Information and Voting: the Wisdom of the Experts versus the Wisdom of the Masses*

Joseph C. McMurray‡

December, 2009

Abstract

This paper analyzes a two-candidate election in which voters share a common objective but possess information of varying quality. The swing voter's curse (Feddersen and Pesendorfer, 1996) dissuades the least-informed citizens from voting, even though voting is costless, but a substantial fraction of the electorate continue to vote, in an effort to aggregate information, even as the number of "experts" grows arbitrarily large. Thus, in addition to explaining roll-off and the empirical correlation between information and voting, this model explains the moderate levels of turnout observed in real-world elections. It also facilitates a deeper analysis of comparative statics than simpler models, yielding novel empirical predictions. Turnout is also socially optimal, implying that society benefits both by allowing abstention and by allowing non-expert citizens to vote.

JEL Classification Numbers: D72, D82

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

---

*Revised version of Wallis Institute of Political Economy Working Paper No. 59

†Brigham Young University Economics Department. Email jmcmurray@byu.edu.

‡This paper is 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 Gabor Virag, Tim Feddersen, John Duggan, Cesar Martinelli, David Austen-Smith, and participants in the Wallis Political Economy workgroup, the NBER Political Economics Group student conference, and the Wallis Conference on Political Economy, for their interest and suggestions.

1 Introduction

In many situations, individuals within a group seek some common objective, but disagree over how that objective can best be accomplished. Though every member of society values national security and economic stability, for example, opinions vary widely regarding how likely a specific policy is to achieve those outcomes. Similarly, citizens may agree on the need to reduce crime, corruption, or poverty, but disagree over what policies will effectively do so. On a much smaller scale, individual members on a medical panel or business committee may disagree over which medical procedure or business strategy to adopt, despite agreeing on the desirability of health outcomes or profits. In situations such as these, majority voting can be a useful mechanism for pooling private information. The centuries-old Condorcet (1785) jury theorem points out that, by doing so, collective decisions can actually be extremely well-informed.

The Condorcet model assumes that individuals are equally well-informed and that everyone votes. In reality, expertise on any given issue varies widely, and in most cases voting is voluntary. In an influential paper, Feddersen and Pesendorfer (1996) consider a voluntary common-value election with informed and uninformed citizens. In equilibrium, uninformed citizens strategically abstain from voting, deferring to those with better judgment; by voting, an uninformed citizen might inadvertantly overturn an informed decision, thereby suffering a "swing voter's curse". As the authors point out, an informational incentive for abstention might explain abstention in costless voting environments, such as *roll-off* (i.e. voting and abstaining in various races on the same ballot, after voting costs are sunk),[1] as well as empirical evidence (reviewed below in Section 2) that information and voter participation are correlated.

Presumably, the number of experts on any issue is small relative to the number of non-experts. If so, the logic of the swing voter's curse implies that only a small, elete fraction of the electorate should vote. This sharply contrasts the logic of the Condorcet jury theorem that everyone should vote, since even poorly informed citizens, by voting in sufficiently large numbers, can together identify the better alternative with arbitrarily high probability. Indeed, the Feddersen-Pesendorfer (hereafter FP) model assumes that citizens are either perfectly informed or perfectly ignorant; perturbing this even slightly can yield an equilibrium with 100% participation.[2] That such similar informational assumptions could yield

[1] As a typical example of abstention and roll-off, Feddersen and Pesendorfer (1996) report that about three out of six million elligible citizens voted in the 1994 Illinois gubernatorial election, but only two million voted on a proposed amendment to the state constitution, listed on the same ballot.

Abstention is also common in small committees, even though voting is costless, usually requiring only the raise of a hand.

[2] If experts comprise only 2% of an electorate, for example, then assuming that informed and uninformed

such opposite behavioral predictions is surprising. It also begs the question of the true relationships between information, voter participation, and election outcomes: for example, low and declining voter turnout, along with general voter ignorance, are commonly viewed as major societal ills, perhaps even undermining the democratic process. Furthermore, voter participation rates in real-world elections are neither as extremely high nor as extremely low as these and other prevailing models predict, suggesting that existing theories of voter motivations are incomplete.

This paper provides a unified framework through which to analyze the competing informational incentives for voter participation. As in the FP and Condorcet models, an unknown state of the world determines which of two alternatives will better achieve some common objective. Individual opinions are represented by informative signals, each correlated with the true state of the world, but otherwise independent of each other. The quality of an individual's signal depends on her level of expertise. The key assumption of this model is that expertise is drawn from a continuous distribution, so that no one is perfectly informed or perfectly ignorant, and no two citizens have precisely the same expertise.[3] In other words, each citizen possesses valuable private information but, with positive probability, also has better-informed peers. The central result is that, in equilibrium, a positive fraction of the electorate votes, and a positive fraction of the electorate abstains. In fact, illustrative examples suggest that both of these fractions are large (i.e. close to half the electorate) for a wide variety of information distributions, even in large electorates. Equilibrium has a simple analytic characterization, and is unique if the distribution of information is sufficiently smooth, facilitating welfare and comparative static analyses that are not feasible in simpler models.

Given that voting is costless, it may seem surprising that citizens abstain, even after receiving informative signals. The logic for this is the swing voter's curse: if everyone voted informatively, the better of two candidates would be more likely to win the election by a single vote than to lose by a single vote, so one additional vote for the inferior candidate would be more likely to change the election outcome than one additional vote for the superior candidate. To avoid making things worse, therefore, a sufficiently uninformed citizen would prefer to abstain. It may seem desirable in this case to make voting mandatory, since voluntary elections fail to utilize nonvoters' private information. In fact, however, mandatory voting can actually make things worse: an optimal voting mechanism would place the greatest weight on signals of the highest quality; voluntary abstention is a crude way of accomplishing this.

citizens can correctly identify the better of two alternatives with probabilities 0.99 and 0.51 yields a unique equilibrium with 100% participation.

[3] Throughout this paper, feminine pronouns refer to voters, and masculine pronouns refer to candidates.

As an electorate grows, so does the expected number of better-informed peers for a citizen of any expertise. Accordingly, a citizen who was formerly indifferent between voting and abstaining now prefers to abstain, and voter turnout declines. By this logic, it may seem that citizens of every expertise level should eventually abstain; if so, only a vanishing and increasingly elite fraction of the electorate should continue to vote. To the contrary, however, turnout remains bounded above zero, and in fact, may be quite high. For a wide variety of information distributions, even some citizens of below-average expertise continue to vote in a large electorate; for these individuals, the fear of lowering average information quality is offset by a desire to increase information quantity. The logic of common-value elections implies that, while voter ignorance is commonly viewed as a threat to democracy, citizens with relatively little expertise on a given issue actually can, by voting, improve election outcomes. As in both the Condorcet and FP models, a large electorate will almost surely select the candidate or alternative that is in fact superior.

Like the FP model, this model predicts a positive relationship between information and voting, consistent with empirical evidence. On the other hand, such a relationship could also arise in a non-strategic model. The richer information distribution in this model, however, provides an additional implication of strategic abstention, which is that the importance of information is *relative*, rather than absolute.[4] That is, being informed, per se, is not what leads an individual to vote; rather, what matters is being better-informed than other members of the electorate. Consistent with this prediction, McMurray (2009) finds strong evidence in American National Election Studies survey data that measures of relative information explain voter turnout better than conventional absolute measures do. For example, whether a citizen votes or not depends less on her education level than on the fraction of her peers whose education levels are lower than her own. The importance of relative information might explain why voter turnout has remained constant or decreased over time, even as education levels have risen. This model also reaffirms the FP prediction that citizens abstain even when voting is costless.

The remainder of this paper is organized as follows. Section 2 begins by reviewing relevant literature, and then Section 3.1 formally introduces the model that has been briefly described above. Sections 3.2 and 3.3 characterize equilibrium behavior in small and large electorates, respectively, and Section 3.4 presents numerical examples. Section 3.5 analyzes the welfare implications of voter turnout, and Section 3.6 analyzes how voter turnout responds to changes in the underlying distribution of information. Section 4 concludes, and proofs of most analytical results are presented in the Appendix.

[4]The FP model includes only two information types, so that absolute and relative information quality are the same.

2 Literature

The ability of majority voting to aggregate private information was observed over two centuries ago by the French mathematician Condorcet (1785), as mentioned in Section 1, and is one of the oldest formal results in political economy. Condorcet's original analysis implicitly assumes that individuals vote informatively. Only recently did Austen-Smith and Banks (1996) point out that this assumption is not innocuous: a citizen's vote will only influence her own utility in the unlikely event that her vote is pivotal, changing the election outcome; if voting is informative, then the event in which her vote is pivotal may carry additional information, which she should utilize in choosing her behavior. Feddersen and Pesendorfer's (1996) swing voter's curse is one of the earliest applications of Austen-Smith and Banks' insight. In a later paper, Feddersen and Pesendorfer (1998) use a similar information structure to compare simple majority rule with unanimity rule, with application to jury voting. Duggan and Martinelli (2001) and Meirowitz (2002) show these results not to be robust, however, to a more generalized information structure. Mechanically, therefore, this paper is similar to those, though the emphasis here is on abstention rather than unanimity rule. In all of these papers, information is distributed exogenously. Martinelli (2006, 2007) analyzes information acquisition before an election; his (2007) model may be viewed as an earlier stage of the voting game considered here.

Empirical research has identified numerous correlations between voter participation and variables related to information. Controlling extensively for covariates, for example, Wolfinger and Rosenstone (1980) find education to be the single best predictor of voter turnout, and Dee (2004) and Milligan, Moretti, and Oreopoulos (2004) find this relationship to be causal. In these studies, voting also increases with age. Palfrey and Poole (1987), Bartels (1996), Degan and Merlo (2007), and Larcinese (2006) find turnout to be correlated with political knowledge, and Lassen (2005) and Strate et al. (1989), respectively, show that controlling for political knowledge reduces the explanatory power of education and age. Wattenberg, McAllister, and Salvanto (2000) find political knowledge to be the most significant factor in explaining roll-off.[5] Ashenfelter and Kelley (1975) show that voter turnout is also high among individuals recently contacted by campaign workers and low among individuals who have recently moved.

Attempts to understand voting behavior have been repeatedly frustrated by the inability to explain moderate levels of turnout. For simplicity, most models simply assume full participation. In the real world, however, many elligible voters choose to abstain. Similarly, game theoretic models that allow abstention invariably predict extremely low voter turnout in

[5] Coupé and Noury (2004) find evidence that information also influences roll-off by survey participants.

equilibrium[6] because, as Downs (1957) observes, the probability of casting a pivotal vote is so miniscule in large elections that even small voting costs deter all but a tiny fraction of citizens from voting. To rationalize voter participation, Riker and Ordeshook (1968) hypothesize that voters enjoy fulfilling a sense of civic duty (or at least maintaining appearances), so that voting costs are actually negative; or that perhaps voters overestimate the true probability of casting a pivotal vote. Margolis (1982) and, more recently, Edlin, Gelman, and Kaplan (2005), suggest instead that voters are altruistic, perceiving immense rewards for contributing (albeit slightly) to the welfare of large numbers of people. In a similar vein, Harsanyi (1980) and Feddersen and Sandroni (2006) posit that voters derive utility from conforming to ethical rules, which are instituted with social welfare in mind.

