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

Joseph McMurray

Working Paper No. 59
July 2008

W. ALLEN WALLIS

Institute of

POLITICAL ECONOMY

UNIVERSITY OF

ROCHESTER

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

Joseph McMurray†

First version: April 2006
This version: July 2008

Abstract

In a common-values election with continuously distributed information quality, the incentive to pool private information conflicts with the swing voter's curse. In equilibrium, therefore, some citizens abstain despite clear private opinions, and others vote despite having arbitrarily many peers with superior information. The dichotomy between one's own and others' information quality can explain the otherwise puzzling empirical relationship between education and turnout, and suggests the importance of relative information variables in explaining turnout, which I verify for U.S. primary elections. Though voluntary elections fail to utilize nonvoters' information, mandatory elections actually do worse; efforts to motivate turnout may actually reduce welfare.

JEL Classification Numbers: D72, D82

Keywords: Voting, Elections, Turnout, Information, Roll-off, Swing Voter's Curse, Jury theorem, Turnout Paradox, Compulsory Voting

*Special thanks to my advisor, Mark Fey, for his generous time and helpful advice. Thanks also to Michael Peress, Gabor Virag, John Duggan, and participants in the Wallis Institute Political Economy workgroup and the NBER Political Economics Group student conference, for their interest and suggestions.

†University of Rochester Economics Department. Email joseph.mcmurray@rochester.edu.

1 Introduction

In every democracy, a large fraction of eligible citizens abstain from voting in public elections. Many of those who do vote participate in some but not all races on the ballot–a phenomenon known as *roll-off.* Low and declining turnout are commonly viewed as threats to democracy. One important reason that is frequently given for not voting is that citizens lack confidence in their understanding of political issues, or their knowledge of political candidates. Consistent with this, Matsusaka (1995) finds that, empirically, participation in U.S. elections is highest among demographic groups that are likely to be well-informed, such as those who are well-educated, older, married, publicly employed, or recently contacted by campaign workers; and low among people who have recently moved. Wolfinger and Rosenstone (1980) find education, in particular, to be the single best predictor of voter participation, and subsequent research (Dee, 2004; Milligan, Moretti, and Oreopoulos, 2004) finds this relationship to be causal. Bartels (1996), Degan and Merlo (2007a), and Larcinese (2006) link voting more directly with information quality per se, and Lassen (2005) concludes that, controlling appropriately for information, education does not otherwise influence turnout. Similarly, Strate et al. (1989) argue that age influences turnout predominantly because life experience develops "civic competence". Wattenberg, McAllister, and Salvanto (2000, p. 245) also find information to be "the most significant factor in explaining the roll-off phenomenon".

To explain the connection between information and voting, Matsusaka (1995) points out that a citizen who is uncertain which of two candidates she[1] prefers expects a lower benefit from voting. As Feddersen and Pesendorfer (1996; hereafter FP) point out, however, this logic leads to abstention only when voting is costly (since the expected benefit of voting is always positive, even for a poorly informed voter), which is not always the case. As an alternative, they identify a strategic incentive for abstention which they call the "swing voter's curse." In their model, voters share common underlying values, so that disagreements arise only because of informational differences; if information were perfect, voting would be unanimous. Informed citizens observe which candidate is superior, and vote for that candidate. An uninformed citizen, uncertain which candidate is better, instead abstains–even when voting is costless, and even if prior beliefs favor one candidate over the other–reasoning that her own vote will influence her payoff only if it is *pivotal* (i.e. either making or breaking a tie), which can only happen when she is voting against the informed voters and therefore for the *wrong* candidate. Thus abstention is actually the *best* way for an uninformed citizen to achieve her desired election outcome. In fact, abstention is socially

[1] Throughout this paper I use feminine pronouns to describe voters, and masculine pronouns to describe candidates.

desirable: the best social outcomes are achieved when informed citizens vote and uninformed citizens abstain.

While attractively simple, the information structure in the FP model is unrealistic: voter heterogeneity is limited to two types, informed voters have perfect information, and uninformed voters are perfectly ignorant. The importance of the latter two restrictions can be seen by comparing the FP model to a related framework, analyzed two centuries earlier by Condorcet (1785). In that model, citizens each observe independent private signals indicating the superior candidate, but these signals are only correct with some (common) probability $q \in \left[\frac{1}{2}, 1\right]$. Even if citizens are only barely informed (e.g. $q = 0.51$), the well-known Condorcet jury theorem states that, as the number of voters grows large, the superior candidate receives a majority of votes with probability approaching one (by the law of large numbers). Thus everyone should vote, and the majority opinion of a sufficiently large number of voters will be better-informed than that of any imperfectly-informed individual. This ability to so effectively aggregate private information has long been regarded as one of the most compelling justifications for the extensive use of majority voting in collective decision-making.

Like FP and Condorcet, I analyze voting behavior in a two-candidate election in which voting is costless, and citizens have common values but imperfect information. As in the Condorcet model, every agent receives a private indication of the superior candidate. Here, however, the quality $Q_i \in \left[\frac{1}{2}, 1\right]$ of a citizen's signal is individual-specific, drawn randomly from some common (and commonly-known) distribution F. In this framework, the Condorcet model corresponds to a degenerate F, and the FP model a discrete F, with positive mass only on the "informed" and "uninformed" extreme types $Q_i = 1$ and $Q_i = \frac{1}{2}$. Of particular interest are continuous F, for which these extreme types are realized with zero probability.[2]

In this framework, the best response to any strategy is characterized by cutpoints on posterior beliefs. Accordingly, equilibrium can be characterized by an information quality threshold T^*, above which agents vote sincerely and below which they abstain. The location of T^* highlights the inherent conflict between the quantity and quality of information: if T^* is high, voters are few but well-informed, as in the FP model; if T^* is low, voters are many but on average less well-informed, as in the Condorcet model. With informative voting, the superior candidate is more likely to be ahead by a single vote than behind by a single vote, so a vote for the correct candidate is strictly less likely to be pivotal than a mistake.

[2] The importance of a generalized information structure is demonstrated by Duggan and Martinelli (2001) and Meirowitz (2002), who overturn a surprising result in FP (1998) that unanimity rule gives jurors poor incentives for truth-telling.

Thus, voters suffer from the swing voter's curse, and in equilibrium some agents abstain (i.e. $T^* > \frac{1}{2}$) even when they view one candidate as better than the other.

As a population grows, T^* rises, because the addition of voters strengthens the swing voter's curse. If F is continuous, then even an agent who might reasonably believe herself to be the best informed member of a small electorate expects to eventually have an arbitrarily large number of better-informed peers, as the population grows. In light of this, intuition suggests that even an extremely well-informed citizen will at some point decide to abstain, deferring to her better-informed peers; if so, the fraction of citizens who votes must fall to 0% (i.e. $T^* \to 1$) as a population grows large. This intuition, however, turns out to be incorrect: a one-vote loss indeed grows less likely, even relative to a one-vote win, but decreasingly so, and T^* asymptotes at some level strictly below one. The precise limit of T^* is determined by the underlying distribution of F, and (for most F) is unique; one benefit of this model is a straightforward formula by which, for any F, turnout can be computed numerically.

As in both the Condorcet and FP models, this model predicts that the superior candidate will be elected with probability approaching one as the electorate grows large. To maximize this probability, however, the Condorcet model suggests that everyone should vote, while the FP model suggests that uninformed citizens should abstain. In this model, a voluntary election is inefficient because, by allowing abstention, it fails to utilize nonvoters' independent private information. Surprisingly, however, mandatory actually only pools *less* information, not more, because it loses information conveyed by the *decision* to vote. A first-best election mechanism would weight votes by voters' underlying confidence levels.

In addition to behavior and welfare predictions, this generalized information structure provides useful comparative static results that are unavailable in simpler models. If everyone's information improves, for example, improves welfare but has two opposing effects on turnout: it lowers abstention by lifting nonvoters above the participation threshold, but also raises abstention by intensifying the swing voter's curse; the net effect depends on which where within the information distribution improvements are concentrated. A mean-preserving decrease in the variance of information lowers abstention in most cases, by shrinking the difference between informed and uninformed voters and thereby weakening the swing voter's curse. An increase in the cost of voting lowers turnout from its social optimum.

Like the original FP model, this model explains the empirical correlations enumerated above between education and turnout, as well as the phenomenon of roll-off, and abstention in other costless voting environments. It further explains what Aldrich (1993) calls "the most important substantive problem in the turnout literature," which is a paradox observed by Brody (1978), that turnout has persistently fallen over recent decades even while education levels have risen, so that education and turnout are positively associated in any cross-sectional study but negatively correlated over time. In the context of this model, one explanation for this paradox is that an individual's tendency to vote is increasing in her own education level but *decreasing* in the education levels of others. A second possibility is that, as the number and scope of laws and policies has proliferated over time, the decisions required of politicians (and therefore voters) have grown more complicated, in effect shifting the entire distribution of information quality downward, thereby swamping the influence within the distribution caused by educational increases. Although information quality cannot be readily observed, voting is sincere and turnout predictions for simple F are similar to actual participation levels in the U.S. The opposite impact of one's own and others' information quality implies that the importance of education, age, and other proxies of information quality should be *relative* rather than *absolute*, a subtle prediction that I confirm empirically, using American National Election Studies (ANES) data on voter participation in political party state primary elections.

The remainder of this paper is organized as follows. I present my model formally in section 2.1, and in sections 2.2 through 2.6 I analyze equilibrium behavior in small and large electorates, welfare, and comparative statics associated both with information and with voting costs. I then present empirical evidence and applications in section 3, and conclude in section 4. In most cases, proofs of analytical results are relogated to the Appendix A.

2 Analysis

2.1 The Model

In an election between two candidates or alternatives, A and B, there is an unknown number N of potential voters, where N has Poisson distribution with mean μ. For a particular realization of N, each citizen is endowed with a privately-known information quality level $Q_i \in \left[\frac{1}{2}, 1\right]$ (with realization q_i), drawn independently from a common and commonly-known distribution F, which has a differentiable density f that is strictly positive between $\frac{1}{2}$ and 1. Before the election, Nature designates one candidate $Z \in \{A, B\}$ as superior to the other. Citizens do not observe nature's choice directly, but know that the candidates will be chosen with equal probability. In addition, each agent receives a private signal $S_i \in \{a, b\}$ (with realization s_i) that corresponds to Z with probability Q_i. That is,

$$\begin{aligned}\Pr(S_i = a|Z = A) &= \Pr(S_i = b|Z = B) = Q_i \\ \Pr(S_i = b|Z = A) &= \Pr(S_i = a|Z = B) = 1 - Q_i\end{aligned}$$

To a perfectly informed agent (i.e. $Q_i = 1$), S_i reveals Z perfectly, and to a perfectly uninformed (i.e. $Q_i = \frac{1}{2}$) agent, the signal provides no information whatsoever; for most agents, of course, S_i is somewhat informative but not perfectly so. Agents' signals are mutually independent (conditional on Z), and signal values are independent of information levels.

An agent may choose to vote for either candidate or to abstain. Her mixed strategy $\sigma_i = \left(\sigma_i^A, \sigma_i^B, \sigma_i^0\right)$ specifies the distribution of the outcome $X_i \in \{A, B, 0\}$ of her vote, where a vote for candidate 0 represents abstention. Agents act simultaneously, and the election winner $X \in \{A, B\}$ is determined by simple majority rule, flipping a coin in the event of a tie. Citizens each receive utility 1 if the superior candidate wins the election and 0 otherwise, so that expected utility–and therefore social welfare–are given merely by the probability $\Pr(X = Z)$.

I restrict attention to symmetric strategies and seek a symmetric Bayesian equilibrium.[3] A symmetric strategy profile σ must specify a mixture for every possible voter type $(q, s) \in \left[\frac{1}{2}, 1\right] \times \{a, b\}$; an individual citizen of type (q_i, s_i) reinterprets σ as a list of her opponents' strategies, and chooses $\sigma_i = \left(\sigma_i^A, \sigma_i^B, \sigma_i^0\right)$ to maximize

$$EU(\sigma_i; \sigma, q_i, s_i) = \Pr(X = Z | \sigma, \sigma_i, q_i, s_i) \tag{1}$$

The strategy σ_i^* that maximizes (1) is a *best response* to σ, and σ^* is a *symmetric Bayesian equilibrium* if $\sigma^*(q, s)$ is a best response to σ^* for every (q, s).

Before proceeding to characterize best responses, it is worth noting that the assumption here that voters share a common objective is less restrictive than it might seem. In the real world, of course, political opinions vary dramatically, both between and within political parties; here, too, posterior beliefs (given below by α_i) regarding the probability of A being superior to B vary continuously between 0 and 1. When individuals in the real world argue over which of two policies to adopt, their disagreement frequently stems from their differing predictions of what effects the competing policies will have; if the effects of two policies could somehow be indisputably predicted, the superiority of one of the two alternatives may well become obvious, so that voters no longer disagree.

