Discussion Paper No. 1195

Abstention in Elections with Asymmetric Information and Diverse Preferences[1]

by
Timothy J. Feddersen[2]
and
Wolfgang Pesendorfer[3]

First Version: October 1996
This Version: October 1997

[1] Formerly, "Two Alternative Elections with Private and Common Values."

[2] J.L. Kellogg Graduate School of Management, Department of Managerial Economics and Decision Sciences, Northwestern University. (e-mail: tfed@nwu.edu)
[3] Department of Economics, Princeton University. (e-mail: pesendor@phoenix.princeton.edu)

Abstract

We analyze a model of a two-candidate election in which voters have asymmetric information and diverse preferences. Voters may costlessly choose to either vote for one of the candidates or abstain. We demonstrate that a strictly positive fraction of the electorate will abstain and, nevertheless, elections effectively aggregate voter's private information. The model also provides an explanation for observed patterns of participation and partisanship.

1 Introduction

Over the last forty years the fraction of the electorate that participates in elections has declined significantly in the United States.[1] The decline in participation has been accompanied by a decrease in partisan attachment raising concerns about the degree to which political outcomes actually enjoy the support of a majority of citizens.[2] These concerns are heightened when it is observed that the decline in participation is particularly pronounced among the less educated and less wealthy.[3] The possibility that the voting electorate is not representative leads naturally to concerns that election outcomes are biased towards the wealthy and better educated.[4]

It is difficult to evaluate the effects of a decline in participation on electoral performance without a model that explains what factors lead to changes in patterns of abstention in the first place. Explaining participation in elections has proven to be a particularly difficult problem for positive political theorists.[5] The cross-sectional comparative static that education level is positively correlated with turnout has been well established empirically.[6] A decision theoretic model that assumes it is less costly for those with better information to participate would seem to be sufficient to explain this cross-sectional variation.[7] However, declining turnout occurs at the same time as education levels have increased. A decision theoretic model will have trouble simultaneously explaining the longitudinal phenomena of declining overall turnout while membership in high turnout categories grows.[8]

In this paper we develop a game-theoretic model that can simultaneously explain the observed cross-sectional and longitudinal variations in participation and partisanship as a function of the quality and distribution of information. The model

[1] See Lijphart (1997) for a review of the literature on turnout.

[2] Niemi and Weisberg (1984 page 483).

[3] Abramson, Aldrich and Rohde (1991, page 102).

[4] Concerns about the possibility of biased outcomes has led to calls for reform e.g., Lijphart (1997) argues in favor of compulsory voting as a solution to the problem of a biased electorate.

[5] Indeed, Morris Fiorina wrote: "turnout is the paradox that ate rational choice theory." This quote was used as a jumping off point by Bernard Grofman (1995) in an essay on rational choice models of participation. Grofman properly observed that the test of a good model is not in providing good point prediction but in the usefulness of the model's comparative statics. He concluded that rational choice models centered on costs to vote do pretty well at predicting changes in participation.

[6] Wolfinger and Rosenstone 1980.

[7] See Matsusaka (1992) for an example of such a theory.

[8] Of course, voting may have become more costly over this period. However, it seems more likely that barriers to voting have been relaxed over this period.

presented here is an extension of a model developed by Feddersen and Pesendorfer (1996), hereafter referred to as FP. FP examine a model in which there are two candidates (candidate 1 and 2) and two states of the world (state 1 and state 2). There are three types of voters: voters who always strictly prefer candidate 1; voters who always strictly prefer 2; and those who strictly prefer candidate 1 in state 1 and candidate 2 in state 2. FP refer to the first two categories of voters as partisans and the latter group as independents. All of the independents are assumed to have identical preferences i.e., if the probability of state 2 is above some threshold all prefer candidate 2. The set of independent voters is further partitioned into those who know the true state and those that are uninformed but share a common knowledge prior. Voting is costless and each voter may vote for either candidate or abstain.

FP showed that equilibrium behavior in large elections has the following features:

1. uninformed independents may suffer the *"swing voter's curse"*[9] and have a strict incentive to abstain even though they have a strict preference between the two candidates;

2. abstention levels may be high even in large electorates;

3. even with strategic abstention, elections satisfy *"full information equivalence"* i.e., the winning candidate is the candidate that would win an election in which all voters were fully informed and voted for their preferred candidate.

Our model generalizes the framework of FP in three ways. First, we introduce a continuum of voter preference types and, hence, in a typical electorate no pair of voters has exactly identical preferences. Preference diversity allows us to analyze participation when information and preferences are correlated. Second, we consider

[9]The following example illustrates the swing voter's curse. Suppose there are three voters who must choose between candidate 1 and 2 using majority rule. There are two states of the world: state 1 and state 2. State 1 is more likely than state 2. All of the voters prefer candidate 1 in state 1 and candidate 2 in state 2. All voters know that exactly one of them knows the true state. It is an equilibrium for those voters who do not know the state to abstain and to let the informed voter decide the outcome. To see why this is the case suppose that in the event the election is tied a coin is tossed to determine the winner. In the proposed equilibrium uninformed voters strictly prefer to abstain because their vote only changes the outcome when they vote for the wrong candidate–the candidate not supported by the perfectly informed voter–and the wrong candidate wins the coin toss.

environments with more than two states. Third, we consider the effects of noisy information.

As in FP we assume that voter preferences depend on a private preference type and on a state which represents the relative quality of the two candidates. For example, suppose that an election is held to decide whether a bridge should be built. Each voter's private preference type is determined by the frequency they would use the bridge. Those that would use the bridge more are more favorable towards building the bridge. The "state" in this example corresponds to the cost of the bridge. As the cost of the bridge increases all voters like the idea of building the bridge less.

Now suppose that voters know their preference type but are uncertain about the state i.e., voters know how frequently they will use the bridge but are unsure about how much the bridge costs. Each state thus determines a fraction of the electorate that prefers candidate 1 to candidate 2 e.g., when the cost of the bridge is low 75% prefer to build the bridge whereas if the cost is high 45% favor the bridge.

In the first part of the paper we present our model and analytical results. We show that the full-information equivalence result in FP is robust i.e., candidate 1 wins with high probability in states in which a majority of the electorate strictly prefers candidate 1 to candidate 2 and loses with high probability otherwise. The most important contribution of this section is to give both necessary and sufficient conditions for significant levels of abstention. We assume the set of states is finite and therefore there is a *critical state* in which the fraction of the electorate that prefers candidate 1 to candidate 2 is closest to 1/2. We show that a strictly positive fraction of the electorate abstains if the fraction of the electorate that prefers candidate 1 to candidate 2 is not equal to 1/2 in the critical state. On the other hand if the fraction that prefers candidate 1 in the critical state is very close to 1/2 then the fraction of the electorate that abstains is very close to zero.

The fact that abstention can occur under some model specifications but not others raises concerns about the swing voter's curse as an explanation for abstention. The concern is as follows: surely there is always some state of the world in which exactly half of any population would prefer one alternative to the other. Our results would then be taken as evidence that the swing voter's curse cannot provide an explanation for abstention. However, such an argument misunderstands what a state represents in our model. A state represents what a very large number of people may learn about the relative quality of the two alternatives. Thus, in the bridge example it may be the case that the true cost of the bridge may range continuously from \$50 to \$50,000,000 but that the most any very large number of people could ever learn about the true cost is that it is above or below \$25,000,000. In this case even though the true cost of the bridge is a continuous random variable it is appropriate to assume there are two states. A more fruitful test of the theory is to examine the comparative statics of the model in settings where significant levels of abstention are predicted.

In the second part of the paper we provide comparative statics results for a series of examples. First, we demonstrate in a simplified version of the model that an increase in the fraction of the electorate that is informed can lead to higher levels of abstention and lower levels of partisanship. Next, we provide examples with noisy signals to illustrate how those with better information may participate more frequently. Moreover, increasing the fraction of the electorate that receives a noisy signal can increase levels of abstention among *both* the informed and uninformed alike. Noisy signals also leads to an electorate that is biased towards the winning candidate. Finally, we illustrate how a highly biased distribution of information may result in an electorate that is dramatically skewed towards one side of the ideological spectrum and in an overall increased level of abstention. However, the biased electorate does not result in biased outcomes. Indeed, in this example the biased distribution of information actually results in a strictly higher probability an election will satisfy full information equivalence than an unbiased distribution of information.

1.1 Related work

There are a variety of papers that focus on the consequences of costs to vote on turnout. See Palfrey and Rosenthal (1983, 1985), Riker and Ordeshook (1968), Feddersen (1992), Morton (1991) and Lohmann (1993a,b). For reviews of both the theoretical and empirical literature on turnout see FP, Aldrich (1993), Grofman (1993) or Matsusaka (1992). Katz and Ghirardato (1997) develop a decision theoretic explanation of roll-off. Their explanation of abstention relies on voters having non-standard utility functions.

Our model assumes that the number of voters is uncertain and distributed according to a Poisson distribution. The idea of using Poisson distributions to analyze large anonymous games is borrowed from Myerson (1997a,b). The reader who is comparing our model with FP will note that FP assumed the number of voters was known. However, they effectively introduced population uncertainty by assuming a positive probability each voter would abstain. The advantage of the Poisson distribution is that it permits easy calculation of equilibrium profiles and examples even in very large electorates.

Feddersen and Pesendorfer (1997) examine voting in two candidate elections with private information and common values without abstention. They assume a continuous state space and demonstrate that large elections satisfy full information equivalence. In this paper we prove the same result for a finite state space. (Proposition 4). We also demonstrate here that if there is a state in which a fully informed electorate splits evenly between the two alternatives then there is no abstention in the limit. In Feddersen and Pesendorfer (1997) this is always the case and hence abstention cannot play a role in that model.

2 Model

We analyze a two alternative election. Alternatives are denoted by $j \in \{1,2\}$. A voter's utility depends on a preference parameter $x \in [-1,1] = X$, the chosen alternative j, and the state $s \in S \subset [0,1]$. We assume the set of states S is finite and that $\{0,1\} \subset S$, i.e., the smallest state is 0 and the largest state is 1.

We denote by $u(j,s,x)$ the utility function of voters. Let

$$v(s,x) \equiv u(2,s,x) - u(1,s,x) \tag{1}$$

denote the utility difference of a voter type x between alternative 2 and alternative 1 in state s.

At the beginning of the game nature selects a state s and an electorate. The electorate is chosen as follows. First, nature selects a number of players according to the Poisson distribution with parameter ν. Thus, the probability that there are n players is given by

$$\Pr(n) = \frac{e^{-\nu}\nu^n}{n!}$$

Second, every player is independently assigned a preference type according to the distribution function $F(x)$. Third, every player receives a signal $m \in \{\emptyset, 1, ..., M\} = M$, where $\emptyset$ describes an uninformative signal. We assume that conditional on s the signal is independently distributed across agents. The probability that an agent receives the signal m in state s is $p(m|s)$.