With costly voting, Matsusaka (1995) explains the connection between information and voting by pointing out that a citizen who is uncertain which of two candidates she prefers expects a lower benefit from voting than another who is more confident, and so may be dissuaded from voting by a smaller voting cost. To avoid eliminating turnout altogether, he adopts the Riker and Ordeshook (1968) assumption that, for some citizens, voting costs are negative. The problem with this reasoning is that, in large electorates, Downs' (1957) logic still implies that citizens with positive voting costs should abstain; citizens with negative voting costs may vote, but this should not depend on information, because citizens with negative voting costs have no reason to abstain. In particular, information should not matter for roll-off; in fact, since voting costs are sunk, roll-off should not occur at all. Admittedly, by assuming that voting is costless, this paper and the FP model essentially side-step the standard paradox of costly voting. Doing so, however, provides a unified explanation for voter abstention and roll-off, consistent with the empirical correlations summarized above. It is also worth pointing out that social motivations such as duty, altruism, ethical concerns, or desires to improve society implicitly endow voters with a common objective, by virtue of their belonging to the same society. Thus, existing explanations for participation in costly elections are most relavent in a common-value environment, such as this.

3 Analysis

3.1 The Model

A group of individual citizens must collectively choose between two candidates or alternatives, A and B, by simple majority voting. One of the alternatives, $Z \in \{A, B\}$, is designated by Nature as superior to the other; each citizen receives utility $U = 1$ if this alter-

[6] See, for example, Palfrey and Rosenthal (1983, 1985).

native is selected and $U = 0$ otherwise. Letting $X \in \{A, B\}$ denote the chosen alternative, therefore, expected utility is given merely by the probability $\Pr(X = Z)$.

Though citizens cannot observe Z directly, it is commonly known that $Z = A$ with probability $\theta = \frac{1}{2}$. The precise number N of potential voters is also unknown, but is commonly known to follow a Poisson distribution with mean μ.[7] For a particular realization of N, each citizen is endowed with information quality $Q_i \in \left[\frac{1}{2}, 1\right]$, representing her level of expertise on the issue at hand, drawn independently from a common distribution F, which has a differentiable density f that is strictly positive between $\frac{1}{2}$ and 1. She then observes a signal $S_i \in \{A, B\}$ that correctly identifies Z with probability $Q_i \equiv \Pr(S_i = A|Z = A) = \Pr(S_i = B|Z = B)$. To the most expert agent (i.e. $Q_i = 1$), for example, S_i reveals Z perfectly; to the least expert agent (i.e. $Q_i = \frac{1}{2}$), S_i reveals nothing. The distribution F of quality levels is assumed to be common knowledge, while Q_i and S_i are observed only privately. Signal values are independent of expertise, independent of the population size, and (conditional on Z) independent of one another.

An individual may choose to vote (at no cost) for either candidate or to abstain. A *symmetric strategy profile* $\sigma : \left[\frac{1}{2}, 1\right] \times \{A, B\} \to \Delta(\{A, B, 0\})$ must therefore specify mixture probabilities $\left(\sigma^A, \sigma^B, \sigma^0\right)$ for each quality type $q \in \left[\frac{1}{2}, 1\right]$ and signal value $s \in \{A, B\}$, where a vote for candidate 0 represents abstention. Let Σ denote the set of such strategies. Since only pure strategies will be relevant in equilibrium, let $\sigma(q, s) = A$ be shorthand notation for the mixed strategy $\sigma(q, s) = (1, 0, 0)$, with similar notation for pure strategies B and 0.

When σ describes her opponents' behavior, an individual may respond according to another symmetric strategy σ_i (defined, for convenience, for an individual who has not yet observed her private information); in general, the probability of the desired election outcome will depend on both of these strategies. σ_{BR} is thus said to be a *best response* to σ if $\sigma_{BR}(q, s)$ maximizes

$$EU(\sigma_i|q, s; \sigma) \equiv \Pr(X = Z|q, s; \sigma, \sigma_i,)$$

for every $(q, s) \in \left[\frac{1}{2}, 1\right] \times \{A, B\}$, and σ^* is a *symmetric Bayesian equilibrium* if it is its own best response.

To facilitate the exposition of results, it is useful to introduce some additional notation, which will later simplify symmetrically in equilibrium. First, given a strategy profile σ, a randomly chosen citizen votes for alternative $x \in \{A, B\}$ in state $z \in \{A, B\}$ with probability

[7] The technical advantages of this assumption are demonstrated by Myerson (1998, 2000) and Bade (2006).

$p_{xz}(\sigma)$, defined as follows:

$$p_{xA}(\sigma) = \int_{1/2}^{1} [q\sigma^x(q,A) + (1-q)\sigma^x(q,B)]\, dF(q) \tag{1}$$

$$p_{xB}(\sigma) = \int_{1/2}^{1} [q\sigma^x(q,B) + (1-q)\sigma^x(q,A)]\, dF(q). \tag{2}$$

By the decomposition property of Poisson random variables (see Myerson, 1998), the numbers $N_{Az}(\sigma)$ and $N_{Bz}(\sigma)$ of votes for candidates A and B in state z are independent Poisson random variables with means $\mu p_{Az}(\sigma)$ and $\mu p_{Bz}(\sigma)$. Accordingly, the probability of any voting outcome is merely the product of Poisson probabilities. For example, the probability $\pi_{wz}(\sigma)$ with which the superior candidate receives exactly w more votes than his opponent is merely the infinite sum of probabilities of $n+w$ and n votes for and against the superior candidate:[8]

$$\pi_{wA}(\sigma) = \sum_{n=\min\{0,w\}}^{\infty} \frac{e^{-\mu p_{AA}(\sigma)}[\mu p_{AA}(\sigma)]^{n+w}}{(n+w)!} \frac{e^{-\mu p_{BA}(\sigma)}[\mu p_{BA}(\sigma)]^{n}}{n!} \tag{3}$$

$$\pi_{wB}(\sigma) = \sum_{n=\min\{0,w\}}^{\infty} \frac{e^{-\mu p_{AB}(\sigma)}[\mu p_{AB}(\sigma)]^{n}}{n!} \frac{e^{-\mu p_{BB}(\sigma)}[\mu p_{BB}(\sigma)]^{n+w}}{(n+w)!}. \tag{4}$$

Of particular interest are the probabilities $\pi_{0z}(\sigma)$, $\pi_{1z}(\sigma)$, and $\pi_{-1z}(\sigma)$ 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. Specifically, an additional vote for the superior candidate would be pivotal when that candidate either ties the election and loses the tie-breaking coin toss, or wins the coin toss but loses the election by exactly one vote. The combined probability of these events is given by $P_z(\sigma)$:

$$P_z(\sigma) = \frac{1}{2}\pi_{0z}(\sigma) + \frac{1}{2}\pi_{-1z}(\sigma). \tag{5}$$

Similarly, a vote for the inferior candidate would be pivotal with probability $\tilde{P}_z(\sigma)$:

$$\tilde{P}_z(\sigma) = \frac{1}{2}\pi_{0z}(\sigma) + \frac{1}{2}\pi_{1z}(\sigma). \tag{6}$$

3.2 Equilibrium

By the environmental equivalence property of Poisson games (see Myerson, 1998), an individual from within the game reinterprets $N_{xz}(\sigma)$ as the number of x votes cast in state

[8]Note that this formulation of π_{wz} accommodates $w < 0$, with the interpretation that Z loses the election by $|w|$ votes.

z by her peers–by voting herself, she can add one to either candidate's total. Accordingly, $\pi_{wz}(\sigma)$ reflect the likelihood of voting outcomes if she herself abstains, and $P_z(\sigma)$ and $\tilde{P}_z(\sigma)$ are the probabilities with which her own vote will change the election outcome, if she votes for or against candidate Z, respectively.

In addition to anticipating the behavior of her peers, an individual must evaluate her own beliefs about the state of the world. For a citizen of type (q, s), let $\beta(q, s) \equiv \Pr(Z = B|Q_i = q, S_i = s)$ denote the posterior probability that B is the better candidate:

$$\beta(q, s) = \begin{cases} 1 - q \text{ if } s = A \\ q \text{ if } s = B \end{cases}. \tag{7}$$

As can be seen in (7), A and B signals lower and raise $\beta(q, s)$, respectively, from the prior of $\frac{1}{2}$. Agents with high Q_i will have the most extreme posterior beliefs. Since f has full support over $\left[\frac{1}{2}, 1\right]$, $\beta(q, s)$ will have full support over $[0, 1]$. Lemma 1 now states that the best response to any strategy profile can be characterized by posterior voting thresholds, defined as follows.

Definition 1 *The symmetric strategy profile σ_{T_A,T_B} is a posterior threshold strategy if there exist thresholds T_A and T_B (with $1-T_A \leq T_B$) such that* $\sigma_{T_A,T_B}(q, s) = \begin{cases} A \text{ if } 1 - \beta(q, s) \geq T_A \\ B \text{ if } \beta(q, s) \geq T_B \\ 0 \text{ otherwise} \end{cases}$ *for every $(q, s) \in \left[\frac{1}{2}, 1\right] \times \{A, B\}$.*

Lemma 1 *σ^* is a best response to σ only if it is a posterior threshold strategy.*

Proof. See Appendix. ■

Given the symmetry of this model, it is natural to direct attention to the special case of a posterior threshold strategy $\sigma_{T,T}$ in which the posterior thresholds coincide. In that case, equation (7) makes clear that voter participation depends only on q, since $1 - \beta(q, A) \geq T$ and $\beta(q, B) \geq T$ are simultaneously true if and only if $q \geq T$. Thus T is merely a quality threshold, and $\sigma_{T,T}$ can be reconceived as a quality threshold σ_T, defined below. Voting under a quality threshold strategy is *sincere* or *informative*, in the sense that a citizen votes A or B only after receiving an A or B signal, respectively.

Definition 2 *σ_T is a quality threshold strategy if* $\sigma_T(q, s) = \begin{cases} s \text{ if } q \geq T \\ 0 \text{ otherwise} \end{cases}$.

Since a quality threshold strategy prescribes symmetric behavior in the two states of the world, expected vote shares (1) and (2) do not depend on the state:

$$p_{AA}(\sigma_T) = p_{BB}(\sigma_T) = \int_T^1 q dF(q) \equiv p_+(\sigma_T) \tag{8}$$

$$p_{AB}(\sigma_T) = p_{BA}(\sigma_T) = \int_T^1 (1-q) dF(q) \equiv p_-(\sigma_T). \tag{9}$$

Consequently, the numbers $N_{AA}(\sigma_T) = N_{BB}(\sigma_T) \equiv N_+(\sigma_T)$ and $N_{AB}(\sigma_T) = N_{BA}(\sigma_T) \equiv N_-(\sigma_T)$ of votes for and against Z also have identical distributions in the two states (denoted by equality here), so that win and pivot probabilities (3) through (6) simplify as well: $\pi_{wA}(\sigma_T) = \pi_{wB}(\sigma_T) \equiv \pi_w(\sigma_T)$, $P_A(\sigma_T) = P_B(\sigma_T) \equiv P(\sigma_T)$, and $\tilde{P}_A(\sigma_T) = \tilde{P}_B(\sigma_T) \equiv \tilde{P}(\sigma_T)$. This in turn leads to symmetric posterior thresholds $T_A(\sigma_T) = T_B(\sigma_T) \equiv T_{BR}(\sigma_T)$, so that the best response to σ_T is another quality threshold strategy $\sigma_{T_{BR}(\sigma_T)}$. By proving the existence of a fixed point $T^* = T_{BR}(\sigma_T)$ of (10), Theorem 3 guarantees the existence of a quality threshold strategy profile that is a symmetric Bayesian equilibrium.