2.2 Best Responses

In state A, an agent with information quality q receives an a signal with probability q and a b signal with probability $1 - q$; in state B, the probabilities are reversed. Given a profile σ, therefore, the probability p_{xz} with which a randomly chosen agent votes for candidate

[3] This restriction is merely for simplicity of exposition: even allowing asymmetric strategies, Theorem 1 guarantees that the best response to any profile is symmetric. Any Bayesian equilibrium, therefore, must be symmetric as well.

$x \in \{A, B, 0\}$ in state $z \in \{A, B\}$ is given by the following.

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

$$p_{xB} = \int_{1/2}^{1} \left[q\sigma^x(q,b) + (1-q)\,\sigma^x(q,a)\right] dF(q) \tag{3}$$

The expected number of votes for candidate x in state z is then given simply by μp_{xz}.

By the decomposition property of Poisson random variables (see Myerson, 1998), the numbers of A, B, and 0 votes in state z are independent Poisson random variables. For any integer w, the probability π_{wz} that the superior candidate wins the election by w votes is simply the infinite sum of probabilities that $k+w$ agents vote for the superior candidate and only k vote for his opponent. Since an agent's vote is *pivotal* when it creates or breaks a tie, of particular interest are the probability π_{0z} of a tie, and the probabilities π_{1z} and π_{-1z} with which the superior candidate wins and loses, respectively, by a single vote. In state A, these probabilities are given by the following, and similar expressions apply to state B.

$$\pi_{0A} = \sum_{k=0}^{\infty} \frac{e^{-(\mu p_{AA})}(\mu p_{AA})^k}{k!} \frac{e^{-(\mu p_{BA})}(\mu p_{BA})^k}{k!} \equiv \sum_{k=0}^{\infty} \psi_{kA} \tag{4}$$

$$\pi_{1A} = \sum_{k=0}^{\infty} \frac{e^{-(\mu p_{AA})}(\mu p_{AA})^{k+1}}{(k+1)!} \frac{e^{-(\mu p_{BA})}(\mu p_{BA})^k}{k!} \equiv \sum_{k=0}^{\infty} \frac{\mu}{k+1} p_{AA}\psi_{kA} \tag{5}$$

$$\pi_{-1A} = \sum_{k=0}^{\infty} \frac{e^{-(\mu p_{AA})}(\mu p_{AA})^k}{k!} \frac{e^{-(\mu p_{BA})}(\mu p_{BA})^{k+1}}{(k+1)!} \equiv \sum_{k=0}^{\infty} \frac{\mu}{k+1} p_{BA}\psi_{kA} \tag{6}$$

where the simplifying notation ψ_{kA} is defined as follows:

$$\psi_{kA} = \frac{e^{-(\mu p_{AA})}(\mu p_{AA})^k}{k!} \frac{e^{-(\mu p_{BA})}(\mu p_{BA})^k}{k!} \tag{7}$$

It is easy to see that $\pi_{1A} = \frac{p_{AA}}{p_{BA}} \pi_{-1A}$, and also that with positive probability no agents vote, resulting in a tie (i.e. $\psi_{0A} > 0$, which implies $\pi_{0A} > 0$).

By the environmental equivalence property of Poisson games (see Myerson, 1998) an individual perceives from within the game that the number of her opponents, like the total number of players from an outside perspective, has Poisson distribution with mean μ, and therefore that the numbers of her fellow citizens who vote for A, B, and 0, respectively, in state z are independent Poisson random variables with means μp_{Az}, μp_{Bz}, and μp_{0z}. She also reinterprets π_{wz} as the probability that Z will win the election by w votes, should she abstain.

Using her private information, an agent can formulate posterior beliefs α_i about the distribution of Z:

$$\alpha_i \equiv \Pr(Z = A|Q_i, S_i) = \begin{cases} Q_i \text{ if } S_i = a \\ 1 - Q_i \text{ if } S_i = b \end{cases} \tag{8}$$

Playing $\sigma_i = \left(\sigma_i^A, \sigma_i^B, \sigma_i^0\right)$ in response to an opponent profile σ will then yield the following expected utility:

$$\begin{aligned} & EU(\sigma_i|\sigma, \alpha_i) \qquad (9) \\ = \quad & \alpha_i \left[\sum_{w=2}^{\infty} \pi_{wA} + \pi_{1A}\left(\sigma_i^A + \sigma_i^0 + \frac{1}{2}\sigma_i^B\right) + \pi_{0A}\left(\sigma_i^A + \frac{1}{2}\sigma_i^0\right) + \pi_{-1A}\left(\frac{1}{2}\sigma_i^A\right)\right] \\ & + (1-\alpha_i)\left[\sum_{w=2}^{\infty} \pi_{wB} + \pi_{1B}\left(\sigma_i^B + \sigma_i^0 + \frac{1}{2}\sigma_i^A\right) + \pi_{0B}\left(\sigma_i^B + \frac{1}{2}\sigma_i^0\right) + \pi_{-1B}\left(\frac{1}{2}\sigma_i^B\right)\right] \end{aligned}$$

The fraction $\frac{1}{2}$ in equation (9) reflects the probability, in the event of a tied election, that a tie-breaking coin toss will favor the superior candidate Z. From (9) it can be seen that a vote for the superior candidate is only *pivotal* (i.e. influences the election outcome) when either the candidates tie but Z loses the coin toss, or Z loses the election by a single vote, when he would have won the coin toss. Let P_z denote the probability, in state z, that one of these occurs:

$$P_z = \frac{1}{2}\pi_{0z} + \frac{1}{2}\pi_{-1z} \tag{10}$$

Similarly, let $\tilde{P}_z$ denote the probability with which a vote for the inferior candidate is pivotal:

$$\tilde{P}_z = \frac{1}{2}\pi_{0z} + \frac{1}{2}\pi_{1z} \tag{11}$$

Using (10) and (11), define the ratios α_A, α_B, and $\hat{\alpha}$ as follows.

$$\alpha_A = \frac{\tilde{P}_B}{P_A + \tilde{P}_B}, \alpha_B = \frac{\tilde{P}_A}{\tilde{P}_A + P_B} \tag{12}$$

$$\hat{\alpha} = \frac{P_B + \tilde{P}_B}{\tilde{P}_A + P_B + P_A + \tilde{P}_B} \tag{13}$$

As Theorem 1 now shows, these thresholds characterize the best response σ_i^* to any strategy profile. Mathematically, $\hat{\alpha}$ must lie between α_A and $1-\alpha_B$: when $\alpha_A \leq \hat{\alpha} \leq 1-\alpha_B$, a best response is to vote A if $\alpha_i \in (\hat{\alpha}, 1]$ and vote B if $\alpha_i \in [0, \hat{\alpha})$; when $1-\alpha_B \leq \hat{\alpha} \leq \alpha_A$, it is to vote A if $\alpha_i \in (\alpha_A, 1]$, vote B if $\alpha_i \in [0, 1-\alpha_B)$, and abstain if $\alpha_i \in (1-\alpha_B, \alpha_A)$. Best responses are in pure strategies except precisely at these cutpoints, which occur with zero probability.

Theorem 1 *Define* α_A*,* $\hat{\alpha}$*, and* α_B *as in (12) and (13) for a symmetric profile* σ*, and let* σ_i^* *denote the best response to* σ *for a citizen with posterior beliefs given by* α_i*, as defined in (8). Then the following must be true.*

(i) If $\alpha_A \leq \hat{\alpha} \leq 1 - \alpha_B$ *then* $\sigma_i^* = \left\{ \begin{array}{c} (1,0,0) \text{ if } \alpha_i > \hat{\alpha} \\ (0,1,0) \text{ if } \alpha_i < \hat{\alpha} \end{array} \right\}$

(ii) If $1 - \alpha_B \leq \hat{\alpha} \leq \alpha_A$ *then* $\sigma_i^* = \left\{ \begin{array}{c} (1,0,0) \text{ if } \alpha_i > \alpha_A \\ (0,0,1) \text{ if } 1 - \alpha_B < \alpha_i < \alpha_A \\ (0,1,0) \text{ if } \alpha_i < 1 - \alpha_B \end{array} \right\}$

Proof. From (9) it is straightforward to verify the following.

$$EU((1,0,0)) - EU((0,0,1)) = \alpha_i P_A - (1-\alpha_i)\tilde{P}_B \quad (14)$$

$$EU((0,1,0)) - EU((0,0,1)) = -\alpha_i \tilde{P}_A + (1-\alpha_i) P_B \quad (15)$$

$$EU((1,0,0)) - EU((0,1,0)) = \alpha_i \left(P_A + \tilde{P}_A\right) - (1-\alpha_i)\left(\tilde{P}_B + P_B\right) \quad (16)$$

(14) is positive if and only if $\alpha_i P_A > (1-\alpha_i)\tilde{P}_B$ or, equivalently, $\alpha_i > \frac{\tilde{P}_B}{P_A + \tilde{P}_B} = \alpha_A$. Similarly, (15) and (16) are positive if and only if $\alpha_i < 1 - \alpha_B$ and $\alpha_i > \hat{\alpha}$, respectively. Part (i) then follows since $\alpha_i > \hat{\alpha} \geq \alpha_A$ implies that both (14) and (16) are positive and $\alpha_i < \hat{\alpha} \leq 1 - \alpha_B$ implies that (15) is positive and (16) is negative. Part (ii) follows since $\alpha_i > \alpha_A \geq \hat{\alpha}$ implies that both (15) and (16) are positive, $1 - \alpha_B < \alpha_i < \alpha_A$ implies that both (15) and (16) are negative, and $\alpha_i < 1 - \alpha_B \leq \hat{\alpha}$ implies that (15) is positive and (16) is negative. ■

One interpretation of the best response thresholds α_A, $\hat{\alpha}$, and α_B is as conditional probabilities. For example, α_A denotes the probability that an A vote is pivotal in the *wrong* direction (i.e. in state B) conditional on its being pivotal at all. If this probability is high, it discourages voters from casting A votes; thus the corresponding A threshold is high, and only those voters with the strongest posterior beliefs vote for candidate A.

2.3 Equilibrium

An equilibrium strategy profile must be its own best response. As Theorem 2 now states, this is only possible when $\alpha_A > \frac{1}{2}$ and $\alpha_B > \frac{1}{2}$. By condition (ii) of Theorem 1, this implies that citizens vote *informatively* (i.e. vote A or B in response to a or b signals, respectively) if sufficiently well-informed (i.e. if $Q_i \geq \alpha_A$ or $Q_i \geq \alpha_B$, respectively) and otherwise abstain. In this model, informative and *sincere* voting (i.e. voting as if directly choosing the election outcome) are equivalent.

Theorem 2 *Let* σ^* *be a symmetric Bayesian equilibrium and let* α_A *and* α_B *be the posterior thresholds defined in (12). Then* $\alpha_A \geq \frac{1}{2}$ *and* $\alpha_B \geq \frac{1}{2}$*.*

Proof. See Appendix A. ■

Given this characterization of equilibrium voting, it is useful to define an *informative-voting participation cutpoint (IPC) strategy* σ_{T_A,T_B} for arbitrary participation thresholds $T_A, T_B \in \left[\frac{1}{2}, 1\right]$, as follows:

$$\sigma_{T_A,T_B}(q,s) = \begin{cases} (1,0,0) \text{ if } s = a \text{ and } q \geq T_A \\ (0,1,0) \text{ if } s = b \text{ and } q \geq T_B \\ (0,0,1) \text{ otherwise} \end{cases} \tag{17}$$

Taken together, Theorems 1 and 2 imply that an equilibrium profile σ^* must be IPC, with participation thresholds $T_A = \alpha_A$ and $T_B = \alpha_B$. Theorem 4 will further show σ^* to be *signal-symmetric* (SIPC), so that the thresholds coincide (i.e. $T_A = T_B = T$).

Under an SIPC strategy, citizens vote informatively if $Q_i \geq T$, and otherwise abstain. Therefore, the vote probabilities from (2) and (3) no longer depend on the state of the world, and simplify to p_+ and p_-, as defined here:

$$p_+ \equiv \int_T^1 q dF(q) = p_{AA} = p_{BB} \tag{18}$$

$$p_- \equiv \int_T^1 (1-q)\, dF(q) = p_{AB} = p_{BA} \tag{19}$$

Similarly, win and pivot probabilities can be written simply as π_w, P, and $\tilde{P}$. Accordingly, best response cutpoints coincide (i.e. $\alpha_A = \alpha_B$) and the best response to $\sigma_{T,T}$ is itself SIPC, with a participation threshold T_{BR} that depends on the participation threshold T, as follows:

$$T_{BR}(T) \equiv \frac{\tilde{P}}{P + \tilde{P}} \tag{20}$$

In general, $T_{BR}(T)$ may be either greater or less than T; for a fixed point $T^* = T_{BR}(T^*)$, the corresponding SIPC profile σ_{T^*,T^*} must be its own best response, and therefore an *SIPC Bayesian equilibrium*. Theorem 3 now states that such a fixed point, and therefore an SIPC Bayesian equilibrium, must always exist. In such an equilibrium $\frac{1}{2} < T^* < 1$, meaning that both voting and abstention must both be positive.

