthermore, $1/2 \leq T^{br} < 1$, with strict inequality if $T < 1$.*

The result that citizens abstain may be surprising in light of the fact that each citizen possesses private information that leads her to strictly prefer one policy over the other. The logic behind Proposition 2 is that voting is informative, so the superior policy likely receives a larger vote share than the inferior policy, and in particular is more likely to win by a single vote than to lose by a single vote. This implies, however, that an additional vote for the superior policy is less likely to be pivotal than a vote for the inferior policy. Conditional on casting a pivotal vote, therefore, an uninformed citizen believes that she has voted for the *wrong* policy, and prefers to have abstained. In other words, she suffers from Feddersen and Pesendorfer's (1996) swing voter's curse even though she is not uninformed. Put differently, a citizen abstains to avoid accidentally overturning the decision of those who are better informed.

Since the best response to a quality threshold strategy σ_T is another threshold strategy $\sigma_{T^{br}}$, the best response threshold T^{br} can be interpreted as a (continuous) mapping from the (compact) set of thresholds into itself. Brouwer's theorem therefore guarantees the existence of a fixed point T^*, which as Theorem 1 now states, characterizes a quality threshold strategy σ_{T^*} that is its own best response, and thus an equilibrium in the voting game. This is illustrated in Figure 1. The additional result that $1/2 < T^* < 1$ follows from Proposition 2, and implies that, in equilibrium, a positive fraction of the electorate votes, and a positive fraction abstains.

Theorem 1. *There exists a quality threshold T^* strictly between $1/2$ and 1 such that σ_{T^*} is a Bayesian Nash equilibrium.*

In stating the existence of an equilibrium quality threshold strategy, Theorem 1 leaves open the possibility of other types of equilibria. As Theorem 2 now states, however, the symmetry of the model rules out this possibility. In other words, if a citizen votes at all in equilibrium, she votes in accordance with her signal; her decision is simply whether to vote or abstain.

Theorem 2. *If σ^* is a Bayesian Nash equilibrium then it is a quality threshold strategy.*

The best response function illustrated in Figure 1 exhibits a unique fixed point, but formally, Theorem 2 does not guarantee a unique equilibrium. Informally, it appears that the equilibrium is indeed unique, as long as the information distribution is well behaved. To see this, note in Figure 1 that multiple equilibria only exist if T^{br} first crosses the 45° line from above, and then rises quickly to cross again from below. This could occur if F were discrete, for example, so that as T increased through an atom of F, a positive fraction of the least-informed voters would suddenly abstain, discontinuously raising the average quality of a vote, and with it the best-response quality threshold. Discrete distributions are ruled out by assumption, of course, but can be approximated arbitrarily closely by continuous distributions. To rule out this possibility, Theorem 4 in Section 4.2 restricts attention to density functions that are log-concave, and states formally that any sequence of equilibrium thresholds approaches a unique limit. Thus, if multiple equilibria exist in finite electorates, they converge as the electorate grows large.

4.2. *Large Elections*

Theorems 1 and 2 in Section 4.1 characterize equilibrium voting behavior for a fixed population size parameter n. This section analyzes voting behavior as n grows large. Lemma 1 begins by showing that the swing voter's curse intensifies as an electorate grows: for any quality threshold strategy σ_T (with $T < 1$) the best response threshold T_n^{br} increases with n. As the entire best-response function increases, so does its fixed point T_n^*.[16] As the participation threshold increases, voter participation decreases.

Lemma 1. *For any $T < 1$, the best-response quality threshold T_n^{br} increases in n.*

One intuition for Lemma 1 is that a poorly informed citizen, who would otherwise abstain, votes when the electorate is small, to hedge against the possibility that her peers' information quality turns out to be even lower than her own. This possibility becomes remote as the electorate grows large, so eventually she abstains. This begs the question, however, of who continues voting in the limit. After all, in a sufficiently large electorate, even a highly expert voter (*e.g.* $q_i = 0.99$) should expect an arbitrarily large number of peers with better information than her own. This might seem to suggest that a citizen of any information level should eventually abstain—or in other words, that T_n^* should approach 1. If so, this would be reminiscent of Feddersen and

16. If there are multiple equilibrium fixed points, the implication of Lemma 1 is that the lowest and highest of these both increase with n.

Pesendorfer's (1996) model (with balanced partisan groups), where uninformed voters abstain with increasing probability until, in the limit, only (perfectly) informed agents continue to vote. Since no one in this model is perfectly informed, that would imply that voter turnout approaches 0 percent.[17]

Assessing the validity of this intuition requires deriving an expression for the limit of the best-response quality threshold $T_n^{br} = \frac{\tilde{P}_n}{P_n + \tilde{P}_n}$. This is facilitated by Myerson's (2000) elegant result that the ratio of pivot probabilities in large elections approaches the square root of the ratio of vote probabilities:

$$\lim_{n\to\infty} \frac{\tilde{P}_n}{P_n} = \sqrt{\frac{v_+}{v_-}}.$$

For any $T<1$, this implies that $T_n^{br}(T)$ approaches $T_\infty^{br}(T)$, defined as follows,

$$\begin{aligned} T_\infty^{br}(T) &= \frac{\sqrt{v_+}}{\sqrt{v_+}+\sqrt{v_-}} \\ &= \frac{\sqrt{E(q_i|q_i \geq T)}}{\sqrt{E(q_i|q_i \geq T)}+\sqrt{1-E(q_i|q_i \geq T)}}, \end{aligned} \tag{13}$$

where $E(q_i|q_i \geq T)$ is the average expertise of a citizen who actually votes.

Since T_n^* is a fixed point of $T_n^{br}(T)$ for every n, a limit point T_∞^* of any sequence of equilibrium quality thresholds must be a fixed point of (13).[18] Equivalently, T_∞^* solves

$$E(q_i|q_i \geq T) = \frac{T^2}{T^2+(1-T)^2}. \tag{14}$$

Theorem 3 now states that such a solution exists, and that $T_\infty^* < 1$.[19] Equation (14) actually has another solution at $T=1$, but $T_n^{br}(1) = \frac{1}{2}$ for any n, so $\lim_{n\to\infty} T_n^{br}(1) = 1/2$. Accordingly, a sequence of equilibrium thresholds cannot converge to one. In fact, T_∞^* is an upper bound on T_n^*, implying that there is no equilibrium in a population of any size for which turnout is lower than $1-F\left(T_\infty^*\right)$. Thus, contrary to the intuition above, moderately informed citizens continue voting, no matter how large the electorate grows, and voter turnout remains positive even in an arbitrarily large electorate.

Theorem 3. *If $\{T_n^*\}$ is a sequence of equilibrium quality thresholds then it has a limit point T_∞^*. Furthermore, if $\lim_{T\to 1} \frac{f'(T)}{f(T)} < \infty$ then $T_\infty^* < 1$.*

A partial intuition for Theorem 3 is that, while a citizen indeed expects a large number of extremely well-informed peers ex ante, she bases her behavior on the *conditional* belief that a surprisingly large fraction have simultaneously erred, thereby rendering her own vote pivotal. Additionally, note that the vote totals N_+ and N_- for the superior and inferior policies are independent Poisson random variables with parameters nv_+ and nv_-, respectively, so that as

17. This would also resemble nonparticipation in capital markets, as noted in Section 2.

18. In discussing sequences, this section implicitly assumes, for simplicity, that n is a natural number.

19. The technical condition that $\lim_{T\to 1} \frac{f'(T)}{f(T)} < \infty$ merely excludes electorates that are arbitrarily close to being perfectly informed, and is sufficient but not necessary for Theorem 3: if $F(q) = 1-\left[1-2\left(q-\frac{1}{2}\right)\right]^{1/\eta}$, with $\eta > 1$, for example, then the condition is violated but nevertheless $T_\infty^* < 1$.

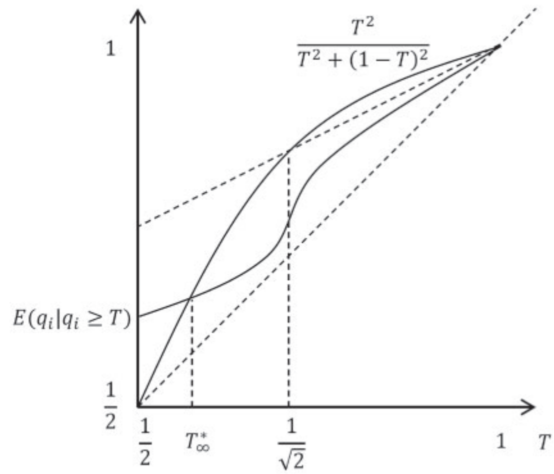

FIGURE 2
Unique convergence of equilibrium threshold

n gets large, the margin of victory $M=N_+-N_-$ can be approximated by a normal distribution with mean $E(M)=n(v_+-v_-)$ and variance $V(M)=n(v_++v_-)$. The ratio of probabilities of a one-vote loss to a one-vote win can thus be approximated by

$$\frac{\phi\left(\frac{-1-E(M)}{\sqrt{V(M)}}\right)}{\phi\left(\frac{1-E(M)}{\sqrt{V(M)}}\right)}=e^{-2\left(\frac{E(M)}{V(M)}\right)},$$

where ϕ is the standard normal pdf. The expected margin of victory $E(M)$ grows unboundedly with the number of voters; by itself, this fact would make a one-vote loss exponentially less likely than a one-vote win, leading even arbitrarily expert citizens to eventually abstain. The variance $V(M)$ of the margin of victory also grows, however, so a difference of two votes becomes less and less meaningful, and the ratio $E(M)/V(M)$ remains constant.[20]

In stating the existence of a limiting equilibrium quality threshold, Theorem 3 (like Theorem 1) makes no claim about uniqueness. This is an important issue, in light of existing game-theoretic literature in which equilibria with high and low voter turnout both exist (*e.g.*, Palfrey and Rosenthal 1983). As Theorem 4 now states, however, the limiting threshold T^*_∞ is indeed unique for the large class of distribution functions that have log-concave densities.[21] This condition is actually stronger than necessary for uniqueness; intuitively, the important thing is that a density is sufficiently "smooth".[22] This logic is illustrated in Figure 2: for Equation (14) to have multiple solutions, $E(q_i|q_i \geq T)$ (which increases in T) must first rise gradually until it intersects $T^2/T^2+(1-T)^2$

20. This observation suggests that the logic of Theorem 3 does not depend on the assumption of population uncertainty; if $N=n$ were fixed and known, for example, N_+ and N_-, together with the number of abstentions, would follow a trinomial distribution. When n is large, therefore, M could still be approximated by a normal distribution with the same mean, but now with variance $V(M)=n[p_+(1-p_+)+p_-(1-p_-)-p_+p_-]$. This also grows at rate n, so the ratio $E(M_+)/V(M_+)$ still remains constant.

21. A density f is *log-concave* if $\ln[f(q)]$ is concave, or equivalently, if $f'(q)/f(q)$ is decreasing in q. As Bagnoli and Bergstrom (2005) show, many of the most familiar distribution functions exhibit this property.

22. All of the densities illustrated below exhibit unique limiting equilibrium thresholds, for example, but not all are log-concave.

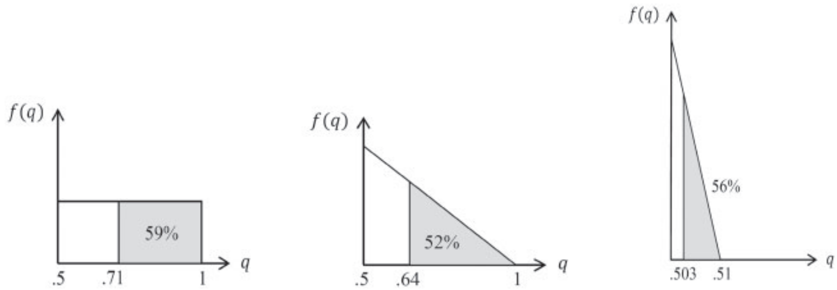

FIGURE 3
Participation thresholds and turnout rates for simple information distributions

from above, and then rise quickly to intersect $T^2/T^2+(1-T)^2$ again from below. This would be possible if, for example, the density of expertise included atoms, or "spikes" of probability, but is not possible when the density is sufficiently "smooth" (*e.g.*, log-concave).

Theorem 4. *If f is log-concave then there exists a unique limit point T^*_∞ such that for any sequence $\{T^*_n\}$ of equilibrium participation thresholds, $T^*_n \to T^*_\infty$.*

The result that the limiting threshold T^*_∞ is uniquely determined by the distribution of expertise on the issue at hand implies that, in large elections, F also determines the precise level of voter turnout τ,

$$\tau = 1 - F\left(T^*_\infty\right), \tag{15}$$

along with the margin of victory μ (as a fraction of the number who voted),

$$\begin{aligned}\mu &= \frac{v_+ - v_-}{\tau} = E\left(q_i | q_i \geq T^*_\infty\right) - \left[1 - E\left(q_i | q_i \geq T^*_\infty\right)\right] \\ &= 2E\left(q_i | q_i \geq T^*_\infty\right) - 1. \end{aligned} \tag{16}$$

Section 4.3 illustrates this by deriving τ and μ for several simple information distributions. Section 4.4 then provides a more general analysis of how turnout and margins of victory respond to changes in the underlying distribution of information.[23]

4.3. *Examples*

Theorem 3 in Section 4.2 states that the equilibrium quality threshold remains bounded below one, so that turnout remains positive in large electorates, but leaves open the theoretical possibility that the bound on turnout is extremely close to one, so that turnout is positive but negligible. To investigate this possibility, this section uses Equation (14) to compute limiting equilibrium thresholds for some simple example distributions. The resulting levels of equilibrium turnout are actually quite high. For example, the first frame of Figure 3 illustrates a uniform distribution

23. Strictly speaking, these labels for τ and μ are misnomers, as turnout $N_+ + N_-/N$ and the margin of victory $|N_+ - N_-|/N_+ + N_-$ are actually random variables. More precisely, τ denotes the probability with which an individual votes, and μ describes how much more likely she is to vote for the superior policy than the inferior policy, conditional on voting at all.

of expertise. This distribution exhibits a unique participation threshold $T^*_\infty = \frac{1}{\sqrt{2}} \approx 0.71$, which from (15) implies that 59 percent of a large electorate votes in equilibrium.