We make the following assumptions:

Assumption 1 *$v(x, s)$ is defined for all $(x, s) \in [-1, 1] \times [0, 1]$; it is continuous in x and strictly increasing in (x, s). Furthermore, $v(-1, s) < 0$ for all s and $v(1, s) > 0$ for all $s \in [0, 1]$.*

Assumption 2 *The state s is chosen according to the probability distribution $g(s)$ where $g(s) > 0$ for all $s \in S$.*

Assumption 3 *The distribution function $F(x)$ has a continuous density $f(x)$ and $f(x)$ is bounded away from zero on $[-1, 1]$.*

Assumption 4 *$p(\emptyset|s) = 1 - q$ for some $0 < q \leq 1$ and all s.*

Assumption 5 *(SMLRP) There is an $\alpha > 0$ such that $p(m|s) \geq \alpha$ for all $(s, m), m \neq \emptyset$. Moreover, if $s' > s$ and $m' > m$, $m, m' \neq \emptyset$, then $p(m|s)p(m'|s') > p(m|s')p(m'|s)$.*

Assumption 1 says that the payoff difference between the two alternatives is increasing both in the preference type and in the state. Thus, higher states make alternative A more attractive. Assumptions (2) and (3) are made for technical convenience. Assumption (4) implies that signal $\emptyset$ does not provide any information about the state s. Assumption (5) says that all other signals are informative and satisfy the strict monotone likelihood ratio property. Informally, SMLRP implies that higher signals are more likely in higher states.

3 Equilibrium

Each player chooses an action $a \in A = \{0, 1, 2\}$ where $a = 0$ denotes abstention, $a = 1$ denotes a vote for alternative 1 and $a = 2$ denotes a vote for alternative 2. By n_a we denote the number of agents who take action a. We assume that the election outcome is determined by majority rule. Alternative 1 is the winner if and only if $n_2 \leq n_1$. This implies that in case of a tie, 1 will be the winner of the election.[10]

A symmetric strategy profile is a measurable function $\sigma : [-1,1] \times M \to \Delta(A)$. The probability that a player with preference parameter x and signal m takes action a is $\sigma_a(x,m)$.

Given a symmetric strategy profile σ the ex-ante probability that a voter takes action a in state s is denoted by $t_a(s)$ where

$$t_a(s) = \sum_{m \in M} p(m|s) \int_{-1}^{1} \sigma_a(x,m) dF(x) \tag{2}$$

Given σ we also define the random variables $n_a(s)$ to be the number of agents who take action a in state s. Since, conditional on s, the probability of receiving any signal is independent across agents and preference types are chosen independently, it follows that $n_a(s), a \in A$, are Poisson random variables with the parameters

$$\nu_a(s) = \nu \cdot t_a(s)$$

We can also consider the random variables $n_j^{-i}(s)$ that describe the number of agents other than agent i who choose action a in state s. It follows from the properties of the Poisson distribution that $n_a(s)$ and $n_a^{-i}(s)$ have the same distribution. In the following we use the notation $n_a(s)$ and n_a both for the description of the whole population and for the description of the behavior of everybody else from the point of view of a particular agent i.

The results in this paper depend on the fact that a voter only influences the outcome of an election when a vote is pivotal. In a voting model with abstention there are two ways in which a vote can be pivotal: a voter can influence the election outcome if either the election is tied ($n_2 = n_1$) or if alternative 2 is ahead by one vote ($n_2 = n_1 + 1$). In the event that $n_2 = n_1 + 1$ voting for 2 and abstaining both lead to the outcome 2 and voting for 1 leads to the outcome 1. We label this event piv_1 and say that a vote is pivotal for candidate 1. In the event that $n_2 = n_1$ voting for 1 and abstaining both lead to the outcome 1 and voting for 2 leads to the outcome 2. We label this event piv_2. We label the event that a vote is pivotal (either piv_1 or piv_2) as piv.

[10] Note that this choice of tie breaking rule is made for technical convenience. All of our results would go through if in the case of a tie alternative 1 wins with probability $\pi \in (0,1)$.

Given state s and a symmetric strategy profile σ the probability a vote is pivotal for alternative 1 is:

$$\Pr\{piv_1|s\} = \sum_{x=0}^{\infty} \frac{e^{-\nu(t_1(s)+t_2(s))}(\nu t_2(s))^{x+1}(\nu t_1(s))^x}{x!(x+1)!}; \tag{3}$$

the probability a vote is pivotal for alternative 2 as:

$$\Pr\{piv_2|s\} = \sum_{x=0}^{\infty} \frac{e^{-\nu(t_1(s)+t_2(s))}(\nu t_2(s))^{x}(\nu t_1(s))^x}{x!x!}; \tag{4}$$

and the probability a vote is pivotal for either alternative as

$$\Pr\{piv|s\} = \Pr\{piv_1|s\} + \Pr\{piv_2|s\}. \tag{5}$$

Since voters have private information about a common value state variable they must form beliefs about the distribution over states conditional on the event a vote is pivotal and their private information. If the profile σ is such that $t_2(s) > 0$ for all s then it follows that $\Pr(piv_j|s) > 0$ for all s. Therefore, the probability distribution over states conditional on being pivotal for alternative $j \in \{1, 2\}$ is well defined and given by

$$\beta(s|piv_j) = \frac{\Pr(piv_j|s)g(s)}{\sum_{w \in S} \Pr(piv_j|w)g(w)}. \tag{6}$$

The probability distribution over states conditional on being pivotal and observing signal m is given by:

$$\beta(s|piv_j, m) = \frac{\Pr(piv_j|s)g(s)p(m|s)}{\sum_{w \in S} \Pr(piv_j|w)g(w)p(m|w)} = \frac{\beta(s|piv_j)p(m|s)}{\sum_{w \in S} \beta(w|piv_j)p(m|w)}. \tag{7}$$

A symmetric strategy profile σ is characterized by cutpoints if for every signal $m \in M$ there is a pair of cutpoints $x_m^1 \leq x_m^2$ such that any voter who receives message m chooses 1 whenever the preference type is smaller than x_m^1; chooses 2 whenever the preference type is larger than x_m^2; and abstains otherwise. We say the profile is characterized by ordered cutpoints when for any $j \in \{1, 2\}$ x_m^j is non-decreasing in m.

Definition 1 *A strategy σ is characterized by* ordered cutpoints *if there are cutpoints $\{(x_m^1, x_m^2)\}_{m \in M}$ with the following properties: for every $m \in M$ $x_m^1 \leq x_m^2$; for every $j \in \{1, 2\}$ $1 > x_1^j \geq .... \geq x_M^j > -1$; and $\sigma_1(x, m) = 1$ if $x < x_m^1$, $\sigma_2(x, m) = 1$ if $x > x_m^2$, and $\sigma_0(x, m) = 1$ if $x_m^1 < x < x_m^2$.*

Observe that if a strategy profile is characterized by ordered cutpoints then voters of type (x, m) with $x < x_M^1$ vote for candidate 1 regardless of their private signal whereas those for whom $x > x_1^2$ vote for candidate 2 regardless of their private signal. We say that voter types with $x \in (x_M^1, x_1^2)$ take *informative action.*

In the following we consider symmetric Nash equilibria. The first proposition demonstrates that every symmetric Nash equilibrium can be characterized by ordered cutpoints and that the probability a voter chooses alternative 1 is strictly decreasing in the state while the probability that a voter chooses 2 is strictly increasing in the state. Proposition 1 is a straightforward consequence of the Strict Monotone Likelihood Ratio Property (SMLRP) and Assumption 1.

Proposition 1 *Suppose Assumptions 1-4 hold. Then a voting equilibrium exists and any voting equilibrium can be characterized by ordered cutpoints.*

Proof. Existence is straightforward. For example, the existence proof in Myerson (1997) applies.

Since there is a strictly positive probability that any player is the only agent it follows that the probability a vote is pivotal for 2 is always strictly positive. Therefore, by Assumption 1 there is an $\varepsilon > 0$ such that all types $x \in [1-\varepsilon, 1]$ must vote for 2 independent of their signal m. By Assumption 3 the probability that a type is in the interval $[1-\varepsilon, 1]$ is strictly positive and hence $t_2(s)$ is strictly positive in any equilibrium. This in turn implies that there is a strictly positive probability that there are no votes for 1 and one vote for 2. Hence the probability that a voter is pivotal for 1 is strictly positive. Therefore, $\beta(s|piv_j, m), j = 1, 2$ and $\beta(s|piv, m)$ are well defined in equilibrium.

Fix a symmetric profile σ and let $E(v(x, s)|piv_j, m)$ denote the expectation of $v(x, s)$ with respect to $\beta(\cdot|piv_j, m)$. A voter of type (x, m) prefers to vote for 2 rather than abstain if

$$E\{v(x,s)|piv_2, m\} = \sum_s v(x,s)\beta(s|piv_2, m) > 0; \tag{8}$$

he prefers to vote for 1 rather than abstain if

$$E\{v(x,s)|piv_1, m\} = \sum_s v(x,s)\beta(s|piv_1, m) < 0, \tag{9}$$

and he prefers to vote for 2 rather than 1 if

$$E\{v(x,s)|piv,m\} = \sum_{s} v(x,s)\beta(s|piv,m) > 0 \tag{10}$$

By Assumption 1 $v(x,s)$ is strictly increasing in x. Thus, if it is a best response for type x to vote for 2 then for every type $x' > x$ the unique best response must be to vote for 2. Similarly, if it is a best response for type x to vote for 1 then the unique best response for every type $x' < x$ must be to vote for 1. It follows that any voting equilibrium can be characterized by cutpoints such that for any $m \in M$ $x_m^1 \leq x_m^2$ and all types with $x < x_m^1$ vote for 1 when they receive signal m and all types with $x > x_m^2$ vote for 2 if they receive signal m. Types in the interval (x_m^1, x_m^2) abstain.

Observe that voters with types x_m^j are either indifferent between voting for j and abstaining (if (x_m^1, x_m^2) is non-empty) or indifferent between voting for 1 and 2 (if $x_m^1 = x_m^2$). As a consequence either

$$E\{v(x_m^j, s)|piv_j, m\} = 0$$

or

$$E\{v(x_m^j, s)|piv, m\} = 0$$

To show that cutpoints are ordered observe that for $m > m'$, such that $m, m' \neq \emptyset$, SMLRP implies that for all x

$$E\{v(x,s)|piv_j, m\} > E\{v(x,s)|piv_j, m'\}$$

and

$$E\{v(x,s)|piv, m\} > E\{v(x,s)|piv, m'\}$$

Therefore, it must be the case that $1 > x_1^j > .... > x_M^j > -1$ for $j = 1, 2$. ■

Proposition 2 *In a symmetric equilibrium then $t_1(s)$ is strictly decreasing in s and $t_2(s)$ is strictly increasing in s.*

Proof. Since a symmetric equilibrium can be characterized by ordered cutpoints we can rewrite Equation (2) as follows:

$$t_1(s) = \sum_{m \in M} p(m|s) F(x_m^1) \tag{11}$$

and

$$t_2(s) = \sum_{m \in M} p(m|s)(1 - F(x_m^2)) \tag{12}$$

Since x_m^j strictly decreasing in m for every $j \in \{1, 2\}$ it follows from a standard property of SMLRP that $t_1(s)$ is strictly decreasing and $t_2(s)$ is strictly increasing in s. ■

Note that the fact that every symmetric equilibrium must have ordered cutpoints does not imply that there are always some voters who abstain. It allows for the case where $x_m^1 = x_m^2$ for all $m \in \{1, 2, ..., M\}$. In particular, consider the case in which there are two states and each voter receives perfect information about the state s. In this case, every voter type that occurs with positive probability (except possibly one type) has a strictly dominant strategy to vote for one of the two alternatives and hence it must be the case that $x_m^1 = x_m^2$ for all m.