Theorem 3 *There exists a threshold T^* strictly between $\frac{1}{2}$ and 1 such that the SIPC profile σ_{T^*,T^*} is an SIPC Bayesian equilibrium.*

Proof. $T_{BR} : \left[\frac{1}{2}, 1\right] \to \left[\frac{1}{2}, 1\right]$ is continuous and $\left[\frac{1}{2}, 1\right]$ is compact, so Brouwer's fixed point theorem gives existence. Following $\sigma_{\frac{1}{2},\frac{1}{2}}$, all citizens vote, in which case $p_+ > p_-$, $\pi_1 > \pi_{-1}$, and therefore $\tilde{P} > P$. In other words, a mistake is strictly more likely to be pivotal than a correct vote. This implies that an agent of type $Q_i = \frac{1}{2}$ should abstain (i.e. $T\left(\frac{1}{2}\right) > \frac{1}{2}$).

Similarly, by $\sigma_{1,1}$ no one votes; in that case, any vote would be pivotal (i.e. $P = \tilde{P} = \frac{1}{2}$), so anyone should vote in response (i.e. $T_{BR}(1) = \frac{1}{2}$). ■

Having shown that an SIPC equilibrium exists, Theorem 4 next shows that, in fact, all equilibria are SIPC. In principle there could be multiple SIPC equilibria, though the proof of Theorem 4 suggests that this is impossible for sufficiently smooth F. Conditions for uniqueness in large electorates are made more precise by Theorem 7 in section 2.4.

Theorem 4 *If σ^* is a Bayesian equilibrium then it is SIPC.*

Proof. See Appendix A. ■

2.4 Large Elections

In this section I analyze asymptotic equilibrium behavior as a population grows large. In doing so, I denote the implicit dependence of the best response and equilibrium participation thresholds T^{μ}_{BR} and T^*_{μ} on the population size parameter μ by a superscript and subscript, respectively. As the number of voters grows large, the superior candidate wins the election with increasing probability. This strengthens the swing voter's curse so that the marginal voter, who had previously been indifferent between voting and abstaining, now strictly prefers to abstain, leaving the election decision in the hands of those with superior information. Accordingly, the best response threshold rises and the participation rate declines. More formally, Theorem 5 states that $T^{\mu}_{BR}(T)$ is increasing in μ for any cutpoint $T < 1$.[4] The highest and lowest fixed points of T^{μ}_{BR} (or the only fixed point, if T^*_{μ} is unique) must therefore rise with μ.

Theorem 5 *For any $T \in \left[\frac{1}{2}, 1\right)$, the best response threshold $T^{\mu}_{BR}(T)$ is increasing in μ.*

Proof. See Appendix A. ■

In the FP model, uninformed voters abstain with increasing probability until, in the limit, they all abstain. Turnout is nevertheless positive, however, because informed citizens always vote. Here, voters are in a way both informed and uninformed. When μ is small, a citizen might reasonably expect to be the best informed voter in the electorate. As $\mu \to \infty$, however, the expected number of better-informed citizens grows arbitrarily large. This may seem to suggest that even extremely well-informed citizens will eventually defer to those with better information (or, equivalently, that $T^*_{\mu} \to 1$), so that turnout approaches 0%. In evaluating this possibility, Lemma 1 derives the function $L(T)$ to which T^{μ}_{BR} converges:

[4] If $T = 1$, no one votes regardless of μ.

$$L(T) = \frac{\sqrt{M(T)}}{\sqrt{M(T)} + \sqrt{1 - M(T)}} \tag{21}$$

where the *mean quality function* $M(T) \equiv E(Q|Q \geq T) = \frac{1}{1-F(T)} \int_T^1 q dF(q) = \frac{p_+}{p_+ + p_-}$ gives the average information quality of citizens who vote. Equivalently, M is the expected vote share of the superior candidate.

Lemma 1 *For any $T \in \left[\frac{1}{2}, 1\right)$, $\lim_{\mu \to \infty} T_{BR}^{\mu} = L(T)$, as defined as in (21).*

Proof. See Appendix A. ■

Since $T_{BR}^{\mu}(T) \to L(T)$, any limit point of a sequence of fixed points of T_{BR}^{μ} must be a point of L. Theorem 6 now shows that such a fixed point must be strictly less than one. In other words, turnout remains bounded strictly above 0%, even as an electorate grows arbitrarily large. The fallacy in the above intuition is that, while a citizen indeed expects both the quantity and the quality of opponents' votes to be high, she bases her behavior not on her expectations, but rather on her *conditional* expectations: when her vote is pivotal, her opponents' votes must be of far lower quantity and quality than she originally expected; the difference in this case between a one-vote win and a one-vote continues to grow, but decreasingly so. By voting rather than abstaining in a finite election, the marginal voter drags down the average quality of votes cast. Her impact on the average quality is reduced, however, as the number of voters grows; in the limit, she contributes to the quantity of votes without reducing the quality at all.

Theorem 6 *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. See Appendix A. ■

In general, it is possible for L to have multiple fixed points between $\frac{1}{2}$ and 1. If the density f is *log-concave* (i.e. $\log(f)$ is concave or, equivalently, $\frac{f'}{f}$ is decreasing), however, as it is for many of the most common distributions,[5] then Theorem 7 rules out this possibility.[6] Uniqueness in the limit, of course, implies a unique participation cutpoint T_{μ}^* for any μ sufficiently large.

[5] 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.

[6] Log-concavity is 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. As the proof of Theorem 7 makes clear, the important things is just that f not have large "spikes" of probability.

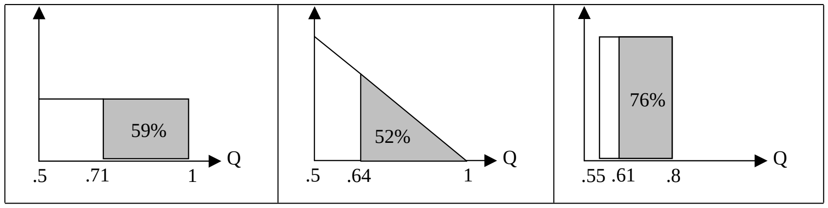

Figure 1: Participation thresholds and turnout rates for simple information distributions

Theorem 7 *If f is log-concave then L has a unique fixed point T_∞^* strictly between $\frac{1}{2}$ and 1.*

Proof. See Appendix A. ■

With a unique equilibrium threshold T_∞^*, the participation rate $1 - F(T_\infty^*)$ is likewise unique. Since these values depend exclusively on the underlying information distribution, both may be computed numerically for any closed-form F, using the formula from (21). To illustrate, Figure 1 displays the equilibrium participation thresholds and turnout rates associated with some simple distributions. If information quality is distributed uniformly between $\frac{1}{2}$ and 1, for example, the limiting cutpoint is $T_\infty^* = 0.71$, and the best-informed 59% of citizens vote.

2.5 Welfare and Election Design

Since voters share common values, social welfare is equivalent to individual utility, and can therefore be measured by the probability $\Pr(X = Z)$ of electing the superior candidate. As $\mu \to \infty$, the law of large numbers implies that candidate Z's actual vote share approaches the expected vote share, $M(T_\infty^*)$. As in the original Condorcet jury theorem, therefore, $\Pr(X = Z) \to 1$. On the other hand, one message of the original Condorcet model 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. While it is difficult to attribute precise meaning to informal arguments, this logic seems to underlie the commonly held concern that low and declining voter turnout is a serious threat to democracy. To prevent voter abstention, a number of democracies (e.g. Australia and several Latin American countries) have made voting compulsory; Lijphart (1997), among others, recommends that the United States do the same.

With common values, however, the socially optimal level of turnout must be achieved in equilibrium; since $T^* > \frac{1}{2}$ in the unique equilibrium, it must be that mandatory voting actually pools less, not more information, than voluntary voting. One intuition for this rather surprising result is that a majority election must weights votes equally. A compulsory election collects a larger number of signals, but collects no information regarding the quality of those signals; in a voluntary election, voters' signals (i.e. signals for which $Q_i \geq T^*$) are given more weight than nonvoters' signals (i.e. signals for which $Q_i < T^*$).[7] The optimal election mechanism would directly ask voters to report both Q_i and S_i,[8] and would weight individual votes by their underlying quality in a maximum likelihood approach, so that candidate A wins if and only if the probability of observing $\{S_i\}_{i=1}^N$ is greater when A is superior to B than when B is superior to A, so that the following inequality holds.[9]

$$\prod_{i:s_i=A} q_i \prod_{j:s_j=B} (1-q_j) > \prod_{i:s_i=A} (1-q_i) \prod_{j:s_j=B} q_j$$

According to Lijphart (1997), John Stuart Mill proposed in 1861 that educated voters be allowed to vote two or more times in an election, and such a system was actually used in Belgium from 1893 to 1919; clearly, the intent of such a policy is similar to the one I have described. A similar method would be to allow voters to rate candidates on a point scale (say 1 to 10), as is common in judging arts and athletic competitions. By awarding two candidates the same number of points, a judge effectively demonstrates that $Q_i = \frac{1}{2}$, while awarding 10 to one and 1 to the other demonstrates Q_i close to one.

2.6 Comparative Statics

Information Quality

The distribution F of information quality that uniquely determines equilibrium voting behavior is itself determined by factors that may vary both regionally and over time, such as voters' education or experience levels, and access to information technology. In this section, therefore, I analyze how turnout responds to changes in F, adding a superscript to denote the reliance on F of the mean quality function M^F, the limiting best response function L^F, and the limiting participation threshold T^F_∞. Participation is given by the *survival function*

[7] Analyzing a model similar to this, Krishna and Morgan (2008) likewise conclude that mandatory voting pools less information than voluntary voting, but for an entirely different reason. In their model, signal quality depends on the state (i.e. a signals are more reliable than b signals). To compensate for this, citizens strategically abstain, if allowed; otherwise, they vote uninformatively, and information is aggregated less effectively.

[8] Since values are common, voters have an incentive to tell the truth: if $S_i = A$ and $Q_i = q$ then claiming $Q_i > q$ just increases the probability of wrongly electing A, and claiming $Q_i < q$ just increases the chance of wrongly electing B.

[9] See Shapley and Grofman (1984). This rule is equivalent to a standard voluntary election if some voters are perfectly uninformed (i.e. $Q_i = \frac{1}{2}$) and others have identical information quality (i.e. $Q_i = q > \frac{1}{2}$), and to a mandatory election if everyone belongs to the latter group. Note that this decision rule is distinct from the Maximum Likelihood rule developed by Kemeny (1959) and discussed by Young (1995), which seeks instead to identify an "average" preference ranking among voters, in an election with multiple candidates.

$\bar{F} \equiv 1 - F$, evaluated at T_∞^F. In what follows, I compare T_∞^F and $\bar{F}\left(T_\infty^F\right)$ with corresponding values T_∞^G and $\bar{G}\left(T_\infty^G\right)$ associated with an alternative distribution G. In preparation for this, Lemma 2 presents a useful condition that is equivalent to $T_\infty^G > T_\infty^F$.

Lemma 2 *Let F and G be continuous, log-concave distributions with strictly positive densities, and let T_∞^F and T_∞^G and M^F and M^G denote the unique limiting participation thresholds and mean quality functions for F and G, respectively. Then $T_\infty^G > T_\infty^F$ if and only if $M^G(T) > M^F(T)$ or, equivalently,*

$$\frac{\bar{F}(T)}{\bar{G}(T)} > \frac{\int_T^1 \bar{F}(q)\,dq}{\int_T^1 \bar{G}(q)\,dq} \tag{22}$$

for $T \in \left\{T_\infty^F, T_\infty^G\right\}$.

To begin, consider a general improvement in information quality. G is said to *first-order stochastically dominate* F (written $G \geq_1 F$) 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 Theorem 8 emphasizes: (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 indirectly (by lowering the average vote quality, thereby weakening the swing voter's curse and lowering T_∞^*).[10] These effects are illustrated with numerical examples in Figure 2, starting from a uniform distribution. Regardless of its effect on turnout, the direct welfare effect of improving information quality is to improve election accuracy, and any strategic response to improved information can only increase welfare further.

Theorem 8 *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 $q \geq M^F\left(T_\infty^F\right)$ 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 $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$.