A uniform distribution is simple, but perhaps unrealistic. A natural by-product of specialization, for example, is that only a small fraction of the electorate possess significant expertise on any particular issue. Accordingly, the second frame of Figure 3 illustrates a distribution that is right-skewed. This produces a limiting quality threshold of $T^*_\infty \approx 0.64$, implying 52 percent turnout, which is slightly lower than for the uniform case.

The first two frames of Figure 3 both presume that some voters are essentially infallible. Another possibility is that policy issues are sufficiently complex that even the best-informed members of the electorate have difficulty predicting which policies will be effective. To treat this possibility, the third frame of Figure 3 depicts an extreme distribution, in which even the most expert citizen can determine the better of two policies with only 51 percent accuracy. In that case, the limiting threshold is $T^*_\infty \approx .503$ and equilibrium turnout is 56 percent in large electorates.

The distributions depicted in Figure 3 differ substantially from one another, but in every case the participation threshold remains bounded well below one. In fact, voter turnout in every case is well within the range observed in actual public elections: approximately half of the electorate votes, while the other half abstains. Perhaps surprisingly, this implies that even citizens of below-average expertise continue to vote. In the case of the first distribution, such citizens vote despite the presence of peers who are arbitrarily well informed. In the third distribution, the marginal citizen votes even though she is only barely better informed than a balanced coin. Figure 2 suggests that the examples depicted in Figure 3 are not particularly special: the function $T^2/T^2+(1-T)^2$ first rises steeply, and then flattens out between $1/\sqrt{2}$ and 1, making it likely that, for a variety of distributions, $E(q_i|q_i \geq T)$ will intersect $T^2/T^2+(1-T)^2$ at some threshold close to 1/2.

From Equation (16), the margins of victory associated with the three example distributions depicted in Figure 3 can be computed as 70 percent, 52 percent, and 1 percent. In real-world public elections, of course, the last of these is most typical. This suggests that, of the three distributions, the last is actually likely to be the most realistic. More generally, rewriting (16) as

$$E\left(q_i|q_i > T^*_\infty\right) = 0.5 + 0.5\mu$$

makes clear that an expected margin of victory smaller than, say, five percent, requires that the average vote quality not exceed 0.525. At the same time, the first two examples illustrate that the margin of victory can also be quite high, when the average vote quality for a particular issue is high—a point emphasized in Sections 4.5 and 6.

In addition to the continuous distributions depicted in Figure 3, Equation (14) applies to discrete distributions, as Section 5.1 emphasizes later. Feddersen and Pesendorfer's (1996) model (without partisans), for example, corresponds to a discrete distribution consisting of two mass points, at $q = 1/2$ and $q = 1$. In that case, $E(q_i|q_i \geq T) = 1$ for any $T > 1/2$, so $T^*_\infty = 1$, and only perfectly informed citizens vote in equilibrium. In contrast, the classic model of Condorcet (1785) corresponds to a degenerate distribution around $q_i = q$, for which $E(q_i|q_i \geq T) = q$ for any $T < q$. In Figure 2, this can be illustrated as a horizontal line at height q, connecting the y-axis with the 45° line, which clearly intersects $T^2/T^2+(1-T)^2$ at $T^*_\infty < q$, implying that everyone votes in equilibrium.

4.4. *Comparative Statics*

Section 4.2 demonstrates that, in large electorates, equilibrium turnout and the margin of victory are uniquely determined by the underlying distribution of information—at least if the density f is log-concave. Maintaining that assumption, this section compares the limiting equilibrium quality

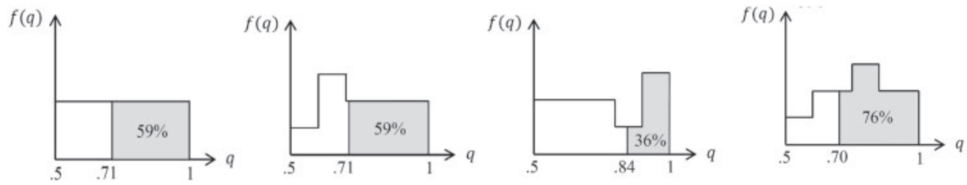

FIGURE 4
The offsetting effects of information improvements on voter turnout

thresholds $T^*_{\infty F}$ and $T^*_{\infty G}$ associated with two information distributions, along with equilibrium turnout τ_F and τ_G and margins of victory μ_F and μ_G. Proposition 3 begins by comparing an electorate F with another G that is better informed, in the sense of first-order stochastic dominance.[24] The result is ambiguous: turnout can be higher in either electorate. This is because improving an individual's information has two offsetting effects: it makes that individual more willing to vote, but makes her peers more willing to abstain. In terms of the model, it lifts nonvoters above the participation threshold, but also causes the participation threshold to rise. The impact on the margin of victory is similarly ambiguous: improving voter expertise increases the superior policy's margin of victory, but as nonvoters cross the participation threshold they become voters of below-average expertise, so the margin of victory falls.

The net effect of improved information depends on whose information improves. Proposition 3 demonstrates this by delineating three cases that are unambiguous. These are illustrated in Figure 4, starting from the uniform distribution of Figure 3: (1) improving nonvoter's information has no effect, if it does not lift them above the participation threshold; (2) improving voter's information raises the margin of victory and strengthens the swing voter's curse, thereby lowering turnout; (3) moderate improvements in nonvoter's information increase turnout by pushing nonvoters above the participation threshold, thereby lowering the margin of victory and weakening the swing voter's curse so that turnout increases even further.[25]

Proposition 3. *Let F and G have log-concave densities f and g, and suppose that G first-order stochastically dominates F. Then the following are true:*

1. If $g(q)=f(q)$ *for all* $q \geq T^*_{\infty F}$ *then* $T^*_{\infty G}=T^*_{\infty F}$, $\tau_G=\tau_F$, *and* $\mu_G=\mu_F$.

2. If $G\left(T^*_{\infty F}\right)=F\left(T^*_{\infty F}\right)$ *then* $T^*_{\infty G}>T^*_{\infty F}$, $\tau_G<\tau_F$, *and* $\mu_G>\mu_F$.

3. If $g(q)\geq f(q)$ *for all q between* $T^*_{\infty F}$ *and* $E_F\left(q_i|q_i \geq T^*_{\infty F}\right)$, *and* $g(q)=f(q)$ *for all* $q \geq E_F\left(q_i|q_i \geq T^*_{\infty F}\right)$, *then* $T^*_{\infty G}<T^*_{\infty F}$, $\tau_G>\tau_F$, *and* $\mu_G<\mu_F$.

Proposition 3 considers a distribution that is better informed than F. Proposition 4 now considers the possibility of a distribution G that is more homogeneous than F. Specifically, let F be a *simple mean-preserving spread* of G, as defined by Diamond and Stiglitz (1974), meaning that the two distributions have the same mean, but the cumulative distribution functions cross only once.[26] As Rothschild and Stiglitz (1970) show, this implies that G has a smaller variance

24. As defined by Haldar and Russell (1969), *G first-order stochastically dominates F* if, for any quality level q, the fraction of citizens with information quality better than q is higher under G than under F (*i.e.*, $1-G(q)\geq 1-F(q)$). Thus, signals are more precise, in the sense of Blackwell (1951).

25. Conversely, reducing nonvoter's information has no effect; small reductions in voter's information increase turnout and reduce the margin of victory; and moderate reductions in voter's information reduce turnout and increase the margin of victory.

26. Diamond and Stiglitz (1974) also call this the *single-crossing property*. More recently, Johnson and Myatt (2006) refer to G as a (counter-clockwise) *rotation* of F.

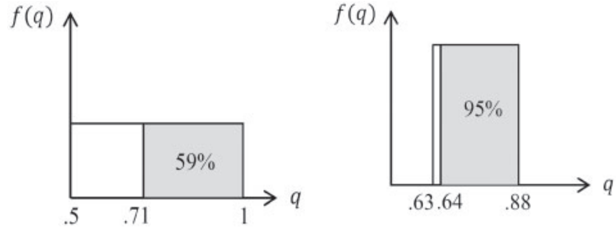

FIGURE 5
The inverse relationship between variance and turnout

than F. The result of Proposition 4, illustrated in Figure 5, is that turnout is higher under G than F, and the margin of victory is lower.

Proposition 4. *Let F and G have log-concave densities f and g and a common mean, and suppose that* $G(q) \le F(q)$ *for all* $q \le \hat{q}$ *and* $G(q) \ge F(q)$ *for all* $q \le \hat{q}$*, for some* $\hat{q} \ge T^*_{\infty F}$*. Then* $T^*_{\infty G} \le T^*_{\infty F}$, $\tau_G > \tau_F$*, and* $\mu_G < \mu_F$.

One intuition for Proposition 4 is that the swing voter's curse is weak when the quality difference between informed and uninformed votes is small, and strong when the quality difference is large. The most extreme case is Condorcet's (1785) model, in which voters are identical (*i.e.*, q_i has zero variance) and turnout is 100 percent.

The statements of Propositions 3 and 4 may seem to indicate a negative correlation between turnout τ and the margin of victory μ, but this relationship does not hold generally. In Figure 3, for example, the third information distribution exhibits a much smaller margin of victory than either of the first two, but at the same time exhibits voter turnout that is lower than in the first electorate and higher than in the second.

4.5. *Welfare*

As Figure 3 illustrates, many citizens continue to vote in large elections, even though their own expertise is extremely limited (*e.g.*, $q_i = 0.503$), and even though their peers possess much better (*e.g.*, almost perfect) information. It may seem likely that such poorly informed votes can only reduce the accuracy of collective decisions, perhaps vindicating efforts to discourage (or prevent) such citizens from voting. On the other hand, since every signal is informative, the opposite intuition that more signals are always better than fewer signals might seem to endorse efforts to increase turnout, for example by instilling a sense of civic duty, or by making voting compulsory. After all, voluntary elections fail to utilize nonvoters' information.

Contrary to either of these intuitions, Theorem 5 states that the probability of the desired election outcome is maximized in equilibrium. Since equilibrium voting necessarily involves some voter abstention, this implies that it is never optimal for all citizens to vote; adding votes beyond the equilibrium level actually reduces welfare, rather than enhancing it. On the

Hadar and Russell's (1969) weaker condition that G second-order stochastically dominates F (*i.e.* $\int_{1/2}^{T} G(q)dq \ge \int_{1/2}^{T} F(q)dq$ for all $T \in \left[\frac{1}{2}, 1\right]$) is not sufficient for Proposition 4. The uniform distribution illustrated in Figure 3 second-order stochastically dominates (and therefore has lower variance than) the discrete distribution given by $P\left(\frac{1}{2}\right) = P(1) = 0.25$ and $P\left(\frac{3}{4}\right) = 0.5$, for example, but implies a higher limiting participation threshold, lower turnout, and a higher margin of victory.

other hand, discouraging participation by nonexperts also reduces welfare: citizens who vote in equilibrium—including those in Figure 3—actually do improve election accuracy. A third implication of Theorem 5 is that improving information can never reduce welfare, even though (as Section 4.4 shows) it may increase or decrease voter turnout: the direct effect of improved information is improved election accuracy, and any equilibrium response by voters only improves welfare further.[27]

Theorem 5. *There exists a strategy σ^{**} that maximizes the probability of electing the superior policy. Furthermore, σ^{**} is a quality threshold strategy, with quality threshold $1/2 < T^{**} < 1$, and constitutes a Bayesian Nash equilibrium.*

Showing the existence of an optimal voting strategy is not trivial, since the space of strategies is not compact under the standard topology. The proof of Theorem 5 proceeds by first demonstrating that every strategy is dominated by a belief threshold strategy. The set of possible belief thresholds is compact, so an optimal strategy within this class is more straightforward. That the optimal strategy constitutes an equilibrium follows from McLennan's (1998) observation that, in common interest games such as this, whatever is optimal for the group is also optimal for each individual. That σ^{**} is a quality threshold strategy with $1/2 < T^{**} < 1$ then follows from Theorem 2 and Proposition 2.

To understand how adding informative votes can reduce the quality of an election outcome, note the inherent trade-off between the quantity and quality of information: lowering the participation threshold increases the number of votes (thereby reducing the variance of the election outcome) but reduces the average quality of a vote (thereby reducing the superior policy's expected margin of victory). It is also useful to recognize that voters actually possess two pieces of private information, s_i and q_i. A first-best election mechanism would obtain both, and weight individual votes by their underlying quality, in a maximum-likelihood approach (Nitzan and Paroush 1982; Shapley and Grofman 1984). A standard election, however, weights votes equally. Abstention allows poorly informed citizens to effectively transfer weight from their own signals to those with higher quality. Put differently, a citizen's vote communicates s_i, while her *decision* to vote communicates information about q_i, that would be lost if all citizens were to vote. Ultimately, some information is lost because signals are continuous but the action space for voters is inherently discrete.[28] This contrasts, for example, with models of capital markets, where infinitely divisible assets enable participants to trade asset quantities precisely in proportion to the quality of their information, and private information is aggregated completely.

Despite this coarsening of information, elections do aggregate information effectively in the limit. Remark 13 states that, just as in Condorcet's (1785) original model, the equilibrium probability of electing the superior policy exceeds 1/2, and approaches 1 as the electorate grows large. The logic is well known: the superior policy's expected vote share exceeds 1/2, and in large electorates, the actual vote share converges to its expectation.

Remark 1 (Jury theorem) *If* $\left\{\sigma_n^*\right\}$ *is a sequence of equilibria then* $\Pr\left(A|\alpha;\sigma_n^*\right) \geq 1/2$ *and* $\Pr\left(B|\beta;\sigma_n^*\right) \geq 1/2$ *for all n, and* $\lim_{n\to\infty}\Pr\left(A|\alpha;\sigma_n^*\right) = \lim_{n\to\infty}\Pr\left(B|\beta;\sigma_n^*\right) = 1$.