Theorem 3 *There exists a threshold T^* strictly between $\frac{1}{2}$ and 1 such that the quality threshold strategy σ_{T^*} is a symmetric Bayesian equilibrium.*

Proof. See Appendix. ■

As the proof of Theorem 3 explains, the formula for $T_{BR}(\sigma_T)$ is as follows:

$$T_{BR}(\sigma_T) \equiv \frac{\tilde{P}(\sigma_T)}{P(\sigma_T) + \tilde{P}(\sigma_T)}. \tag{10}$$

The right hand side of (10) can be interpreted as a conditional probability with which a vote *against* Z is pivotal, given that a vote–either for or against Z–is pivotal at all. If a vote against Z is much more likely to be pivotal than a vote for Z then the swing voter's curse is strong: only individuals with high quality information will cast votes. As Theorem 3 points out, the fact that individual opinions are more often right than wrong implies that Z is more likely to lead by one vote than to trail by one vote, and so that an additional vote for Z's opponent is more likely to be pivotal than an additional vote for Z. In equilibrium, therefore, a positive (expected) fraction of the electorate abstain.

Theorem 3 does not guarantee the uniqueness of equilibrium; in general, there may be multiple equilibrium thresholds. Theorem 6 demonstrates uniqueness in the limit, however, provided that the density f is log-concave. I conjecture that uniqueness for any population size parameter would be guaranteed by a similar smoothness condition. In any case, Theorem 4 next shows that no other equilibria exist. That is, every equilibrium can be characterized by a quality threshold. If there is indeed a unique equilibrium quality threshold, Theorem 4 implies that this is the unique symmetric Bayesian equilibrium for the game.

Theorem 4 *σ^* is a symmetric Bayesian equilibrium only if it is a quality threshold strategy.*

Proof. See Appendix. ■

3.3 Large Elections

Theorems 3 and 4 in section 3.2 characterize equilibrium voting behavior for a fixed (expected) population size μ. The purpose of this section is to analyze voting behavior as μ grows large. Lemma 2 first shows that the swing voter's curse intensifies as an electorate grows: the best response threshold increases with μ. This implies that (if it is unique) the equilibrium threshold T_μ^* also increases with μ, so that voter participation falls.[9]

Lemma 2 *For any $T < 1$, the best response threshold $T_{BR}(\sigma_T; \mu)$ is increasing in μ.*

Proof. See Appendix. ■

The result that turnout declines as less-informed agents increasingly delegate to those with better information is reminiscent of the original FP model, in which uninformed voters abstain with increasing probability until, in the limit, only (perfectly) informed agents continue to vote. This begs the question of how exclusive in this model voter participation eventually becomes. One possible intuition is that T_μ^* might approach one as $\mu \to \infty$ because a well-informed citizen votes as long as she reasonably expects to be among the best-informed members of a small electorate, but eventually abstains as the number of better-informed peers grows arbitrarily large. Since no one in this model possesses perfect information, this would imply voter turnout approaching 0%.

As theorem 5 next demonstrates, however, this is not the case: T_μ^* remains bounded below one, so that voter turnout remains positive even in an arbitrarily large electorate. While a citizen indeed expects a large number of extremely well-informed votes, *ex ante*, she bases her behavior on the *conditional* expectation that a surprisingly large fraction of her peers have simultaneously erred, thereby rendering her own vote pivotal.[10] The probability of a one-vote win continues to grow relative to the probability of a one-vote loss, but at an ever slowing pace.[11] The condition that $\lim_{T\to 1} \frac{f'(T)}{f(T)} < \infty$ merely requires that $f(T)$ not

[9] When multiple thresholds support equilibria for a given μ, Lemma 2 implies that the lowest and highest of these thresholds both increase with μ.

[10] When her vote is pivotal, a voter infers her peers to be fewer in number (i.e. low N), less well-informed (i.e. low Q), and less accurate (i.e. high fraction $S_i \neq Z$, given Q) than she had originally expected.

[11] The expected margin of victory $E(N_+ - N_-)$ increases with μ. By itself, this would cause a one-vote win to become exponentially more likely than a one-vote loss. At the same time, however, the variance $Var(N_+ - N_-)$ of possible election outcomes grows by the same factor, reducing the distinction between the two events. Note that this would still be the case if, say, N were fixed and known, so that N_+ and N_- had binomial distribution instead of Poisson.

grow infinite too rapidly as $T \to 1$, reinforcing the assumption that no one possesses perfect information.[12]

Theorem 5 *Let* $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ *be a sequence of equilibrium participation thresholds for a sequence* $\{\mu_k\}_{k=1}^{\infty}$ *of population parameters such that* $\mu_k \to \infty$*, and let* T^*_∞ *be a limit point of* $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$*. If* $\lim_{T\to 1} \frac{f'(T)}{f(T)} < \infty$ *then* $T^*_\infty < 1$*.*

Proof. See Appendix. ■

The final theorem of this section states that the limiting quality threshold T^*_∞ is unique, provided only that the density f is *log-concave* (i.e. $\log(f)$ is concave or, equivalently, $\frac{f'}{f}$ is decreasing), as is true for many of the most common distributions.[13] As the proof of Theorem 6 makes clear, log-concavity is actually a stronger condition than necessary: T^*_∞ may easily be unique, for example, if F is bimodal, though f is not log-concave in that case. What is important for uniqueness is that f not have sudden "spikes" of probability, which log-concavity rules out. The uniqueness of T^*_∞ does not by itself imply a unique T^*_μ for even large μ, but it does imply that participation thresholds converge to T^*_∞, so that turnout in a large electorate is uniqueely determined.

Theorem 6 *Let* $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ *be a sequence of equilibrium participation thresholds for a sequence* $\{\mu_k\}_{k=1}^{\infty}$ *of population parameters such that* $\mu_k \to \infty$*. If* f *is log-concave then* $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ *approaches a unique limit* $T^*_\infty < 1$*.*

Proof. See Appendix. ■

3.4 Examples

The asymptotic uniqueness of the equilibrium quality threshold implies that voter participation in large electorates is completely determined by the underlying distribution F of information. Furthermore, equation (39) in the proof of Theorem 5 provides a simple formula for determining T^*_∞ for a given distribution function, either analytically or numerically. This allows us to generate precise examples of equilibrium behavior, using simple distribution functions. Figure 1 displays three such examples. After computing the limiting equilibrium quality threshold T^*_∞, voter participation is simply given by the *survival function* $\bar{F} \equiv 1 - F$, evaluated at T^*.

---

[12] If $\lim_{T\to 1} \frac{f'(T)}{f(T)} = \infty$ then all higher-order derivatives are unbounded as $T \to 1$.

[13] Bagnoli and Bergstrom (2005) show the following distributions to have log-concave densities: uniform, normal, logistic, extreme value, chi-squared, chi, exponential, Laplace, Weibull (for some parameter values), power function, gamma, and beta. Also, any truncation, linear transformation, or mirror-image of a log-concave density is log-concave.

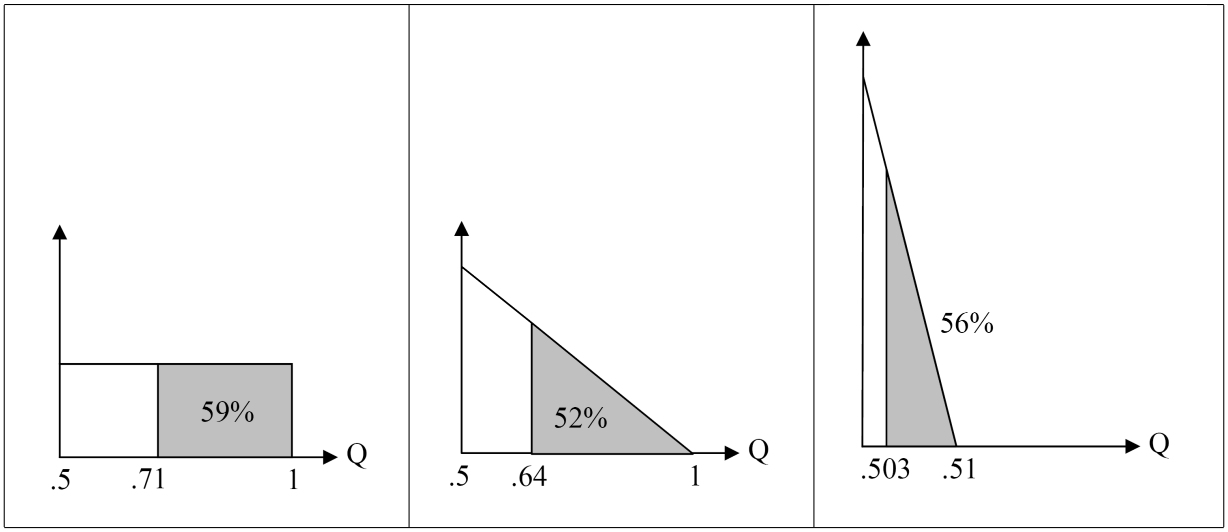

Figure 1: Participation thresholds and turnout rates for simple information distributions

If information quality is distributed uniformly between $\frac{1}{2}$ and 1, as in the first frame of Figure 1, then $T_\infty^* = 0.71$, implying that the best-informed 59% of citizens vote. If the distribution of information is skewed so that experts are outnumbered by non-experts, as in the second frame of Figure 1, then the quality threshold falls to $T_\infty^* = 0.64$ and turnout falls to 59%. If, as in the third frame of Figure 1, even the most expert members of society have only 51% accurate information, the best-informed 56% of the electorate will vote in equilibrium.

Theorem 5 in the previous section states that the equilibrium quality threshold remains bounded, so that turnout remains positive even in a large electorate. Strictly speaking, however, this guarantees only that the turnout rate exceeds some positive fraction $\varepsilon > 0$ of the electorate, leaving open the theoretical possibility that turnout is essentially zero in large elections. The examples in Figure 1 demonstrate, however, that turnout may remain remarkably high: for each of the example distributions, even citizens with below-average information quality continue to vote, so that predicted turnout exceeds 50% of the electorate. As discussed in the Introduction, this level of voter turnout is unusual in strategic voting models, which often predict extremely high or low voter participation.

3.5 Welfare and Election Design

By the logic of the original Condorcet model, equilibrium (i.e. informative) voting by a large electorate almost certainly elects the superior candidate (i.e. $\lim_{\mu\to\infty} \Pr(X = Z|\sigma_{T^*};\mu) = 1$). On the other hand, one central message of the original Condorcet jury theorem is that election decisions are best made by utilizing the independent information of as many voters as possible, even if that information is of low quality; by failing to utilize nonvoters' valuable information, equilibrium voting in this model is in one sense inefficient. This begs the question of whether voting should be made compulsory, as it is in a number of democracies (e.g. Australia and several Latin American countries), and as has been recommended for the United States (e.g. Lijphart, 1997). In answer, however, corollary 7 states that the socially optimal threshold strategy is in an equilibrium, which, as the proofs of theorems 3 and 4 explain, necessarily involves some voter abstention.[14] Thus, adding votes beyond the equilibrium level may actually reduce welfare, rather than enhance it.

Corollary 7 *Let* $\sigma_o \equiv \arg\max_{\sigma\in\Sigma} \Pr(X = Z|\sigma)$ *denote the optimal voting strategy. Then*

1. σ_o *is a symmetric Bayesian equilibrium, and can be characterized by an equilibrium threshold* $T^* \in \left[\frac{1}{2}, 1\right]$.

2. $\Pr(X = Z|\sigma_o) > \Pr\left(X = Z|\sigma_{\frac{1}{2}}\right)$.