We now show that if there are any messages that are not *perfectly informative* a strictly positive fraction of the electorate will abstain. We call a signal $m \in M$ perfectly informative if there is a unique state s_m in which the signal can be received with strictly positive probability i.e., $p(m|s) > 0$ if and only if $s = s_m$.

Proposition 3 *Suppose Assumption 1 holds. Suppose σ is an equilibrium strategy and let $(x_m^j), j = 1, 2$ be the corresponding cutpoints. If m is not perfectly informative then $x_m^2 - x_m^1 > 0$.*

Proof. Equations (3) and (4) imply that

$$\begin{aligned} \frac{\Pr\{piv_1|s\}}{\Pr\{piv_2|s\}} &= \frac{\sum_{x=0}^{\infty} \frac{t_1(s)^x t_2(s)^{x+1}}{x!(x+1)!}}{\sum_{x=0}^{\infty} \frac{t_1(s)^x t_2(s)^x}{x!x!}} \\ &\equiv L(t_1(s), t_2(s)). \end{aligned}$$

In the appendix (Lemma 6) we demonstrate that

$$\frac{\partial}{\partial t_1} L(t_1, t_2) < 0, \frac{\partial}{\partial t_2} L(t_1, t_2) > 0$$

and therefore, since t_1 is strictly decreasing in s and t_2 is strictly increasing in s it follows that $L(t_1(s), t_2(s))$ is strictly increasing in s.

Let $\beta(s|Y)$ denote the probability of state s conditional on the event Y. We can interpret $\beta(s|piv_j, m)$ as the distribution that is achieved by updating $\beta(s|m) =$

$\frac{g(s)p(m|s)}{\sum_m g(s)p(m|s)}$ with the signal $piv_j \in \{piv_1, piv_2\}$ which satisfies SMLRP since $L(t_1(s), t_2(s))$ is strictly increasing in s. If m is not perfectly informative then it is non-zero for two or more states. Then by a standard property of SMLRP (see Milgrom (1979)) it follows that $\beta(s|piv_1, m)$ strictly first order stochastically dominates $\beta(s|piv_2, m)$. Since v is strictly increasing in s this allows us to conclude that

$$E\{v(x,s)|piv_1, m\} > E\{v(x,s)|piv_2, m\}. \tag{13}$$

Note that piv is the union of the events piv_1 and piv_2 which both occur with strictly positive probability and therefore

$$E\{v(x,s)|piv_1, m\} > E\{v(x,s)|piv, m\} > E\{v(x,s)|piv_2, m\}. \tag{14}$$

Consider a voter $(\hat{x}, m)$ who is indifferent between voting for 1 and voting for 2, that is,

$$E\{v(x,s)|piv, m\} = 0.$$

Such a voter exists by Assumption 1. By inequality 14 this voter strictly prefers to abstain since conditional on piv_1, i.e., if a vote for 1 is decisive the voter strictly prefers alternative 2 and conditional on piv_2 (if a vote for 2 is decisive) the voter strictly prefers alternative 1. Since v is continuous it follows that there is an interval of voters who strictly prefer to abstain whenever m is not perfectly informative. ■

To gain an intuition for Proposition 3 consider a (not perfectly informed) voter who is indifferent between voting for alternative 2 and abstaining. This implies that conditional on the event that $n_1 = n_2$ the voter is indifferent between the two candidates. Such a voter strictly prefers abstaining to voting for alternative 1. To see this note that voting for 1 instead of abstaining only makes a difference in the event that $n_2 = n_1 + 1$. Because the expected vote share of alternative 2 is strictly increasing in s the probability distribution over states conditional on $n_2 = n_1 + 1$ puts more weight on states that are favorable to candidate 2.[11] Thus, alternative 2 is even more desirable in the event that $n_2 = n_1 + 1$ than in the event $n_2 = n_1$ and the voter strictly prefers to abstain rather than vote for alternative 1. As a consequence there is an interval of preference types who prefer abstaining to voting for either candidate.

[11] Technically, the distribution over states conditional on $n_2 = n_1 + 1$ first order stochastically dominates the probability distribution over states conditional on $n_2 = n_1$.

Proposition 3 does not exclude the possibility that the expected fraction of voters who choose to abstain in equilibrium converges to zero as the population goes to infinity. Thus, to conclude that our model predicts that a positive fraction of the electorate abstains we need a stronger result: we need to demonstrate that the fraction of voters who abstain stays bounded away from zero for all k.

4 Equilibrium Characterization for Large Populations

In the following we will consider a sequence of equilibria corresponding to a sequence of expected number of voters ν_k where $\nu_k \rightarrow \infty$. Along the sequence we fix the information structure defined by $g(s), p(m|s)$ and $F(x)$.

4.1 Preliminary Results

In this section we demonstrate four technical Lemmas that will allow us to prove our main results on abstention and information aggregation.

We define the ratio of expected vote shares for the two alternatives given state $s \in S$ and profile σ as:

$$\rho(s) = \frac{t_1(s)}{t_2(s)} \tag{15}$$

We use the subscript k to indicate the dependence of the equilibrium on the parameter ν_k. Thus, e.g., (σ_k) denotes a sequence of voting equilibria and $\rho_k(s)$ describes the ratio of probabilities of voting for 1 and voting for 2.

Lemma 1 says that for large k the probability of the event $n_1 = n_2$ differs from the probability of the event $n_1 = n_2 + 1$ by a factor that is approximately equal to $\sqrt{\rho_k(s)}$. Lemma 1 is a consequence of the assumption that the number of voters is a Poisson random variable and was first proven by Myerson (1997b). The proof can be found in the Appendix.

Lemma 1 *Suppose that $t_{1k}(s) > 0$ and $t_{2k}(s) > 0$ for all k. Then*

$$\left| \frac{\Pr_k\{piv_1|s\}}{\Pr_k\{piv_2|s\}} - \sqrt{\rho_k(s)} \right| \rightarrow 0$$

as $k \rightarrow \infty$.

Proof. see Appendix A. ■

If $t_{1k}(s) > t_{2k}(s)$, i.e., the expected vote share of alternative 1 is larger than the expected vote share of alternative 2 then the Lemma implies that the event $n_1 = n_2$ is more likely than the event $n_1 = n_2 + 1$. In addition, the relative likelihood of these two events converges to the square root of the expected vote shares.

The next three lemmas are concerned with the distribution over states conditional on a vote being pivotal. Since every voter behaves *as if* a vote is pivotal characterizing this distribution is essential for our results.

Let

$$T = \left\{ \arg\min \left| \sqrt{t_{1k}(s)} - \sqrt{t_{2k}(s)} \right| \right\}, \tag{16}$$

i.e., the states that minimize the difference between the square roots of the expected vote shares for each alternative. The next lemma shows that in a large election conditional on a vote being pivotal almost all probability weight must be on states in T.

Lemma 2 *Suppose that* $\left(x^j_{1k} - x^j_{Mk}\right) > \epsilon > 0$ *for* $j = 1$ *or* $j = 2$ *and all* k. *Then as* $k \to \infty$ $\sum_{s \in T} \beta_k(s|piv_j) \to 1$ *for* $j = 1, 2$ *and* $\sum_{s \in T} \beta_k(s|piv) \to 1$.

Proof. see Appendix A. ■

Let $x(s)$ denote the preference type who is indifferent between the two alternatives in state s. Thus $x(s)$ is defined by the equation

$$v(x(s), s) = 0 \tag{17}$$

Assumption 1 implies that $x(s)$ is well defined for every $s \in S$. All voters with preference types $x < x(0)$ prefer alternative 1 in every state and hence have a strictly dominant strategy to vote for 1. Similarly, voters with preference types $x > x(1)$ always vote for alternative 2. If there are many voters and if $F(x(1)) > 1/2$ ($F(x(0)) < 1/2$) then with high probability the majority of voters prefers alternative 1 (alternative 2) irrespective of their information and hence alternative 1 (alternative 2) must win the election with high probability in every state. The following assumption rules out these trivial cases.

Assumption 6 $F(x(0)) > 1/2, F(x(1)) < 1/2$.

We define s^1 to be the largest state in which 1 is elected if all voters know the state, i.e.,

$$s^1 = \max\{s : F(x(s)) \geq 1/2\}$$

Similarly, we define s^2 to be the smallest state in which 2 is elected if all voters know the state, i.e.,

$$s^2 = \min\{s : F(x(s)) < 1/2\}$$

Clearly $s^1 < s^2$ and (s^1, s^2) is a pair of consecutive states. Also note that s^1, s^2 are well defined if Assumption 6 holds.

The following Lemma says that in a voting equilibrium with a large electorate beliefs over states conditional on a voter being pivotal are concentrated on s^1 and s^2.

Lemma 3 *Suppose Assumptions 1-6 hold. Consider a sequence of voting equilibria. Then $\beta_k(s^1|piv) + \beta_k(s^2|piv) \rightarrow 1$ as $k \rightarrow \infty$. Also, for $j = 1, 2$, $\beta_k(s^1|piv_j) + \beta_k(s^2|piv_j) \rightarrow 1$ as $k \rightarrow \infty$.*

Proof. see Appendix A. ■

To gain some intuition for Lemma 3 note that there is set of states for which the ratio of the expected equilibrium vote shares of the two alternatives is closest to one. Since, from Proposition 1 the equilibrium expected vote share of candidate 1(2) is a strictly de(in)creasing function of s it follows that this set consists either of a unique state or of a pair of consecutive states. Call this consecutive pair of states $\{s', s''\}$. If it is not the case that either s^1 or s^2 is in this set then suppose for e.g., that $s^1 < s'$ and $s^1 < s''$. The fact that all beliefs are concentrated on the set $\{s', s''\}$ implies that $t_{1k}(s) > F(x(s^1)) > .5$ for all s. But then by Lemma 2 all beliefs must be concentrated on state $s = 0 < s'$ i.e., under such a profile voter's beliefs conditioned on a vote being pivotal must be concentrated on the state which gives the lowest expected vote share for candidate 1.

In the following Lemma we refine the characterization of the limit support of the distribution $\beta_k(s|piv)$: we give conditions under which both $s^1$and s^2 have positive probability conditional on a voter being pivotal. Let

$$\bar{M} = \{m : p(m, s^1)p(m, s^2) = 0 \text{ and } p(m, s^1) + p(m, s^2) > 0\}$$

denote the set of signals that allow an agent to distinguish states s^1 and s^2 with certainty. In other words, an agent who receives a signal $m \in \bar{M}$ and knows that

$s \in \{s^1, s^2\}$ can determine the state with certainty. The expression $(F(x(s^1)) - F(x(s^2)) \sum_{m\in\bar{M}} p(m|s)$ denotes the expected fraction of voters who receive signals in $\bar{M}$ in state s and who have state dependent preferences over the states in $\{s^1, s^2\}$. Following FP we call such voters perfectly informed swing voters.

The following assumption requires that the probability a voter is a perfectly informed swing voter in state $s \in \{s^1, s^2\}$ must be less than twice the difference between $\frac{1}{2}$ and the fully informed vote share in states $s^i, i = 1, 2$.

Assumption 7 $2|\frac{1}{2} - F(x(s^i))| > (F(x(s^1)) - F(x(s^2)) \sum_{m\in\bar{M}} p(m|\bar{s})$ *for* $i = 1, 2$.