27. By similar reasoning, welfare also increases with the size of the electorate. As note 14 emphasizes, another important implication of Theorem 5 is that the competing assumptions of strategic and ethical voting are behaviorally equivalent.

28. One way to make voting less discrete is to allow voters to cast multiple votes. Another is to score alternatives numerically, as judges do in athletic or music competitions. In such systems, a voter can partially defer to her peers by voting for multiple policy alternatives, or awarding similar scores.

The limiting result of Remark 13 does not depend crucially on the exact level of voter participation, but in finite electorates, Theorem 5 provides important policy guidance for maximizing the likelihood of desired election outcomes. Specifically, efforts to increase voter participation, or to dissuade citizens with limited information from voting, may both be misguided.[29] Improving nonvoter's information can improve election outcomes, but only if sufficient learning takes place to push these individuals above the participation threshold. Improving voter's information unambiguously improves election outcomes, but will cause voter participation to fall. Thus, voter turnout seems a less useful measure of the quality of election decisions than is commonly perceived.

A better gauge of quality might be the margin of victory μ, which in this model is a linear function of the average quality $E\left(q_i|q_i \geq T^*_\infty\right)$ of a vote, and is therefore largest when decisions are obvious, or voters are quite well informed.[30] In this light, supermajority voting rules in settings such as criminal jury verdicts and constitutional amendments can be viewed as precautions against acting on the basis of insufficient information. In a similar vein, McMurray (2012) proposes a model in which informative margins of victory in public elections convey electoral "mandates" to political officials.

5. ROBUSTNESS

5.1. *Information Structure*

In Section 3, citizens receive two independent pieces of private information: a continuous quality variable q_i and a binary signal s_i. A more standard formulation of private information instead posits a single, continuous, signal. These approaches may appear quite different, but they are actually quite similar. Specifically, the present formulation imposes one restriction on the more general framework, which is that the distribution of signals is symmetric, so that posteriors θ and $1-\theta$ are equally likely. In that case, θ_i can be decomposed according to its strength q_i and direction s_i, as follows: if θ_i exceeds $1/2$, let $s_i = b$ and $q_i = \theta_i$; otherwise, let $s_i = a$ and $q_i = 1-\theta_i$. It is then straightforward to verify that, as in the model above, q_i is independent of s_i (and of ω). This decomposition is useful for expressing the comparative statics results of Section 4.4.

In addition to symmetric signals, the model above assumes symmetric prior beliefs $\Pr(\alpha) = \Pr(\beta)$ and utility gains $u(A|\alpha) - u(B|\alpha) = u(B|\beta) - u(A|\beta)$ in the two states, as well as a symmetric tie-breaking rule. All of this symmetry serves to simplify the model's exposition, but is not essential to the structure of equilibrium. Relaxing symmetry, Proposition 1 would still characterize best response voting as a belief threshold strategy, and Brouwer's theorem would imply the existence of a pair of equilibrium belief thresholds, even if those thresholds are no longer symmetric around $1/2$, so that equilibrium voting does not exactly reduce to a quality threshold strategy. When almost indifferent between the two policies, a citizen would still prefer to abstain.[31] The logic of Theorem 3 would then apply to both thresholds, so that turnout in the

29. The result that equilibrium turnout is socially optimal stems partly from the assumptions that voting and acquiring private information are both costless. With costly voting, turnout is inefficiently low, so raising turnout slightly can improve welfare. Even then, however, Theorem 5 implies that such efforts go too far if they raise turnout beyond the level that arises in equilibrium in costless environments; for example, voluntary voting is superior to mandatory voting in the costly environment of Krishna and Morgan (2010). In his model of endogenous information acquisition, Triossi (2011) does not analyze the welfare associated with a voluntary voting equilibrium, but does note that full participation is not optimal.

30. This, too, is imperfect, however. When nonvoter's information improves slightly, for example, as in part 3 of Proposition 3, welfare improves but $E\left(q_i|q_i \geq T^*_\infty\right)$ and μ fall.

31. McMurray (2012) treats the case of asymmetric utility, and Krishna and Morgan (2010) treat the case of asymmetric signal quality (with homogeneous voters). Other types of asymmetry could be treated similarly.

limit would remain positive (in both states). As a conjecture, it seems that smoothness conditions analogous to those of Theorem 4 would also imply that the limiting pair of equilibrium thresholds is unique.

Beyond symmetry, Section 3 assumes a continuous distribution of expertise. Among other things, this implies that signals may be arbitrarily informative. In other economic settings, such as auctions (Wilson 1977; Milgrom 1979), (Duggan and Martinelli 2001; Meirowitz 2002), and social learning models (Smith and Sorensen 2000), this is a key requirement for information aggregation, because it makes an individual willing to trust her own information even above a unanimous consensus among her peers.[32] In the majority election setting of this article, however, this condition is unimportant: if q_i were bounded above by $\bar{q}$ then the logic of Theorem 3 would imply that $T_\infty^* < \bar{q}$ instead of $T_\infty^* < 1$, so a positive fraction of the electorate would still continue voting in the limit. The assumption of full support merely emphasizes the main result that citizens of only moderate expertise continue voting in large electorates, *even* when their peers are arbitrarily well informed.

A continuous F also excludes the possibility of mass points, but this, too, is purely for simplicity. Above and below the participation threshold, mass points merely add a mass of voters or nonvoters.[33] Intuitively, it may seem that an exception must be made for a mass point at $q=1$, since surely citizens should defer in the limit to the positive fraction of the electorate who are perfectly informed. Even in that case, however, Theorem 3 remains valid, implying that moderately informed citizens continue voting, even in the limit. To see this, recall that T_∞^* remains bounded below 1 in the proof of Theorem 3 because $T^2/T^2+(1-T)^2$ approaches 1 with a slope approaching 0 while $E(q_i|q_i \geq T)$ approaches 1 more gradually, with slope approaching 1/2. A distribution with a mass point at one can be thought of as a weighted average of the continuous distribution analyzed earlier and a unit mass of perfectly informed citizens, so the slope of $E(q_i|q_i \geq T)$ is simply a weighted average of 0 and 1/2, which is still positive. If expertise were uniformly distributed for 80 percent of an electorate but the other 20 percent were infallible, for instance, then the left-hand side of (14) would be $E(q_i|q_i \geq T) = 0.8\frac{T+1}{2} + 0.2$, implying that $T_\infty^* = 0.75$. In other words, half of those with noisy signals would still vote, implying 60 percent turnout. Intuitively, moderately informed citizens vote in finite electorates to hedge against the possibility that perfectly informed types are not realized, and then continue voting in the limit because not all equilibrium voters are infallible.

As a final note, the model of Section 3 assumes that each citizen knows her precise position within the distribution of expertise, which may seem implausible. However, it is straightforward to reformulate posterior beliefs $\theta(\omega|q,s)$ or vote probabilities $v_x(\omega)$ to accommodate uncertainty regarding either one's own q_i or the prevailing distribution F of expertise, without substantively altering the subsequent analysis. Incorrect beliefs (*e.g.*, overconfidence) regarding one's own or others' expertise can also be accommodated, though this requires careful assumptions about a citizen's perception of the beliefs of her peers (and beliefs about beliefs, etc.). If a citizen is aware

32. Without this condition, Feddersen and Pesendorfer (1998) predict erroneous jury verdicts, because a juror trusts the unanimous "guilty" verdict of her peers above her own private opinion that a defendant is innocent. Similarly, Banerjee (1992) and Bikhchandani *et al.* (1992) predict "herding" to a (possibly erroneous) permanent consensus, because individuals ignore their private signals, and follow previous movers. In large auctions, as Pesendorfer and Swinkels (1997) emphasize, the winning bid only approximates the auction item's true value if a bidder can trust her own estimation even when it exceeds every other bidder's.

33. If a threshold and mass point coincide, equilibrium requires mixed strategies. Also, violating the smoothness condition of Theorem 4 can induce multiple equilibria.

of others' overconfidence but not her own (and this is common knowledge), for example, then the nature of equilibrium is unchanged, but participation is higher.[34]

5.2. *Heterogeneous Preferences*

The model of Section 3 makes the strong assumption that voter preferences are perfectly correlated, so that if informational differences could somehow be resolved, voting would be unanimous. This section extends the model to allow an imperfect (but still positive) correlation. To this end, suppose that in state α a fraction $\rho \geq 1/2$ of the electorate prefer policy A, and in state β the same fraction prefer policy B.[35] The model of Section 3 is then the special case of $\rho = 1$. Previously, the superior policy was elected if $v_+ > v_-$, which also guaranteed that citizens abstain in equilibrium. As formulated in (6) and (7), however, v_+ and v_- now reflect the probabilities with which citizen i votes for and against her *own* preferred policy. Two individuals now only share a common interest with probability $\rho^2 + (1-\rho)^2 \geq 1/2$, so citizen i votes for and against j's preferred policy with probabilities v_{j+} and v_{j-}, defined as follows.

$$v_{j+} = \left[\rho^2 + (1-\rho)^2\right] v_+ + [2\rho(1-\rho)] v_-$$
$$v_{j-} = \left[\rho^2 + (1-\rho)^2\right] v_- + [2\rho(1-\rho)] v_+.$$

If $\rho < 1$, the difference between v_{j+} and v_{j-} is smaller than the difference between v_+ and v_-, implying that the election is closer and abstention is lower. As long as $\rho > 1/2$, however, $v_+ > v_-$ implies that $v_{j+} > v_{j-}$, so abstention and information aggregation occur, just as before.

The assumption that preferences are positively correlated reflects an assumption that policies provide public goods, which by definition benefit large fractions of the electorate. Mueller (2003, ch. 3, 14) argues that even seemingly zero-sum issues may be viewed as public goods, as altruism or insurance motivations mitigate otherwise obvious conflicts of interest. Wealthy individuals may favor redistribution, for example, either out of concern for the poor, or for fear that they themselves may become poor in the future.[36] In Mueller's formulation, an altruist's objective function $U_i = u_i + \lambda \sum_{j \neq i} u_j$ places positive weight $\lambda > 0$ on the utility u_j of each of her peers, in addition to her own utility u_i. This can be rewritten as a weighted average $U_i = (1-\lambda) u_i + \lambda n \bar{u}$ of her own utility and the average utility $\bar{u} = \sum_{j=1}^{n} u_j$ in the population. If all citizens are altruists, the second term of this function is common, thereby inducing a correlation between U_i and U_j. In fact, this correlation may be quite high even if λ is close to zero, because the common term is proportional to the size n of the electorate.[37]

Existing literature provides several alternative specifications of preference heterogeneity. In every case, the common theme is that conflicts of interest reduce a citizen's willingness to defer

34. This can be seen from Figure 1: if citizens of type q believe themselves to have expertise $q^*(q) > q$ then equilibrium occurs where the best-response function intersects $q^*(q)$, which by definition is above the 45° line, and therefore to the left of the original T^*.

35. With this formulation, private signals are *exchangeable*, as in Milgrom and Weber (1982).

36. Mueller (2003, ch. 14) cites several examples of groups that vote contrary to their own interests, narrowly defined. Anecdotally, expenditures on public education are quite popular even among families with no school-aged children, and programs such as food aid and disaster relief remain quite popular, despite servicing only small segments of the electorate.

37. Edlin *et al.* (2007), Evren (2010), and Faravelli and Walsh (2011) use similar formulations of altruism to explain voting in costly elections. For a summary of experimental evidence of altruism, see Fehr and Schmidt (2006).

to the judgment of her peers, but that some abstention still occurs as long as some homogeneity is maintained. For example, Feddersen and Pesendorfer (1996) consider partisan supporters for both policies, in addition to the independents whose preferences depend on the state. This increases turnout, both because partisans all vote, and because some uninformed independents vote against the majority partisan group, to mitigate its influence. Some independents still abstain, however, if partisan groups are sufficiently small (and balanced). Kim and Fey (2007) add adversarial voters, who prefer policy B in state α and policy A in state β, with similar effect. Feddersen and Pesendorfer (1999) consider a model with multiple states of the world, and with a continuum of private biases, the most extreme of which behave essentially as partisans from their earlier model. If biases are large, relative to the utility differences between states (or, equivalently, if the state space is sufficiently fine), then abstention vanishes as the electorate grows, because citizens increasingly vote on the basis of private biases instead of abstaining, relying on their peers to provide commonly valued information. If biases are small (or the state space is coarse) then abstention occurs as before. By comparison to these various models, the condition above that abstention occurs as long as $\rho > 1/2$ is quite mild.

Admittedly, seemingly partisan behavior such as voting along political party lines, with little attention to the details of particular issues or candidates, is commonplace in public elections. On the other hand, McMurray (2012) points out that "partisan" behavior and "ideological" differences can also be interpreted within the context of information: if A and B are *liberal* and *conservative* policies, for example, then citizens who believe strongly that one of the two policies is superior (*i.e.* q_i close to one) may be viewed as strongly liberal or conservative, while those with weaker opinions (*i.e.* q_i close to zero) may be viewed as weak partisans, or even independents.[38] An informational view of ideology can explain why positions drift over time, and sometimes even shift abruptly in response to new information or epiphanies.[39] Informational considerations can also explain persuasive efforts such as debate, policy research, endorsements, and advertising, which are only worthwhile if a citizen expects her peers (once informed) to adopt her own positions. The references in Goeree and Grosser (2007) also document an empirical tendency for voters on both sides of an issue to believe that they belong to the majority, consistent with the formulation of this section.[40]

6. EVIDENCE AND APPLICATIONS

The central empirical prediction of this article is that turnout remains bounded at moderate levels, even in large electorates, contrary to other models, which notoriously predict either that turnout should be quite low, or that everyone should vote. This section discusses a number of additional predictions that are consistent with empirical observation. In some cases, these provide insight into trends that have previously been viewed as puzzles. In other instances, they provide alternatives to existing explanations, implying a need for more detailed empirical analysis. Evidence in Section 6.1 relates to a citizen's use of her own information; evidence in Section 6.2 relates to a strategic response to the information of others.

38. The description of independents as slightly partisan is consistent with Flanigan and Zingale's (1998) finding that self-declared independents who "lean" in favor of one candidate or the other resemble weak partisans as closely as strong and weak partisans resemble each other.