Proof. First, it is important to demonstrate that σ_o is well-defined. This is guaranteed by the Weierstrass extreme value theorem because welfare $\Pr(X = Z|\sigma)$ is a continuous function of σ on the compact set $\Sigma = \left\{\sigma : \left[\frac{1}{2}, 1\right] \times \{A, B\} \to \Delta(\{A, B, 0\})\right\}$ of symmetric strategies (under the sup norm topology). In a common interest game such as this, social welfare and private utility are identical; as McLennan (1998) shows, this implies that the social optimum is also individually optimal, and thus constitutes an equilibrium. By Theorem 4, an equilibrium can be characterized by a quality threshold T^*. This establishes part 1. Part 2 follows simply because, as the proof of Theorem 3 shows, $\sigma_{\frac{1}{2}}$ is not an equilibrium, and so must be inferior to σ_o. ■

It may seem surprising that adding informative votes can reduce the quality of an election outcome. One way to understand this is that voters actually possess two pieces of private information: S_i and Q_i. An ideal election mechanism would obtain both, and would weight individual votes by their underlying quality.[15] A compulsory election collects a larger number of signals than a voluntary election, but must weigh them equally; in a voluntary election, the *decision* to vote conveys information about Q_i, allowing greater weight on high-quality than on low-quality signals (i.e. positive instead of zero weight).[16]

[14] The result from Theorem 6 that equilibrium turnout approaches a unique limit implies that, in large electorates, the equilibrium level of turnout is arbitrarily close to the social optimum.

[15] Shapley and Grofman (1984) show that the optimal use of information would be a maximum likelihood approach. A similar mechanism that is more common is scoring, as in art and athletic competitions. When a judge is unable to distinguish between contestants, she can award similar scores, essentially delegating to the other judges.

[16] In a model similar to this, Krishna and Morgan (2008) likewise conclude that mandatory voting can lower welfare, but for a different reason. In their model, the accuracy of private information is asymmetric

Corollary 7 formally demonstrates the social benefit of voter abstention. Informally, popular literature has long denounced uninformed voting, calling for poorly informed citizens to stay home on election day, leaving elections to those who are better informed. Lijphart (1997) notes that this was even codified at the turn of the twentieth century in Belgium, where educated citizens were allowed to cast multiple votes in public elections. The examples in section 3.4, however, show that socially optimal (i.e. equilibrium) voter participation extends strikingly far down the information distribution, including many who are by all standards non-experts. Thus, efforts to influence voter turnout should seek neither of the extremes of full and exclusive voter participation.

3.6 Comparative Statics

By Theorem 6, equilibrium behavior in a large electorate is uniquely determined by the underlying information distribution F. This, in turn, is determined by factors that may vary both regionally and over time, such as voters' education or experience levels, and access to information technology. Accordingly, I denote in this section the limiting equilibrium quality threshold and its associated level of voter turnout by T_{∞}^{F} and $\bar{F}\left(T_{\infty}^{F}\right)$ for a given distribution function F, and compare these with T_{∞}^{G} and $\bar{G}\left(T_{\infty}^{G}\right)$ of an alternative distribution G.

Proposition 1 first analyzes a general improvement in information quality. G is said to *first-order stochastically dominate* F (denoted $G \geq_1 F$) 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. $\bar{G}(q) \geq \bar{F}(q)$ for all q). In general, moving from F to G has two opposite effects: turnout increases as nonvoters are lifted above the participation threshold, but decreases as improved voter information strengthens the swing voter's curse. Which of these two effects dominates depends primarily on whose information quality improves most, as Proposition 1 delineates: (1) below T_{∞}^{*}, small information improvements have no effect because citizens do not vote; (2) above T_{∞}^{*}, information improvements lower turnout by strengthening the swing voter's curse; (3) moderate improvements in nonvoters' information increase turnout, both directly by pushing nonvoters above T_{∞}^{*} and by (lowering the average vote quality, thereby) weakening the swing voter's curse.[17] These three effects are illustrated in Figure 2, starting from a uniform distribution. Note that, regardless of its effect on turnout,

with respect to the state (i.e. A signals are more reliable than B signals). When abstention is allowed, voters compensate for this by abstaining in different proportions in response to the two signals. When abstention is not allowed, they instead compensate by voting uninformatively (with some probability), with the consequence that less information is revealed and welfare is reduced.

[17] Symmetrically, information reductions below T have no effect, small reductions above T increase turnout, and moderate reductions above T reduce turnout.

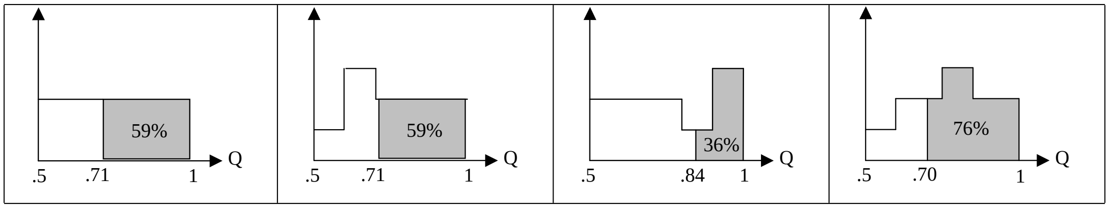

Figure 2: When information improves for different segments of the electorate, turnout may remain the same, decrease, or increase

improving information quality can only improve social welfare.[18]

Proposition 1 *Let F and G be continuous, log-concave distributions with strictly positive densities, and suppose $G \geq_1 F$. Then the following must be true:*

1. If $G(q) = F(q)$ for all $q \geq T_\infty^F$ then $T_\infty^G = T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) = \bar{F}\left(T_\infty^F\right)$.

2. If $G(q) = F(q)$ for all $q \leq T_\infty^F$ then $T_\infty^G \geq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \leq \bar{F}\left(T_\infty^F\right)$.

3. If $G(q) = F(q)$ for all $q \geq E_F\left(Q|Q \geq T_\infty^F\right)$ and $\bar{G}\left(T_\infty^F\right) \geq \bar{F}\left(T_\infty^F\right)$ then $T_\infty^G \leq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$.

Proof. See Appendix. ∎

Proposition 1 considers the case in which one electorate is better informed than another. Proposition 2 considers a second possibility, which is that one population is more homogeneous than the other. G and F are said to satisfy the *single-crossing property* (denoted $G \geq_{sc} F$) if there exists $\hat{q} \in \left[\frac{1}{2}, 1\right]$ such that $F(q) \geq G(q)$ for all $q \leq \hat{q}$ and $F(q) \leq G(q)$ for all $q \geq \hat{q}$. When F and G share a common mean, the single-crossing property implies that G has a smaller variance than F.[19] In that case, (as long as G crosses F above T_∞^F) Proposition 2 states that turnout is higher under G than under F, as illustrated in Figure 3. Intuitively, this is because the swing voter's curse is weak when the quality difference between informed and uninformed votes is small; the most extreme case is the Condorcet model, in which voters are identical and turnout is 100%.

Proposition 2 *Let F and G be continuous, log-concave distributions with strictly positive densities and a common mean, such that G crosses F at $\hat{q}$. If $T_\infty^F \leq \hat{q}$ then $T_\infty^G \leq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$.*

Proof. See Appendix. ∎

[18] The direct effect of improving information quality is to improve election accuracy; in this common value environment, any strategic response to improved information only improves welfare further.

[19] For distributions with a common mean, the single-crossing property is stronger than (i.e. implies) second-order stochastic dominance, but weaker than (i.e. is implied by) first-order stochastic dominance.

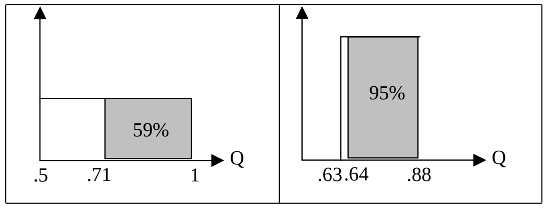

Figure 3: A mean-preserving decrease in the variance of information quality raises turnout.

4 Conclusion

The influential models of Condorcet (1785) and Feddersen and Pesendorfer (1996) predict opposite reactions to imperfect information: in the FP model, poorly informed citizens abstain from voting, deferring to informed voters and thereby avoiding the swing voter's curse; in the Condorcet model, poorly informed citizens vote in as large numbers as possible. By allowing a more general distribution of expertise, this papers blends the above models to unify the analysis of information and voting. In equilibrium, competing incentives balance so that citizens both vote and abstain, above and below an information threshold. Even in large electorates, this threshold is bounded such that the fractions of citizens who continue to vote and abstain remain substantial.

Like the FP model, this model's equilibrium structure is consistent with extensive empirical evidence that information is a central determinant of voter participation. If citizens possess expertise only on certain issues, it may also explain the common phenomenon of roll-off, by which citizens vote and abstain on the same ballot. Assuming that expertise is correlated accross issues, it is also consistent with empirical evidence that information reduces roll-off. Standard voting models fail to explain roll-off, since costless voting environments give citizens no reason to abstain. They invariably predict either extremely high or extremely low voter participation, in contrast with the moderate turnout observed in actual elections. In such models, information should matter only when voting is costly.

The general information structure in this model enables analysis of the impact on voter turnout and welfare of changes to the underlying distribution of information. This is impossible in simpler models that allow only one or two information levels. The generalized information structure also highlights an implication of strategic abstention that is not apparent in simpler models. That is, in deciding whether to vote or abstain, a citizen must consider not only her own expertise, but also the expertise of others in the electorate, who will make the collective decision without her if she abstains. In other words, the importance of information for voting is relative, rather than absolute. Consistent with this prediction, McMurray (2009) finds evidence in American National Election Studies data that citizens with well-informed peers are less likely to vote, both in general elections and in presidential primaries. Among those who vote in the presidential election, those with well-informed peers are also less likely to participate in senate or gubernatorial races on the same ballot.

Voter abstention and ignorance are commonly viewed as major threats to democracy. The first of these concerns can be justified by the Condorcet jury theorem, since election decisions are best made by incorporating as much private information as possible, suggesting a useful role for get-out-the-vote efforts, or even mandatory voting laws. The concern of voter ignorance can be justified by the swing voter's curse, since uninformed citizens might inadvertantly overturn an informed collective decision. As discussed in Section 3.5, however, the logic of common values implies that the socially optimal level of voter participation is achieved in equilibrium. Voluntary abstention, therefore, can actually improve election outcomes, by placing greater weight on higher quality signals. At the same time, election results can be improved even by votes from citizens with relatively little expertise. Efforts to reduce voting costs or improve voter information can improve welfare, but are unlikely to eliminate voter abstention, and may even reduce voter turnout. Voter participation *per se* is therefore less useful as an indication of the quality of election outcomes than is commonly believed.

A Proofs

Whenever meaning is clear in the following notation, I supress the argument (σ) from p_{xz}, π_{wz}, P_z, $\tilde{P}_z$, T_A, T_B, $\hat{\beta}_{BA}$, $\hat{\beta}_{B0}$, and $\hat{\beta}_{A0}$.