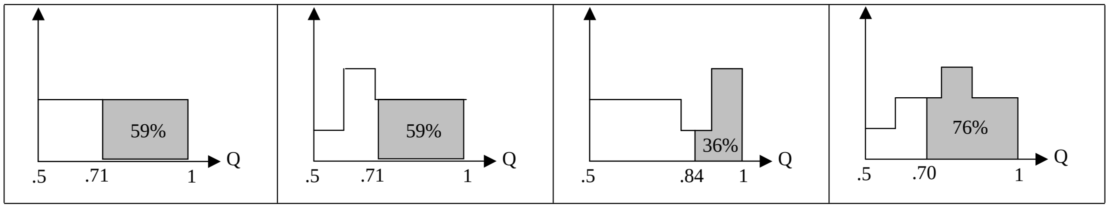

Figure 2: When information improves for different segments of the electorate, turnout may remain the same, decrease, or increase

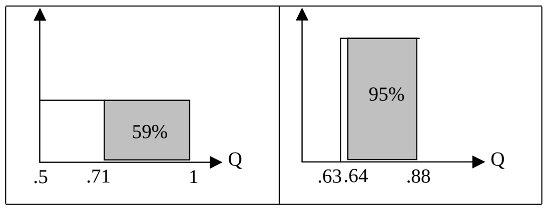

Figure 3: A mean-preserving decrease/increase in the variance of information quality raises/lowers turnout.

Proof. See Appendix A. ■

Of course, it may be that neither of two distributions first-order stochastically dominates the other. A related but weaker condition is that $\int_{1/2}^{T} \bar{G}(q)\,dq \geq \int_{1/2}^{T} \bar{F}(q)\,dq$ for every T, or that G *second-order stochastically dominates* F (written $G \geq_2 F$). For distributions with a common mean, $G \geq_2 F$ implies that G has a smaller variance than F. In this case, provided that T_∞^F lies below the common mean (as numerical examples suggest is typical), Theorem 9 states that turnout is higher under G than under F. Intuitively, this is because the swing voter's curse is weak when the quality difference between informed and uninformed votes is small; in the extreme case, voters all have identical information quality and turnout is 100%, as in the Condorcet model. The welfare difference between F and G is ambiguous; though the number of votes is higher in G, the average quality of votes is higher in F. The association between variance and turnout is illustrated for a uniform distribution in Figure 3.

Theorem 9 *Let F and G be continuous, log-concave distributions with strictly positive densities and a common mean m, such that $G \geq_2 F$ and $T_\infty^F \leq m$. 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 A. ■

[10] Symmetrically, information reductions below T have no effect, small reductions above T increase turnout, and moderate reductions above T reduce turnout both directly and indirectly.

Voting Costs and Benefits

While the swing voter's curse is particularly notable for its explanation of abstention in costless environments, the cost of voting in the majority of real-world election settings is clearly positive.[11] If voters pay a cost $C > 0$ and obtain a benefit B for improving the election outcome, the expected benefit of voting will exceed the cost if and only if the following (equivalent) inequalities hold:

$$\begin{aligned} Q_i P B + (1 - Q_i)\, \tilde{P}\,(-B) &> C \qquad (23) \\ \left[Q_i\left(P + \tilde{P}\right) - \tilde{P}\right] B &> C \\ Q_i &> \frac{\tilde{P} + \frac{C}{B}}{P + \tilde{P}} \qquad (24) \end{aligned}$$

With the addition of this $\frac{C}{B}$ term, this new participation threshold is higher–and turnout is lower–than before. Since the former level of turnout maximizes social welfare (as discussed in section 2.5), the new level is inefficiently low. The extent of this distortion depends on the size of the benefit term; fortunately, in an "important" election (i.e. when B is large) the distortion term is small.[12]

The prediction that voting costs discourage turnout raises the question of why citizens nevertheless participate in large elections? As Downs (1957) observes, pivot probabilities shrink asymptotically to zero as the number of voters grows large, so the inequality in (23) eventually fails as $\mu \to \infty$, and voting appears irrational. In response to this "turnout paradox", Riker and Ordeshook (1968) hypothesize that voters are motivated by a sense of civic duty (or respond to social pressure), incurring a private cost for the benefit of society.[13] This explanation has limited justification when preferences are not common (since, as Borgers 2004 shows, turnout is too high in that case, rather than too low[14]), but in this common-values setting a sense of duty is completely natural.

To formalize the notion of civic duty, Riker and Ordeshook (1968) add a benefit D to the left hand side of (23), independent of the election outcome. Accordingly, even as pivot

[11] Knack (1994), for instance, confirms the conventional wisdom that inclement weather slightly lowers turnout.

[12] Interestingly, when voting is costless the importance of an election does not influence voters' participation decisions: a high B renders a correct vote more valuable, but also makes an incorrect vote more painful. Thus, important races on a ballot may attract voters to polls, but will not influence voters' roll-off decisions once they arrive (and voting costs are sunk).

[13] Dowding (2005) and Geys (2006) review numerous attempts to rationalize voting behavior, but none has endured longer than Riker and Ordeshook's (1968) duty hypothesis.

[14] If turnout is too high, civic duty should lead citizens to abstain, rather than to vote, and society should recognize nonvoters, rather than voters, as being good citizens.

probabilities tend to zero, (23) remains satisfied as long as $D > C$. With this addition, (24) becomes the following:

$$Q_i > \frac{\tilde{P} + \frac{C-D}{B}}{P + \tilde{P}} \tag{25}$$

From (25) it is clear that D reduces the turnout distortion generated by voting costs, thereby mitigating the externality. If $D > C$, however, then the externality actually *reverses*: since welfare was uniquely maximized for the original threshold $\frac{\tilde{P}}{P+\tilde{P}}$, any threshold *higher or lower* than that reduces welfare. Thus, while the notion of civic duty is compelling in this common values setting, the D term formulation used by Riker and Ordeshook (1968) is problematic: $D > C$ actually reduces welfare, while $D < C$ is too small to motivate turnout when pivot probabilities tend to zero. One alternative formulation for civic duty that can only be welfare-improving is an inflated B term, though only an infinite B would completely eliminate the turnout distortion in (24).[15]

3 Evidence and Applications

3.1 Sincere Voting, Moderate Turnout, and Roll-off

Matsusaka (1993) questions whether voters are sufficiently sophisticated to be able to calculate and respond to miniscule pivot probabilities. Recreating the incentives of the original FP model in an experimental setting, Battaglini, Morton, and Palfrey (2006) confirm that players can and indeed do respond to the swing voter's curse, but leave open the question of whether citizens vote strategically in the real world. One prediction of many game-theoretic voting models, beginning with Austen-Smith and Banks (1996), is that equilibrium voting is *insincere*. Using panel data on voting behavior, however, Degan and Merlo (2007b) fail to reject the hypothesis that voters merely vote sincerely. Sincere voting, however, is perfectly consistent with this model: strategic considerations guide voters' participation decisions, but citizens vote sincerely if they vote at all (Theorems 1 and 2).

It is difficult to test this model's validity formally, since information quality cannot be directly observed. Informally, however, Figure 1 illustrates how predicted turnout rates can be computed for simple distributions, using equation (21). Roughly speaking, turnout rates predicted for these distributions are similar to actual participation rates in U.S. state and national elections. No matter the distribution, the result from Theorem 6 that $\frac{1}{2} < T^*_\infty < 1$ guarantees that both turnout and abstention will be positive, even in a large electorate.

[15] Edlin, Gelman, and Kaplan (2005) hypothesize that voters are motivated by altruism. If voters receive a benefit β for each fellow-citizen that a policy benefits in addition to the private benefit B, then the total benefit $B + \mu\beta$ indeed approaches infinity as the population grows large.

Avoiding some of the criticisms of strategic models, Matsusaka (1995) explains the connection between education and voting in a model that is similar to this but decision-theoretic: the expected benefit of voting is increasing in Q_i, so uninformed citizens are more easily dissuaded by voting costs. As FP rightly point out, however, this fails to explain the ubiquitous phenomenon of roll-off voting, since, for a citizen inside the voting booth, voting costs are sunk. Other examples of costless voting environments include committee voting, which requires only the raise of a hand, and mail-in ballots (available in most states and used exclusively in the state of Oregon), which allow citizens to vote conveniently from home. These examples of abstention even when voting is costless can be accommodated in this model, since for Q_i sufficiently low, the expected benefit of voting is actually negative.

3.2 Turnout and Education

Another evidence of the validity of this model is the strong empirical connection between education and voting, as mentioned in the Introduction. To reiterate, Wolfinger and Rosenstone (1980) find education to be the best available predictor of voting. The connection between education and information quality is intuitive; as mentioned earlier, several authors (e.g. Bartels, 1996; Degan and Merlo, 2007a; and Larcinese, 2006) also find direct empirical connections between information quality and voting, and Lassen (2005) finds that, controlling appropriately for information, education does not otherwise influence turnout.

Even to the extent that alternative theories can and do explain the positive association between education and voting, they inevitably fail to explain the paradoxical observation by Brody (1978) that voter participation has persistently fallen over recent decades even while education levels have risen, so that education and turnout are positively correlated in any given election, but *negatively* correlated over time. Surveying the literature, Aldrich (1993) discusses problems with existing explanations of this phenomenon, ultimately finding it to be "the most important substantive problem in the turnout literature."

In contrast with other models, this model provides at least two plausible explanations for the Brody (1978) paradox. First, note that while an increase in a citizen's own information quality makes her more likely to vote, an increase in her peers' information makes a citizen *less* likely to vote, by strengthening the swing voter's curse.[16] To illustrate this, compare the first two panels of Figure 1, and note that a citizen with information quality $Q_i = 0.7$ will vote in the second electorate, but will abstain from voting in the first, because her peers' information quality is higher. Thus, as education levels increase generally, it is entirely possible that citizens who formerly voted should now abstain.

[16] FP (1999) demonstrate this possibility for certain parameter values, but the larger class of information distributions here facilitates a much more detailed analysis.

A second possibility is that, even though education levels have risen for certain demographics, information quality in general has declined, because of other factors. For example, as the number and scope of laws and policies have proliferated over time, the decisions required of politicians (and therefore voters) may have grown more complicated, in effect shifting the entire distribution of information quality downward. A citizen who, decades ago, might have felt confident in her opinions on most political issues, may find now that many issues extend beyond her training and expertise.

Admittedly, these two explanations are somewhat contradictory: the first asserts that improved information lowers turnout by strengthening the swing voter's curse, and the second that reduced information quality lowers turnout by dragging voters below the participation threshold; an alternative possibility is that improved information raises turnout by lifting nonvoters above the participation threshold, or that reduced information quality raises turnout by weakening the swing voter's curse. It is conceivable that information quality has improved for certain segments of society but decreased overall, but whether this will raise or lower turnout depends on the distribution, as described by Theorem 8. Therefore, which of the two explanations is at work–or which is dominant, if both are at work–are questions better left for future research; the primary purpose of this discussion is to illustrate how easily this model accommodates an empirical phenomenon which has previously been difficult to explain.

3.3 Absolute vs. Relative Information

The result that an individual's own information quality makes her more likely to vote while others' information quality makes her less likely to vote implies that the importance of information quality is somewhat *relative*, not absolute. Put differently, a citizen is most likely to vote when her own information quality is high, but when the information quality of others in her electorate is low. This provides a subtle difference between strategic and decision-theoretic motivations for abstention, which can be tested empirically, using established information proxies such as education and age. In this section I present statistical evidence from the American National Election Studies (ANES). During presidential election years between 1972 and 1992, the ANES asked respondents whether or not they had participated in their state's political primary elections, the previous spring. Though preferences within a party may be quite heterogeneous, the common values assumption is likely a good approximation here, since the over-arching goal within a political party is to select candidates who will be most likely to triumph in the upcoming general election.

Table 1 displays results from a probit model, where the dependent variable is equal to one if an individual voted in her party's state primary election, and zero otherwise. The

Table 1-----Information and Voting in Primary Elections

Probit: 1 = voted in primary election

	1a	1b	2a	2b	3a	3b
Education (level)	0.0252***	-0.0001	0.0283***	0.0048	0.0340***	0.0101
	(5.87)	(-0.01)	(6.35)	(0.36)	(6.03)	(0.71)
Education (%)	-	0.0017**	-	0.0015*	-	0.0015*
	-	(2.45)	-	(1.89)	-	(1.80)
Age (level)	0.0057***	0.0018	0.0016***	0.0020	0.0073***	0.002
	(15.07)	(1.47)	(15.48)	(1.27)	(14.86)	(1.29)
Age (%)	-	0.0026***	-	0.0026***	-	0.0034***
	-	(3.45)	-	(2.76)	-	(3.58)
Income (level)	0.0384***	0.0082	0.0396	0.0024	0.0431***	-0.0225
	(6.16)	(0.43)	(6.13)	(0.10)	(5.23)	(-0.89)
Income (%)	-	0.0010	-	0.0012	-	0.0023**
	-	(1.49)	-	(1.50)	-	(2.52)
Information (level)	0.0707***	0.0207	0.0750***	0.0322	0.0681***	0.0065
	(10.16)	(1.1)	(10.36)	(1.35)	(7.43)	(0.25)
Information (%)	-	0.0019***	-	0.0015*	-	0.0023**
	-	(2.87)	-	(1.85)	-	(2.45)
Year dummies	yes	yes	-	-	yes	yes
State-Party dummies	yes	yes	-	-	yes	yes
Year-State-Party dummies	-	-	yes	yes	-	-
# of Observations	6614	6614	6507	6507	4125	4125
Years	1972-92	1972-92	1972-92	1972-92	1980-92	1980-92
Pseudo R-squared	0.123	0.132	0.151	0.153	0.142	0.148

Notes : Table entries are marginal effects, with z-statistics in parenthesis. All specifications include controls (not shown) for gender and race. Data source is American National Election Studies (1972, 1976, 1980, 1988, 1992). Percentile variables are computed within year-state-party groups. *, **, ***, indicate significance at the 10%, 5%, and 1% levels, respectively.

independent variables are standard proxies of information quality: education (in seven categories), age (in years), income[17] (in five categories), and a subjective measure of information quality (in five categories) made by the interviewer at the time of an interview.[18] For each absolute information proxy, I also compare an individual with the others from the same state, year, and party,[19] to generate a percentile variable that indicates an individual's relative position within the distribution of her peers. Her education percentile, for example, indicates the fraction of her peers with education levels lower than or equal to than her own. Non-strategic models of abstention, such as Matsusaka (1995), may predict a positive relationship between absolute information and voting, but provide no reason to expect relative information measures to have any coefficient other than zero. This model, in contrast, predicts that relative information quality should have a positive relationship with voting, while making no predictions regarding the importance of absolute information.[20]

[17] Income can approximate information quality if both are influenced by, say, analytical thinking skills.