39. See the references cited in McMurray (2012).

40. The same pattern is evident in Fischer (1999).

6.1. *Information*

The central feature of equilibrium in the model above is that the strength of a citizen's opinion determines whether she votes or abstains. This is consistent with a growing body of empirical evidence that finds information variables to be key determinants of voter participation. In Delhi, India, for example, Banerjee *et al.* (2010) report that individuals who receive candidate "report cards" are more likely to vote, especially when incumbent politicians have performed poorly. Lassen (2005) reports higher turnout in a referendum on government restructuring in Copenhagen, Denmark, in districts randomly chosen to experiment with the policy prior to the referendum. McMurray (2011) cites numerous studies that find participation to be correlated with political knowledge, education, age, access to news media, and contact from campaign workers, among others. Palfrey and Poole (1987) and Delli Carpini and Keeter (1989) also find that citizens with strong ideological stances vote more frequently, and are on average better informed, than those with weaker positions, consistent with the informational interpretation of ideology described in Section 5.2.

The traditional explanation for abstention is that voting requires time, which is costly. Voting costs also provide a simpler explanation for the empirical connection between information and voting, since as Matsusaka (1995) points out, uncertainty reduces the expected benefit of voting, making it less likely to be worth the cost. As Feddersen and Pesendorfer (1996) point out, however, abstention is not limited to costless environments. For example, voters frequently "roll off" by voting in some races, but skipping others on the same ballot, even though voting costs have already been paid.[41] Empirically Battaglini *et al.* (2008, 2009) demonstrate strategic abstention in a laboratory environment, and Wattenberg *et al.* (2000) confirm that citizens who continue voting are better informed, on average, than citizens who roll off. Another challenge to the costly voting theory is Downs' (1957) well-known observation that voting costs should dissuade all but a small fraction of the electorate from voting. A common assumption is that voters are motivated by a sense of civic duty, as Riker and Ordeshook (1968) suggest, but as McMurray (2011) explains, the theoretical link between information and voting breaks down in that case: regardless of information, dutiful citizens should all vote, which should lower the pivot probability so that others all abstain. By contrast, as Section 4.3 emphasizes, participation rates in this model resemble those of actual public elections.

Another result of the informational nature of voting in this model is that, as Sections 4.3 and 5.2 illustrate and emphasize, citizens' votes are correlated with one another, so margins of victory may be quite large. Specifically, margins are large when issues are in some sense obvious (*i.e.*, average vote quality is high), such as ballot initiatives to update archaic government procedures or constitutional language, or elections to retain incumbent candidates who have clear records of high-quality performance. Empirically, margins are high in precisely these situations: historically, for example, the average margin of victory for U.S. gubernatorial elections was 23 percent (Mueller 2003, ch. 11); for many offices, popular incumbents even run for reelection unopposed. Coate *et al.* (2008) report large margins of victory in local referenda over Texas liquor laws, but interpret this as evidence against rational voting models, because existing models often predict exact ties in expectation.[42]

41. Abstention is also commonplace in committees, where voting requires only the raise of a hand.

42. This occurs in Krasa and Polborn (2009), for example, even if supporters of one policy far outnumber supporters of the other, because free-riding incentives are more severe in the majority group.

6.2. *Strategic Effects*

An equilibrium prediction that is more subtle, but still quite central to the logic of this article, is that while her own information makes a citizen more willing to vote, the information of her peers makes a citizen more willing to abstain. Direct evidence of this strategic delegation can be seen in the laboratory experiments of Morton and Tyran (2011). In public elections, this prediction provides a possible explanation for Brody's (1978) otherwise puzzling observation that turnout in U.S. national elections has declined during decades in which education levels were rising.[43] Having forged a theoretical link between information and voting, for example, both Matsusaka (1995) and Feddersen and Pesendorfer (1996) predict that improving education should improve information, and therefore raise voter turnout. In this model, by contrast, improving information can also cause turnout to fall, as Proposition 3 demonstrates.

Another implication of strategic voting is that, as Lemma 1 states, citizens increasingly abstain as the electorate grows large, in deference to those with better information. Consistent with this prediction, Geys (2006a) reports that about 60 percent of empirical studies find turnout and election size to be negatively correlated. The traditional interpretation of this finding is that citizens are unwilling to pay voting costs in large elections, because a vote is less likely to be pivotal. The alternative explanation of strategic abstention, however, also provides a possible explanation for the 40 percent of studies that find no significant correlation, since turnout rates in this model converge quickly to the limit. Farber (2010) finds evidence consistent with this in firm-level votes on whether workers should unionize or not: turnout rates first decline with the number of eligible voters, but then flatten out as elections continue to grow.[44] The strategic theory also implies the same pattern in costless environments, such as roll off, suggesting a possible avenue for future empirical work.

Like small elections, turnout is also higher in close elections, at least in two-thirds of the studies reviewed by Geys (2006a). This is traditionally interpreted as further evidence that a citizen only votes when her vote is sufficiently likely to be pivotal, but again, the analysis of this article suggests a new possibility: for distributions with a high equilibrium threshold, turnout is low, but average vote quality—and therefore the margin of victory—are high; for distributions with a low threshold, the reverse is true. Thus, in the statements of Propositions 3 and 4, turnout and the margin of victory move in opposite directions. While this seems to be the general rule, it is not universal, as Figure 3 illustrates, perhaps explaining why more studies do not find a positive correlation. Also, while a more standard theory would seem to predict that the correlation between closeness and voter turnout should diminish as an electorate grows large, since in that case pivot probabilities are small regardless of closeness, Farber (2010) finds exactly the opposite. Strategic effects provide a possible explanation for this finding, since the comparative statics results that are derived in Section 4.4 for large electorates are likely muted in small electorates, where the need for information quantity outweighs the need for quality, so a citizen votes regardless of the average vote quality of her peers (and thus regardless of the margin of victory).

The above applications highlight the ambiguous relationship between information and voter participation (or margins of victory) at the aggregate level. In some cases, however, macro evidence exhibits much less ambiguity: voter turnout is consistently higher in national elections than in state and local elections, for example, where candidates are more obscure, and is higher in general elections than in party primaries, where voters cannot infer candidate positions from party

43. Aldrich (1993) calls this observation "the most important substantive problem in the turnout literature". McDonald and Popkin (2001) argue that declining turnout is an illusory effect of decreasing voter elligibility, but the puzzle remains that education levels rose during periods that turnout did not.

44. A similar pattern can be seen in Coate *et al.* (2008).

affiliations. The positive correlation between information and turnout at the macro level is not inconsistent with Proposition 3, of course, but suggests the need for deeper analysis. To this end, the nuance of Propositions 3 and 4 may be instructive: in local elections, for example, the variance of information seems likely to be high, because candidates may be personally acquainted with some voters and completely unknown to others, so low participation is to be expected. Similarly, the informational difference between primary and general elections seems largest for those who do *not* participate in the primaries, suggesting that part 3 of Proposition 3 applies, and higher turnout in the general election is thus the natural prediction.

7. CONCLUSION

The model analyzed in this article provides a deeper and richer analysis of information aggregation in elections than has previously been available. Most basically, it corroborates Condorcet's (1785) classic insight that election outcomes may be quite well informed, even if individuals are not. Because of this, however, a citizen with relatively little expertise has an equilibrium incentive to ignore her own information, and abstain from voting, in deference to those with better information. This is socially optimal, because it increases the impact of the best-informed votes. At the same time, however, citizens with only limited expertise continue voting, even in large electorates. This reduces average vote quality, but also increases the quantity of votes, thereby reducing the likelihood of a collective mistake. Thus, improving information increases welfare, but may cause turnout either to rise or to fall, and efforts to increase voter turnout, or to dissuade or prevent citizens with limited expertise from voting, may both be misguided.

Understandably, the dual message that voter participation can be either too high or too low may leave a citizen feeling conflicted as to whether she should vote according to her tentative private opinion, or abstain in deference to those with more complete information. Assuming that the rest of the electorate votes optimally, a simple rule of thumb could be for a citizen to anticipate voter turnout, and compare herself to the marginal voter. If she expects 75 percent voter participation, for example, she should vote as long as she is better informed than the bottom quartile of the electorate; otherwise, she should abstain. If she expects 50 percent turnout, she should vote if her expertise is above the median, and otherwise abstain.

Beyond the analysis of this article, there are a number of important directions in which the model above should be extended and explored. Examples include the analysis of costly voting, and a more complete treatment of preference heterogeneity—issues that are central to political-economic research, but typically not analyzed in combination with commonly valued private information.[45] Another example is the analysis of voters' preliminary decisions to acquire (costly) private information, which in the present model is provided exogenously.[46] Other possibilities include heterogeneous beliefs or correlated private signals, which could arise if voters commit cognitive errors in processing available information, or communicate prior to an election.[47] Also, the model could be extended to allow multiple states of the world, reflecting the possibility that the quality difference between the superior and inferior policies may be either large or small.[48]

45. Krishna and Morgan (2010, 2011) make progress in these directions.

46. Martinelli (2006, 2007), Triossi (2011), and Oliveros (2011b) show that information costs may prevent elections from identifying good policies. Oliveros (2011a) shows that information acquisition decisions can lead to a nonmonotonic relationship between information and voter participation.

47. A number of recent papers study how dependence influences information aggregation (*e.g.* see Peleg and Zamir 2010), but not how it influences other election outcomes, such as participation.

48. Feddersen and Pesendorfer (1999) allow this possibility in their extension to private values; McMurray (2012) considers a model in which the optimal policy is drawn from a continuum of alternatives.

Ultimately, extensions such as these are unlikely to change the fundamental results that a citizen's own information makes her more willing to vote, while the information of her peers makes her more willing to abstain, so that at the macro level, voter turnout and the margin of victory are determined entirely by the *distribution* of expertise. As Section 6 emphasizes, these predictions are broadly consistent with—and provide novel explanations for—much of the available evidence from actual public elections. In particular, this can explain why voter turnout remains at moderate levels, even in large elections. Explaining empirical patterns is useful in its own right, and also suggests that voters do value the expertise of their fellow voters. This then corroborates Condorcet's (1785) informational view of elections, as mechanisms for pooling collective wisdom to identify good policies and candidates.

APPENDIX

A. PROOFS

Proposition 2 (Swing voter's curse) *σ_i^* is a best response to the quality threshold strategy σ_T only if σ_i^* is a quality threshold strategy $\sigma_{T^{br}}$, with quality threshold $T^{br}=\tilde{P}/P+\tilde{P}$. Furthermore, $1/2\leq T^{br}<1$, with strict inequality if $T<1$.*

Proof That the best response to σ_T is another quality threshold strategy, with threshold T^{br} defined in (12), is argued in the text preceding the statement of the proposition. If $T=1$ then σ_T prescribes abstention for citizens of all types (*i.e.* $v_+=v_-=0$), which guarantees a tie (*i.e.* $\pi_0=1$), implying that a vote for either policy is equally likely to be pivotal (*i.e.*, $P=\tilde{P}=1/2$), and a citizen of any information quality should vote in response (*i.e.* $T^{br}=1/2$). If instead $T<1$, a positive fraction of the electorate votes. Since voters honestly report their signals, which are positively correlated with the truth, the superior policy receives a larger expected vote share than the inferior policy (*i.e.* $v_+-v_-=\int_T^1(2q-1)\,dF(q)>0$), and is therefore more likely to be ahead by a single vote than behind by a single vote (*i.e.* $\pi_1/\pi_{-1}=v_+/v_->1$). This implies, however, that an additional vote for the superior policy is less likely to be pivotal than a vote for the inferior policy (*i.e.* $\tilde{P}-P=1/2(\pi_1-\pi_{-1})>0$), and the result that $T^{br}>1/2$ follows from (12). That $T^{br}<1$ for any T follows from (12) because P is strictly positive, as $N=0$ with positive probability, resulting in a tie. ‖

Lemmas A1 and A2 are useful in preparation for the proof of Theorem 2.

Lemma A1. *If $T^{br}\geq T$ then $\frac{\partial}{\partial T}T^{br}(T)\geq 0$.*

Proof For any quality threshold strategy σ_T, let ψ_k denote the probability of precisely k votes for each policy:

$$\psi_k=\frac{e^{-nv_+}(nv_+)^k}{k!}\frac{e^{-nv_-}(nv_-)^k}{k!}=\frac{n^{2k}}{k!k!}e^{-n(v_++v_-)}(v_+v_-)^k. \tag{A.1}$$

With this notation, win probabilities can be rewritten as $\pi_0=\sum_{k=0}^{\infty}\psi_k$, $\pi_1=\sum_{k=0}^{\infty}\psi_k\frac{nv_+}{k+1}$, and $\pi_{-1}=\sum_{k=0}^{\infty}\psi_k\frac{nv_-}{k+1}$. Also, noting that $\pi_1=\frac{v_+}{v_-}\pi_{-1}$, define the ratio γ as follows, so that $\pi_1=\pi_0\gamma v_+$ and $\pi_{-1}=\pi_0\gamma v_-$.

$$\gamma=\frac{1}{v_+}\frac{\pi_1}{\pi_0}=\frac{1}{v_-}\frac{\pi_{-1}}{\pi_0}. \tag{A.2}$$

In terms of γ, $T^{br}\equiv\frac{\tilde{P}}{P+\tilde{P}}$ can be written as follows.

$$\begin{aligned}T^{br}&=\frac{\frac{1}{2}(\pi_0+\pi_0\gamma v_+)}{\frac{1}{2}(\pi_0+\pi_0\gamma v_-)+\frac{1}{2}(\pi_0+\pi_0\gamma v_+)}\\&=\frac{1+\gamma v_+}{2+\gamma v_-+\gamma v_+}.\end{aligned} \tag{A.3}$$