Lemma 1 *σ^* is a best response to σ only if it is a posterior threshold strategy.*

Proof. In state B, a vote for B (the superior candidate) is pivotal with probability $P_B(\sigma)$, and yields a utility benefit of $+1$. In state A, a vote for B (the inferior candidate) is pivotal with probability $\tilde{P}_A(\sigma)$, and yields a utility penalty of -1. For an individual with private information (q, s), therefore, the expected benefit of voting for candidate B is given by the following difference:

$$EU(B|q,s;\sigma) - EU(0|q,s;\sigma) = \beta(q,s) P_B(\sigma) - (1-\beta(q,s)) \tilde{P}_A(\sigma),$$

which is positive if and only if $\beta(q,s) \geq \frac{\tilde{P}_A(\sigma)}{P_B(\sigma)+\tilde{P}_A(\sigma)} \equiv \hat{\beta}_{B0}(\sigma)$. By similar calculations, the expected benefit $EU(A|q,s;\sigma) - EU(0|q,s;\sigma)$ of voting for A is positive if and only if $\beta(q,s) \leq \frac{P_A(\sigma)}{P_A(\sigma)+\tilde{P}_B(\sigma)} \equiv \hat{\beta}_{A0}(\sigma)$, and the expected benefit $EU(B|q,s;\sigma) - EU(A|q,s;\sigma)$ of voting B instead of A is positive if and only if $\beta(q,s) \geq \frac{P_A(\sigma)+\tilde{P}_A(\sigma)}{P_A(\sigma)+\tilde{P}_A(\sigma)+P_B(\sigma)+\tilde{P}_B(\sigma)} \equiv \hat{\beta}_{BA}(\sigma)$. Algebraically, it must be that either $\hat{\beta}_{B0}(\sigma) \leq \hat{\beta}_{BA}(\sigma) \leq \hat{\beta}_{A0}(\sigma)$, or else that $\hat{\beta}_{A0}(\sigma) \leq$

$\hat{\beta}_{BA}(\sigma) \leq \hat{\beta}_{B0}(\sigma)$. In the first case, no agent prefers to abstain, so the best response to σ can be characterized by posterior thresholds $T_A(\sigma) = 1 - \hat{\beta}_{BA}(\sigma)$ and $T_B(\sigma) = \hat{\beta}_{BA}(\sigma)$. In the second case, abstention is preferable to voting for $\beta(q,s)$ between $\hat{\beta}_{A0}(\sigma)$ and $\hat{\beta}_{B0}(\sigma)$, so $T_A(\sigma) = 1 - \hat{\beta}_{A0}(\sigma)$ and $T_B(\sigma) = \hat{\beta}_{B0}(\sigma)$ characterize the best response. ■

Theorem 8 *There exists a threshold T^* strictly between $\frac{1}{2}$ and 1 such that the quality threshold strategy σ_{T^*} is a symmetric Bayesian equilibrium.*

Proof. Since $Q_i \geq \frac{1}{2}$ by assumption, equations (8) and (9) make clear that $p_+(\sigma_T) > p_-(\sigma_T)$ for any $T < 1$, which implies that $\pi_1(\sigma_T) > \pi_{-1}(\sigma_T)$, and therefore that $P(\sigma_T) < \tilde{P}(\sigma_T)$. The proof of Lemma 1 characterizes best response posterior thresholds using ratios which, for a quality threshold strategy σ_T, simplify to the following: $\hat{\beta}_{A0}(\sigma_T) = \frac{P(\sigma_T)}{P(\sigma_T)+\tilde{P}(\sigma_T)}$, $\hat{\beta}_{B0}(\sigma_T) = \frac{\tilde{P}(\sigma_T)}{P(\sigma_T)+\tilde{P}(\sigma_T)}$, and $\hat{\beta}_{BA}(\sigma_T) = \frac{1}{2}$. $P(\sigma_T) < \tilde{P}(\sigma_T)$ implies that $\hat{\beta}_{A0}(\sigma_T) < \hat{\beta}_{BA}(\sigma_T) < \hat{\beta}_{B0}(\sigma_T)$, so that the best response is characterized by $T_{BR}(\sigma_T) = \hat{\beta}_{B0}(\sigma_T)$, consistent with (10). It also implies that $T_{BR}(\sigma_T) \geq \frac{1}{2}$ for any $T \leq 1$, with strict inequality whenever $T < 1$. In particular, $T_{BR}\left(\frac{1}{2}\right) > \frac{1}{2}$ and $T_{BR}(1) < 1$. Since $T_{BR}(\sigma_T)$ is also continuous in T, the existence of a fixed point $T^* = T_{BR}(\sigma_T)$ strictly between $\frac{1}{2}$ and 1 follows from the intermediate value theorem. ■

Lemmas A1 through A3 are useful in preparation for the proof of Theorem 4.

Lemma A1 *A posterior threshold strategy with no abstention (i.e. σ_{T_A,T_B} with $1-T_A = T_B$) is not a symmetric Bayesian equilibrium.*

Proof. As the proof of Lemma 1 explains, a best response will include some abstention if $\hat{\beta}_{A0} < \hat{\beta}_{B0}$ or, equivalently, if $\tilde{P}_A\tilde{P}_B > P_AP_B$. As shown below, this inequality holds for σ_{T_A,T_B} with $1 - T_A = T_B > \frac{1}{2}$, implying that σ_{T_A,T_B} is not its own best response. If $T_A = 1 - T_B > \frac{1}{2}$ then a symmetric argument applies, and $T_A = 1 - T_B = \frac{1}{2}$ only if $\sigma^* = \sigma_{1/2}$, which is shown not to be an equilibrium in the proof of Theorem 3.

When $1-T_A = T_B$, well-informed citizens vote informatively and poorly-informed citizens merely vote A. In this case, expressions for p_{xz} reduce from (1) and (2), as follows:

$$\begin{array}{ll} p_{AA} = F(T_B) + \int_{T_B}^1 q dF(q) & p_{BA} = \int_{T_B}^1 (1-q) dF(q) \\ p_{AB} = F(T_B) + \int_{T_B}^1 (1-q) dF(q) & p_{BB} = \int_{T_B}^1 q dF(q). \end{array} \tag{11}$$

These expressions, together with (1) through (6), imply the following:

$$
\begin{aligned}
& \tilde{P}_A \tilde{P}_B - P_A P_B \\
= \; & \frac{1}{2}\left(\pi_{0A} + \pi_{1A}\right) \frac{1}{2}\left(\pi_{0B} + \pi_{1B}\right) - \frac{1}{2}\left(\pi_{0A} + \pi_{-1A}\right) \frac{1}{2}\left(\pi_{0B} + \pi_{-1B}\right) \\
= \; & \frac{1}{4} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{e^{-\mu} \mu^{2j+2k}}{j!k!} p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k} \left[\left(1 + \frac{\mu p_{AA}}{j+1}\right)\left(1 + \frac{\mu p_{BB}}{k+1}\right) - \left(1 + \frac{\mu p_{BA}}{j+1}\right)\left(1 + \frac{\mu p_{AB}}{k+1}\right) \right] \\
= \; & \frac{e^{-\mu}}{4} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{\mu^{2j+2k}}{j!k!} p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k} \left[\frac{\mu\left(p_{AA} - p_{BA}\right)}{j+1} - \frac{\mu\left(p_{AB} - p_{BB}\right)}{k+1} + \frac{\mu^2\left(p_{AA}p_{BB} - p_{BA}p_{AB}\right)}{(j+1)(k+1)} \right] \\
> \; & \frac{e^{-\mu}}{4} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{\mu^{2j+2k+1}}{j!k!} p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k} \left[\frac{F\left(T_B\right)}{j+1} - \frac{F\left(T_B\right)}{k+1} \right] \\
= \; & \frac{e^{-\mu}}{4} F\left(T_B\right) \left[\sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} \frac{\mu^{2j+2k+1} p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k}}{(j+1)!(k+1)!} (k-j) + \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} \frac{\mu^{2j+2k+1} p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k}}{(j+1)!(k+1)!} (k-j) \right] \\
= \; & \frac{e^{-\mu}}{4} F\left(T_B\right) \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} \frac{\mu^{2j+2k+1}}{(j+1)!(k+1)!} \left(p_{AA}^{j} p_{BA}^{j} p_{AB}^{k} p_{BB}^{k} - p_{AA}^{k} p_{BA}^{k} p_{AB}^{j} p_{BB}^{j} \right) (k-j) \\
= \; & \frac{e^{-\mu}}{4} F\left(T_B\right) \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} \frac{\mu^{2j+2k+1} p_{AA}^{j} p_{BA}^{j} p_{AB}^{j} p_{BB}^{j}}{(j+1)!(k+1)!} \left(p_{AB}^{k-j} p_{BB}^{k-j} - p_{AA}^{k-j} p_{BA}^{k-j} \right) (k-j) \\
> \; & 0.
\end{aligned}
$$

■

Lemma A2 *If* $T_{BR}\left(\sigma_T\right) \geq T$ *then* $\frac{\partial}{\partial T} T_{BR}\left(\sigma_T\right) \geq 0$.

Proof. For notational brevity, let $\psi_k\left(\sigma_T\right)$ denote the probability of precisely k votes for Z and k votes for his opponent, given the quality threshold strategy σ_T,

$$
\psi_k \equiv \frac{e^{-\mu p_+}\left(\mu p_+\right)^k}{k!} \frac{e^{-\mu p_-}\left(\mu p_-\right)^k}{k!} = \frac{\mu^{2k}}{k!k!} e^{-\mu\left(p_+ + p_-\right)} \left(p_+ p_-\right)^k, \tag{12}
$$

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

$$
\gamma \equiv \frac{1}{p_+} \frac{\pi_1}{\pi_0} = \frac{1}{p_-} \frac{\pi_{-1}}{\pi_0}. \tag{13}
$$

Then the best response quality threshold function $T_{BR}\left(\sigma_T\right) \equiv \frac{\tilde{P}}{P + \tilde{P}}$ can be written in terms of γ, and differentiated, as follows:

$$
\begin{aligned}
T_{BR}\left(\sigma_T\right) &= \frac{\frac{1}{2}\left(\pi_0 + \pi_0 \gamma p_+\right)}{\frac{1}{2}\left(\pi_0 + \pi_0 \gamma p_-\right) + \frac{1}{2}\left(\pi_0 + \pi_0 \gamma p_+\right)} \\
&= \frac{1 + \gamma p_+}{2 + \gamma p_- + \gamma p_+}. \qquad (14)
\end{aligned}
$$

Now let primed variables denote derivatives with respect to the underlying quality threshold T (e.g. $T'_{BR} \equiv \frac{\partial}{\partial T} T_{BR}(\sigma_T)$). The derivative T'_{BR} depends on the derivatives of all the variables defined earlier, which can be expressed as follows:

$$p'_+ = -Tf \tag{15}$$
$$p'_- = -(1-T)f \tag{16}$$
$$\psi'_k = \psi_k(\mu f + kG) \tag{17}$$
$$\pi'_0 = \sum\nolimits_{k=0}^{\infty} \psi_k(\mu f + kG) \tag{18}$$
$$\pi'_1 = \sum\nolimits_{k=0}^{\infty} \frac{\mu}{k+1}\left(\psi'_k p_+ + \psi_k p'_+\right) \tag{19}$$
$$\pi'_{-1} = \sum\nolimits_{k=0}^{\infty} \frac{\mu}{k+1}\left(\psi'_k p_- + \psi_k p'_-\right) \tag{20}$$
$$\gamma' = \frac{A_1}{(p_+\pi_0)^2} \tag{21}$$
$$P' = \frac{1}{2}\left(\pi'_0 + \pi'_{-1}\right) \tag{22}$$
$$\tilde{P}' = \frac{1}{2}\left(\pi'_0 + \pi'_1\right) \tag{23}$$
$$T'_{BR} = \frac{A_2}{(2+\gamma p_- + \gamma p_+)^2} \tag{24}$$

where $G = \frac{p'_+}{p_+} + \frac{p'_-}{p_-}$ in (17) and the numerators A_1 and A_2 in (21) and (24) are given by the following:

$$\begin{aligned}
A_1 &\equiv \pi'_1(p_+\pi_0) - \pi_1\left(p'_+\pi_0 + p_+\pi'_0\right) \\
&= p_+\pi_0\pi'_1 - p'_+\pi_0\pi_1 - p_+\pi'_0\pi_1 \\
&= p_+\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\frac{\mu}{k+1}\left(\psi'_k p_+ + \psi_k p'_+\right) - p'_+\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\psi_k\frac{\mu p_+}{k+1} - p_+\sum_{j=0}^{\infty}\psi'_j\sum_{k=0}^{\infty}\psi_k\frac{p_+\mu}{k+1} \\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(\mu f + kG)\frac{p_+^2\mu}{k+1} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(\mu f + jG)\frac{p_+^2\mu}{k+1} \\
&= p_+^2\mu G\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k(k-j)\frac{1}{k+1}.
\end{aligned} \tag{25}$$

$$\begin{aligned}
A_2 &= \left(\gamma' p_+ + \gamma p'_+\right)\left(2 + \gamma p_- + \gamma p_+\right) - \left(1 + \gamma p_+\right)\left(\gamma' p_+ + \gamma p'_+ + \gamma' p_- + \gamma p'_-\right) \\
&= \left(\gamma' p_+ + \gamma p'_+\right)\left(1 + \gamma p_-\right) - \left(1 + \gamma p_+\right)\left(\gamma' p_- + \gamma p'_-\right) \\
&= \gamma p'_+ \left(1 + \gamma p_-\right) - \left(1 + \gamma p_+\right)\gamma p'_- + \gamma' \left(p_+ - p_-\right) \\
&= -\gamma T f \frac{2}{\pi_0} P + \gamma \left(1 - T\right) f \frac{2}{\pi_0} \tilde{P} + \gamma' \left(p_+ - p_-\right) \\
&= \gamma f \frac{2}{\pi_0} \left(P + \tilde{P}\right)\left(\frac{\tilde{P}}{P + \tilde{P}} - T\right) + \gamma' \left(p_+ - p_-\right) \qquad (26)
\end{aligned}$$