[18] Zaller (1985) finds interviewers' impressions to be the most useful information variable for explaining turnout in the ANES. Interviews were conducted both before and after the November general elections; I use the pre-election information measure, but the post-election measure yields similar results.

[19] State-year-party groups with fewer than 15 observations are discarded. Other size cutoffs yield similar results.

[20] As discussed in section 2.6, comparative statics results regarding the absolute level of information quality are ambiguous.

Specifications I and II include controls for gender and race, as well as fixed year effects and fixed effects for each of the sixty-two state-party pairs. Specification I includes only the absolute information proxies, each of which exhibits a strong and positive relationship with turnout. The cell entries in Table 1 are marginal effects and z-statistics; other things equal, increasing one education, income, or information quality level, respectively, makes a citizen 2.5%, 3.8%, and 7.1% more likely to vote. To this, specification II adds the percentile proxies of relative information quality. With this addition, the marginal effects of absolute variables are each reduced by over two thirds. None remains significant at conventional levels, and the point estimate of the marginal effect of education actually becomes slightly negative. The relative information proxies, on the other hand, are all positive and strongly significant (with the exception of income, which is statistically insignificant, but more significant than the absolute measure of income, with a z-statistic of 1.49 versus 0.43).

The magnitudes of the marginal effects of absolute and relative information proxies cannot be compared directly, since the former is the marginal effect of increasing one category of education, income, or information quality, or one year of age; while the latter is the marginal effect of rising by one percentage point within a distribution. However, the two can be compared by considering a transition from the bottom of the distribution to the top. For example, other things equal, rising from the bottom to the top of the education distribution makes a citizen $100 \times 0.17\% = 17\%$ more likely to vote. Rising from the bottom to the top of the age, income, and information distributions makes her 26%, 10%, and 19% more likely to vote, respectively. In contrast, moving from age 17 to age 99 makes a citizen only $82 * 0.18\% = 14.76\%$ more likely to vote, and moving from income or information category 1 to category 5 makes her $4 \times 0.82\% = 3.28\%$ and $4 \times 2.07\% = 8.28\%$ more likely to vote, respectively. The absolute level of education has no (or even negative) impact on turnout, and the 17% increase in voting propensity because of moving from the bottom to the top of the education distribution is even larger than the $6 \times 2.52\% = 15.12\%$ increase associated with moving from education category 1 to category 7, in specification I.

The remaining columns of Table 1 are similar to columns I and II. Instead of fixed year and separate state-party effects, specifications III and IV include fixed effects for each of the 189 state-year-party groups with at least 15 observations. Specifications V and VI use the original fixed effects, but only use the most recent data, beginning in 1980. In both cases, the qualitative results of specifications I and II are repeated: excluding the percentile variables, regressions III and V exhibit large and significant marginal absolute effects. When percentiles are included in IV and VI, these absolute variables lose significance and the marginal relative effects are instead positive and significant. The consistent importance of relative information quality, together with the consistent loss of significance of absolute information quality, suggest that voters indeed respond (either deliberately or instinctively) to the information quality of those around them, by choosing not to vote.

4 Conclusion

In a common values election setting, the influential models of Condorcet (1785) and Feddersen and Pesendorfer (1996) predict opposite reactions to imperfect information: in the FP model, the swing voter's curse leads poorly informed citizens to abstain from voting; in the Condorcet model everyone votes, so that the election decision will be based on as much (albeit low-quality) information as possible. By allowing an arbitrary distribution of information quality, this model becomes a natural but insightful blend of its predecessors: below the unique equilibrium participation threshold T^*, citizens opt not to express their own opinions, deferring instead to those with better information; above T^*, citizens vote even when the number of better-informed peers grows infinitely large.

While turnout decisions are made strategically, voting is sincere (and in pure strategies). Though information quality cannot be readily observed, simple distributions yield turnout predictions similar to actual participation levels in U.S. state and national elections, and voting and abstention both remain positive, even in large electorates. Strategic considerations can explain abstention in costless voting environments, such as roll-off, and comparative statics can explain the empirical association between voting and education or information, as well as the simultaneous rise and fall of education and turnout. Also, the empirical importance of education, age, and other reasonable proxies of information quality in explaining participation in party primary elections appears to be relative, rather than absolute.

One prediction of this model that fails empirically is that elections should tend to result in landslide victories. The expected margin of victory (as a fraction of the total number of voters) is $MV = 2\left[E\left(Q|Q \geq T^*\right) - \frac{1}{2}\right]$,[21] which is extremely high for many distributions: for the uniform distribution illustrated in Figure 1, for example, Z gets over 85% of the votes! Related to this is the unrealistic assumption that citizens cannot communicate their information to each other prior to an election. As Coughlan (2000) points out in response to FP (1998), if citizens were allowed to communicate freely, asymmetric information would be eliminated and voting would be unanimous. Without completely abandoning the basic framework of this model, two possible approaches for avoiding these unrealistic results are to relax the assumptions of common values or common prior beliefs; I discuss these possibilities in McMurray (2008).

[21] This expression is given by the difference between the expected numbers $E\left(Q_i|i \text{ votes}\right)$ and $E\left(1 - Q_i|i \text{ votes}\right) = 1 - E\left(Q_i|i \text{ votes}\right)$ of votes for and against Z.

Like both the FP and Condorcet models, this model predicts that, as the population grows large, the best candidate will be elected with probability approaching one. To maximize social welfare, however, the FP model implies that poorly informed citizens should abstain, while the Condorcet model implies they should vote. Here, because poorly informed citizens are not perfectly uninformed, abstention is inefficient. Compulsory voting laws, however, only exacerbate the problem, losing information conveyed by the turnout decision itself; a better approach would be to weight voters' opinions with their accuracy levels. If voting is costly, it generates a positive externality. A sense of civic duty could mitigate this externality, but is an unsatisfactory resolution for the "paradox of turnout" in large elections, as a large sense of duty actually reverses the voting externality.

While voter abstention has traditionally been viewed as a serious threat to democracy, this model demonstrates that turnout can also be too high. Rather than targeting turnout per se, therefore, government policy and social activism should focus on eliminating barriers to voting (e.g. simplify registration requirements, provide rides to polls, institute internet voting, etc.). Even if barriers were completely eliminated, 100% turnout is unlikely, but the resulting equilibrium level of turnout would be socially optimal. Recent declines in voter turnout may or may not be cause for alarm; as mentioned above, two somewhat opposite possibilities are that falling turnout reflects either a strategic response to educational improvements, or a general rise in the complexity of government. Improving education and voter information and simplifying, limiting, or otherwise reducing the complexity of government will have ambiguous effects on turnout, but can only improve social welfare.

References

[1] Abramson, Paul R., John H. Aldrich, Phil Paolino, and David W. Rhode. 1992. "'Sophisticated' Voting in the 1988 Presidential Primaries." *The American Political Science Review*, 86(1): 55-69.

[2] Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science*, 37(1): 246-278.

[3] 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.

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

[5] Battaglini, Marco, Rebecca Morton, and Thomas Palfrey. 2006. "The Swing Voter's Curse in the Laboratory." CEPR Discussion Paper 5458.

[6] Borgers, Tilman. 2004. "Costly Voting." *The American Economic Review*, 94(1): 57-66.

[7] Brody, Richard A. 1978. "The Puzzle of Political Participation in America." In *The New American Political System*, ed. Anthony King. Washington DC: American Enterprise Institute.

[8] Chamberlain, G., and Rothchild, M. 1981. "A Note on the Probability of Casting a Decisive Vote." *Journal of Economic Theory*, 25: 152-162.

[9] Condorcet, Marquis de. 1785. *Essai sur l'application de l'analyse a la probabilite' des decisions rendues a la pluralite' des voix.* Paris: De l'imprimerie royale. Trans. Iain McLean and Fiona Hewitt. 1994.

[10] Coughlan, Peter J. 2000. "In Defense of Unanimous Jury Verdicts." *American Political Science Review* 94(2): 375-94.

[11] Crain, Mark W., Donald R. Leavens, and Lynn Abbot. 1987. "Voting and Not Voting at the Same Time." *Public Choice* 53, pp. 221-229.

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

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

[14] Degan, Arianna and Antonio Merlo. 2007b. "Do Voters Vote Sincerely?" NBER Working Paper 12922.

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

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

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

[18] 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.

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

[20] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1997. "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, 65(5): 1029-1058.

[21] 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.

[22] Feddersen, Timothy J. and Wolfgang Pesendorfer. 1999. "Abstention in Elections with Asymmetric Information and Diverse Preferences." *The American Political Science Review*, 93(2): 381-398.

[23] Geys, Benny. 2006. "'Rational' Theories of Voter Turnout: A Review." *Political Studies Review*, 4: 16-35.

[24] Kemeny, John. 1959. "Mathematics Without Numbers." *Daedalus*, 88: 571-591.

[25] Knack, Stephen. 1994. "Does Rain Help the Republicans?: Theory and Evidence on Turnout and the Vote." *Public Choice*, 79: 187-209.

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

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

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

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

[30] Matsusaka, John. 1993. "Election Closeness and Voter Turnout: Evidence from California Ballot Propositions." *Public Choice*, 76: 313-334.

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

[32] McKelvey, R. and T. Page. 1986. "Common Knowledge, Consensus, and Aggregate Information," *Econometrica*, 54: 109-127.

[33] McMurray, Joseph C. 2008. "Information, Ideology, and Voting." Unpublished.

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

[35] 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.

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

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

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

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

[40] 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.

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

[42] 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.

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

[44] Peyton Young. 1995. "Optimal Voting Rules." *The Journal of Economic Perspectives*, 9(1): 51-64.