Assumption 7 requires that $F(x(s^i)) \neq \frac{1}{2}$ for $i = 1, 2$. In words this means that the under conditions of full information it cannot be the case that exactly half the voter are expected to prefer candidate 1 to candidate 2. Furthermore, if $F(x^1) = 1 - F(x^2)$ i.e., the expected vote shares for each candidate in states s^1 and s^2 are symmetric around $\frac{1}{2}$ then Assumption 7 always holds whenever there is a strictly positive probability that an agent receives a signal that does not allow him to perfectly discriminate between s^1 and s^2.

The next Lemma demonstrates that when Assumption 7 holds then conditional on being pivotal a voter believes both states s^1 and s^2 have probability bounded away from zero.

Lemma 4 *Suppose Assumptions 1-7 hold. Then there is an* $\epsilon > 0$ *and a* $k' < \infty$ *such that* $\beta_k(s^1|piv) \geq \epsilon, \beta_k(s^2|piv) \geq \epsilon$ *for all* $k \geq k'$.

To understand why Assumption 7 is needed in Lemma 4 consider the following example:

Example Suppose there are two states $s = 1, 2$. Further suppose that $F(x(1)) = 0.75$, i.e., the probability that a randomly drawn voter prefers 1 in state 1 is 0.75 and $F(x(2)) = 0.4$, i.e., the probability that a randomly drawn voter prefers 2 in state 2 is 0.6. Thus $2|.5 - F(x(2))| = 0.2$ and $F(x(0)) - F(x(1)) = .35$.

First consider the case where all voters know the state. Clearly this implies that in a large electorate alternative 1 gets close to 75% of the vote in state 1 and close to 40% of the vote in state 2 with probability close to one. Conditional on a vote being pivotal the probability of state 2 converges to one in this case. Also note that Assumption 7 is violated in this case since $.35 > 2$.

Now assume that some voters do not know s, i.e., do not get a perfectly informative signal. In a large electorate, if there is a small fraction of uninformed voters, these voters must behave *as if* the state is $s = 2$. Thus an uninformed voter will vote for candidate 1 with probability 0.4 and for candidate 2 with probability 0.6. If q is the probability that a voter is informed then the expected vote share for candidate 1 in state 1 is now

$$t_1(s = 1) = q \cdot 0.75 + (1 - q) \cdot 0.4$$

and the expected vote share for candidate 2 in state 2 is

$$t_2(s = 2) = q \cdot 0.6 + (1 - q) \cdot 0.6 = 0.6.$$

Thus the described equilibrium strategies and the resulting limit distribution over states conditional on a vote being pivotal is valid as long as

$$q \cdot 0.75 + (1 - q) \cdot 0.4 > 0.6$$

or

$$q > \frac{0.2}{0.35}$$

For $q < 0.2/.35$ Assumption 7 is satisfied and the Lemma shows that both states $s = 1$ and $s = 2$ must have strictly positive probability in the limit distribution over states conditional on a vote being pivotal.

4.2 Full Information Equivalence

In FP and in a follow up paper, (Feddersen and Pesendorfer 1997), it was demonstrated that elections have a property called *full information equivalence*: the election outcome under private and asymmetric information converges in probability to the election outcome that would occur if all the voters knew the true state and voted for their preferred candidate.

The following Theorem demonstrates that this property also holds for the present model. In fact, the result follows as a corollary to Lemma 3.

Proposition 4 *Suppose Assumptions 1-5 hold and* $F(x(s^1)) > 1/2$. *Then, as* $k \to \infty$, *the probability that alternative* 1 *is elected converges to* 1 *for* $s \leq s^1$ *and to* 0 *for* $s > s^1$.

Proof. By Step 1 of the proof of Lemma 2 it follows from $F(x(s^1)) > 1/2$ that there is an $\varepsilon > 0$ such that $(x^j_{mk} - x^j_{m'k}) > \varepsilon$ for $m' > m$ and all k. As a consequence, there is an $\varepsilon' > 0$ such that

$$|t_j(s) - t_j(s')| > \varepsilon'. \tag{18}$$

Hence t_1 is strictly decreasing in s with slope bounded away from zero and t_2 are strictly increasing in s with slope bounded away from zero. By the proof of step 2 of Lemma 2 we know that for any convergent subsequence

$$\begin{aligned} & \lim \left(\sqrt{t_1(s^1)} - \sqrt{t_2(s^1)}\right) \\ \geq \quad & 0 \geq \lim \left(\sqrt{t_1(s^2)} - \sqrt{t_2(s^2)}\right). \end{aligned}$$

If the above two inequalities are strict then (18) together with the strong law of large numbers implies the Theorem.

In the remainder of the proof we show that indeed both of these inequalities must be strict. To see this first note that (18) implies that at least one of the two inequalities is strict. Thus suppose that $\lim \left(\sqrt{t_1(s^1)} - \sqrt{t_2(s^1)}\right) = 0$. Then by the argument given in step 2 of the proof of Lemma 2 $\beta(s^1|piv_j) \to 1$. This implies that $t_1(s^1) \to F(x(s^1)) > 1/2$ (by Assumption) which contradicts the hypothesis that $\lim \left(\sqrt{t_1(s^1)} - \sqrt{t_2(s^1)}\right) = 0$. An analogous argument shows that the second inequality is strict. ■

4.3 Abstention

In this section we demonstrate that Assumption 7 guarantees that the fraction of voters who abstain stays bounded away from zero as $k \to \infty$. This assumption guarantees abstention because it ensures that as the population size grows imperfectly informed voters must place positive weight on both states s^1, s^2 conditional on a vote being pivotal.

Proposition 5 *Suppose Assumptions 1-7 hold. Suppose there exists a signal m such that $p(m|s) > 0$ for any $s \in \{s^1, s^2\}$. Then there is an $\alpha > 0$ and a k' such that the expected fraction of voters who abstain in equilibrium is larger than α for all $k > k'$.*

Proof. see Appendix A. ■

In the following we provide an intuition for Proposition 5: From Lemma 3 we know that

$$\beta_k(s^1|piv_1) \approx \frac{g(s^1)\Pr_k(piv_1|s^1)}{g(s^1)\Pr_k(piv_1|s^1) + g(s^2)\Pr_k(piv_1|s^2)}$$

Recall that $\rho(s)$ is strictly decreasing in s since $t_2(s)$ is strictly increasing in s and $t_1(s)$ is strictly decreasing in s. Using the approximation formula derived in Lemma 1 we can compute:

$$\beta_k(s^1|piv_2) \approx \frac{g(s^1)\frac{\Pr_k(piv_1|s^1)}{\sqrt{\rho_k(s^1)}}}{g(s^1)\frac{\Pr_k(piv_1|s^1)}{\sqrt{\rho_k(s^1)}} + g(s^2)\frac{\Pr_k(piv_1|s^2))}{\sqrt{\rho_k(s^2)}}} \tag{19}$$

Conditioning on the event piv_2 can therefore be interpreted as conditioning on the event piv_1 and an additional conditionally independent signal, where the probability of observing the additional signal in state s is given by

$$\frac{1/\sqrt{\rho_k(s)}}{\sum 1/\sqrt{\rho_k(s)}}$$

If both $\beta_k(s^1|piv_1)$ and $\beta_k(s^2|piv_1)$ are bounded away from zero and if $\rho_k(s)$ is strictly decreasing in s then there is an $\epsilon' > 0$ such that

$$\beta_k(s^1|piv_2) < \beta_k(s^1|piv_1) - \epsilon',$$

for all k, i.e., the two probability distributions are different even in the limit. But this is enough to allow us to make the argument given in Proposition 3 for an interval of preference type with length uniformly bounded away from zero and hence Proposition 5 follows.

The previous proposition demonstrated that a strictly positive fraction of the electorate will always abstain but it did not provide a sense of how large that fraction may be. The next proposition provides a bound on the fraction of agents who abstain in equilibrium. Define a *critical state* to be an $\hat{s} \in \arg\min_s |F(x(s)) - 1/2|$ i.e., a state in which the fraction of the electorate that prefers candidate 1 to candidate 2 is closest to 1/2. It follows from the definition of s^1 and s^2 that $\hat{s} \in \{s^1, s^2\}$. We now show that the fraction of the electorate that abstains in a large election goes to zero as $|F(x(\hat{s})) - 1/2|$ goes to zero. Note that $\hat{s}$ is a state that would lead to the closest election outcome if all agents know the state of the world. Thus, if there is a state in which the election is expected to be very close under complete information, then there is very little abstention.

Proposition 6 *Suppose Assumptions 1-6 hold. For every* $\varepsilon > 0$ *there is an* $\eta > 0$ *such that* $|F(x(\hat{s})) - 1/2| \leq \eta$ *implies that* $\limsup_k (1 - t_{1k}(s) - t_{2k}(s)) \leq \varepsilon$ *for all* $s \in S$*, i.e., the expected fraction of voters who abstain is bounded above by* ε *for sufficiently large* k.

The intuition is that if $|F(x(\hat{s})) - 1/2|$ is very small then the equilibrium vote share of each alternative must be close to $1/2$ in both s^1 *and* s^2. But (by Lemma 1) this implies that a single vote provides very little information about the state since voters are expected to vote for either candidate with close to equal probability in both s^1 and s^2. Hence conditioning on the event piv_1 provides very similar information to conditioning on the event piv_2 and a small fraction of the electorate abstains.[12]

As a corollary this implies that if the state space is "fine", i.e., if the utility variation between the states s^1 and s^2 is small then the level of abstention is small.

Corollary 1 *Suppose Assumptions 1-6 hold. Then for every* $\epsilon > 0$ *there is an* ν *and a* k' *such that if* $\max_x |v(x, s^2) - v(x, s^1)| < \nu$ *then the expected fraction of voters who abstain is less than* ϵ *for all* $k > k'$.

We can use Propositions 5 and 1 to relate the level of abstention to the "aggregate" level of information. Consider the following example.

Example Consider the bridge example from the introduction and suppose voters are uncertain about the true cost of the bridge. Specifically, let the true cost of the bridge be a continuous random variable $c \in [0, 1]$ that is drawn by nature according to some probability distribution. Suppose that there are K television stations who each do an independent investigation into the cost of the bridge and televise a news report that reports only if the bridge is very expensive or not too expensive. Thus,

[12] Assumption 6 implies that there cannot be a state in which the expected vote shares are exactly equal under full information. In this case the fraction of voters who abstain converges to zero. We excluded it here because it would require a different proof. The reason is that in this case the fraction of voters who use their private information may converge to zero. Therefore, a more delicate argument is needed to establish full information equivalence (see Feddersen and Pesendorfer (1997)) which is needed to prove Proposition 5.

the only information a very large number of voters can learn about the cost is from a TV report. The probability that a television station reports the bridge costs as high is equal to $r(c)$ where $r(c)$ is strictly increasing in c and $r(c) > 0$ for all c i.e., as the true cost of the bridge increases a news report is more likely to report that the bridge is expensive. Each voter watches at most one television news report but some do not watch at all. The probability a voter watches a report is q. Let $s_k = E(c|k)$ denote the expected cost of the bridge conditional on k out of K television stations reporting that the bridge is expensive. Finally, suppose that the expected payoff of agents given k endorsements for alternative 1 is given by

$$E\{v(x,t)|k\} = v(x, s_k)$$

where $v(x, s_k)$ satisfies Assumption 1.