From (8) and (9) note that $p'_+ < p'_- < 0$ for any $T < 1$, implying that $G < 0$. Without the fraction $\frac{1}{k+1}$, the double sum in (25) would equal zero: though $\psi_j \psi_k (k - j)$ is positive whenever $k > j$, 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 G is also negative, the sign of A_1, and therefore of γ', must therefore be positive. The second term of the sum in (26) is positive since γ' is positive; when $\frac{\tilde{P}}{P+\tilde{P}} \geq T$, the first term is positive as well, implying as desired that $T'_{BR} > 0$. ■

Lemma A3 *Let σ_{T_A,T_B} be a posterior threshold strategy profile, with thresholds $T_A, T_B \in \left[\frac{1}{2}, 1\right]$. Then the following must be true:*

1. *If $T_A > T_B$ then $\hat{\beta}_{B0}\left(\sigma_{T_A,T_B}\right) > T_{BR}\left(\sigma_{T_A}\right)$ and $\hat{\beta}_{A0}\left(\sigma_{T_A,T_B}\right) < T_{BR}\left(\sigma_{T_B}\right)$*
2. *If $T_A < T_B$ then $\hat{\beta}_{B0}\left(\sigma_{T_A,T_B}\right) < T_{BR}\left(\sigma_{T_A}\right)$ and $\hat{\beta}_{A0}\left(\sigma_{T_A,T_B}\right) > T_{BR}\left(\sigma_{T_B}\right)$*

   where $T_{BR}\left(\sigma_T\right)$ is the best response threshold to the quality threshold strategy σ_T.

Proof. Under σ_{T_A,T_B}, (1) and (2) simplify to the following:

$$\begin{aligned}
p_{AA} &= \int_{T_A}^{1} q dF\left(q\right) &(27)\\
p_{AB} &= \int_{T_A}^{1} \left(1 - q\right) dF\left(q\right) &(28)\\
p_{BA} &= \int_{T_B}^{1} \left(1 - q\right) dF\left(q\right) &(29)\\
p_{BB} &= \int_{T_B}^{1} q dF\left(q\right) &(30)
\end{aligned}$$

For the closely-related quality threshold strategy σ_{T_A}, the corresponding probabilities are given by (27) and (28) alone: $p_{AA}\left(\sigma_{T_A}\right) = p_{BB}\left(\sigma_{T_A}\right) = p_{AA}$ and $p_{AB}\left(\sigma_{T_A}\right) = p_{BA}\left(\sigma_{T_A}\right) = p_{AB}$. Similarly, the probabilities for σ_{T_B} are given by (30) and (29) alone: $p_{AA}\left(\sigma_{T_B}\right) = p_{BB}\left(\sigma_{T_B}\right) = p_{BB}$ and $p_{AB}\left(\sigma_{T_B}\right) = p_{BA}\left(\sigma_{T_B}\right) = p_{BA}$.

An equivalent condition to $\hat{\beta}_{B0}(\sigma_{T_A,T_B}) > T_{BR}(\sigma_{T_A})$ is that $\tilde{P}_A(\sigma_{T_A,T_B}) P_B(\sigma_{T_A}) > \tilde{P}_A(\sigma_{T_A}) P_B(\sigma_{T_A,T_B})$. When $T_A > T_B$, this inequality must hold, as can be seen below:

$$\begin{aligned}
& \tilde{P}_A(\sigma_{T_A,T_B}) P_B(\sigma_{T_A}) - \tilde{P}_A(\sigma_{T_A}) P_B(\sigma_{T_A,T_B}) \qquad (31)\\
= \; & \frac{1}{4}\sum_{j=0}^{\infty} \frac{\mu^{2j} e^{-\mu(p_{AA}+p_{BA})}}{j!j!} p_{AA}^j p_{BA}^j \left(1+\frac{\mu p_{AA}}{j+1}\right) \sum_{k=0}^{\infty} \frac{\mu^{2k} e^{-\mu(p_{AB}+p_{AA})}}{k!k!} p_{AB}^k p_{AA}^k \left(1+\frac{\mu p_{AB}}{k+1}\right) \\
& -\frac{1}{4}\sum_{j=0}^{\infty} \frac{\mu^{2j} e^{-\mu(p_{AA}+p_{AB})}}{j!j!} p_{AA}^j p_{AB}^j \left(1+\frac{\mu p_{AA}}{j+1}\right) \sum_{k=0}^{\infty} \frac{\mu^{2k} e^{-\mu(p_{AB}+p_{BB})}}{k!k!} p_{AB}^k p_{BB}^k \left(1+\frac{\mu p_{AB}}{k+1}\right) \\
> \; & \frac{e^{-\mu(2p_{AA}+p_{BA}+p_{AB})}}{4} \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\mu^{2j+2k} p_{AA}^j p_{BA}^j p_{AB}^k p_{AA}^k}{j!j!k!k!} \left(1+\frac{\mu p_{AA}}{k+1}\right)\left(1+\frac{\mu p_{AB}}{j+1}\right)\left[1-\left(\frac{p_{AB}}{p_{BA}}\right)^j \left(\frac{p_{BB}}{p_{AA}}\right)^k\right] \qquad (32)
\end{aligned}$$

Whenever $j \leq k$, the final product in (32) is less than one, so that the bracketed difference is positive.[20] For any negative term, therefore, it must be that $j > k$; in that case, the corresponding (k,j) term has otherwise equal magnitude, but receives greater weight (i.e. $(p_{BA}p_{AA})^j (p_{AA}p_{AB})^k$ instead of $(p_{BA}p_{AA})^k (p_{AA}p_{AB})^j$). Thus the sign of (31) is positive.

A symmetric derivation reveals that $T_A > T_B$ also implies that $\tilde{P}_B(\sigma_{T_A,T_B}) P_A(\sigma_{T_B}) < \tilde{P}_B(\sigma_{T_B}) P_A(\sigma_{T_A,T_B})$ or, equivalently, that $\hat{\beta}_{A0} < T_{BR}(\sigma_{T_B})$, establishing part 1. Part 2 follows from identical reasoning, for the case in which $T_A < T_B$. ■

Theorem 4 *σ^* is a symmetric Bayesian equilibrium only if it is a quality threshold strategy.*

Proof. If σ^* is a symmetric Bayesian equilibrium then Lemma 1 $\sigma^* = \sigma_{T_B,T_B}$ is a posterior threshold strategy for appropriate choice of T_A and T_B. Lemma A1 shows that $1 - T_A < T_B$, implying that abstention is positive in equilibrium. The proof of Lemma 1 therefore characterizes the best response to σ^* by the posterior thresholds $T_A(\sigma^*) = \hat{\beta}_{A0}(\sigma^*)$ and $T_B(\sigma^*) = \hat{\beta}_{B0}(\sigma^*)$.

I next claim that $T_A, T_B > \frac{1}{2}$, implying that voting is *sincere* or *informative* in the sense that citizens vote A and B only in response to A and B signals, respectively. This is because $T_A < \frac{1}{2}$ (i.e. $\frac{1}{2} < 1 - T_A < T_B$) would imply that individuals who receive B signals vote B only if they are sufficiently well-informed (i.e. $q \geq T_B$), while those who are poorly informed (i.e. $q \leq 1 - T_B$), together with everyone who receive A signals, instead vote A. In that

[20] This can be easily seen by rewriting the final term of the difference as

$$\left(\frac{p_{AB}}{p_{BA}}\right)^k \left(\frac{p_{BB}}{p_{AA}}\right)^j = \left(\frac{p_{AB}}{p_{BA}}\right)^{k-j} \left(\frac{p_{AB}p_{BB}}{p_{BA}p_{AA}}\right)^j$$

case, (1) and (2) simplify to

$$\begin{aligned} p_{AA} &= F(1-T_A) + \int_{T_A}^{1} q dF(q) & p_{BA} &= \int_{T_B}^{1} (1-q)\, dF(q) \\ p_{AB} &= F(1-T_A) + \int_{T_A}^{1} (1-q)\, dF(q) & p_{BB} &= \int_{T_B}^{1} q dF(q), \end{aligned} \tag{33}$$

implying that $\pi_{0B} > \pi_{0A}$ and $\pi_{1B} > \pi_{-1A}$, and therefore that $\tilde{P}_B > P_A$ and $T_A(\sigma_{T_A,T_B}) > \frac{1}{2}$. By symmetric reasoning, $T_B < \frac{1}{2}$ would imply that $T_B(\sigma_{T_A,T_B}) > \frac{1}{2}$.

The final step of this proof is to show that $T_A = T_B$; in other words, σ^* can be characterized by a single quality threshold. The logic of this step is to suppose by way of contradiction that $T_A > T_B$, and then (using Lemmas A3 and A2) compare the best responses to σ_{T_A,T_B} and to the closely-related quality threshold strategies σ_{T_A} and σ_{T_B} (symmetric reasoning applies, of course, if instead $T_A < T_B$). There are three relevant cases to consider:

Case 1: $T_{BR}(\sigma_{T_B}) \leq T_B$. In this case, by Lemma A3, $\hat{\beta}_{A0} < T_{BR}(\sigma_{T_B}) \leq T_B < T_A$, so σ_{T_A,T_B} is not an equilibrium.

If $T_{BR}(\sigma_{T_B}) > T_B$ then, since $T_{BR}(\sigma_1) = \frac{1}{2}$, there exists (by the Intermediate Value Theorem) an equilibrium with a participation threshold strictly greater than T_B. Let T^* denote this threshold–or, if there are more than one such equilibria, let T^* denote the lowest equilibrium threshold (i.e. the threshold closest to T_B). This threshold distinguishes the remaining two cases.

Case 2: $T_{BR}(\sigma_{T_B}) > T_B$ and $T_A \geq T^*$. By Lemma A2, $T_{BR}(\sigma_T)$ is increasing between T_B and T^*, which implies that $T_{BR}(\sigma_{T_B}) < T_{BR}(\sigma_T) = T^*$. By Lemma A3, this implies that $\hat{\beta}_{A0} < T_{BR}(\sigma_{T_B}) < T^* \leq T_A$, so σ_{T_A,T_B} is not an equilibrium.