Implicitly, (6) through (11) depend on the underlying quality threshold T. Accordingly, let primed variables denote derivatives with respect to T. The derivative $\frac{\partial}{\partial T}T^{br}(T)$ depends on all of these derivatives, as follows,

$$v'_+ = -Tf \tag{A.4}$$

$$v'_- = -(1-T)f \tag{A.5}$$

$$\psi'_k = \psi_k(nf+kc_1) \tag{A.6}$$

$$\pi'_0 = \sum\nolimits_{k=0}^{\infty}\psi_k(nf+kc_1) \tag{A.7}$$

$$\pi'_1 = \sum\nolimits_{k=0}^{\infty}\frac{n}{k+1}\left(\psi'_k v_+ + \psi_k v'_+\right) \tag{A.8}$$

$$\pi'_{-1} = \sum\nolimits_{k=0}^{\infty}\frac{n}{k+1}\left(\psi'_k v_- + \psi_k v'_-\right) \tag{A.9}$$

$$\gamma' = \frac{c_2}{(v_+\pi_0)^2} \tag{A.10}$$

$$P' = \frac{1}{2}\left(\pi'_0+\pi'_{-1}\right) \tag{A.11}$$

$$\tilde{P}' = \frac{1}{2}\left(\pi'_0+\pi'_1\right) \tag{A.12}$$

$$\frac{\partial}{\partial T}T^{br}(T) = \frac{c_3}{(2+\gamma v_- + \gamma v_+)^2}, \tag{A.13}$$

where f is the density of expertise, $c_1 = \frac{v'_+}{v_+}+\frac{v'_-}{v_-}$, and c_2 and c_3 are given by

$$\begin{aligned}
c_2 &\equiv \pi'_1(v_+\pi_0) - \pi_1\left(v'_+\pi_0 + v_+\pi'_0\right)\\
&= v_+\pi_0\pi'_1 - v'_+\pi_0\pi_1 - v_+\pi'_0\pi_1\\
&= v_+\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\frac{n}{k+1}\left(\psi'_k v_+ + \psi_k v'_+\right)\\
&\quad -v'_+\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\psi_k\frac{nv_+}{k+1} - v_+\sum_{j=0}^{\infty}\psi'_j\sum_{k=0}^{\infty}\psi_k\frac{nv_+}{k+1}\\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(nf+kc_1)\frac{v_+^2 n}{k+1} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(nf+jc_1)\frac{v_+^2 n}{k+1}\\
&= v_+^2 nc_1\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(k-j)\frac{1}{k+1}
\end{aligned} \tag{A.14}$$

and

$$\begin{aligned}
c_3 &= \left(\gamma' v_+ + \gamma v'_+\right)(2+\gamma v_- + \gamma v_+) - (1+\gamma v_+)\left(\gamma' v_+ + \gamma v'_+ + \gamma' v_- + \gamma v'_-\right)\\
&= \left(\gamma' v_+ + \gamma v'_+\right)(1+\gamma v_-) - (1+\gamma v_+)\left(\gamma' v_- + \gamma v'_-\right)\\
&= \gamma v'_+(1+\gamma v_-) - (1+\gamma v_+)\gamma v'_- + \gamma'(v_+ - v_-)\\
&= -\gamma Tf\frac{2}{\pi_0}P + \gamma(1-T)f\frac{2}{\pi_0}\tilde{P} + \gamma'(v_+ - v_-)\\
&= \gamma f\frac{2}{\pi_0}\left(P+\tilde{P}\right)\left(\frac{\tilde{P}}{P+\tilde{P}} - T\right) + \gamma'(v_+ - v_-).
\end{aligned} \tag{A.15}$$

From (A.4) and (A.5) note that $v'_+ < v'_- < 0$ for any $T > \frac{1}{2}$, implying that $c_1 < 0$. Without the fraction $\frac{1}{k+1}$, the double sum in (A.14) would equal zero, because $\psi_j\psi_k(k-j)$ is positive whenever $k > j$, but the term with reversed indexes is negative and of equal magnitude. Dividing by $k+1$ places greater weight on negative than positive terms, so the double sum must be negative. Since c_1 is also negative, c_2 and γ' are both positive, and so is the second term of the sum in (A.15). When $\frac{\tilde{P}}{P+\tilde{P}} = T^{br} \geq T$, the first term is positive as well, implying that (A.13) is positive, as claimed. ‖

Lemma A2. *Let σ_{T_A,T_B} be a belief threshold strategy profile, with thresholds $T_A \le \frac{1}{2} \le T_B$. If $1-T_A > T_B$ then $\hat{\theta}_{0B} > T^{br}(1-T_A)$ and $\hat{\theta}_{A0} > 1-T^{br}(T_B)$; otherwise, both inequalities are reversed.*

Proof When $T_A \le \frac{1}{2} \le T_B$, citizens with a and b signals only vote A and B (or abstain), respectively, so $v_x(\omega;\sigma_{T_A,T_B})$ reduces from (2) to the following.

$$v_A(\alpha) = \int_{1-T_A}^{1} q\,dF(q) \qquad v_A(\beta) = \int_{1-T_A}^{1}(1-q)\,dF(q)$$
$$v_B(\alpha) = \int_{T_B}^{1}(1-q)\,dF(q) \qquad v_B(\beta) = \int_{T_B}^{1} q\,dF(q).$$

In that case, it is straightforward to verify that $v_A(\beta) < v_A(\alpha) < v_B(\beta)$, that $v_A(\beta) < v_B(\alpha) < v_B(\beta)$, and that $v_A(\beta) + v_B(\beta) > v_A(\alpha) + v_B(\alpha)$. Note that $v_A(\alpha)$ and $v_A(\beta)$ correspond to the vote probabilities $v_+(\sigma_{1-T_A})$ and $v_-(\sigma_{1-T_A})$ associated with the quality threshold strategy σ_{1-T_A}, while $v_B(\beta)$ and $v_B(\alpha)$ correspond to the vote probabilities $v_+(\sigma_{T_B})$ and $v_-(\sigma_{T_B})$ associated with σ_{T_B}. Therefore, the pivot probabilities $P(\sigma_{1-T_A})$ and $\tilde{P}(\sigma_{1-T_A})$ associated with σ_{1-T_A} can be written in terms of $v_A(\alpha)$ and $v_A(\beta)$, and the pivot probabilities $P(\sigma_{T_B})$ and $\tilde{P}(\sigma_{T_B})$ associated with σ_{T_B} can be written in terms of $v_B(\alpha)$ and $v_B(\beta)$. For example,

$$\begin{aligned}
P(\sigma_{1-T_A}) &= \frac{1}{2}\left[\pi_0(\beta;\sigma_{1-T_A}) + \pi_{-1}(\beta;\sigma_{1-T_A})\right] \\
&= \frac{1}{2}\sum_{k=0}^{\infty} \frac{n^{2k} e^{-n\left[v_-(\sigma_{1-T_A}) + v_+(\sigma_{1-T_A})\right]}}{k!k!}\left[v_-(\sigma_{1-T_A})v_+(\sigma_{1-T_A})\right]^k \left[1 + \frac{nv_-(\sigma_{1-T_A})}{k+1}\right] \\
&= \frac{1}{2}\sum_{k=0}^{\infty} \frac{n^{2k} e^{-n[v_A(\beta)+v_A(\alpha)]}}{k!k!}\left[v_A(\beta)v_A(\alpha)\right]^k \left(1 + \frac{nv_A(\beta)}{k+1}\right).
\end{aligned}$$

An equivalent condition to $\hat{\theta}_{0B} = \frac{\tilde{P}(\alpha;\sigma_{T_A,T_B})}{\tilde{P}(\alpha;\sigma_{T_A,T_B}) + P(\beta;\sigma_{T_A,T_B})} > T^{br}(1-T_A) = \frac{\tilde{P}(\sigma_{1-T_A})}{P(\sigma_{1-T_A}) + \tilde{P}(\sigma_{1-T_A})}$ is that $\tilde{P}(\alpha;\sigma_{T_A,T_B})P(\sigma_{1-T_A}) > \tilde{P}(\sigma_{1-T_A})P(\beta;\sigma_{T_A,T_B})$. Writing these probabilities in terms of $v_x(\omega;\sigma_{T_A,T_B})$ reveals that this inequality holds, given the inequalities above:

$$\begin{aligned}
&\tilde{P}(\alpha)P(\sigma_{1-T_A}) - \tilde{P}(\sigma_{1-T_A})P(\beta) \qquad \text{(A.16)}\\
&= \frac{1}{2}\sum_{j=0}^{\infty} \frac{n^{2j} e^{-n[v_A(\alpha)+v_B(\alpha)]}}{j!j!} v_A^j(\alpha)v_B^j(\alpha)\left(1 + \frac{nv_A(\alpha)}{j+1}\right) \times \\
&\quad \frac{1}{2}\sum_{k=0}^{\infty} \frac{n^{2k} e^{-n[v_A(\beta)+v_A(\alpha)]}}{k!k!} v_A^k(\alpha)v_A^k(\beta)\left(1 + \frac{nv_A(\beta)}{k+1}\right) \\
&\quad -\frac{1}{2}\sum_{j=0}^{\infty} \frac{n^{2j} e^{-n[v_A(\alpha)+v_A(\beta)]}}{j!j!} v_A^j(\alpha)v_A^j(\beta)\left(1 + \frac{nv_A(\alpha)}{j+1}\right) \times \\
&\quad \frac{1}{2}\sum_{k=0}^{\infty} \frac{n^{2k} e^{-n[v_A(\beta)+v_B(\beta)]}}{k!k!} v_A^k(\beta)v_B^k(\beta)\left(1 + \frac{nv_A(\beta)}{k+1}\right) \\
&> \frac{e^{-n[2v_A(\alpha)+v_B(\alpha)+v_A(\beta)]}}{4}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\left(\begin{array}{c} \frac{n^{2j+2k}}{j!j!k!k!} v_A^j(\alpha)v_A^k(\beta)\left(1+\frac{nv_A(\alpha)}{j+1}\right)\left(1+\frac{nv_A(\beta)}{k+1}\right) \\ \times\left[v_B^j(\alpha)v_A^k(\alpha) - v_A^j(\beta)v_B^k(\beta)\right]\end{array}\right) \\
&= \frac{e^{-n(2v_A(\alpha)+v_B(\alpha)+v_A(\beta))}}{4} \times \\
&\quad \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\left(\begin{array}{c} \frac{n^{2j+2k} v_A^j(\alpha)v_B^j(\alpha)v_A^k(\beta)v_A^k(\alpha)}{j!j!k!k!}\left(1+\frac{nv_A(\alpha)}{k+1}\right)\left(1+\frac{nv_A(\beta)}{j+1}\right) \\ \times\left[1 - \left(\frac{v_A(\beta)}{v_B(\alpha)}\right)^{k-j}\left(\frac{v_A(\beta)v_B(\beta)}{v_B(\alpha)v_A(\alpha)}\right)^j\right]\end{array}\right) \qquad \text{(A.17)}\\
&> 0.
\end{aligned}$$

The final inequality holds because, whenever $j \le k$, the final product in (A.17) is less than one, so that the bracketed difference is positive. For any negative term, therefore, it must be that $j > k$; in that case, however, the corresponding

(k,j) term has otherwise equal magnitude but receives greater weight (*i.e.*, $v_B^j(\alpha)v_A^j(\alpha)v_A^k(\alpha)v_A^k(\beta)$ instead of $v_B^k(\alpha)v_A^k(\alpha)v_A^j(\alpha)v_A^j(\beta)$). Thus the sign of (A.16) is positive.

Maintaining the assumption that $1-T_A>T_B$, a similar derivation reveals further that $P(\alpha)\tilde{P}(\sigma_{T_B})>\tilde{P}(\beta)P(\sigma_{T_B})$ or, equivalently, that $\hat{\theta}_{A0}>1-T^{br}(T_B)$, as claimed. Identical derivations for the case of $1-T_A<T_B$ yield the reverse inequalities. ‖

Theorem 2. *If σ^* is a Bayesian Nash equilibrium then it is a quality threshold strategy.*

Proof As Section 3 points out, a Bayesian Nash equilibrium σ^* in a game of Poisson population uncertainty is necessarily symmetric. By Proposition 1, it is therefore also a belief threshold strategy. To see that equilibrium thresholds must be symmetric around $1/2$, consider first the best response to a belief threshold strategy for which the thresholds $1/2<T_A\leq T_B$ both exceed $1/2$. In that case, individuals who receive b signals vote B only if they are sufficiently well informed (*i.e.* $q\geq T_B$), while those who receive a signals all vote A, along with some who received b signals but are poorly informed (*i.e.* $q\leq T_A$) instead vote A. Thus, (2) reduces to

$$\begin{aligned} v_A(\alpha)&=F(T_A)+\int_{T_A}^{1} q\,dF(q) & v_B(\alpha)&=\int_{T_B}^{1}(1-q)\,dF(q)\\ v_A(\beta)&=F(T_A)+\int_{T_A}^{1}(1-q)\,dF(q) & v_B(\beta)&=\int_{T_B}^{1} q\,dF(q). \end{aligned} \tag{A.18}$$

From these, it is clear that $v_A(\beta)>v_B(\alpha)$ and $v_A(\alpha)>v_B(\beta)$, implying that $\pi_0(\beta)>\pi_0(\alpha)$ and $\pi_1(\beta)>\pi_{-1}(\alpha)$ and therefore $\tilde{P}(\beta)>P(\alpha)$. It then follows that $\hat{\theta}_{A0}=\frac{P(\alpha)}{P(\alpha)+\tilde{P}(\beta)}<1/2<T_A$, implying that σ_{T_A,T_B} is not its own best response. By symmetric reasoning, σ_{T_A,T_B} is not its own best response if $T_A\leq T_B<1/2$.

Next, consider a belief threshold strategy for which $1-T_A>T_B$. Using Lemmas A1 and A2, the best response to this strategy can be compared with the best responses to the quality threshold strategies σ_{1-T_A} and σ_{T_B}. There are three cases to consider:

Case 1: $T^{br}(T_B)\leq T_B$. In this case, Lemma A2 implies that $\hat{\theta}_{A0}>1-T^{br}(T_B)\geq 1-T_B>T_A$, so σ_{T_A,T_B} is not its own best response.