This example can be represented in our model as a finite state space model in which the state space is $S = \{s_0, ..., s_K\}$ and the probability a voter observes report $m \in \{\emptyset, l, h\}$ in state s_k is

$$p(m|s_k) = \begin{cases} \frac{qk}{K} & \text{if} \quad m = h \\ q(1 - \frac{k}{K}) & \text{if} \quad m = l \\ (1-q) & \text{if} \quad m = \emptyset \end{cases}$$

When $K = 1$ there are two types of agents, those who are perfectly informed and those who are uninformed. When $K > 1$ there are no perfectly informed agents. When $K \to \infty$ the state space approximates a continuous state space $[0, 1]$.

Clearly, if the number of independent investigations by television stations is large then $s_{k+1} - s_k$ is small. More precisely, for every $\nu > 0$ there is a K such that $s_{k+1} - s_k < \nu$ uniformly for all $k = 0, 1, ..., K$. Thus, many independent reports leads to better aggregate information and a smaller fraction of the population abstaining in equilibrium as long as v is continuous in s.

5 Partisanship, Information and Participation

In this section we use a simplified version of our model to demonstrate how private information combined with preference diversity can result in two seemingly contradictory comparative statics: (1) more informed voters participate with higher probability than less informed voters; and (2) increasing the fraction of the electorate that is better informed results in increased abstention. This version of the model also demonstrates that partisanship and information are linked. In particular, we illustrate why a better informed electorate may be less partisan than a less informed electorate.

Assume there are two states $s = 1, 2$ and

$$v(x,s) = \begin{cases} x - 1 \text{ if } s = 1 \\ x + 1 \text{ if } s = 2. \end{cases} \tag{20}$$

The agents' prior assigns probability 1/2 to either state. Agents are distributed uniformly on $[-1, 1]$. There are three signals $M = \{1, 2, \emptyset\}$. Signals 1 and 2 are informative while $\emptyset$ provides no information. Voters receive an informative signal with probability q. The conditional probability of observing signal $m \in M$ given state s is $p(m|s)$ where:

$$p(m|s) = \begin{cases} pq & \text{if} & m = s \\ (1-p)q & \text{if} & m \neq s, m \neq \emptyset \\ (1-q) & \text{if} & m = \emptyset \end{cases}$$

and $p > 1/2$.

An equilibrium in this model is characterized by the vector of cutpoints $(x^1_\emptyset, x^2_\emptyset, x^1_2, x^2_2, x^1_1, x^2_1)$ where x^j_m is the voter type who is indifferent between voting for candidate j and abstaining conditional upon observing signal m. Let $\beta(1|piv_j, m)$ be the probability of state 1 given piv_j and message m.

$$\beta(1|piv_j, m) = \frac{1}{1 + \frac{p(m|2)}{p(m|1)} \frac{\Pr\{piv_j|2\}}{\Pr\{piv_j|1\}}} \tag{21}$$

It follows from (20) that for any $j \in \{1, 2\}$ and $m \in M$ the cutpoint x^j_m is given by:

$$x^j_m = 2\beta(1|piv_j, m) - 1 \tag{22}$$

The expected vote share for candidate 1 in state s is $t_1(s)$ and, from Equation (11), can be written

$$t_1(s) = .5\left(p(1|s)(1 + x^1_1) + p(2|s)(1 + x^1_2) + p(\emptyset|s)(1 + x^1_\emptyset)\right) \tag{23}$$

Similarly the expected vote share for candidate 2 in state s is

$$t_2(s) = .5\left(p(1|s)(1 - x^2_1) + p(2|s)(1 - x^2_2) + p(\emptyset|s)(1 - x^2_\emptyset)\right). \tag{24}$$

We analyze equilibrium behavior in large elections. Lemma 1 requires that in the limit (as $\nu \to \infty$):

$$\frac{\Pr\{piv_1|2\}}{\Pr\{piv_1|1\}} = \frac{\sqrt{t_2(2)}\sqrt{t_1(1)}}{\sqrt{t_1(2)}\sqrt{t_2(1)}} \frac{\Pr\{piv_2|2\}}{\Pr\{piv_2|1\}} \tag{25}$$

Lemma 2 implies:

$$\sqrt{t_1(1)} - \sqrt{t_1(2)} = \sqrt{t_2(2)} - \sqrt{t_2(1)} \tag{26}$$

The system of equations (22)-(26) may be solved to find a limit equilibrium strategy profile.

5.1 Perfectly informative signals

Suppose the signal $m \in \{1, 2\}$ is perfectly informative i.e., $p = 1$. In this example the only variable is the fraction of the electorate that is perfectly informed (q). This assumption combined with the symmetry of the setting greatly simplifies the system of equations and permits an analytical solution. Observe that $x_1^j = -1$ and $x_2^j = 1$ for any $j \in \{1, 2\}$ i.e., all voters who observe a perfectly informative signal vote for candidate 1 if they observe signal 1 and candidate 2 otherwise. By the assumption of a symmetric distribution of preference types it must be the case that $t_1(1) = t_2(2)$, $t_1(2) = t_2(1)$ and $x_\emptyset^1 = -x_\emptyset^2$. A little bit of algebra can be used to show that the ratio of pivot probabilities reduces to

$$\frac{\Pr\{piv_2|2\}}{\Pr\{piv_2|1\}} = \sqrt{\frac{t_1(2)}{t_1(1)}}$$

$$\frac{\Pr\{piv_1|2\}}{\Pr\{piv_1|1\}} = \sqrt{\frac{t_1(1)}{t_1(2)}}$$

We can now find the symmetric limit equilibrium by solving the following system of equations for the cutpoint $x_\emptyset^1$ and the expected vote shares $t_1(1)$ and $t_1(2)$:

$x_\emptyset^1 = \frac{2}{1+\sqrt{\frac{t_1(1)}{t_1(2)}}} - 1$
$t_1(1) = q + .5(1-q)(1+x_\emptyset^1)$
$t_1(2) = .5(1-q)(1+x_\emptyset^1)$

When we solve this system we get:

$x_\emptyset^1 = \frac{q}{q-2}$
$t_1(1) = \frac{1}{2-q}$
$t_1(2) = \frac{(1-q)^2}{2-q}$

(27)

Abstention as a function of the fraction of the electorate that is perfectly informed is given by the equation:

$$1 - t_1(1) - t_1(2) = q\frac{1-q}{2-q}$$

Figure 1 plots this function and illustrates that when the fraction of the electorate that is perfectly informed grows the result may sometimes be more abstention. Since perfectly informed voters never abstain one might imagine that increasing the fraction of such voters in the electorate would automatically reduce abstention. However, as the fraction of informed voters grows so does the informativeness of the election result. As a consequence the uninformed voters become more willing to abstain. When the fraction of the informed electorate is small this equilibrium effect dominates and abstention increases. When the fraction of the informed electorate is large the additional abstention of the uninformed is outweighed by the additional participation by the informed.

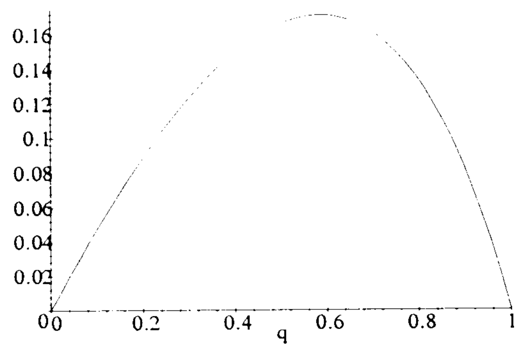

Figure 1.

This model can also be used to illustrate the relationship between information and partisanship. One natural definition of a partisan is a voter who always votes for the same candidate independent of the state. The set of voters with preference types below $x_\emptyset^1$ and who do not receive an informative signal always vote for candidate 1 while those with types above $-x_\emptyset^1$ always vote for candidate 2. Thus there is a set of voters in this model who behave like partisans. The expected fraction of the electorate that behaves like partisans is just the fraction of uninformed voters who do not abstain. This fraction is given below:

$$(1-q)(1-\frac{q}{2-q})$$

Figure 2 plots this fraction and demonstrates that as the fraction of the electorate that is informed increases the fraction of voters behaving like partisans decreases.

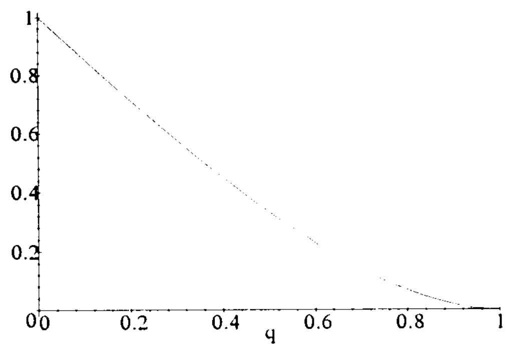

Figure 2.

5.2 Noisy Signals

In this section we examine a set of examples in which no voter has perfect information. We say that a signal that is informative but not perfectly informative i.e., $p \in (.5, 1)$ is noisy. The motivation for considering noisy signals is that changes in either the quality of information or in the fraction of the electorate that is informed can effect the decision of all the voters. In particular, if voters only have noisy signals then as the fraction of the electorate that is informed increases it may be the case that abstention increases among both informed an uninformed alike. The examples also illustrate how noisy signals lead to a biased electorate without introducing a bias in outcomes. We say that the electorate is biased in some state s if the mean of the set of voters expected to vote in state s is not 0.

Assume the same structure as in the model above except those who are informed observe a noisy signal, i.e., $p \in (.5, 1)$. There are now three cutpoints that must be calculated. Even those who are informed learn something from the event a vote is pivotal therefore $x_1^1 \neq x_2^1$. However, the symmetry of the example still implies that $x_\emptyset^1 = -x_\emptyset^2$, $x_1^1 = -x_2^2$ and $x_2^1 = -x_1^2$.[13]

In Table 1 below the entry A is the total level of abstention, A_i is the fraction of the informed voters who abstain and A_u is the fraction of uninformed voters who abstain. In the last row we look at the case in which there are no uninformed voters.

[13]The system of equations used to compute the following examples may be found in appendix B. We compute examples here because we are unable to solve the system analytically.

In the last column we look at the result when the signal is perfectly informative.

Three things stand out in these examples. First in every cell it is the case that abstention by informed voters is strictly lower than abstention by the uninformed. Second, increasing either the informativeness of the private signal or the fraction of the electorate that is informed always results in higher levels of abstention by the uninformed and often by informed voters as well. Third, increasing either p or q always results in a larger increase in abstention among the uninformed than among the informed.[14]

Table 1.

	$p = .55$	$p = .7$	$p = 1$
$q = .2$	$A = .001$ $A_i = .001$ $A_u = .001$	$A = .015736$ $A_i = .013656$ $A_u = .016256$	$A_u = .09$
$q = .4$	$A = .001996$ $A_i = .001984$ $A_u = .002004$	$A = .030915$ $A_i = .027747$ $A_u = .033027$	$A_u = .15$
$q = .6$	$A = .00299$ $A_i = .00297$ $A_u = .003$	$A = .045488$ $A_i = .042275$ $A_u = .050308$	$A_u = .17$
$q = .8$	$A = .00398$ $A_i = .00398$ $A_u = .00402$	$A = .05941$ $A_i = .0573$ $A_u = .0681$	$A_u = .13$
$q = 1$	$A_i = .004975$	$A_i = .072632$	$A = 0$

To see how noisy signals can generate the full spectrum of partisans and a biased electorate consider the following two examples.