Case 3: $T_{BR}(\sigma_{T_B}) > T_B$ and $T_A < T^*$. Since $T_A \in [T_B, T^*]$, an interval in which (by Lemma A2) T_{BR} is increasing, $T_{BR}(\sigma_{T_A}) > T_{BR}(\sigma_{T_B})$. Lemma A3 then implies that $\hat{\beta}_{B0} > T_{BR}(\sigma_{T_A}) > T_{BR}(\sigma_{T_B}) > T_B$, again implying that σ_{T_A,T_B} is not an equilibrium. ∎

Proposition 2 *For any $T < 1$, the best response threshold $T_{BR}(\sigma_T; \mu)$ is increasing in μ.*

Proof. Vote probabilities p_+ and p_- do not depend on μ. For a fixed threshold T, the probability ψ_k from (12) that each candidate receives k votes depends only on μ. The same is true, therefore, of win and pivot probabilities π_w, P, and $\tilde{P}$. Differentiate ψ_k and π_w with respect to μ, as follows:

$$\begin{aligned}
\frac{\partial \psi_k}{\partial \mu} &= \frac{(p_+ p_-)^k}{k!k!}\left[2k\mu^{2k-1}e^{-\mu(p_+ + p_-)} - (p_+ + p_-)\,\mu^{2k}e^{-\mu(p_+ + p_-)}\right] \\
&= \psi_k\left[\frac{2k}{\mu} - (p_+ + p_-)\right] \qquad (34)\\
\frac{\partial \pi_0}{\partial \mu} &= \sum_{k=0}^{\infty}\psi_k\left[\frac{2k}{\mu} - (p_+ + p_-)\right] \qquad (35)\\
\frac{\partial \pi_1}{\partial \mu} &= \sum_{k=0}^{\infty}\left(\frac{\partial \psi_k}{\partial \mu}\frac{\mu p_+}{k+1} + \psi_k\frac{p_+}{k+1}\right)\\
&= \sum_{k=0}^{\infty}\psi_k\left\{\left[\frac{2k}{\mu} - (p_+ + p_-)\right]\frac{\mu p_+}{k+1} + \frac{p_+}{k+1}\right\} \qquad (36)\\
\frac{\partial \pi_{-1}}{\partial \mu} &= \sum_{k=0}^{\infty}\left(\frac{\partial \psi_k}{\partial \mu}\frac{\mu p_-}{k+1} + \psi_k\frac{p_-}{k+1}\right)\\
&= \sum_{k=0}^{\infty}\psi_k\left\{\left[\frac{2k}{\mu} - (p_+ + p_-)\right]\frac{\mu p_-}{k+1} + \frac{p_-}{k+1}\right\} \qquad (37)
\end{aligned}$$

From these, differentiate the ratio $\frac{\tilde{P}}{P}$ of pivot probabilities by the quotient rule:

$$\frac{\partial}{\partial \mu}\frac{\tilde{P}}{P} = \frac{1}{P^2}\left(P\frac{\partial \tilde{P}}{\partial \mu} - \tilde{P}\frac{\partial P}{\partial \mu}\right) \qquad (38)$$

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

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

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

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

The parenthesis term $P\frac{\partial \tilde{P}}{\partial \mu} - \tilde{P}\frac{\partial P}{\partial \mu}$ from (38) then simplifies to

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

which is positive since $\pi_1 > \pi_{-1}$ and $\pi_0 > 0$. Thus $\frac{\partial}{\partial \mu}\frac{\tilde{P}}{P} > 0$, and therefore $\frac{\partial}{\partial \mu}\left(\frac{\tilde{P}}{P+\tilde{P}}\right) > 0$, which is equivalent to the desired result. ■

Theorem 5 *Let $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ be a sequence of equilibrium participation thresholds for a sequence μ_k of population parameters such that $\mu_k \to \infty$ as $k \to \infty$, and let T^*_∞ be a limit point of $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$. Then $T^*_\infty < 1$.*

Proof. For any $T < 1$, Myerson (2000) derives the limiting ratio of pivot probabilities $\lim_{\mu\to\infty}\frac{\tilde{P}}{P} = \frac{\sqrt{p_+}}{\sqrt{p_-}}$ as merely the ratio of square roots of the expected vote shares of the two candidates. This implies that the best response threshold $T_{BR}(\sigma_T;\mu)$ approaches $L(T)$, defined as follows:

$$L(T) = \frac{\sqrt{E(Q|Q\geq T)}}{\sqrt{E(Q|Q\geq T)} + \sqrt{1 - E(Q|Q\geq T)}}, \tag{39}$$

where $E(Q|Q\geq T)$ is the average information quality of agents above the voting threshold T. Since $T^*_{\mu_k}$ is a fixed point of $T_{BR}(\sigma_T;\mu_k)$ for every μ_k, a limit point of $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ must be a fixed point of $L(T)$. Equivalently, it must be that $E(Q|Q\geq T) = \frac{T^2}{T^2+(1-T)^2} \equiv \Gamma(T)$.

As Figure 4 illustrates, both $E(Q|Q\geq T)$ and $\Gamma(T)$ increase to $E(Q|Q\geq 1) = \Gamma(1) = 1$. As $T \to 1$, however, the slope of $\Gamma(T)$ approaches zero while the slope of $E(Q|Q\geq T)$ approaches $\frac{1}{2}$ (or greater, as demonstrated below). Thus, for T sufficiently close to one, $E(Q|Q\geq T) < \Gamma(T)$ or, equivalently, $L(T) < T$. If it were the case that $T^*_{\mu_k} \to 1$, there would be a k sufficiently high that $T^*_{\mu_k} > L\left(T^*_{\mu_k}\right)$ or, equivalently, that $T_{BR}\left(\sigma_{T^*_{\mu_k}};\mu_k\right) > L\left(T^*_{\mu_k}\right)$, which cannot be since (by Lemma 2) L is an upper bound on $T_{BR}(\sigma_{T^*};\mu)$.

To see that $\lim_{T\to 1}\frac{\partial}{\partial T}E(Q|Q\geq T) \geq \frac{1}{2}$, differentiate $E(Q|Q\geq T)$ using the quotient rule:

$$\frac{\partial}{\partial T}E(Q|Q\geq T) = \frac{-Tf(1-F) + p_+ f}{(1-F)^2} = h\left[E(Q|Q\geq T) - T\right], \tag{40}$$

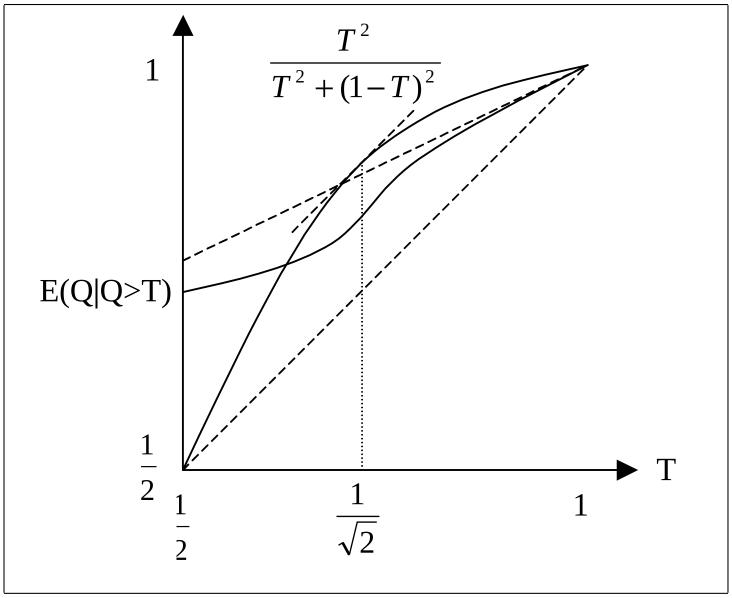

Figure 4: $M(T)$ is convex and then concave, with slope approaching $\frac{1}{2}$ as $T \to 1$, and so must intersect $\Gamma(T)$ exactly once. This implies the existence of a unique fixed point $T = L(T)$ of the limiting best response function, and therefore a unique equilibrium in a large electorate.

where $h(T) = \frac{f(T)}{1-F(T)}$ is the *hazard function* of F, 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 $\lim_{T\to 1} \frac{h'}{h^2} = \lim_{T\to 1} \frac{1}{h}\frac{f'}{f} + 1 = 1$, so L'Hospital's rule applied to equation (40) yields

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

If $\lim_{T\to 1} f(T) < \infty$ then L'Hospital's rule instead implies:

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

In either case, $\lim_{T\to 1} \frac{\partial}{\partial T} E(Q|Q \geq T) \geq \frac{1}{2}$ (with equality if $\lim_{T\to 1} f(T) > 0$). ■

Theorem 6 *Let $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ be a sequence of equilibrium participation thresholds for a sequence $\{\mu_k\}_{k=1}^{\infty}$ of population parameters such that $\mu_k \to \infty$. If f is log-concave then $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ approaches a unique limit $T^*_\infty < 1$.*

Proof. The proof of Theorem 5 shows that any limit point of $\left\{T^*_{\mu_k}\right\}_{k=1}^{\infty}$ solves $E\left(Q|Q \geq T\right) = \Gamma\left(T\right) \equiv \frac{T^2}{T^2+(1-T)^2}$. Such a point surely exists, since $E\left(Q|Q \geq \frac{1}{2}\right) > \frac{1}{2} = \Gamma\left(\frac{1}{2}\right)$ and $E\left(Q|Q \geq T\right) < \Gamma\left(T\right)$ for T sufficiently close to one (by the proof of Theorem 5, since $\frac{f'}{f}$ decreasing implies $\lim_{T\to 1}\frac{f'}{f} < \infty$).

Differentiating (40) yields the second derivative of $E\left(Q|Q \geq T\right)$,

$$\begin{aligned}\frac{\partial^2 E\left(Q|Q \geq T\right)}{\partial T^2} &= h'\left[E\left(Q|Q \geq T\right) - T\right] + h\left[\frac{\partial E\left(Q|Q \geq T\right)}{\partial T} - 1\right] \\ &= \frac{f'}{f}\frac{\partial E\left(Q|Q \geq T\right)}{\partial T} + 2h\left[\frac{\partial E\left(Q|Q \geq T\right)}{\partial T} - \frac{1}{2}\right], \end{aligned} \tag{41}$$

which is positive if and only if

$$\frac{\partial}{\partial T}E\left(Q|Q \geq T\right) \geq \frac{1}{2 + \frac{1}{h}\frac{f'}{f}}. \tag{42}$$

A log-concave density f first increases and then decreases with T, below and above some maximizer $T^{**} \in \left[\frac{1}{2}, 1\right]$. Thus the right-hand side of (42) is greater than $\frac{1}{2}$ when $T < T^{**}$ and less than $\frac{1}{2}$ when $T > T^{**}$. Bagnoli and Bergstrom (2005) demonstrate that when f is log-concave h is increasing in T, which implies that the right hand side of (42) is also increasing in T on the interval $\left[\frac{1}{2}, T^{**}\right]$. They also show that the *mean residual lifetime* function $E\left(Q - T|Q \geq T\right)$ is decreasing in T, implying that $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right) \leq 1$.

As $T \to 1$, the proof of Theorem 5 shows that $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right) \to \frac{1}{2}$. This implies that $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right) \geq \frac{1}{2}$ on the entire interval $[T^{**}, 1]$, since if $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right) = \frac{1}{2} - \varepsilon$ for some $T > T^{**}$ and $\varepsilon > 0$ then (42) would fail and $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right)$ would remain bounded above by $\frac{1}{2} - \varepsilon$. By similar reasoning, $E\left(Q|Q \geq T\right)$ must be convex (i.e. (42) must hold) on the entire interval $\left[\frac{1}{2}, T^{**}\right]$; otherwise, once the inequality in (42) failed to hold, the left-hand and right-hand sides of (42) would decrease and increase with T, respectively, so that $\frac{\partial}{\partial T}E\left(Q|Q \geq T\right)$ would continue to decline and would remain bounded below $\frac{1}{2}$.