[45] Zaller, John. 1985. "Pre-Testing Information Items on the 1986 NES Pilot Survey." *Report to the National Election Study Board of Overseers.*

Appendices

A Proofs

Theorem 2 *Let σ^* be a symmetric Bayesian equilibrium and let α_A and α_B be the posterior thresholds defined in (12). Then $\alpha_A > \frac{1}{2}$ and $\alpha_B > \frac{1}{2}$.*

Proof. I first show that $\alpha_A \geq 1 - \alpha_B$ by supposing to the contrary that $\alpha_A < 1 - \alpha_B$. According to condition (i) of Theorem 1, then, the best response σ_{BR} to σ^* consists of voting for A whenever $\alpha_i > \hat{\alpha}$ and voting B otherwise, never abstaining. I assume here that $\hat{\alpha} \geq \frac{1}{2}$; otherwise a symmetric argument applies. When $\hat{\alpha} \geq \frac{1}{2}$, all agents of quality $q < \hat{\alpha}$ vote B; above $\hat{\alpha}$, agents vote informatively (i.e. $\sigma_{BR}^A(q,a) = \sigma_{BR}^B(q,b) = 1$). Therefore, the vote probabilities from (2) and (3) simplify to the following,

$$\begin{array}{ll} p_{AA} = \int_{\hat{\alpha}}^1 q dF(q) & p_{BA} = F(\hat{\alpha}) + \int_{\hat{\alpha}}^1 (1-q)\, dF(q) \\ p_{AB} = \int_{\hat{\alpha}}^1 (1-q)\, dF(q) & p_{BB} = F(\hat{\alpha}) + \int_{\hat{\alpha}}^1 q dF(q) \end{array}$$

implying the following inequalities:

$$p_{BB} > p_{BA} > p_{AB} \tag{26}$$

$$p_{BB} > p_{AA} > p_{AB} \tag{27}$$

$$p_{AA}p_{BA} > p_{AB}p_{BB} \tag{28}$$

Given (26) through (28), I now argue that $P_A P_B < \tilde{P}_A \tilde{P}_B$ or, equivalently, that $\alpha_A > 1 - \alpha_B$, which is the desired contradiction:

$$\begin{aligned} & P_A P_B - \tilde{P}_A \tilde{P}_B \\ = \; & \frac{1}{4}(\pi_{0A} + \pi_{-1A})(\pi_{0B} + \pi_{-1B}) - \frac{1}{4}(\pi_{0A} + \pi_{1A})(\pi_{0B} + \pi_{1B}) \\ = \; & \frac{1}{4}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{e^{-\mu(p_{AA}+p_{AB}+p_{BA}+p_{BB})}}{j!k!} (p_{AA}p_{BA})^j (p_{AB}p_{BB})^k \times \qquad (29) \\ & \left[\left(1 + \frac{\mu}{j+1}p_{BA}\right)\left(1 + \frac{\mu}{k+1}p_{AB}\right) - \left(1 + \frac{\mu}{j+1}p_{AA}\right)\left(1 + \frac{\mu}{k+1}p_{BB}\right)\right] \qquad (30) \end{aligned}$$

Whenever $j \geq k$, the bracketed term in (30) is negative, since

$$\begin{aligned} & \left[\left(1 + \frac{\mu}{j+1}p_{BA}\right)\left(1 + \frac{\mu}{k+1}p_{AB}\right) - \left(1 + \frac{\mu}{j+1}p_{AA}\right)\left(1 + \frac{\mu}{k+1}p_{BB}\right)\right] \\ = \; & \left[\frac{\mu}{k+1}(p_{AB} - p_{BB}) + \frac{\mu}{j+1}(p_{BA} - p_{AA}) + \frac{\mu}{k+1}\frac{\mu}{j+1}(p_{BA}p_{AB} - p_{AA}p_{BB})\right] \\ < \; & \left[\frac{\mu}{j+1}(p_{AB} - p_{AA} + p_{BA} - p_{BB})\right] \\ < \; & 0 \end{aligned}$$

For any (j,k) term of the series in (29) that is positive, therefore, the corresponding (k,j) term is negative. This only occurs when $k < j$, in which case (28) implies that the weight $(p_{AA}p_{BA})^k (p_{AB}p_{BB})^j$ on the negative term exceeds the weight $(p_{AA}p_{BA})^j (p_{AB}p_{BB})^k$ on the positive term. Thus the series must be negative.

Given that $\alpha_A \geq 1-\alpha_B$, suppose next that $\alpha_B < \frac{1}{2}$, so that agents who receive b signals all vote B. Agents who receive a signals will vote A if they are sufficiently well-informed (i.e. $q \geq \alpha_A$), vote for B if they are sufficiently uninformed (i.e. $q \leq 1-\alpha_B$), and abstain otherwise. The vote probabilities from (2) and (3) then simplify to the following:

$$\begin{array}{ll} p_{AA} = \int_{\alpha_A}^1 q dF(q) & p_{BA} = F(1-\alpha_B) + \int_{1-\alpha_B}^1 (1-q)\, dF(q) \\ p_{AB} = \int_{\alpha_A}^1 (1-q)\, dF(q) & p_{BB} = F(1-\alpha_B) + \int_{1-\alpha_B}^1 q dF(q) \end{array}$$

These once again imply inequalities (26) through (28), which then imply the following.

$$\begin{aligned} \tilde{P}_A - P_B &= \frac{1}{2}\left(\pi_{0A} + \pi_{1A} - \pi_{0B} - \pi_{-1B}\right) \\ &= \frac{1}{2}\sum_{k=0}^{\infty} e^{-\mu(p_{AA}+p_{BA})}\left(\mu^2 p_{AA}p_{BA}\right)^k \left[1 + \frac{\mu}{k+1}p_{AA}\right] \\ &\quad -\frac{1}{2}\sum_{k=0}^{\infty} e^{-\mu(p_{AB}+p_{BB})}\left(\mu^2 p_{AB}p_{BB}\right)^k \left[1 + \frac{\mu}{k+1}p_{AB}\right] \\ &\geq \frac{1}{2}\sum_{k=0}^{\infty} e^{-\mu(p_{AA}+p_{BA})}\left(\mu^2 p_{AA}p_{BA}\right)^k \left[\frac{\mu}{k+1}\left(p_{AA} - p_{AB}\right)\right] \\ &\geq 0 \end{aligned}$$

The result that $\tilde{P}_A \geq P_B$ is equivalent to $\alpha_B \geq \frac{1}{2}$; thus in equilibrium it cannot be that $\alpha_B < \frac{1}{2}$ or, by a symmetric argument, that $\alpha_A < \frac{1}{2}$. ■

Lemmas A1 and A2 are useful in preparation for the proof of Theorem 4.

Lemma A1 *If* $T_{BR}(T) \geq T$ *then* $T'_{BR}(T) \geq 0$.

Proof. The computation of T'_{BR} is rather involved since T_{BR} depends on P and $\tilde{P}$, and therefore on π_0, π_1, π_{-1}, p_+, and p_-, as well as the simplifying notation ψ_k defined in (7).

Differentiate each of these component parts, as follows:

$$
\begin{aligned}
p'_+ &= -Tf && (31)\\
p'_- &= -(1-T)f && (32)\\
\psi'_k &= \frac{\mu^{2k}}{k!k!}\left\{\mu f e^{-\mu(1-F)}\left(p_+p_-\right)^k + e^{-\mu(1-F)}k\left(p_+p_-\right)^{k-1}\left(p'_+p_- + p_+p'_-\right)\right\}\\
&= \frac{\mu^{2k}}{k!k!}e^{-\mu(1-F)}\left(p_+p_-\right)^k\left[\mu f + k\frac{p'_+p_- + p_+p'_-}{p_+p_-}\right]\\
&= \psi_k\left(\mu f + kG\right) && (33)\\
\pi'_0 &= \sum_{k=0}^{\infty}\psi_k\left(\mu f + kG\right) && (34)\\
\pi'_1 &= \sum_{k=0}^{\infty}\frac{\mu}{k+1}\left(\psi'_kp_+ + \psi_kp'_+\right) && (35)\\
\pi'_{-1} &= \sum_{k=0}^{\infty}\frac{\mu}{k+1}\left(\psi'_kp_- + \psi_kp'_-\right) && (36)\\
P' &= \frac{1}{2}\left(\pi'_0 + \pi'_{-1}\right) && (37)\\
\tilde{P}' &= \frac{1}{2}\left(\pi'_0 + \pi'_1\right) && (38)
\end{aligned}
$$

where $G = \frac{p'_+}{p_+} + \frac{p'_-}{p_-}$ in (33) and where, for notational convenience, I suppress the argument T (e.g. writing f instead of $f(T)$). Note that $p'_+ \leq p'_- \leq 0$ and therefore that $G \leq 0$.

It is also useful, remembering that $\pi_1 = \frac{p_+}{p_-}\pi_0$, to define the ratio γ as follows, so that $\pi_1 = \pi_0\gamma p_+$ and $\pi_{-1} = \pi_0\gamma p_-$.

$$\gamma \equiv \frac{1}{p_+}\frac{\pi_1}{\pi_0} = \frac{1}{p_-}\frac{\pi_{-1}}{\pi_0} \qquad (39)$$

The derivative of γ with respect to T is

$$\gamma' = \frac{A_1}{\left(p_+\pi_0\right)^2} \qquad (40)$$

where A_1 is defined as follows.

$$\begin{aligned}
A_1 &\equiv \pi_1'\left(p_+\pi_0\right) - \pi_1\left(p_+'\pi_0 + p_+\pi_0'\right) \\
&= p_+\pi_0\pi_1' - p_+'\pi_0\pi_1 - p_+\pi_0'\pi_1 \\
&= p_+\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\frac{\mu}{k+1}\left(\psi_k'p_+ + \psi_k p_+'\right) - p_+'\sum_{j=0}^{\infty}\psi_j\sum_{k=0}^{\infty}\psi_k\frac{\mu p_+}{k+1} - p_+\sum_{j=0}^{\infty}\psi_j'\sum_{k=0}^{\infty}\psi_k\frac{p_+\mu}{k+1} \\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\left(\mu f + kG\right)\frac{p_+^2\mu}{k+1} - \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\left(\mu f + jG\right)\frac{p_+^2\mu}{k+1} \\
&= p_+^2\mu G\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_j\psi_k\left(k-j\right)\frac{1}{k+1} \qquad (41)
\end{aligned}$$

Without the fraction $\frac{1}{k+1}$, the double sum in (41) would equal zero: though $\psi_j\psi_k\left(k-j\right)$ 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 be positive.

Writing $T_{BR}\left(T\right) \equiv \frac{\tilde{P}}{P+\tilde{P}}$ in terms of γ gives the following:

$$\begin{aligned}
T_{BR}\left(T\right) &= \frac{\frac{1}{2}\left(\pi_0 + \pi_0\gamma p_+\right)}{\frac{1}{2}\left(\pi_0 + \pi_0\gamma p_-\right) + \frac{1}{2}\left(\pi_0 + \pi_0\gamma p_+\right)} \\
&= \frac{1+\gamma p_+}{2+\gamma p_- + \gamma p_+} \qquad (42)
\end{aligned}$$

Differentiating (42) yields the following, by the quotient rule:

$$T_{BR}'\left(T\right) = \frac{A_2}{\left(2+\gamma p_- + \gamma p_+\right)^2}$$

where

$$\begin{aligned}
A_2 &= \left(\gamma' p_+ + \gamma p_+'\right)\left(2+\gamma p_- + \gamma p_+\right) - \left(1+\gamma p_+\right)\left(\gamma' p_+ + \gamma p_+' + \gamma' p_- + \gamma p_-'\right) \\
&= \left(\gamma' p_+ + \gamma p_+'\right)\left(1+\gamma p_-\right) - \left(1+\gamma p_+\right)\left(\gamma' p_- + \gamma p_-'\right) \\
&= \gamma p_+'\left(1+\gamma p_-\right) - \left(1+\gamma p_+\right)\gamma p_-' + \gamma'\left(p_+ - p_-\right) \\
&= -\gamma T f\frac{2}{\pi_0}P + \gamma\left(1-T\right)f\frac{2}{\pi_0}\tilde{P} + \gamma'\left(p_+ - p_-\right) \\
&= \gamma f\frac{2}{\pi_0}\left(P+\tilde{P}\right)\left(\frac{\tilde{P}}{P+\tilde{P}} - T\right) + \gamma'\left(p_+ - p_-\right) \qquad (43)
\end{aligned}$$

The second term of the sum in (43) is positive since γ' is positive; when $\frac{\tilde{P}}{P+\tilde{P}} \geq T$, the first term is positive as well, so $T_{BR}'\left(T\right) > 0$. ■

Lemma A2 *Let* σ_{T_A,T_B}, σ_{T_A,T_A}, *and* σ_{T_B,T_B} *be IPC strategy profiles, with participation 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* $\alpha_B > T_{BR}(T_A)$ *and* $\alpha_A < T_{BR}(T_B)$
2. *If* $T_A < T_B$ *then* $\alpha_B < T_{BR}(T_A)$ *and* $\alpha_A > T_{BR}(T_B)$

Proof. Since σ_{T_A,T_B} is IPC, the probabilities (2) and (3) of voting correctly and incorrectly in each state simplify to the following:

$$p_{AA} = \int_{T_A}^{1} q dF(q) \tag{44}$$

$$p_{AB} = \int_{T_A}^{1} (1-q) dF(q) \tag{45}$$

$$p_{BA} = \int_{T_B}^{1} (1-q) dF(q) \tag{46}$$

$$p_{BB} = \int_{T_B}^{1} q dF(q) \tag{47}$$

If $T_A > T_B$, it is straightforward to verify that inequalities (26) through (28) hold, as does the following.[22]

$$p_{AB} + p_{BB} > p_{AA} + p_{BA} \tag{48}$$

In contrast, consider the vote probabilities associated with the signal-symmetric profile σ_{T_A,T_A}: p_{AA} and p_{BB} are both equivalent to (44), and p_{AB} and p_{BA} are both equivalent to (45). Similarly, for σ_{T_B,T_B} both p_{AA} and p_{BB} are equivalent to (47) and both p_{AB} and p_{BA} are equivalent to (46). Using these, I now compare the best response cutpoint $\alpha_B = \frac{\tilde{P}_A}{\tilde{P}_A + P_B}$ under σ_{T_A,T_B} and σ_{T_A,T_A}. An equivalent condition to $\alpha_B > T_{BR}(T_A)$ is that $\tilde{P}_A(\sigma_{T_A,T_B}) P_B(\sigma_{T_A,T_A}) > \tilde{P}_A(\sigma_{T_A,T_A}) P_B(\sigma_{T_A,T_B})$, which I show here must be the case.