Table 2.

	x_2^1	x_2^2	$x_\emptyset^1$	$x_\emptyset^2$	x_1^1	x_1^2	
$q = .6$ $p = .95$	$-.95$	$-.82$	$-.32$	$.32$	$.82$	$.95$	$A = .164$ $A_i = .065$ $A_u = .31$
$q = .8$ $p = .95$	$-.96$	$-.76$	$-.45$	$.45$	$.76$	$.96$	$A = .17$ $A_i = .10$ $A_u = .45$

[14]The reader may wonder if there is a theorem that might be proved to the effect that those with better information always participate more frequently and that abstention among the uniformed always increases faster as the electorate as a whole becomes more informed. While this pattern appears to be robust for a variety of examples it is possible to construct a counterexample in which those with better information abstain with higher probability than those with worse information. The example requires the introduction of considerable asymmetry in preferences. Details are available upon request from the authors.

These cutpoints define seven intervals. All voters with preferences in the interval $(-1, x_2^1)$ and $(x_1^2, 1)$ are strong partisans for candidates 1 and 2 respectively in the sense that they always vote for their candidate independent of their private information. Those in the intervals (x_2^1, x_2^2) and (x_1^1, x_1^2) are weak partisans in that they either vote for their candidate or abstain. Voters in the interval (x_2^2, x_1^1) are independents i.e., they may vote for either candidate depending on their information. However, we may want to introduce a further distinction among the independents and, following the voting behavior literature, call those in the intervals $(x_2^2, x_\emptyset^1)$ and $(x_\emptyset^2, x_1^1)$ independent leaners because e.g., those in the interval $(x_2^2, x_\emptyset^1)$ vote for candidate 1 when they observe signal 1 and when they observe $\emptyset$ whereas those in the interval $(x_\emptyset^1, x_1^1)$ only vote for 1 when they observe signal 1.

Increasing the fraction of the electorate that receives the noisy signal from .6 to .8 has the effect of reducing the fraction of strong partisans and independent leaners while increasing the fraction of weak partisans and pure independents. Abstention only increases by 1% however abstention among both the informed and uninformed increases by more than 50% among both groups. The overall abstention figures mask the fact that the composition of the abstainers changes dramatically in each state. Consider the case in which $p = .95$ and $q = .8$. In state 1 the probability of observing message 1 is $p(1|1) = .76$ and the probability of observing message 2 is $p(2|1) = .04$. Thus in state 1 76% of the weak partisans for candidate 2 will abstain (i.e., voters in the interval (x_1^1, x_1^2)) while only 4% of the weak partisans for candidate 1 will abstain. Since all of our analytical results continue to hold in this example we know that in state 1 candidate 1 almost surely wins and in state 2 candidate 2 almost surely wins. Noisy signals generate an electorate that is biased for the winner. Even though the observed electorate is biased the outcome of the election is never biased.

6 Bias and Abstention

Another stylized fact about participation in elections is that those on the left always participate less frequently than those on the right. The explanation for this appears to be the correlation between education (a proxy for information) and political preferences. The correlation may work as follows. Those with higher levels of education enjoy higher incomes and this pushes their political preferences to the right. The following example demonstrates that a biased distribution of information in our model can lead to an electorate that is always biased towards the more informed end of the ideological spectrum. While the bias in the electorate may be dramatic this does not result in a biased outcome. Indeed, in this example the biased distribution of information results in strictly higher probabilities the election satisfies full information equivalence than an unbiased distribution of information.

For purposes of exposition call those with preference types below 0 leftists and those above 0 rightists. Assume the same structure as in example 1 above with the proviso that all rightists observe the perfectly informative signal while leftists receive the uninformative signal. Thus the distribution of information is maximally biased towards those on the right. It is clear that any abstention due to the swing voter's curse will occur only among leftists guaranteeing that, independent of the state, the portion of the electorate that votes will be skewed to the right. Because the rightists all vote for 1 when they observe signal 1 and vote for 2 otherwise it only remains to determine the behavior of the leftists. The cutpoints, expected vote shares and expected level of abstention by those on the left (A_L) are given in Table 3.[15]

All leftist voters with preference types below $x^1_\emptyset$ are partisans for candidate 1 (the leftist), voters in the interval $(x^1_\emptyset, x^2_\emptyset)$ always abstain and those in interval $(x^2_\emptyset, 0)$ are partisans for candidate 2 (the right wing candidate). The level of abstention among the leftists is extreme at 30% nevertheless candidate 1 wins in a landslide in state 1 as does candidate 2 in state 2, indeed the margin of victory is the same in each state. Thus the pattern of abstention may be dramatically skewed towards one end of the ideological spectrum without biasing the outcome of the election. Of course in this example while voters have different preferences they have common values in the sense that all prefer candidate 1 in state 1 and candidate 2 in state 2. It is a simple matter to generalize this example, indeed any of the above examples, to include voters who always prefer one candidate or the other independent of the state.[16]

To get a sense of the effect that a skewed distribution has on equilibrium behavior

[15]The system of equations that was solved to generate this example may be found in Appendix B.

[16]It is also the case that in actual elections the level of abstention changes from one election to the next and that abstention among those on the left is higher in elections the democrats lose than when the democrats win. We have constructed examples with symmetric distributions of voters in which those on the left always abstain more frequently and abstention is higher when the left candidate loses. Details are available from the authors upon request.

consider example 1 above where $q = .5$ and there is no bias in the distribution of information. See column 2 of Table 3 below. In this case the fraction of the electorate that is informed is the same as in the above example. The difference is that a leftist is as likely to receive an informative signal as a rightist.

Table 3.

	Bias	No Bias
$x^1_\emptyset$	$-.8$	$-1/3$
$x^2_\emptyset$	$-.2$	$1/3$
$t_1(1)$	.6	2/3
$t_2(1)$	.1	1/6
$t_1(2)$	.1	1/6
$t_2(2)$	.6	2/3
A	.3	1/6

Abstention is lower in this example meaning that a biased distribution of information increases abstention. It also follows that a skewed distribution of information decreases partisanship since abstention only occurs among the uninformed and the only partisans in both examples are those who are uninformed and vote. Finally, while both the biased and the unbiased electorate choose candidate 1 in state 1 and candidate 2 in state 2 with high probability, in large finite elections the biased distribution of information actually result in *lower* probabilities of error than the unbiased example. This follows from the fact that all the informed voters vote correctly in each setting. The uninformed voters only introduce noise thus the fact that abstention is higher with the biased distribution of information marginally improves the performance of the electoral mechanism.

7 Conclusion

In the first section of the paper we demonstrated that abstention due to the swing voter's curse is robust to preference diversity and variable information environments. However, the dependence of the size of abstention on the state space raises some questions of interpretation. Under what circumstances is a coarse state space an appropriate model and when is a fine state space more appropriate? It seems to us that the coarse state space model is more appropriate under conditions in which the quality of the aggregate information available to voters is low. Congressional, state and local elections along with ballot initiatives are more appropriately described as low information affairs than are presidential or senate elections.

In the second section of the paper we presented a set of examples that demonstrate the range of phenomena that are consistent with the model. In particular we demonstrated that a biased electorate may be caused by a biased distribution of information without creating a biased outcome. The fact that a large range of behavior may be supported as equilibrium in some information environment complicates the task of using this model as a predictive tool. However, the normative results that elections work well as information aggregation mechanisms even in the face of what would appear to be a systematically biased electorate and large scale abstention is highly robust. At a minimum this should give pause to those who would argue that a persistently biased electorate inevitably leads to biased outcomes.

8 Appendix A

The modified Bessel function $I_v(x)$ is defined as

$$I_v(x) = \left(\frac{1}{2}x\right)^v \sum_{k=0}^{\infty} \frac{(\frac{1}{4}x^2)^k}{k!(v+k)!}$$

Note the following facts (see Abramowitz and Stegum (1970) p. 376):

$$\begin{aligned} I_0'(x) &= I_1(x) \\ I_1'(x) &= I_0(x) - \frac{1}{x}I_0(x) \\ I_2'(x) &= I_1(x) - \frac{2}{x}I_1(x) \\ I_0(0) &= 1, I_1(0) = 0, I_2(0) = 0 \\ I_0(x) - I_2(x) &= \frac{2}{x}I_1(x) \end{aligned} \tag{28}$$

Also note that from equation 9.7.1. in Abramowitz and Stegum (p. 377) it follows that

$$\lim_{x\to\infty} \frac{\frac{e^x}{\sqrt{2\pi x}}}{I_\nu(x)} = 1 \tag{29}$$

Clearly, the first two equations of 28 imply (together with the initial conditions $I_0(0) = 1, I_1(0) = 0$) that

$$I_0(x) > I_1(x) \tag{30}$$

since $I_1'(x) - I_0'(x) < I_0(x) - I_1(x)$ and hence $I_0(x) - I_1(x)$ cannot change sign (or be equal to zero) for $x > 0$.

Similarly, the second and the third equations imply that $I_2'(x) - I_1'(x) < \frac{1}{x}(I_1(x) - I_2(x))$ and together with the initial conditions this again implies that $I_1(x) > I_2(x)$ for all $x > 0$. Also note that

$$I_0(x) - I_2(x) = \frac{2}{x}I_1(x)$$

This, together with $I_0(x) - I_2(x) > I_0(x) - I_1(x)$ implies that

$$0 < \frac{I_0(x)}{I_1(x)} - 1 < \frac{I_0(x) - I_2(x)}{I_1(x)} = \frac{2}{x}$$

Therefore it follows that

$$\frac{I_0(x)}{I_1(x)} \to 1 \tag{31}$$

Definition 1 *Define the function* $L(v,\omega) = \frac{\sum_{x=0}^{\infty} \frac{v^x \omega^{x+1}}{x!(x+1)!}}{\sum_{x=0}^{\infty} \frac{v^x \omega^x}{x!x!}}$ *for any* $v > 0$, $\omega > 0$.

Lemma 5 $L(v,\omega) = \frac{\sqrt{\omega}}{\sqrt{v}} \frac{I_1(2\sqrt{v\omega})}{I_0(2\sqrt{v\omega})}$.

Proof. Note that

$$
\begin{aligned}
I_0(2\sqrt{v\omega}) &= 1 + \frac{(2\sqrt{v\omega})^2}{2^2\,(1!)^2} + \frac{(2\sqrt{v\omega})^4}{2^4(2!)^2} + \frac{(2\sqrt{v\omega})^6}{2^6(3!)^2} + \ldots \\
&= 1 + \frac{(v\omega)^1}{(1!)^2} + \frac{(v\omega)^2}{(2!)^2} + \frac{(v\omega)^3}{(3!)^2} + \ldots \\
&= \sum_{x=0}^{\infty} \frac{v^x \omega^x}{x!x!} \\
\frac{\sqrt{\omega}}{\sqrt{v}} I_1(2\sqrt{v\omega}) &= \frac{\sqrt{\omega}}{\sqrt{v}} \left(\frac{(2\sqrt{v\omega})}{2} + \frac{(2\sqrt{v\omega})^3}{2^3 1!2!} + \frac{(2\sqrt{v\omega})^5}{2^5 2!3!} + \ldots \right) \\
&= \frac{\sqrt{\omega}}{\sqrt{v}} \left(\frac{(\sqrt{v\omega})}{1} + \frac{(\sqrt{v\omega})^3}{1!2!} + \frac{(\sqrt{v\omega})^5}{2!3!} + \ldots \right) \\
&= \left(\frac{\omega}{1} + \frac{v\omega^2}{1!2!} + \frac{v^2\omega^3}{2!3!} + \ldots \right) = \sum_{x=0}^{\infty} \frac{v^x \omega^{x+1}}{x!(x+1)!}
\end{aligned}
$$

Lemma 6 $\partial L(v,\omega)/\partial v < 0, \partial L(v,\omega)/\partial \omega > 0$.