The importance of these results is illustrated in Figure 4: in the interval $[T^{**}, 1]$, $E\left(Q|Q \geq T\right)$ lies below the dotted line of slope $\frac{1}{2}$, and therefore cannot intersect $\Gamma\left(T\right)$ to the right of $\frac{1}{\sqrt{2}}$ (i.e. the point at which $\Gamma\left(T\right)$ intersects the dotted line with slope $\frac{1}{2}$). The two functions can have at most one intersection point below $\frac{1}{\sqrt{2}}$, since the slopes of $\Gamma\left(T\right)$ and $E\left(Q|Q \geq T\right)$ are bounded above and below one, respectively. In the interval $\left[\frac{1}{2}, T^{**}\right]$, $E\left(Q|Q \geq T\right)$ is convex, and so can intersect the concave function $\Gamma\left(T\right)$ only once. Thus, if $E\left(Q|Q \geq T\right)$ and $\Gamma\left(T\right)$ intersect above T^{**} then there is a unique intersection point, and if they do not intersect above T^{**} then there is a unique intersection point. ■

Proposition 1 *Let F and G be continuous, log-concave distributions with strictly positive densities, and suppose $G \geq_1 F$. Then the following must be true:*

1. If $G(q) = F(q)$ for all $q \geq T_\infty^F$ then $T_\infty^G = T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) = \bar{F}\left(T_\infty^F\right)$.

2. If $G(q) = F(q)$ for all $q \leq T_\infty^F$ then $T_\infty^G \geq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \leq \bar{F}\left(T_\infty^F\right)$.

3. If $G(q) = F(q)$ for all $q \geq E_F\left(Q|Q \geq T_\infty^F\right)$ and if $\bar{G}\left(T_\infty^F\right) \geq \bar{F}\left(T_\infty^F\right)$ then $T_\infty^G \leq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$.

Proof. Denote by L_F and L_G the limiting best response functions from (39) for the distributions F and G. The proof of Theorem 5 makes clear that $L_F(T)$ is greater than, equal to, and less than T, respectively, when T is less than, equal to, or greater than T_∞^F, and that corresponding relationships hold for $L_G(T)$ and T_∞^G. Therefore, $T_\infty^F \leq T_\infty^G$ if and only if $L^G(T) \geq L^F(T)$ for every T between T_∞^F and T_∞^G. Equivalently, $E_G(Q|Q \geq T) \geq E_F(Q|Q \geq T)$, or (integrating by parts),

$$\frac{\bar{F}(T)}{\bar{G}(T)} \geq \frac{\int_T^1 \bar{F}(q)\,dq}{\int_T^1 \bar{G}(q)\,dq}. \tag{43}$$

1. If $G(q) = F(q)$ for all $q \geq T_\infty^F$ then at $T = T_\infty^F$ the left and right hand sides of (43) are both equal to one, implying that $T_\infty^G = T_\infty^F$ and therefore $\bar{G}\left(T_\infty^G\right) = \bar{F}\left(T_\infty^F\right)$.

2. $G\left(T_\infty^F\right) = F\left(T_\infty^F\right)$ implies that at $T = T_\infty^F$ the left hand side of (43) is equal to one. The right hand side, however, is less than or equal to one for any T. Therefore, $T_\infty^G \geq T_\infty^F$, implying that $G\left(T_\infty^G\right) \geq G\left(T_\infty^F\right) = F\left(T_\infty^F\right)$ or, equivalently, $\bar{G}\left(T_\infty^G\right) \leq \bar{F}\left(T_\infty^F\right)$.

3. Since $G(q) = F(q)$ for all $q \geq E_F\left(Q|Q \geq T_\infty^F\right)$,

$$\begin{aligned}
\int_{T_\infty^F}^1 q\,dG(q) &= \int_{T_\infty^F}^1 q\,dF(q) + \int_{T_\infty^F}^{E_F\left(Q|Q\geq T_\infty^F\right)} q\left[dG(q) - dF(q)\right] \\
&< \int_{T_\infty^F}^1 q\,dF(q) + \int_{T_\infty^F}^{E_F\left(Q|Q\geq T_\infty^F\right)} E_F\left(Q|Q \geq T_\infty^F\right)\left[dG(q) - dF(q)\right] \\
&= E_F\left(Q|Q \geq T_\infty^F\right)\left\{\int_{T_\infty^F}^1 dF(q) + \int_{T_\infty^F}^{E_F\left(Q|Q\geq T_\infty^F\right)} \left[dG(q) - dF(q)\right]\right\},
\end{aligned}$$

and similarly

$$\int_{T_\infty^F}^1 dG(q) = \int_{T_\infty^F}^1 dF(q) + \int_{T_\infty^F}^{E_F\left(Q|Q\geq T_\infty^F\right)} \left[dG(q) - dF(q)\right].$$

Therefore, $E_G\left(Q|Q \geq T_\infty^F\right) = \frac{\int_{T_\infty^F}^1 q\,dG(q)}{\int_{T_\infty^F}^1 dG(q)} < E_F\left(Q|Q \geq T_\infty^F\right)$, implying that $T_\infty^G < T_\infty^F$. Since $G \geq_1 F$, $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$. ■

Proposition 2 *Let F and G be continuous, log-concave distributions with strictly positive densities and a common mean, such that G single-crosses F at $\hat{q}$. If $T_\infty^F \leq \hat{q}$ then $T_\infty^G \leq T_\infty^F$ and $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$.*

Proof. The common mean m of F and G can be written as $m = \frac{1}{2} + \int_{1/2}^{1} \bar{F}(q)\,dq = \frac{1}{2} + \int_{1/2}^{1} \bar{G}(q)\,dq$,[21] implying that $\int_{1/2}^{T} \bar{G}(q)\,dq + \int_{T}^{1} \bar{G}(q)\,dq = \int_{1/2}^{T} \bar{F}(q)\,dq + \int_{T}^{1} \bar{F}(q)\,dq = m - \frac{1}{2}$. Since G single-crosses F at $\hat{q}$, $T_{\infty}^{F} \leq \hat{q}$ implies that $\int_{1/2}^{T_{\infty}^{F}} \bar{G}(q)\,dq \geq \int_{1/2}^{T_{\infty}^{F}} \bar{F}(q)\,dq$, which therefore implies that $\int_{T_{\infty}^{F}}^{1} \bar{G}(q)\,dq \leq \int_{T_{\infty}^{F}}^{1} \bar{F}(q)\,dq$. The right hand side of (43) is therefore greater than one at T_{∞}^{F}, while the left hand side of (43) is less than one. Thus (43) does not hold, imploying that $T_{\infty}^{G} \leq T_{\infty}^{F}$ and therefore $\bar{G}\left(T_{\infty}^{G}\right) \geq \bar{G}\left(T_{\infty}^{F}\right) \geq \bar{F}\left(T_{\infty}^{F}\right)$. ■

References

[1] Ashenfelter, Orley, and Stanley Kelley, Jr. 1975. "Determinants of Participation in Presidential Elections." *Journal of Law and Economics*, 18: 695-733.

[2] Austen-Smith, David and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *The American Political Science Review*, 90(1): 34-45.

[3] Bade, Sophie. 2006. "A Simple Model of the Swing Voter's Curse." Unpublished, Penn State University. 2006.

[4] Bagnoli, Mark, and Ted Bergstrom. 2005. "Log-Concave Probability and Its Applications." *Economic Theory*, 26(2): 445-469.

[5] Condorcet, Marquis 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.

[6] Coupé, Tom, and Abdul Noury. 2004. "Choosing Not to Choose: on the Link Between Information and Abstention." *Economics Letters*, 84: 261-265.

[7] Dee, Thomas. 2004. "Are there Civic Returns to Education?" *Journal of Public Economics*, 88(9-10): 1697-1720.

[8] Degan, Arianna and Antonio Merlo. 2007. "A Structural Model of Turnout and Voting in Multiple Elections." Penn Institute for Economic Research Working Paper 07-011.

[9] Dowding, Keith. 2005. "Is it Rational to Vote? Five Types of Answer and a Suggestion." *Political Studies Association*, 7: 442-459.

[21] This can be seen by intergrating these expressions by parts.

[10] Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

[11] Duggan, John and Cesar Martinelli. 2001. "A Bayesian Model of Voting in Juries." *Games and Economic Behavior*, 37: 259-294.

[12] Edlin, Aaron, Andrew Gelman, and Noah Kaplan. 2005. "Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others." *Rationality and Society*, 19(2): 225-246.

[13] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *The American Economic Review*, 86(3): 408-424.

[14] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1998. "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *The American Political Science Review*, 92(1): 23-35.

[15] Feddersen, Timothy J. and Alvaro Sandroni. 2006. "A Theory of Participation in Elections." *American Economic Review*, 96(4): 1271-1282.

[16] Harsanyi, John C. 1980. "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior." *Theory and Decision*, 12(2): 115-33.

[17] Krishna, Vijay, and John Morgan. 2008. "On the Benefits of Costly Voting." http://ssrn.com/abstract=1143564.

[18] Larcinese, Valentino. 2006. "Information Acquisition, Ideology, and Turnout: Theory and Evidence from Britain." STICERD Political Economy and Public Policy series 18.

[19] Lassen, David. 2005. "The Effect of Information on Voter Turnout: Evidence from a Natural Experiment." *American Journal of Political Science*, 49(1): 103-118.

[20] Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *The American Political Science Review*, 91(1): 1-14.

[21] Margolis, Howard. 1982. *Selfishness, Altruism, and Rationality.* Chicago: University of Chicago Press.

[22] Martinelli, Cesar. 2006. "Would Rational Voters Acquire Costly Information?" *Journal of Economic Theory*, 129: 225-251.

[23] Martinelli, Cesar. 2007. "Rational Ignorance and Voting Behavior." *International Journal of Game Theory*, 35:315-355.

[24] Matsusaka, John. 1995. "Explaining Voter Turnout Patterns: An Information Theory." *Public Choice*, 84: 91-117.

[25] McLennan, Andrew. 1998. "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review*, 92(2): 413-418.

[26] McMurray, Joseph C. 2009. Empirical Evidence of Strategic Voter Abstention, unpublished.

[27] Meirowitz, Adam. 2002. "Informative Voting and Condorcet Jury Theorems with a Continuum of Types." *Social Choice and Welfare*, 19: 219-236.

[28] Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics*, 88(9-10): 1667-1695.

[29] Myerson, Roger. 1998. "Population Uncertainty and Poisson Games." *International Journal of Game Theory*, 27: 375-392.

[30] Myerson, Roger. 2000. "Large Poisson Games." *Journal of Economic Theory*, 94: 7-45.

[31] Palfrey, Thomas R. and Keith T. Poole. 1987. "The Relationship between Information, Ideology, and Voting Behavior." *American Journal of Political Science*, 31(3): 511-530.

[32] Palfrey, Thomas R. and Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice*, 41(1): 7-53.

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

[34] Riker, William H. and Peter C. Ordeshook. 1968. "A Theory of The Calculus of Voting." *American Political Science Review*, 62: 25-42.

[35] Shapley, Lloyd and Bernard Grofman. 1984. "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies." *Public Choice*, 43: 329-343.

[36] Strate, John M., Charles J. Parrish, Charles D. Elder, and Coit Ford. 1989. "Life Span Civic Development and Voting Participation." *American Political Science Review*, 82(2): 443-464.

[37] Suen, Wing. 2008. "Second Order Stochastic Dominance." http://www.econ.hku.hk/~wsuen/uncertainty/ssd.pdf (accessed July 7, 2008).

[38] Wattenberg, Martin, Ian McAllister, and Anthony Salvanto. 2000. "How Voting is Like Taking an SAT Test: An Analysis of American Voter Rolloff." *American Politics Research*, 28: 234-250.

[39] Wolfinger, Raymond E. and Rosenstone, Stephan J. 1980. *Who Votes?* New Haven: Yale University Press.