If $T^{br}(T_B)>T_B$ then, since $T^{br}(1)<1$, there exists at least one fixed point of T^{br} between T_B and 1 (by the intermediate value theorem). Let $\bar{T}^*$ denote the lowest (*i.e.* the closest to T_B) of these fixed points. This threshold distinguishes the remaining two cases:

Case 2: $T^{br}(T_B)>T_B$ and $1-T_A\geq\bar{T}^*$. By Lemma A1, T^{br} is increasing between T_B and $\bar{T}$, which implies that $T^{br}(T_B)<T^{br}(\bar{T}^*)=\bar{T}^*$. By Lemma A2, this implies that $\hat{\theta}_{A0}\geq 1-T^{br}(T_B)>1-\bar{T}^*\geq T_A$, so σ_{T_A,T_B} is not its own best response.

Case 3: $T^{br}(T_B)>T_B$ and $1-T_A<\bar{T}^*$. In this case $T^{br}(1-T_A)>T^{br}(T_B)$, since $T_B\leq 1-T_A\leq\bar{T}^*$ and T^{br} is increasing (by Lemma A1) between T_B and $\bar{T}^*$. Lemma A2 then implies that $\hat{\theta}_{0B}>T^{br}(1-T_A)>T^{br}(T_B)>T_B$, so σ_{T_A,T_B} is not its own best response.

Symmetric reasoning applies, of course, if $1-T_A<T_B$. Thus, σ^* is its own best response only if it is a belief threshold strategy for which $1-T_A=T_B$, which is equivalent to a quality threshold strategy. ‖

Lemma 1. *For any $T<1$, the best-response quality threshold T_n^{br} increases in n.*

Proof For a given quality threshold strategy σ_T, vote shares v_+ and v_- for the two policies do not depend on the number n of citizens. Outcome probabilities do implicitly depend on the size of the electorate, however, and can be differentiated with respect to n, as follows.

$$\begin{aligned} \frac{\partial\psi_k}{\partial n} &= \frac{(v_+v_-)^k}{k!k!}\left[2kn^{2k-1}e^{-n(v_++v_-)}-(v_++v_-)n^{2k}e^{-n(v_++v_-)}\right]\\ &=\psi_k\left[\frac{2k}{n}-(v_++v_-)\right] \end{aligned} \tag{A.19}$$

$$\frac{\partial\pi_0}{\partial n}=\sum_{k=0}^{\infty}\psi_k\left[\frac{2k}{n}-(v_++v_-)\right] \tag{A.20}$$

$$\begin{aligned}\frac{\partial \pi_1}{\partial n} &= \sum_{k=0}^{\infty}\left(\frac{\partial \psi_k}{\partial n}\frac{nv_+}{k+1}+\psi_k\frac{v_+}{k+1}\right)\\ &= \sum_{k=0}^{\infty}\psi_k\left\{\left[\frac{2k}{n}-(v_++v_-)\right]\frac{nv_+}{k+1}+\frac{v_+}{k+1}\right\}\end{aligned} \tag{A.21}$$

$$\begin{aligned}\frac{\partial \pi_{-1}}{\partial n} &= \sum_{k=0}^{\infty}\left(\frac{\partial \psi_k}{\partial n}\frac{nv_-}{k+1}+\psi_k\frac{v_-}{k+1}\right)\\ &= \sum_{k=0}^{\infty}\psi_k\left\{\left[\frac{2k}{n}-(v_++v_-)\right]\frac{nv_-}{k+1}+\frac{v_-}{k+1}\right\}.\end{aligned} \tag{A.22}$$

The ratio $\frac{\tilde{P}}{P}$ of pivot probabilities can then also be differentiated, by the quotient rule:

$$\frac{\partial}{\partial n}\frac{\tilde{P}}{P}=\frac{1}{P^2}\left(P\frac{\partial \tilde{P}}{\partial n}-\tilde{P}\frac{\partial P}{\partial n}\right) \tag{A.23}$$

where $P\frac{\partial \tilde{P}}{\partial n}$ is given by

$$\begin{aligned}P\frac{\partial \tilde{P}}{\partial n} &= \frac{1}{4}(\pi_0+\pi_{-1})\left(\frac{\partial \pi_0}{\partial n}+\frac{\partial \pi_1}{\partial n}\right)\\ &= \frac{1}{4}\sum_{j=0}^{\infty}\psi_j\left(1+\frac{nv_-}{j+1}\right)\sum_{k=0}^{\infty}\psi_k\left\{\left[\frac{2k}{n}-(v_++v_-)\right]\left(1+\frac{nv_+}{k+1}\right)+\frac{v_+}{k+1}\right\}\\ &= \frac{1}{4}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\\ &\quad\times\left\{\left[\frac{2k}{n}-(v_++v_-)\right]\left(1+\frac{nv_-}{j+1}\right)\left(1+\frac{nv_+}{k+1}\right)+\frac{v_+}{k+1}+n\frac{v_-}{j+1}\frac{v_+}{k+1}\right\}\end{aligned}$$

and similarly $\tilde{P}\frac{\partial P}{\partial n}$ is given by

$$\begin{aligned}\tilde{P}\frac{\partial P}{\partial n} &= \frac{1}{4}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\\ &\quad\times\left\{\left[\frac{2k}{n}-(v_-+v_+)\right]\left(1+\frac{nv_+}{j+1}\right)\left(1+\frac{np_-}{k+1}\right)+\frac{p_-}{k+1}+n\frac{v_+}{j+1}\frac{p_-}{k+1}\right\}\end{aligned}$$

The expression $P\frac{\partial \tilde{P}}{\partial n}-\tilde{P}\frac{\partial P}{\partial n}$ in (A.23) then simplifies to

$$\begin{aligned}P\frac{\partial \tilde{P}}{\partial n}-\tilde{P}\frac{\partial P}{\partial n} &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\frac{(v_+-v_-)}{k+1}\\ &= \sum_{j=0}^{\infty}\psi_j\left[\sum_{k=0}^{\infty}\psi_k\frac{v_+}{k+1}-\sum_{k=0}^{\infty}\psi_k\frac{v_-}{k+1}\right]\\ &= \pi_0\frac{1}{n}(\pi_1-\pi_{-1})\end{aligned}$$

which is positive for any T, since $\pi_1>\pi_{-1}$ and $\pi_0>0$. Thus $\frac{\partial}{\partial n}\frac{\tilde{P}}{P}>0$, and therefore $\frac{\partial}{\partial n}\left(\frac{\tilde{P}}{P+\tilde{P}}\right)>0$, which is the desired result. ‖

Lemma A3 is a useful preparation for Theorems 3 and 4.

Lemma A3. *If* $\lim_{T\to 1}f'(T)/f(T)<\infty$ *then* $\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i\geq T)\geq 1/2$.

Proof Differentiate $E(q_i|q_i \geq T)$ using the quotient rule and (A.4) as follows,

$$\begin{aligned}\frac{\partial}{\partial T}E(q_i|q_i \geq T) &= \frac{\partial}{\partial T}\frac{v_+}{1-F}\\ &= \frac{-Tf(1-F)+v_+f}{(1-F)^2}\\ &= h[E(q_i|q_i \geq T)-T],\end{aligned} \tag{A.24}$$

where $h = f/1-F$ is the *hazard function* of F (and v_+, F, f, and h are evaluated at T), with derivative given by

$$h' = \frac{f'(1-F)+f^2}{(1-F)^2} = h\left(\frac{f'}{f}+h\right).$$

If $\lim_{T\to 1} f(T) = \infty$ then $\lim_{T\to 1} h = \infty$. Given the assumption that $\lim_{T\to 1} f'(T)/f(T) < \infty$, therefore, $\lim_{T\to 1} h'/h^2 = \lim_{T\to 1} \frac{1}{h}\frac{f'}{f} + 1 = 1$, so L'Hopital's rule applied to equation (A.24) yields

$$\begin{aligned}\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T) &= \lim_{T\to 1}\frac{\frac{\partial}{\partial T}E(q_i|q_i \geq T)-1}{-h'/h^2}\\ &= 1-\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T),\end{aligned}$$

implying that $\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T) = \frac{1}{2}$. If $\lim_{T\to 1} f(T) < \infty$ then, again using the assumption that $\lim_{T\to 1}\frac{f'(T)}{f(T)} < \infty$, L'Hopital's rule instead yields the following,

$$\begin{aligned}\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T) &= \lim_{T\to 1}\frac{f'[E(q_i|q_i \geq T)-T]+f\left[\frac{\partial}{\partial T}E(q_i|q_i \geq T)-1\right]}{-f}\\ &= \lim_{T\to 1}\left\{1-\frac{\partial}{\partial T}E(q_i|q_i \geq T)-\frac{f'}{f}[E(q_i|q_i \geq T)-T]\right\}\\ &\geq 1-\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T),\end{aligned}$$

implying that $\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T) \geq \frac{1}{2}$ (with equality unless $\lim_{T\to 1}\frac{f'(T)}{f(T)} = -\infty$), as claimed. ||

Theorem 3. *If* $\{T_n^*\}$ *is a sequence of equilibrium quality thresholds then it has a limit point* T_∞^*. *Furthermore, if* $\lim_{T\to 1}\frac{f'(T)}{f(T)} < \infty$ *then* $T_\infty^* < 1$.

Proof The left- and right-hand sides of (14) both equal one when $T = 1$, but approach this value with different slopes: Lemma A3 states that $\lim_{T\to 1}\frac{\partial}{\partial T}E(q_i|q_i \geq T) \geq \frac{1}{2}$, but $\lim_{T\to 1}\frac{\partial}{\partial T}\frac{T^2}{T^2+(1-T)^2} = 0$. This implies that, for T close to one, the right-hand side of (14) exceeds the left-hand side. The opposite is true when $T = 1/2$, however, so (14) has at least one interior solution between 1/2 and 1, by the Intermediate Value theorem. Let T^{**} denote the largest such solution, implying that $T_\infty^{br}(T) < T$ for all T between T^{**} and 1. By Lemma 1, however, T_∞^{br} is an upper bound on T_n^{br}, implying that $T_n^{br}(T) < T$ for all T between T^{**} and 1, as well. Thus, if T exceeds T^{**} then T does not constitute an equilibrium quality threshold for any n. Accordingly, any sequence $\{T_n^*\}$ of equilibrium quality thresholds is bounded above by T^{**}. Such a sequence is increasing by Lemma 1, however, so a limit point $T_\infty^* \leq T^{**} < 1$ exists, as claimed. ||

Theorem 4. *If* f *is log-concave then there exists a unique limit point* T_∞^* *such that for any sequence* $\{T_n^*\}$ *of equilibrium participation thresholds,* $T_n^* \to T_\infty^*$.

Proof A log-concave density f first increases and then decreases with T, below and above some maximizer $\hat{T} \in [1/2, 1]$. Above $\hat{T}$, it must be the case that $\frac{\partial}{\partial T}E(q_i|q_i \geq T) \geq 1/2$. To see this, first differentiate (A.24) to obtain the second derivative of $E(q_i|q_i \geq T)$,

$$\begin{aligned}\frac{\partial^2 E(q_i|q_i \geq T)}{\partial T^2} &= h'[E(q_i|q_i \geq T)-T]+h\left[\frac{\partial E(q_i|q_i \geq T)}{\partial T}-1\right]\\ &= \frac{f'}{f}\frac{\partial E(q_i|q_i \geq T)}{\partial T}+2h\left[\frac{\partial E(q_i|q_i \geq T)}{\partial T}-\frac{1}{2}\right],\end{aligned}$$

which is positive if and only if

$$\frac{\partial}{\partial T}E(q_i|q_i \geq T) \geq \frac{1}{2+\frac{1}{h}\frac{f'}{f}}. \tag{A.25}$$

Above $\hat{T}$, the right-hand side of (A.25) is greater than $1/2$ because f is decreasing, so f' is negative. If it were the case for some $T > \hat{T}$ that $\frac{\partial}{\partial T} E(q_i|q_i \geq T) < \frac{1}{2}$, therefore, then the inequality in (A.25) would not be satisfied, implying that $E(q_i|q_i \geq T)$ is locally concave. But then $\frac{\partial}{\partial T} E(q_i|q_i \geq T)$ would decrease further with T, implying that $\lim_{T\to 1} \frac{\partial}{\partial T} E(q_i|q_i \geq T) < 1/2$. Since $\lim_{T\to 1} \frac{f'(T)}{f(T)} < \infty$ for a log-concave density, this would contradict Lemma A3.

Below $\hat{T}$, $E(q_i|q_i \geq T)$ is necessarily convex. To see this, first note that the right-hand side of (A.25) is less than $1/2$, because f is increasing, so f' is positive. If the inequality in (A.25) were violated for some $T < \hat{T}$, therefore, then there would necessarily exist some $\varepsilon > 0$ such that $\frac{\partial}{\partial T} E(q_i|q_i \geq T) \leq 1/2 - \varepsilon$. However, the right-hand side of (A.25) also increases with T on the interval $\left[1/2, \hat{T}\right]$, because f' is positive, $\frac{f'}{f}$ is decreasing (by log-concavity) and as Bagnoli and Bergstrom (2005) demonstrate, h is increasing in T. As T increases, therefore, the left- and right-hand sides of (A.25) would decrease and increase, respectively. This implies that the inequality would never be reversed, so $\frac{\partial}{\partial T} E(q_i|q_i \geq T)$ would continue to decline, and would be bounded above by $1/2 - \varepsilon$. In particular, $\frac{\partial}{\partial T} E\left(q_i|q_i \geq \hat{T}\right) \leq \frac{1}{2} - \varepsilon$, contradicting the above result that $\frac{\partial}{\partial T} E(q_i|q_i \geq T) \geq \frac{1}{2}$ for all $T \geq \hat{T}$.