Proof.

$$
\partial L/\partial v = \frac{\left(\sum_{x=0}^{\infty} \frac{x v^{x-1}\omega^{x+1}}{x!(x+1)!}\right)\left(\sum_{x=0}^{\infty} \frac{v^x\omega^x}{x!x!}\right) - \left(\sum_{x=0}^{\infty} \frac{x v^{x-1}\omega^x}{x!x!}\right)\left(\sum_{x=0}^{\infty} \frac{v^x\omega^{x+1}}{x!(x+1)!}\right)}{\left(\sum_{x=0}^{\infty} \frac{v^x\omega^x}{x!x!}\right)^2}
$$

$$
\begin{aligned}
&= \left(\sum_{x=0}^{\infty} \frac{v^{x-1}\omega^{x+1}}{(x-1)!(x+1)!}\right)\left(\sum_{x=0}^{\infty} \frac{v^x\omega^x}{x!x!}\right) - \left(\sum_{x=0}^{\infty} \frac{v^{x-1}\omega^x}{(x-1)!x!}\right)\left(\sum_{x=0}^{\infty} \frac{v^x\omega^{x+1}}{x!(x+1)!}\right) \\
&= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} \frac{v^{x-1}\omega^{x+1}}{(x-1)!(x+1)!}\frac{v^y\omega^y}{y!y!} - \frac{v^{x-1}\omega^x}{(x-1)!x!}\frac{v^y\omega^{y+1}}{y!(y+1)!} \\
&= \sum_{x=0}^{\infty}\sum_{y\le x} \frac{v^{x+y-1}\omega^{x+y+1}}{(x-1)!x!(y-1)!y!}\left[\frac{1}{(x+1)y} - \frac{1}{(y+1)y} + \frac{1}{(y+1)x} - \frac{1}{(x+1)x}\right]
\end{aligned}
$$

A straightforward calculation shows that

$\left[\frac{1}{(x+1)y} - \frac{1}{(y+1)y} + \frac{1}{(y+1)x} - \frac{1}{(x+1)x}\right] = \frac{2xy - x^2 - y^2}{(x+1)y(y+1)x} < 0$

From Lemma 5 we know that

$$
L(v,\omega) = \frac{\sqrt{\omega}}{\sqrt{v}} \frac{I_1(2\sqrt{v\omega})}{I_0(2\sqrt{v\omega})}
$$

Therefore using equations (28)

$$\begin{aligned}\frac{\partial L}{\partial \omega} &= \frac{1}{2\sqrt{v\omega}}\frac{I_1(2\sqrt{v}\sqrt{\omega})}{I_0(2\sqrt{v}\sqrt{\omega})} + \\ &\quad \frac{I_0(2\sqrt{v}\sqrt{\omega}) - \frac{1}{2\sqrt{v\omega}}I_1(2\sqrt{v}\sqrt{\omega})}{I_0(2\sqrt{v}\sqrt{\omega})} - \frac{I_1(2\sqrt{v}\sqrt{\omega})^2}{I_0(2\sqrt{v}\sqrt{\omega})^2} \\ &= \frac{I_0(2\sqrt{v\omega})^2 - I_1(2\sqrt{v\omega})^2}{I_0(2\sqrt{v\omega})^2}\end{aligned}$$

Since $I_0(x) > I_1(x) > 0$ for $x > 0$ (Equation 30) the result follows. ■

Lemma 1. Suppose that $t_{1k}(s) > 0$ and $t_{2k}(s) > 0$ for all k. Then

$$\left|\frac{\Pr_k\{piv_1|s\}}{\Pr_k\{piv_2|s\}} - \sqrt{\rho_k(s)}\right| \to 0$$

as $k \to \infty$.

Proof. First note that by Lemma 5

$$\begin{aligned}\frac{\Pr_k\{piv_2|s\}}{\Pr_k\{piv_1|s\}} &= \frac{\sqrt{\nu_k t_{1k}(s)}}{\sqrt{\nu_k t_{2k}(s)}}\frac{I_0\left(2\nu_k\sqrt{t_{1k}(s)t_2(s)}\right)}{I_1\left(2\nu_k\sqrt{t_{1k}(s)t_{2k}(s)}\right)} \\ &= \sqrt{\rho_k(s)}\frac{I_0\left(2\nu_k\sqrt{t_{1k}(s)t_{2k}(s)}\right)}{I_1\left(2\nu_k\sqrt{t_{1k}(s)t_{2k}(s)}\right)}\end{aligned}$$

By Assumption 1 there is an $\epsilon > 0$ such that for all $x \in [-1, -1+\epsilon)$ an agent strictly prefers alternative 1 for all states s and similarly for all $x \in (1-\epsilon, 1]$ an agent strictly prefers alternative 2. This implies that $t_{jk}(s)$ is bounded away from zero for all k and all s. Now the result follows from Equation 31 ■

Lemma 2. Suppose that $\left(x^j_{1k} - x^j_{Mk}\right) > \epsilon > 0$ for $j = 1$ or $j = 2$ and all k. Then as $k \to \infty$ $\sum_{s\in T}\beta_k(s|piv_j) \to 1$ for $j = 1, 2$ and $\sum_{s\in T}\beta_k(s|piv) \to 1$.

Proof. By the assumption of strict SMLRP it follows that $t_1(s)$ is strictly decreasing while $t_{2k}(s)$ is strictly increasing in s with

$$\begin{aligned}t_{1k}(s) - t_{1k}(s') &> \epsilon' > 0 \qquad (32)\\ t_{2k}(s') - t_{2k}(s) &> \epsilon' > 0\end{aligned}$$

for $s' > s$ and for all k. Consider a convergent subsequence and let $t_j(s) = \lim_{k\to\infty} t_{jk}(s)$. Further, let

$$T = \left\{\arg\min\left|\sqrt{t_1(s)} - \sqrt{t_2(s)}\right|\right\} \qquad (33)$$

It follows from 32 that T consists either of a singleton or a pair of consecutive states. We now show that agents must believe that conditional on a vote being pivotal the state is almost certainly in T.

Using Lemma 5 we write

$$\begin{aligned}\frac{\Pr_k\{piv_2|s'\}}{\Pr_k\{piv_2|s\}} &= \frac{\sum_{x=0}^{\infty}\frac{e^{-\nu(t_{1k}(s')+t_{2k}(s'))}(\nu t_{2k}(s'))^x(\nu t_{1k}(s'))^x}{x!x!}}{\sum_{x=0}^{\infty}\frac{e^{-\nu(t_{1k}(s)+t_{2k}(s))}(\nu t_{2k}(s))^x(\nu t_{1k}(s))^x}{x!x!}} \\ &= \frac{e^{-\nu(t_{1k}(s')+t_{2k}(s'))}I_0\left(2\nu\sqrt{t_{1k}(s')t_{2k}(s')}\right)}{e^{-\nu(t_{1k}(s)+t_{2k}(s))}I_0\left(2\nu\sqrt{t_{1k}(s)t_{2k}(s)}\right)} \\ &= e^{\nu(t_{1k}(s)-t_{1k}(s')+t_{2k}(s)-t_{2k}(s'))}\frac{I_0\left(2\nu\sqrt{t_{1k}(s')t_{2k}(s')}\right)}{I_0\left(2\nu\sqrt{t_{1k}(s)t_{2k}(s)}\right)}\end{aligned}$$

Note that from Equation (29) it follows that

$$\frac{\Pr_k\{piv_2|s'\}}{\Pr_k\{piv_2|s\}} \longrightarrow e^{\nu_k(t_{1k}(s)-t_{1k}(s')+t_{2k}(s)-t_{2k}(s'))}\frac{\frac{e^{2\nu_k\sqrt{t_{1k}(s')t_{2k}(s')}}}{\sqrt{2\pi 2\nu_k\sqrt{t_{1k}(s')t_{2k}(s')}}}}{\frac{e^{2\nu_k\sqrt{t_{1k}(s')t_{2k}(s')}}}{\sqrt{2\pi 2\nu_k\sqrt{t_{1k}(s)t_{2k}(s)}}}} \tag{34}$$

$$= \frac{\sqrt{\sqrt{t_{1k}(s')t_{2k}(s')}}}{\sqrt{\sqrt{t_{1k}(s)t_{2k}(s)}}}e^{\nu_k\left(\left(\sqrt{t_{1k}(s)}-\sqrt{t_{2k}(s)}\right)^2-\left(\sqrt{t_{1k}(s')}-\sqrt{t_{2k}(s')}\right)^2\right)}$$

It follows from Lemma 1 that

$$\frac{\Pr_k\{piv_2|s'\}}{\Pr_k\{piv_2|s\}} \longrightarrow 0$$

if

$$\left(\sqrt{t_1(s')}-\sqrt{t_2(s')}\right) > \left(\sqrt{t_1(s)}-\sqrt{t_2(s)}\right). \tag{35}$$

Since 35 holds for any $s \in T$ and $s' \notin T$ it follows that

$$\frac{\sum_{s\in T}\Pr_k\{piv_1|s'\}}{\sum_{s\in S}\Pr_k\{piv_1|s\}} \longrightarrow 1.$$

Now Lemma 1 implies that $\frac{\sum_{s\in T}\Pr_k\{piv_2|s'\}}{\sum_{s\in S}\Pr_k\{piv_2|s\}} \longrightarrow 1$ and hence

$$\frac{\sum_{s\in T}\Pr_k\{piv|s'\}}{\sum_{s\in S}\Pr_k\{piv|s\}} \longrightarrow 1$$

Since

$$\beta_k(s|piv_j) = \frac{g(s)\Pr_k\{piv_j|s'\}}{\sum_{s\in S}g(s)\Pr_k\{piv_j|s'\}}$$

the Lemma follows ■

Lemma 3. Suppose Assumptions 1-6 hold. Consider a sequence of voting equilibria. Then $\beta_k(s^1|piv) + \beta_k(s^2|piv) \to 1$ as $k \to \infty$. Also, for $j = 1, 2$, $\beta_k(s^1|piv_j) + \beta_k(s^2|piv_j) \to 1$ as $k \to \infty$.

Proof. Consider a sequence of voting equilibria $(x^j_{mk}), j = 1, 2$. Consider a convergent subsequence of (x^j_{mk}).

Step 1: Suppose $x^j_{1k} - x^j_{Mk} \to 0$ for $j = 1, 2$. Then $\beta_k(s^1|piv) \to 1$ as $k \to \infty$ and $F(x(s^1)) = 1/2$.