[22] Inequality (48) results from the following:

$$\begin{aligned} p_{AB} + p_{BB} - p_{AA} - p_{BA} &= \int_{q_A}^{1} (1-q) dF(q) + \int_{q_B}^{1} q dF(q) - \int_{q_A}^{1} q dF(q) - \int_{q_B}^{1} (1-q) dF(q) \\ &= \int_{q_B}^{q_A} q dF(q) - \int_{q_B}^{q_A} (1-q) dF(q) \\ &> 0 \end{aligned}$$

Consider the sign of the following difference.

$$
\begin{aligned}
& \tilde{P}_A\left(\sigma_{T_A,T_B}\right) P_B\left(\sigma_{T_A,T_A}\right) - \tilde{P}_A\left(\sigma_{T_A,T_A}\right) P_B\left(\sigma_{T_A,T_B}\right) && (49) \\
= \; & \frac{1}{4}\sum_{k=0}^{\infty}\frac{\mu^{2k}}{k!k!}e^{-\mu(p_{BA}+p_{AA})}\left(p_{BA}p_{AA}\right)^k\left[1+\frac{\mu}{k+1}p_{AA}\right]\times \\
& \sum_{j=0}^{\infty}\frac{\mu^{2j}}{j!j!}e^{-\mu(p_{AB}+p_{AA})}\left(p_{AB}p_{AA}\right)^j\left[1+\frac{\mu}{j+1}p_{AB}\right] \\
& -\frac{1}{4}\sum_{k=0}^{\infty}\frac{\mu^{2k}}{k!k!}e^{-\mu(p_{AB}+p_{AA})}\left(p_{AB}p_{AA}\right)^k\left[1+\frac{\mu}{k+1}p_{AA}\right]\times \\
& \sum_{j=0}^{\infty}\frac{\mu^{2j}}{j!j!}e^{-\mu(p_{BB}+p_{AB})}\left(p_{BB}p_{AB}\right)^j\left[1+\frac{\mu}{j+1}p_{AB}\right] \\
= \; & \frac{1}{4}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{\mu^{2k}}{k!k!}\frac{\mu^{2j}}{j!j!}e^{-\mu(p_{AB}+2p_{AA}+p_{BA})}\left(p_{BA}p_{AA}\right)^k\left(p_{AA}p_{AB}\right)^j\times \\
& \left(1+\frac{\mu}{k+1}p_{AA}\right)\left(1+\frac{\mu}{j+1}p_{AB}\right)\left[1-e^{-\mu(p_{AB}+p_{BB}-p_{AA}-p_{BA})}\left(\frac{p_{AB}}{p_{BA}}\right)^k\left(\frac{p_{BB}}{p_{AA}}\right)^j\right] \\
> \; & \frac{1}{4}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{\mu^{2k}}{k!k!}\frac{\mu^{2j}}{j!j!}e^{-\mu(p_{AB}+2p_{AA}+p_{BA})}\left(p_{BA}p_{AA}\right)^k\left(p_{AA}p_{AB}\right)^j\times && (50) \\
& \left(1+\frac{\mu}{k+1}p_{AA}\right)\left(1+\frac{\mu}{j+1}p_{AB}\right)\left[1-\left(\frac{p_{AB}}{p_{BA}}\right)^k\left(\frac{p_{BB}}{p_{AA}}\right)^j\right] && (51)
\end{aligned}
$$

The inequality in (50) follows from (48). Whenever $j \leq k$, inequalities (27) and (28) imply that the final product in (51) is strictly less than one, so that the bracketed difference is positive.[23] 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. $\left(p_{BA}p_{AA}\right)^j\left(p_{AA}p_{AB}\right)^k$ instead of $\left(p_{BA}p_{AA}\right)^k\left(p_{AA}p_{AB}\right)^j$). Thus the sign of (49) is positive.

A symmetric derivation reveals that $\tilde{P}_B\left(\sigma_{T_A,T_B}\right) P_A\left(\sigma_{T_B,T_B}\right) < \tilde{P}_B\left(\sigma_{T_B,T_B}\right) P_A\left(\sigma_{T_A,T_B}\right)$ or, equivalently, that $\alpha_A < T_{BR}\left(T_B\right)$, establishing part (i). Part (ii) follows from identical reasoning, for the case in which $T_A < T_B$. ■

Theorem 4 *If σ^* is a Bayesian equilibrium then it is SIPC.*

Proof. Together, Theorems 1 and 2 imply that σ^* must be IPC, with participation thresholds $T_A = \alpha_A$ and $T_B = \alpha_B$. To show further that it must be signal-symmetric, suppose

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

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

by way of contradiction that $T_A > T_B$ (a symmetric argument will apply to the case of $T_A < T_B$). The logic of this proof is to compare the best response thresholds to $\sigma^* = \sigma_{T_A,T_B}$ with the best responses to the closely-related signal-symmetric strategies σ_{T_A,T_A} and σ_{T_B,T_B}. There are three relevant cases to consider:

Case 1: $T_{BR}(T_B) \leq T_B$. By Lemma A2, $\alpha_A < T_{BR}(T_B) \leq T_B < T_A$, so σ_{T_A,T_B} is not an equilibrium.

If $T_{BR}(T_B) > T_B$ then, since $T(\sigma_{1,1}) = \frac{1}{2}$, there exists (by the Intermediate Value Theorem) some SIPC equilibrium with a participation threshold strictly above T_B. 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. By Lemma A1, T_{BR} is increasing between T_B and T^*.

Case 2: $T_{BR}(T_B) > T_B$ and $T_A \geq T^*$. Since T_{BR} is increasing between T_B and T^*, clearly $T_{BR}(T_B) < T^*$. By Lemma A2, this implies that $\alpha_A < T_{BR}(T_B) < T^* \leq T_A$, so σ_{T_A,T_B} is not an equilibrium.

Case 3: $T_{BR}(T_B) > T_B$ and $T_A < T^*$. Since $T_A \in [T_B, T^*]$, an interval in which T_{BR} is increasing, $T_{BR}(T_A) > T_{BR}(T_B)$. Lemma A2 then implies that $\alpha_B > T_{BR}(T_A) > T_{BR}(T_B) > T_B$, again ensuring that σ_{T_A,T_B} is not an equilibrium. ■

Theorem 5 *For any $T \in \left[\frac{1}{2}, 1\right]$, the best response threshold $T^{\mu}_{BR}(T)$ 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 (7) 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:

$$\begin{aligned}
\frac{\partial \psi_k}{\partial \mu} &= \frac{(p_+ p_-)^k}{k!k!}\left[2k\mu^{2k-1}e^{-\mu(1-F)} - (1-F)\,\mu^{2k}e^{-\mu(1-F)}\right] \\
&= \psi_k \left[\frac{2k}{\mu} - (1-F)\right] \qquad (52) \\
\frac{\partial \pi_0}{\partial \mu} &= \sum_{k=0}^{\infty} \psi_k \left[\frac{2k}{\mu} - (1-F)\right] \qquad (53) \\
\frac{\partial \pi_1}{\partial \mu} &= \sum_{k=0}^{\infty} \left(\frac{\partial \psi_k}{\partial \mu}\frac{\mu p_+}{k+1} + \psi_k \frac{p_+}{k+1}\right) \\
&= \sum_{k=0}^{\infty} \psi_k \left\{\left[\frac{2k}{\mu} - (1-F)\right]\frac{\mu p_+}{k+1} + \frac{p_+}{k+1}\right\} \qquad (54) \\
\frac{\partial \pi_{-1}}{\partial \mu} &= \sum_{k=0}^{\infty} \left(\frac{\partial \psi_k}{\partial \mu}\frac{\mu p_-}{k+1} + \psi_k \frac{p_-}{k+1}\right) \\
&= \sum_{k=0}^{\infty} \psi_k \left\{\left[\frac{2k}{\mu} - (1-F)\right]\frac{\mu p_-}{k+1} + \frac{p_-}{k+1}\right\} \qquad (55)
\end{aligned}$$

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

$$\frac{\partial\left(\tilde{P}/P\right)}{\partial \mu} = \frac{1}{P^2}\left(P\frac{\partial \tilde{P}}{\partial \mu} - \tilde{P}\frac{\partial P}{\partial \mu}\right) \qquad (56)$$

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

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

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

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

The parenthesis term $P\frac{\partial\tilde{P}}{\partial\mu}-\tilde{P}\frac{\partial P}{\partial\mu}$ from (56) 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 $\frac{\partial}{\partial\mu}\left(\frac{\tilde{P}}{P}\right)>0$, and therefore $\frac{\partial}{\partial\mu}\left(\frac{\tilde{P}}{P+\tilde{P}}\right)>0$, which is equivalent to the desired result. ■

Lemma 1 *For $T\in\left[\frac{1}{2},1\right)$, $\lim_{\mu\to\infty}T_{BR}^{\mu}=L(T)$, as defined as in (21).*

Proof. For any $T\in\left[\frac{1}{2},1\right)$, Myerson (2000) derives the limiting ratio of pivot probabilities as merely the ratio of the square roots of the expected numbers of agents taking either action:

$$\lim_{\mu\to\infty}\frac{\tilde{P}}{P}=\frac{\sqrt{p_+}}{\sqrt{p_-}}$$

From this, it follows immediately that

$$\begin{aligned}
\lim_{\mu\to\infty}\frac{\tilde{P}}{P+\tilde{P}} &= \frac{\sqrt{p_+}}{\sqrt{p_+}+\sqrt{p_-}} \\
&= \frac{\sqrt{M(T)}}{\sqrt{M(T)}+\sqrt{1-M(T)}}
\end{aligned}$$

where $M(T)\equiv E(Q|Q\geq T)=\frac{p_+}{p_++p_-}$. ■

Theorem 6 *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 every μ_k, $T_{\mu_k}^*$ must be a fixed point of $T_{BR}^{\mu_k}$. Since $T_{BR}^{\mu_k}$ converges pointwise to L, a limit point of $T_{\mu_k}^*$ must therefore be a fixed point of L. Solving (21) for $M(T)$, it can easily be shown that $L(T)$ is greater than, equal to, or less than T if and only if $M(T)$ is (respectively) greater than, equal to, or less than $\Gamma(T)\equiv\frac{T^2}{T^2+(1-T)^2}$, which strictly increases from $\Gamma\left(\frac{1}{2}\right)=\frac{1}{2}$ to $\Gamma(1)=1$, and is strictly concave, as illustrated in Figure 4. The slope of Γ approaches zero as $T\to 1$:

$$\begin{aligned}
\lim_{T\to 1}\Gamma'(T) &= \lim_{T\to 1}\frac{2T\left[T^2+(1-T)^2\right]-T^2\left[2T-2(1-T)\right]}{\left[T^2+(1-T)^2\right]^2} \qquad (57)\\
&= \frac{2(1+0)-1(2-0)}{(1+0)^2}=0
\end{aligned}$$

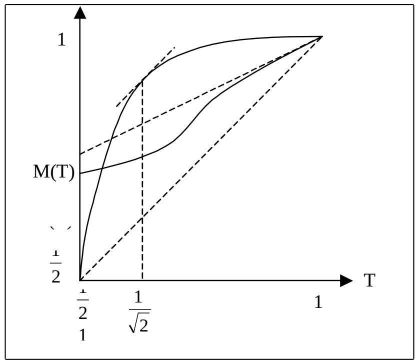

Figure 4: T=L(T) is a fixed point of the limiting best response function if and only if M(T)=Gamma(T)

Like Γ, M also increases to $M(1) = 1$. The slope of M, however, approaches $\frac{1}{2}$. This can be seen as follows using the quotient rule, remembering that $\frac{dp_+}{dT} = -Tf$, as in equation (31) in the proof of Proposition A1. Though I suppress the notation, p_+, f, F, M, and the *hazard function*, or failure rate $h = \frac{f}{1-F}$, are evaluated at T.

$$\begin{aligned} M'(T) &= \frac{-Tf(1-F) + p_+ f}{(1-F)^2} \\ &= \frac{f}{1-F}\left(\frac{p_+}{1-F} - T\right) \\ &= h\left[M(T) - T\right] \end{aligned} \tag{58}$$

Case 1: $\lim_{T\to 1} f(T) < \infty$. As $T \to 1$, the difference $M(T) - T$ approaches zero. The condition that $\lim_{T\to 1} f(T) < \infty$ implies that $f(M - T)$ also converges to zero, so by L'Hospital's rule

$$\begin{aligned} \lim_{T\to 1} M'(T) &= \lim_{T\to 1} \frac{f'\left[M(T) - T\right] + f\left[M'(T) - 1\right]}{-f} \\ &= \lim_{T\to 1} \left\{-\frac{f'}{f}\left[M(T) - T\right] - M'(T) + 1\right\} \end{aligned}$$

This equation is satisfied if and only if

$$\lim_{T\to 1} M'(T) = \frac{1}{2} - \lim_{T\to 1} \frac{f'}{f}\left[M(T) - T\right]$$

Since f is a density function (and must integrate to one) it must be that $\lim_{T\to 1} \frac{f'}{f} < \infty$, and so the second term of this difference goes to zero, leaving $\lim_{T\to 1} M'(T) = \frac{1}{2}$.