The importance of these results on $\partial/\partial T E(q_i|q_i \geq T)$ is illustrated in Figure 2: in the interval $\left[\hat{T}, 1\right]$, $E(q_i|q_i \geq T)$ lies below the dotted line of slope $1/2$, and therefore cannot intersect $T^2/T^2 + (1-T)^2$ to the right of $1/\sqrt{2}$. Below $1/\sqrt{2}$, however, the two functions can have at most one intersection point, because the slopes of $T^2/T^2 + (1-T)^2$ and $E(q_i|q_i \geq T)$ are bounded above and below one, respectively.[49] Thus, $E(q_i|q_i \geq T)$ and $T^2/T^2 + (1-T)^2$ can intersect at most once between $\hat{T}$ and 1. If there is such an intersection point, then there can be no intersection in $\left[1/2, \hat{T}\right]$ because $E(q_i|q_i \geq T)$ is convex and $T^2/T^2 + (1-T)^2$ is concave; if there is no intersection point in $\left[\hat{T}, 1\right]$ then there is exactly one intersection point in $\left[1/2, \hat{T}\right]$ for the same reason. Thus $E(q_i|q_i \geq T)$ and $T^2/T^2 + (1-T)^2$ intersect exactly once between $1/2$ and 1, implying a unique limit point T^*_∞ for any sequence $\left\{T^*_n\right\}$ of equilibrium quality thresholds. ‖

Proposition 3. *Let F and G have log-concave densities f and g, and suppose that G first-order stochastically dominates F. Then the following are true:*

1. *If $g(q) = f(q)$ for all $q \geq T^*_{\infty F}$ then $T^*_{\infty G} = T^*_{\infty F}$, $\tau_G = \tau_F$, and $\mu_G = \mu_F$.*
2. *If $G\left(T^*_{\infty F}\right) = F\left(T^*_{\infty F}\right)$ then $T^*_{\infty G} > T^*_{\infty F}$, $\tau_G < \tau_F$, and $\mu_G > \mu_F$.*
3. *If $g(q) \geq f(q)$ for all q between $T^*_{\infty F}$ and $E_F\left(q_i|q_i \geq T^*_{\infty F}\right)$, and $g(q) = f(q)$ for all $q \geq E_F\left(q_i|q_i \geq T^*_{\infty F}\right)$, then $T^*_{\infty G} < T^*_{\infty F}$, $\tau_G > \tau_F$, and $\mu_G < \mu_F$.*

Proof As a preliminary step, note the equivalence, for any T, of the following inequalities.

$$E_G(q_i|q_i \geq T) \geq E_F(q_i|q_i \geq T) \tag{A.26}$$

$$\frac{\int_T^1 [1-G(q)]dq}{1-G(T)} \geq \frac{\int_T^1 [1-F(q)]dq}{1-F(T)} \tag{A.27}$$

$$T^{br,\sigma_T}_{\infty G} \geq T^{br,\sigma_T}_{\infty F} \tag{A.28}$$

$$\mu_G(T) \geq \mu_F(T). \tag{A.29}$$

The equivalence of (A.26) and (A.27) can be seen by integrating by parts; that (A.28) and (A.29) are equivalent to (A.26) follow immediately from Equations (13) and (16).

1. If $g(q) = f(q)$ for all $q \geq T^*_{\infty F}$ then (A.27), and therefore each of the above inequalities, is satisfied with equality at $T^*_{\infty F}$. In particular, (A.28) implies that $T^*_{\infty G} = T^*_{\infty F}$. Then $\tau_G = \tau_F$ and $\mu_G = \mu_F$ follow from (15) and (16).

2. $G\left(T^*_{\infty F}\right) = F\left(T^*_{\infty F}\right)$ implies that the denominators on either side of (A.27) are equal at $T^*_{\infty F}$. The numerator on the left-hand side strictly exceeds the numerator on the right-hand side, however, because $G\left(T^*_{\infty F}\right) = F\left(T^*_{\infty F}\right)$ implies that $G(q) < F(q)$ (*i.e.*, $1-G(q) > 1-F(q)$) for some $q > T^*_{\infty F}$ and a surrounding neighborhood. Thus the inequalities in (A.26) through (A.29) are all strict at $T^*_{\infty F}$, implying that $T^*_{\infty G} > T^*_{\infty F}$ and therefore $\tau_G = 1-G\left(T^*_{\infty G}\right) < 1-G\left(T^*_{\infty F}\right) = 1-F\left(T^*_{\infty F}\right) = \tau_F$. Also, $E^*_G > E_G\left(q_i|q_i \geq T^*_{\infty F}\right) > E^*_F$, where $E^*_F = E_F\left(q_i|q_i \geq T^*_{\infty F}\right)$ and $E^*_G = E_G\left(q_i|q_i \geq T^*_{\infty G}\right)$ denote the average expertise of voters in distributions F and G, respectively, implying that $\mu_G = 2E^*_G - 1 > 2E^*_F - 1 = \mu_F$.

49. That $\partial/\partial T E(q_i|q_i \geq T) \leq 1$ is equivalent to Bagnoli and Bergstrom's (2005) result that for distributions with log-concave densities the *mean residual lifetime* function $E(q_i - T|q_i \geq T)$ decreases in T.

3. Since $g(q) = f(q)$ for all $q \geq E_F^*$,

$$\begin{aligned}\int_{T_{\infty F}^*}^1 qg(q)dq &= \int_{T_{\infty F}^*}^1 qf(q)dq + \int_{T_{\infty F}^*}^{E_F^*} q[g(q)-f(q)]dq \\ &< \int_{T_{\infty F}^*}^1 qf(q)dq + E_F^* \int_{T_{\infty F}^*}^{E_F^*} [g(q)-f(q)]dq \\ &= E_F^* \left\{ \int_{T_{\infty F}^*}^1 f(q)dq + \int_{T_{\infty F}^*}^{E_F^*} [g(q)-f(q)]dq \right\},\end{aligned}$$

and similarly

$$\int_{T_{\infty F}^*}^1 g(q)dq = \int_{T_{\infty F}^*}^1 f(q)dq + \int_{T_{\infty F}^*}^{E_F^*} [g(q)-f(q)]dq.$$

Therefore, $E_G(q_i | q_i \geq T_{\infty F}^*) = \int_{T_{\infty F}^*}^1 qg(q)dq / \int_{T_{\infty F}^*}^1 g(q)dq < E_F^*$, so that inequality (A.26) does not hold, and therefore neither do inequalities (A.27) through (A.29). This implies that $T_{\infty G}^* < T_{\infty F}^*$, and therefore that $\tau_G = 1 - G(T_{\infty G}^*) > 1 - F(T_{\infty G}^*) > 1 - F(T_{\infty F}^*) = \tau_F$ and $\mu_G = 2E_G^* - 1 < 2E_G(q_i | q_i \geq T_{\infty F}^*) - 1 < 2E_F^* - 1 = \mu_F$, as claimed. ‖

Proposition 4. *Let F and G have log-concave densities f and g and a common mean, and suppose that $G(q) \leq F(q)$ for all $q \leq \hat{q}$ and $G(q) \geq F(q)$ for all $q \leq \hat{q}$, for some $\hat{q} \geq T_{\infty F}^*$. Then $T_{\infty G}^* \leq T_{\infty F}^*$, $\tau_G > \tau_F$, and $\mu_G < \mu_F$.*

Proof Using integration by parts, the common mean of F and G can be written as $1/2 + \int_{1/2}^1 [1-F(q)]dq = 1/2 + \int_{1/2}^1 [1-G(q)]dq$, implying that $\int_{1/2}^T [1-F(q)]dq + \int_T^1 [1-F(q)]dq = \int_{1/2}^T [1-G(q)]dq + \int_T^1 [1-G(q)]dq$ for any $T \in [1/2, 1]$. Since $\hat{q} \geq T_{\infty F}^*$ it must be that $1 - G(T_{\infty F}^*) \geq 1 - F(T_{\infty F}^*)$ and $\int_{1/2}^{T_{\infty F}^*} [1-G(q)]dq \geq \int_{1/2}^{T_{\infty F}^*} [1-F(q)]dq$, implying that $\int_{T_{\infty F}^*}^1 [1-G(q)]dq \leq \int_{T_{\infty F}^*}^1 [1-F(q)]dq$. Thus, the inequality in (A.27) does not hold, and therefore neither do (A.28) and (A.29). This implies that $T_{\infty G}^* \leq T_{\infty F}^*$, and in turn that $\tau_G = 1 - G(T_{\infty G}^*) \geq 1 - G(T_{\infty F}^*) \geq 1 - F(T_{\infty F}^*) = \tau_F$ and $\mu_G = 2E_G^* - 1 \leq 2E_G(q_i | q_i \geq T_{\infty F}^*) - 1 \leq 2E_F^* - 1 = \mu_F$. ‖

Theorem 5. *There exists a strategy σ^{**} that maximizes the probability of electing the superior policy. Furthermore, σ^{**} is a quality threshold strategy, with quality threshold $\frac{1}{2} < T^{**} < 1$, and constitutes a Bayesian Nash equilibrium.*

Proof The logic of this proof is to demonstrate that a strategy σ only maximizes welfare if it is a belief threshold strategy. A globally optimal strategy then exists by the Weierstrass extreme value theorem, because the probability of a desired election outcome is a continuous function over the compact set $\{(T_A, T_B): 0 \leq T_A \leq T_B \leq 1\}$ of belief thresholds. That the optimal strategy also constitutes a Bayesian Nash equilibrium follows from McLennan's (1998) observation that, in common interest games such as this, the socially optimal strategy profile is optimal for each individual. Theorem 2 rules out the possibility of asymmetric belief thresholds, reducing the optimal strategy to a quality threshold strategy. That $1/2 < T^{**} < 1$ then follows from Proposition 2.

The proof that every strategy σ is welfare dominated by a belief threshold strategy σ_{T_A, T_B} proceeds by construction. As a preliminary step, let $\tilde{f}(\theta)$ denote the density of posterior beliefs $\theta(\beta | q, s)$ induced by the density f of expertise,

$$\tilde{f}(\theta) = \begin{cases} \frac{1}{2} f(1-\theta) & \text{if } \theta < \frac{1}{2} \\ \frac{1}{2} f(\theta) & \text{if } \theta \geq \frac{1}{2} \end{cases}.$$

The conditional densities $\tilde{f}(\theta | \alpha)$ and $\tilde{f}(\theta | \beta)$ can then be written in terms of $\tilde{f}(\theta)$, as follows,

$$\begin{aligned}\tilde{f}(\theta | \alpha) &= \begin{cases} (1-\theta) f(1-\theta) & \text{if } \theta < \frac{1}{2} \\ (1-\theta) f(\theta) & \text{if } \theta \geq \frac{1}{2} \end{cases} \\ &= 2(1-\theta)\tilde{f}(\theta), \\ \tilde{f}(\theta | \beta) &= \begin{cases} \theta f(1-\theta) & \text{if } \theta < \frac{1}{2} \\ \theta f(\theta) & \text{if } \theta \geq \frac{1}{2} \end{cases} \\ &= 2\theta \tilde{f}(\theta).\end{aligned}$$

Also, since her posterior belief $\theta(\beta | q, s)$ uniquely summarizes a citizen's private information, a voting strategy $\sigma: [0,1] \longrightarrow \{A, B, 0\}$ can be redefined as a function of $\theta(\beta | q, s)$, instead of as a function of (q, s).

With this notation, define σ_{T_A,T_B} so that the (unconditional) probabilities of voting for either policy are the same as under σ:

$$\int_0^{T_A} \tilde{f}(\theta)d\theta = \int_0^1 I_{[\sigma(\theta)=A]}\tilde{f}(\theta) \tag{A.30}$$

$$\int_{T_B}^1 \tilde{f}(\theta)d\theta = \int_0^1 I_{[\sigma(\theta)=B]}\tilde{f}(\theta). \tag{A.31}$$

According to the strategy σ, the probability $v_B(\beta;\sigma)$ of voting for policy B in state β is

$$\begin{aligned} v_B(\beta;\sigma) &= \int_0^1 I_{[\sigma(\theta)=B]}\tilde{f}(\theta|\beta) \\ &= \int_0^1 2\theta I_{[\sigma(\theta)=B]}\tilde{f}(\theta) \\ &= 2E[\theta|\sigma(\theta)=B]\int_0^1 I_{[\sigma(\theta)=B]}\tilde{f}(\theta). \end{aligned}$$

According to the strategy σ_{T_A,T_B}, this same probability is

$$\begin{aligned} v_B(\beta;\sigma_{T_A,T_B}) &= \int_{T_B}^1 \tilde{f}(\theta|\beta)d\theta \\ &= \int_{T_B}^1 2\theta\tilde{f}(\theta)d\theta \\ &= 2E(\theta|\theta>T_B)\int_{T_B}^1 \tilde{f}(\theta)d\theta. \end{aligned}$$

Since $\sigma(\theta)=B$ for values of θ other than those closest to one, $E[\theta|\sigma(\theta)=B]<E(\theta|\theta>T_B)$. Given (A.31), this implies that $v_B(\beta;\sigma)<v_B(\beta;\sigma_{T_A,T_B})$. By analogous derivations, $v_A(\alpha;\sigma)<v_A(\alpha;\sigma_{T_A,T_B})$, $v_B(\alpha;\sigma)>v_B(\alpha;\sigma_{T_A,T_B})$, and $v_A(\beta;\sigma)>v_A(\beta;\sigma_{T_A,T_B})$. In other words, relative to σ, the belief threshold strategy σ_{T_A,T_B} specifies more A voting and less B voting in state α, and more B voting and less A voting in state β.

In state β, the probability $\Pr(N_B>N_A|\beta)$ that the number of B votes exceeds the number of A votes can be written as

$$\begin{aligned} \Pr(N_B>N_A|\beta) &= \sum_{j=0}^{\infty}\sum_{k=0}^{j-1}\frac{e^{-nv_B(\beta)}}{j!}[nv_B(\beta)]^j\frac{e^{-nv_A(\beta)}}{k!}[nv_A(\beta)]^k \\ &= \sum_{j=0}^{\infty}\frac{e^{-nv_B(\beta)}}{j!}[nv_B(\beta)]^j\frac{\Gamma[j-1,nv_A(\beta)]}{(j-1)!}, \end{aligned} \tag{A.32}$$

where $\Gamma[j-1,nv_A(\beta)]/(j-1)!=\sum_{k=0}^{j-1}e^{-nv_A(\beta)}/k![nv_B(\beta)]^k$ is the cumulative distribution function for a Poisson random variable, expressed in terms of the *upper incomplete gamma function* $\Gamma(s,\xi)=\int_\xi^\infty t^{s-1}e^{-t}dt$ (see Abramowitz and Stegun, 1964). This cdf is decreasing in the error probability $v_A(\beta)$,

$$\frac{\partial}{\partial v_A(\beta)}\frac{\Gamma[j-1,nv_A(\beta)]}{(j-1)!}=\frac{-ne^{-nv_A(\beta)}[nv_A(\beta)]^{j-2}}{(j-1)!}<0,$$

implying that $\Pr(N_B>N_A|\beta)$ is decreasing in $v_A(\beta)$, as well. Rewritten as

$$\begin{aligned} \Pr(N_B>N_A|\beta) &= \sum_{j=0}^{\infty}\sum_{k=j+1}^{\infty}\frac{e^{-nv_A(\beta)}}{j!}[nv_A(\beta)]^j\frac{e^{-nv_B(\beta)}}{k!}[nv_B(\beta)]^k \\ &= \sum_{j=0}^{\infty}\frac{e^{-nv_A(\beta)}}{j!}[nv_A(\beta)]^j\left[1-\frac{\Gamma[j-1,nv_B(\beta)]}{(j-1)!}\right], \end{aligned}$$

(A.32) is also increasing in $v_B(\beta)$. Since σ_{T_A,T_B} increases $v_B(\beta)$ and decreases $v_A(\beta)$ relative to σ, therefore, it increases the total probability $\Pr(B|\beta)$ of electing policy B in state β, and reduces the probability $\Pr(A|\beta)$ of an error. By a similar derivation, it also increases $\Pr(A|\alpha)$ and decreases $\Pr(B|\alpha)$, implying that it increases the overall probability $1/2\Pr(A|\alpha)+1/2\Pr(B|\beta)$ of the desired election outcome. ‖

Acknowledgments. This article was part of my dissertation at the University of Rochester. Special thanks to my advisor, Mark Fey, for his generous time and helpful advice. Thanks also to Marco Ottaviani and the anonymous referees, as well as to Gábor Virág, Tim Feddersen, Cesar Martinelli, David Austen-Smith, John Duggan, John Morgan, Val Lambson, Brennan Platt, Becky Morton, and participants of the Wallis Conference on Political Economy, for their interest and suggestions.

REFERENCES

ABRAMOWITZ, M. and IRENE A.S. (1964), " *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 10th ed.*" (New York: Dover).

ABRAMSON, P.R., ALDRICH, J.H., PAOLINO, P. and Rhode, D.W. (1992), "'Sophisticated' Voting in the 1988 Presidential Primaries", *The American Political Science Review*, **86**, 55–69.

ALDRICH, J.H. (1993), "Rational Choice and Turnout", *American Journal of Political Science*, **37**, 246–278.

ALDRICH, J.H. (1997), "When is it Rational to Vote?" in *Perspectives on Public Choice, a Handbook*, Dennis Mueller (ed.). (Cambridge: Cambridge University Press) 379–390.

ALVAREZ, R.M. and NAGLER, J. (2000), "A New Approach for Modelling Strategic Voting in Multiparty Elections", *British Journal of Political Science*, **30**, 57–75.

BAGNOLI, M. and BERGSTROM, T. (2005), "Log-Concave Probability and Its Applications", *Economic Theory*, **26**, 445–469.

BANERJEE, A.V. (1992), "A Simple Model of Herd Behavior", *The Quarterly Journal of Economics*, **107**, 797–817.

BANERJEE, A.V., KUMAR, S. PANDE, R. and SU, F. (2010), "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India". Manuscript, NBER Political Economy Meeting papers. http://www.nber.org/confer/2010/POLf10/Pande_Banerjee_Kumar_Su.pdf.

BATTAGLINI, M., MORTON, R. and PALFREY, T. (2008), "Information Aggregation and Strategic Abstention in Large Laboratory Elections", *The American Economic Review: Papers & Proceedings*, **98**, 194–200.

BATTAGLINI, M, MORTON, R. and PALFREY, T. (2009), "The Swing Voter's Curse in the Laboratory", *Review of Economic Studies*, **77**, 61–89.

BIKHCHANDANI, S., HIRSHLEIFER, D. and WELCH, I. (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *The Journal of Political Economy*, **100**, 992–1026.

BLACKWELL, D. (1951), "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman. Berkely, (CA: University of Berkeley Press) 93–102.

BRODY, R.A. (1978), "The Puzzle of Political Participation in America", In *The New American Political System*, ed. Anthony King. (Washington DC: American Enterprise Institute).

COATE, S., CONLIN, M. and MORO, A. (2008), "The Performance of the Pivotal-Voter Model in Small-Scale Elections: Evidence from Texas Liquor Referenda", *Journal of Public Economics*, **92**, 582–596.

CONDORCET, M. de. (1785), *Essay on the Application of Analysis to the Probability of Majority Decisions*, (Paris: De l'imprimerie royale) (Trans. Iain McLean and Fiona Hewitt. 1994).

DELLI CARPINI, M.X. and KEETER, S. (1996), "*What Americans Know About Politics and Why it Matters*" (New Haven, CT: Yale University Press).

DIAMOND, P.A. and STIGLITZ, J.E. (1974), "Increases in Risk and Risk Aversion", *Journal of Economic Theory*, **8**, 337–360.

DOWNS, A. (1957), "*An Economic Theory of Democracy*", (New York: Harper and Row).

DUGGAN, J. and MARTINELLI, C. (2001), "A Bayesian Model of Voting in Juries." *Games and Economic Behavior*, **37**, 259–294.

DUVERGER, M. (1954), "*Political Parties*", (New York: Wiley).

EDLIN, A., GELMAN, A. and KAPLAN, N. (2007), "Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others", *Rationality and Society*, **19**, 225–246.

ESPONDA, I. and VESPA, E.I. (2011), "Hypothetical Thinking and Information Extraction: Pivotal Voting in the Laboratory", Manuscript, New York University.

EVREN, Ö. (2010), "Altruism and Voting: A Large-Turnout Result That Does not Rely on Civic Duty or Cooperative Behavior", Manuscript, New York University.

FARAVELLI, M. and WALSH, R. (2011), "Smooth Politicians and Paternalistic Voters: A Theory of Large Elections". NBER Working paper # 17397.

FARBER, H. (2010), "Rational Choice and Voter Turnout: Evidence from Union Representation Elections", Manuscript, Princeton University.

FEDDERSEN, T.J. (2004), "Rational Choice Theory and the Paradox of Not Voting", *Journal of Economic Perspectives*, **18**, 99–112.

FEDDERSEN, T.J. and PESENDORFER, W. (1996), "The Swing Voter's Curse", *The American Economic Review*, **86**, 408–424.

FEDDERSEN, T.J. and PESENDORFER, W. (1998), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting", *The American Political Science Review*, **92**, 23–35.

FEDDERSEN, T.J. and PESENDORFER, W. (1999), "Abstention in Elections with Asymmetric Information and Diverse Preferences", *The American Political Science Review*, **93**, 381–398.

FEDDERSEN, T.J. and SANDRONI, A. (2006), "A Theory of Participation in Elections", *The American Economic Review*, **96**, 1271–1282.

FEHR, E. and SCHMIDT, K.M. (2006), "The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories", in *Handbook of Reciprocity, Gift-Giving and Altruism*, vol. 1, eds. by Serge-Christophe Kolm and Jean Mercier Ythier (Amsterdam: North-Holland Publishing Co.) 615–691.

FISCHER, A.J. (1999), "The Probability of Being Decisive", *Public Choice*, **101**, 267–283.

FLANIGAN, W.H. and ZINGALE, N.H. (1998), *Political Behavior of the American Electorate*, 9th ed. (Washington D.C.: Congressional Quarterly Press).

FUJIWARA, T. (2011), "A Regression Discontinuity Test of Strategic Voting and Duverger's Law", *Quarterly Journal of Political Science*, **6**, 197–233.
GEYS, B. (2006a), "Explaining Voter Turnout: A Review of Aggregate-level Research", *Electoral Studies*, **25**, 637–663.
GEYS, B. (2006b), "'Rational' theories of Voter Turnout: a Review", *Political Studies Review*, **4**, 16–35.
GOEREE, J.K. and GROSSER, J. (2007), "Welfare Reducing Polls", *Economic Theory*, **31**, 51–68.
GROSSMAN, S.J. and STIGLITZ, J.E. (1976), "Information and Competitive Price Systems", *The American Economic Review*, **66**, 246–253.
HADAR, J. and Russell, W.R. (1969), "Rules for Ordering Uncertain Prospects", *The American Economic Review*, **59**, 25–34.
JOHNSON, J.P. and MYATT, D.P. (2006), "On the Simple Economics of Advertising, Marketing, and Product Design", *The American Economic Review*, **96**, 756–784.
KIM, J. and FEY, M. (2007), "The Swing Voter's Curse with Adversarial Preferences", *Journal of Economic Theory*, **135**, 236–252.
KRASA, S. and POLBORN, M. (2009), "Is Mandatory Voting better than Voluntary Voting?", *Games and Economic Behavior*, **66**, 275–291.
KRISHNA, V. and MORGAN, J. (2010), "On the Benefits of Costly Voting", Manuscript. http://ssrn.com/abstract=1143564.
KRISHNA, V. and MORGAN, J. (2011), "Overcoming Ideological Bias in Elections", *Journal of Political Economy*, **119**, 183–211.
LASSEN, D. (2005), "The Effect of Information on Voter Turnout: Evidence from a Natural Experiment", *American Journal of Political Science*, **49**, 103–118.
MARTINELLI, C. (2006), "Would Rational Voters Acquire Costly Information?" *Journal of Economic Theory*, **129**, 225–251.
MARTINELLI, C. (2007), "Rational Ignorance and Voting Behavior", *International Journal of Game Theory*, **35**, 315–355.
MATSUSAKA, J. (1995), "Explaining Voter Turnout Patterns: An Information Theory", *Public Choice*, **84**, 91–117.
McDONALD, M.P. and POPKIN, S.L. (2001), "The Myth of the Vanishing Voter", *The American Political Science Review*, **95**, 963–974.
McLENNAN, A. (1998), "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents", *American Political Science Review*, **92**, 413–418.
McMURRAY, J.C. (2012), "Ideology as Opinion: A Spatial Model of Common-value Elections", Manuscript, Brigham Young University.
McMURRAY, J.C. (2011). "The Paradox of Information and Voter Turnout", Manuscript, Brigham Young University.
MEIROWITZ, A. (2002), "Informative Voting and Condorcet Jury Theorems with a Continuum of Types", *Social Choice and Welfare*, **19**, 219–236.
MILGROM, P.R. (1979), "A Convergence Theorem for Competitive Bidding with Differential Information", *Econometrica*, **47**, 670–688.
MILGROM, P.R. and STOKEY, N. "Information, Trade and Common Knowlegde", *Journal of Economic Theory*, **26**, 17–27.
MILGROM, P.R. and WEBER, R.J. (1982), "A Theory of Auctions and Competitive Bidding", *Econometrica*, **50**, 1089–1122.
MORTON, R.B. and TYRAN, J.-R. (2011), "Let the Experts Decide? Asymmetric Information, Abstention, and Coordination in Standing Committees", *Games and Economic Behavior*, **72**, 485–509.
MUELLER, D.C. (2003), *Public Choice III* (New York: Cambridge University Press).
MYERSON, R. (1998), "Population Uncertainty and Poisson Games", *International Journal of Game Theory*, **27**, 375–392.
MYERSON, R. (2000), "Large Poisson Games", *Journal of Economic Theory*, **94**, 7–45.
NITZAN, S. and PAROUSH, J. (1982), "Optimal Decision Rules in Uncertain Dichotomous Choice Situations", *International Economic Review*, **23**, 289–297.
OLIVEROS, S. (2011a), "Abstention, Ideology and Information Acquisition", Manuscript, University of California-Berkeley.
OLIVEROS, S. (2011b), "Aggregation of Endogenous Information in Large Elections", Manuscript, University of California-Berkeley.
OTTAVIANI, M. and SORENSEN, P.N. (2010), "Noise, Information, and the Favorite-Longshot Bias in Parimutuel Predictions", *American Economic Journal: Microeconomics*, **2**, 58–85.
PALFREY, T.R. and POOLE, K.T. (1987), "The Relationship between Information, Ideology, and Voting Behavior", *American Journal of Political Science*, **31**, 511–530.
PALFREY, T.R. and ROSENTHAL, H. (1983), "A Strategic Calculus of Voting". *Public Choice*, **41**, 7–53.
PELEG, B. and ZAMIR, S. (2010), "Extending the Condorcet Jury Theorem to a General Dependent Jury", Manuscript, Hebrew University.
PESENDORFER, W. and SWINKELS, J.M. (1997), "The Loser's Curse and Information Aggregation in Common Value Auctions", *Econometrica*, **65**, 1247–1281.
RIKER, W.H. and ORDESHOOK, P.C. (1968), "A Theory of The Calculus of Voting", *American Political Science Review*, **62**, 25–42.

ROTHSCHILD, M. and STIGLITZ, J.E. (1970), "Increasing Risk: I. A Definition", *Journal of Economic Theory*, **2**, 225–243.

SHAPLEY, L. and B. GROFMAN (1984), "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies", *Public Choice*, **43**, 329–343.

SMITH, L. and SORENSEN, P. (2000), "Pathological Outcomes of Observational Learning", *Econometrica*, **68**, 371–398.

TRIOSSI, M. (2011), "Costly Information Acquisition. Is it Better to Toss a Coin?" Working paper, University of Chile.

WATTENBERG, M. McALLISTER, I. and SALVANTO, A. (2000), "How Voting is Like Taking an SAT Test: An Analysis of American Voter Rolloff", *American Politics Research*, **28**, 234–250.

WILSON, R. (1977), "A Bidding Model of Perfect Competition", *Review of Economic Studies*, **44**, 511–518.

YOUNG, H.P. (1988), "Condorcet's Theory of Voting", *American Political Science Review*, **82**, 1231–1244.