Case 2: $\lim_{T\to 1} f(T) = \infty$. By the quotient rule, the derivative h' of the hazard function is as follows:

$$\begin{aligned} h' &= \frac{f'(1-F)+f^2}{(1-F)^2} \\ &= \frac{f}{1-F}\left(\frac{f'}{f}+\frac{f}{1-F}\right) \\ &= h\left(\frac{f'}{f}+h\right) \end{aligned} \tag{59}$$

Since f is a density function $\lim_{T\to1}\frac{f'}{f}<\infty$, and the condition that $\lim_{T\to1}f(T)=\infty$ implies that $\frac{1}{h}\to0$, so the following must be true:

$$\begin{aligned} \lim_{T\to1}\frac{h'}{h^2} &= \lim_{T\to1}\frac{h\frac{f'}{f}+h^2}{h^2} \\ &= \lim_{T\to1}\frac{f'}{f}\frac{1}{h}+1 \\ &= 1 \end{aligned}$$

Therefore, applying L'Hospital's rule to equation (58) yields

$$\begin{aligned} \lim_{T\to1}M'(T) &= \lim_{T\to1}\frac{M'(T)-1}{-h'/h^2} \\ &= 1-\lim_{T\to1}M'(T) \end{aligned}$$

which implies $\lim_{T\to1}M'(T)=\frac{1}{2}$.

Thus $\Gamma'(T)\to0$ but $M'(T)\to\frac{1}{2}$; for T sufficiently close to one, therefore, $M(T)<\Gamma(T)$ or, equivalently, $L(T)<T$. If it were the case that $T^*_{\mu_k}\to1$, there would be a k sufficiently high that $T^*_{\mu_k}>L\left(T^*_{\mu_k}\right)$ or, equivalently, that $T^{\mu_k}_{BR}\left(T^*_{\mu_k}\right)>L\left(T^*_{\mu_k}\right)$. This cannot be, however, since L is an upper bound on T^{μ}_{BR}. ■

Theorem 7 *If f is log-concave then L has a unique fixed point between $\frac{1}{2}$ and 1.*

Proof. Existence of a fixed point between $\frac{1}{2}$ and 1 follows from the Intermediate Value Theorem, since L is continuous and $L\left(\frac{1}{2}\right)>\frac{1}{2}$ but, as shown in the proof of Theorem 6, $M'(T)>0$ for T sufficiently close to 1, which implies that $L(T)<T$. To see uniqueness, it is useful first to review two useful implications of the log-concavity of f, demonstrated by Bagnoli and Bergstrom (2005):

- (P1) $\frac{f'}{f}$ is monotonically decreasing.
- (P2) The *hazard* function $h=\frac{f}{1-F}$ is monotonically increasing.

- (P3) The *mean residual lifetime* function $\int_T^1 (q-T)\,dF(q) = M(T) - T$ is monotonically decreasing or, equivalently, $M'(T) < 1$.

P1 implies that $\lim_{T\to 1} \frac{f'}{f} < \infty$, and guarantees the existence of a maximizer $\hat{T}$, such that f is increasing on $\left[\frac{1}{2}, \hat{T}\right]$ and decreasing on $\left[\hat{T}, 1\right]$. In what follows, Step 1 shows that if M intersects Γ in the interval $\left[\hat{T}, 1\right)$ then there can be no other intersection point in $\left[\frac{1}{2}, 1\right)$, and Step 2 shows that if M does not intersect Γ in $\left[\hat{T}, 1\right)$ there will be a unique intersection point in $\left[\frac{1}{2}, \hat{T}\right]$.

Step 1. For this step, I first show that $M'(T) \geq \frac{1}{2}$ when $T > \hat{T}$. To see this, first differentiate (58) to obtain the second derivative of M:

$$\begin{aligned} M''(T) &= h'(M-T) + h(M'-1) \\ &= h\left(\frac{f'}{f} + h\right)(M-T) + h(M'-1) \\ &= M'\left(\frac{f'}{f} + h\right) + h(M'-1) \\ &= \frac{f'}{f}M' + 2h\left(M' - \frac{1}{2}\right) \end{aligned} \qquad (60)$$

where the second equality follows from (59) and the third follows from (58). M is convex if and only if (60) is positive or, equivalently, if the following inequality holds:

$$M'(T) \geq \frac{1}{2 + \frac{1}{h}\frac{f'}{f}} \qquad (61)$$

For $T \geq \hat{T}$, the right hand side of (61) is greater than $\frac{1}{2}$ because f' is negative. If it were the case for $\varepsilon > 0$ at any $T \geq \hat{T}$ that $M'(T) = \frac{1}{2} - \varepsilon$, therefore, then M must be concave, implying that $M'(T)$ decreases further as T increases, so that $\lim_{T\to 1} M'(T)$ is bounded above by $\frac{1}{2} - \varepsilon$. This contradicts the result from the proof of Theorem 6, however, that $\lim_{T\to 1} M'(T) = \frac{1}{2}$; thus, for $T \geq \hat{T}$ it must be that $M'(T) \geq \frac{1}{2}$.

The importance of the result that $M'(T) \geq \frac{1}{2}$ is illustrated in Figure 4: M must lie between the dotted lines of slopes 1 and $\frac{1}{2}$, and therefore cannot intersect Γ to the right of

$\frac{1}{\sqrt{2}}$ (which is the point at which Γ intersects the dotted line with slope $\frac{1}{2}$).[24] To the left of $\frac{1}{\sqrt{2}}$, however, the slope of Γ exceeds 1; since log-concavity (P3) implies $M'(T) < 1$ for all T, M and Γ can intersect only once in this region. Thus, if M intersects Γ at a point $T^* > \hat{T}$ then $T^* < \frac{1}{\sqrt{2}}$, and T^* must be the unique intersection point of M and Γ, and therefore the unique fixed point of L.

Step 2. In the interval $\left[\frac{1}{2}, \hat{T}\right]$, f' must be positive, which implies that the right hand side of (61) is less than $\frac{1}{2}$. M must be convex in this interval, because if (61) fails then $M'(T) = \frac{1}{2} - \varepsilon$ for some $\varepsilon > 0$. This cannot be, however, since then as T increases concavity would imply that M' decreases, while (P1) and (P2) together imply that the right hand side of (61) increases. Thus M would continue concave until T reaches 1, and M' would be bounded above by $\frac{1}{2} - \varepsilon$, contradicting the result from the proof of Theorem 6, that $\lim_{T\to 1} M'(T) = \frac{1}{2}$. Since M is convex on $\left[\frac{1}{2}, \hat{T}\right]$ and Γ is strictly concave, there can be only one intersection point in this interval. ■

Lemma 2 *Let F and G be continuous, log-concave distributions with strictly positive densities, and let T_∞^F and T_∞^G and M^F and M^G denote the unique limiting participation thresholds and mean quality functions for F and G, respectively. Then $T_\infty^G > T_\infty^F$ if and only if $M^G(T) > M^F(T)$ or, equivalently,*

$$\frac{\bar{F}(T)}{\bar{G}(T)} > \frac{\int_T^1 \bar{F}(q)\,dq}{\int_T^1 \bar{G}(q)\,dq} \tag{62}$$

for $T \in \left\{T_\infty^F, T_\infty^G\right\}$.

Proof. The fixed points T_∞^F and T_∞^G of L^F and L^G are such that $L^F(T) > T$ and $L^G(T) > T$ for any threshold T below T_∞^F and T_∞^G, respectively, and similarly $L^F(T) < T$ and $L^G(T) < T$ for any T above T_∞^F and T_∞^G. Clearly, then, $T_\infty^G > T_\infty^F$ if and only if $L^G(T) > L^F(T)$ for both fixed points $T \in \left\{T_\infty^F, T_\infty^G\right\}$. As is made clear by equation (21), $L^G(T) > L^F(T)$ if and

[24] This point is determined algebraically as follows:

$$\begin{aligned}
\frac{T^2}{T^2 + (1-T)^2} &= \frac{1}{2}(1+T) \\
2T^2 &= T^2(1+T) + (1-T)^2(1+T) \\
T^2(1-T) &= (1-T)\left(1-T^2\right) \\
T^2 &= 1 - T^2 \\
T^2 &= \frac{1}{2}
\end{aligned}$$

only if $M^G(T) > M^F(T)$. Rewriting this final expression and integrating by parts yields (22). For example,

$$\begin{aligned} M^F(T) &= \frac{1}{\bar{F}(T)} \int_T^1 q dF(q) \\ &= \frac{1}{\bar{F}(T)} \int_T^1 \bar{F}(q)\, dq \end{aligned}$$

■

Theorem 8 *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 $\bar{G}(T_\infty^G) = \bar{F}(T_\infty^F)$.

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

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

Proof. 1. If $G(q) = F(q)$ for all $q \geq T_\infty^F$ then the left and right hand sides of (22) are both equal to one, and it follows immediately that $T_\infty^G = T_\infty^F$ and therefore $\bar{G}(T_\infty^G) = \bar{F}(T_\infty^F)$.

2. $G \geq_1 F$ implies that the right hand side of (22) is strictly less than one, but that the left hand side is equal to one since $\bar{G}(T_\infty^F) = \bar{F}(T_\infty^F)$. This implies that $T_\infty^G \geq T_\infty^F$, and therefore $G(T_\infty^G) \geq G(T_\infty^F) = F(T_\infty^F)$ or, equivalently, $\bar{G}(T_\infty^G) \leq \bar{F}(T_\infty^F)$.

3. The following derivation shows also that $M^G(T_\infty^F) < M^F(T_\infty^F)$.

$$\begin{aligned} \bar{G}(T_\infty^F) M^G(T_\infty^F) &= \int_{T_\infty^F}^1 q g(q)\, dq \\ &= \int_{T_\infty^F}^1 q f(q)\, dq + \int_{T_\infty^F}^{M^F(T_\infty^F)} q[g(q) - f(q)]\, dq \\ &< \int_{T_\infty^F}^1 q f(q)\, dq \\ &\quad + M^F(T_\infty^F) \int_{T_\infty^F}^{M^F(T_\infty^F)} [g(q) - f(q)]\, dq \\ &= M^F(T_\infty^F)\left[1 - F(T_\infty^F)\right] + M^F(T_\infty^F)\left[G(M^F) - F(M^F) - G(T_\infty^F) + F(T_\infty^F)\right] \\ &= \bar{G}(T_\infty^F) M^F(T_\infty^F) \end{aligned}$$

By Lemma 2, $M^G(T_\infty^F) < M^F(T_\infty^F)$ implies $T_\infty^G < T_\infty^F$, and so $\bar{F}(T_\infty^G) \geq \bar{F}(T_\infty^F)$. Since $G \geq_1 F$, $\bar{G}(T_\infty^G) \geq \bar{F}(T_\infty^G) \geq \bar{F}(T_\infty^F)$. ■

Theorem 9 *Let F and G be continuous, log-concave distributions with strictly positive densities and a common mean m, such that $G \geq_2 F$ and $T_\infty^F \leq m$. Then $T_\infty^G \leq T_\infty^F$ and $\bar{G}(T_\infty^G) \geq \bar{F}(T_\infty^F)$.*

Proof. Since F and G have a common mean, $G \geq_2 F$ implies that the variance of Q is smaller under G than under F. With $T_\infty^F \leq m$, it must therefore be that $\bar{G}\left(T_\infty^F\right) \geq \bar{F}\left(T_\infty^F\right)$, and so the left hand side of (22) is less than one for $T = T_\infty^F$. The right-hand side of (22) can be rewritten as

$$\frac{\int_T^1 \bar{F}\left(q\right) dq}{\int_T^1 \bar{G}\left(q\right) dq} = \frac{\int_{1/2}^1 \bar{F}\left(q\right) dq - \int_{1/2}^T \bar{F}\left(q\right) dq}{\int_{1/2}^1 \bar{G}\left(q\right) dq - \int_{1/2}^T \bar{G}\left(q\right) dq}$$

where the numerator and denominator contain equivalent expressions for the common mean $\int_{1/2}^1 \bar{F}\left(q\right) dq = \int_{1/2}^1 \bar{G}\left(q\right) dq = m$. By definition, $G \geq_2 F$ implies $\int_{1/2}^T \bar{G}\left(q\right) dq \geq \int_{1/2}^T \bar{F}\left(q\right) dq$; since the left-hand side is less than one and the right-hand side is greater than one, the inequality in (22) does not hold, implying that $T_\infty^G \leq T_\infty^F$. This implies that $\bar{G}\left(T_\infty^G\right) \geq \bar{G}\left(T_\infty^F\right)$ and, together with the fact that $\bar{G}\left(T_\infty^F\right) \geq \bar{F}\left(T_\infty^F\right)$, implies $\bar{G}\left(T_\infty^G\right) \geq \bar{F}\left(T_\infty^F\right)$. ■