Proof. Since

$$\begin{aligned} E(v(x^2_{1k}, s)|piv_1, 1) &= E(v(x^2_{Mk}, s)|piv_1, M) \\ E(v(x^1_{1k}, s)|piv_2, 1) &= E(v(x^1_{Mk}, s)|piv_2, M) \end{aligned}$$

it follows from the continuity of v that for $x \in [x^2_{Mk}, x^2_{1k}]$ and $x' \in [x^1_{Mk}, x^1_{1k}]$

$$\begin{aligned} E(v(x, s)|piv_1, 1) - E(v(x, s)|piv_1, M) &\to 0 \\ E(v(x', s)|piv_2, 1) - E(v(x', s)|piv_2, M) &\to 0 \end{aligned} \tag{36}$$

Since the signal satisfies Assumption 2 it follows that there exists an s, s' such that $\Pr_k(s = s|piv_1) \to 1$ and $\Pr_k(s = s'|piv_2) \to 1$. By equation 36 and the fact that $t_{jk}(s)$ is bounded away from zero it follows that $s = s'$.

We now show that it must be the case that $F(x(s')) = 1/2$. Suppose this were false. $F(x(s')) < 1/2$. Then the fraction of voters who prefer 2 in state s is less than $1/2 - \epsilon$ for some $\epsilon > 0$. Since $x^j_{1k} - x^j_{Mk} \to 0$ for $j = 1, 2$ it follows that the fraction of voters who vote for 2 is less than $1/2 - \epsilon/2$ for sufficiently large k for all s. This in turn implies that the probability that $n_2 = n_1$ is maximized at $\max_{s \in S} s$ contradicting the assumption that $F(x(s')) < 1/2$.

Step 2: Suppose $\lim \left(x^j_{1k} - x^j_{Mk}\right) > \epsilon > 0$ for $j = 1$ or $j = 2$. Then $\beta_k(s^1|piv) + \beta_k(s^2|piv) \to 1$ as $k \to \infty$.

Proof. By Lemma 2 we know that $\sum_T \beta_k(s|piv) \to 1$. Thus it is sufficient to show that $T \subseteq \{s^1, s^2\}$.

Consider a convergent subsequence and let $t_j(s) = \lim_{k \to \infty} t_{jk}(s)$. From the definition of T and the monotonicity of $t(s)$ it follows that if $t_1(s^1) \geq t_2(s^1)$ and $t_1(s^1) \leq t_2(s^1)$, then $T \subseteq \{s^1, s^2\}$. To see that this is indeed the case suppose

$\bar{s} = \arg\max_T s$ and $\bar{s} < s^1$(an analogous argument can be made if there is an $s'' \in T$ with $s'' > s^2$). By (32) we have

$$\epsilon' < \sqrt{t_2(s^1)} - \sqrt{t_1(s^1)}. \tag{37}$$

From the definition of T (2) we know $s^1 \notin T$ which implies

$$\beta_k(s \geq \bar{s}|piv) \to 0. \tag{38}$$

Consider any m such that $p(m|s^1) > 0$ and let

$$x_k(m) := \{x : E\{v(x,s)|piv, m\} = 0\}.$$

If $p(m|s) > 0$ for some $s \in T$ then (38) implies that a voter who receives signal m and conditions on the event that a vote is pivotal believes the probability of state $s \geq s^1$ is close to zero for large k. Thus if $p(m|s^1) > 0$ then for all $\epsilon > 0$ there is a k' such that for $k > k'$

$$x_k(m) > x(s^1) - \epsilon$$

But this in turn implies that $t_{1k}(s^1) - 1/2$ converges to a non-negative number and hence $\sqrt{t_1(s^1)} - \sqrt{t_2(s^1)} > 0$ contradicting (37) ■

Lemma 4. Suppose Assumptions 1-7 hold. Then there is an $\epsilon > 0$ and a $k' < \infty$ such that $\beta_k(s^1|piv) \geq \epsilon, \beta_k(s^2|piv) \geq \epsilon$ for all $k \geq k'$.

Proof. Suppose contrary to the Proposition that there is a subsequence such that

$$\beta_k(s^1|piv) \to 0$$

This implies that

$$\beta_k(s^2|piv) \to 1$$

and hence $t_{1k}(s^2) \to F(x(s^2))$ and $t_{2k}(s^2) \to 1 - F(x(s^2))$. Moreover, since all voters who receive a noisy signal behave as if state 2 occurred,

$$t_{1k}(s^1) \to F(x(s^2)) + (F(x(s^1)) - F(x(s^2)) \sum_{\bar{M}} p(m|s^1).$$

and

$$t_{2k}(s^1) \to 1 - F(x(s^2)) - (F(x(s^1)) - F(x(s^2)) \sum_{\bar{M}} p(m|s^1).$$

Let $\Delta = (F(x(s^1)) - F(x(s^2)) \sum p(m|s^1)$ and observe that

$$\left(\sqrt{F(x(s^2)) + \Delta} - \sqrt{1 - F(x(s^2)) - \Delta}\right)^2 < \left(\sqrt{F(x(s^2))} - \sqrt{1 - F(x(s^2))}\right)^2 \quad (39)$$

whenever $0 < \Delta < 1 - 2F(x(s^2))$.

Assumption 7 says that $0 \leq \Delta < 1 - 2F(x(s^2))$. Hence 39 holds with a strict inequality whenever $\Delta > 0$ and therefore we have that for sufficiently large k

$$\left(\sqrt{t_{1k}(s^1)}) - \sqrt{t_{2k}(s^1)}\right)^2 < \left(\sqrt{t_{1k}(s^2)}) - \sqrt{t_{2k}(s^2)})\right)^2. \quad (40)$$

To see that (40) also holds (for large k) when $\Delta = 0$ observe that in this case $t_{1k}(s^j) \to F(x(s^2)) < 1/2$ and $t_{2k}(s^j) \to 1 - F(x(s^2)) > 1/2$ since every agent behaves as if state s^2 has occurred.. Thus we have that for large k

$$\begin{aligned} t_{1k}(s^2) &< t_{1k}(s^1) < 1/2 \\ t_{2k}(s^2) &> t_{2k}(s^1) > 1/2 \end{aligned}$$

and (40) follows for large k.

From the proof of Lemma 2 we know that

$$\frac{\Pr_k\{piv_j|s^1\}}{\Pr_k\{piv_j|s^2\}} \to \frac{\sqrt{\sqrt{t_{1k}(s^1)t_{2k}(s^1)}}}{\sqrt{\sqrt{t_{1k}(s^2)t_{2k}(s^2)}}} e^{\nu_k\left(\left(\sqrt{t_{1k}(s^2)} - \sqrt{t_{2k}(s^2)}\right)^2 - \left(\sqrt{t_{1k}(s^1)} - \sqrt{t_{2k}(s^1)}\right)^2\right)} \quad (41)$$

Inequality (40) then implies that for every $\epsilon > 0$ there is a k' such that for $k > k'$

$$\beta_k(s^1|piv) \geq \frac{\sqrt{\sqrt{t_{1k}(s^1)t_{2k}(s^1)}}}{\sqrt{\sqrt{t_{1k}(s^2)t_{2k}(s^2)}}} \frac{g(s^1)}{g(s^2)} \beta_k(s^2|piv) - \epsilon$$

Since $g(s)$ and $t_{jk}(s)$ are bounded away from zero for all s this in turn contradicts $\beta_k(s^1|piv) \to 0$. (The argument for the case where $\beta_k(s^2|piv) \to 0$ is analogous.) ■

Proof of Proposition 5: Since Assumption 7 holds we can apply Lemma 3 to establish that $\beta_k(s^1|piv) > \epsilon > 0$ and $\beta_k(s^2|piv) > \epsilon > 0$ for large k. By Assumption 1 this in turn implies that $x^j_{1k} - x^j_{Mk} > \epsilon > 0$ for $j = 1$ or $j = 2$. for large k. Thus, there exists an $\epsilon' > 0$ such that $\rho_k(s) - \rho_k(s') \geq \epsilon'$ for all $s < s'$. From Lemma 1 we know that

$$\beta_k(s^1|piv_2) - \frac{g(s^1)\frac{\Pr_k(piv_1|s^1)}{\sqrt{\rho_k(s^1)}}}{g(s^1)\frac{\Pr_k(piv_1|s^1)}{\sqrt{\rho_k(s^1)}} + g(s^2)\frac{\Pr_k(piv_1|s^2))}{\sqrt{\rho_k(s^2)}}} \to 0$$

and hence there is an k' and an ϵ'' such that

$$\beta_k(s^1|piv_2) < \beta_k(s^1|piv_1) - \epsilon'' \tag{42}$$

for all $k > k'$. Consider an agent who receives a signal m with $p(m|s^1) > 0$ and $p(m|s^2) > 0$. Since (42) holds it follows that there is an $\epsilon''' > 0$ and a k' such that for $k > k'$

$$E\{v(x,s)|piv_1,m\} < E\{v(x,s)|piv_2,m\} - \epsilon'''$$

Since $x^1_{m,k}$ is given by $E\{v(x^1_{m,k},s)|piv_1,m\} = 0$ and $x^2_{m,k}$ is given by $E\{v(x^2_{m,k},s)|piv_2,m\} = 0$ the result now follows. ■

Proof of Proposition 6: *Step 1:* First we demonstrate that the fraction of players who abstain and who receive a signal that allows them to exclude a state in the set $\{s^1, s^2\}$ converges to zero. I.e., we show that the fraction of voters who abstain and who receive a signal with the property $p(m|s^1)p(m|s^2) = 0$ converges to zero.

To prove Step 1 we distinguish two cases:

Case 1: $p(m|s^1) > 0$ and $p(m|s^2) = 0$ or $p(m|s^1) = 0$ and $p(m|s^2) > 0$. Suppose W.L.O.G that $p(m|s^1) > 0$ and $p(m|s^2) = 0$. In this case receiving message m permits the agent to perfectly discriminate between states s^1 and s^2. Now it follows that

$$E(v(x,s)|piv_1,m) - E(v(x,s)|piv_2,m) \to 0$$

since

$$\begin{aligned} E(v(x,s)|piv_1,m) &\to v(x,s^1) \\ E(v(x,s)|piv_2,m) &\to v(x,s^1) \end{aligned}$$

Case 2. $p(m|s^1) = p(m|s^2) = 0$. In this case receiving the message m tells the agent that the state is neither s^1 or s^2. It follows from SMLRP that either $p(m|s) = 0$ for any $s < s^1$ or for any $s > s^2$. W.L.O.G let $p(m|s) = 0$ for any $s < s^1$. Let $\bar{s} < s^1$ be the largest state such that $p(\bar{s}|m) > 0$. Now it follows from $\rho(s)$ strictly decreasing in s that

$$\begin{aligned} E(v(x,s)|piv_1,m) &\to v(x,\bar{s}) \\ E(v(x,s)|piv_2,m) &\to v(x,\bar{s}) \end{aligned}$$

and the fraction of types receiving message m that abstain goes to zero.

Step 2. In Step 2 we demonstrate that the fraction of voters who receive a signal with the property $p(m|s^1)p(m|s^2) > 0$ and who abstain can be made arbitrarily small if η is sufficiently small. These are voters whose signal does not allow them to exclude a state in the set $\{s^1, s^2\}$. To prove this we demonstrate that for small η the beliefs conditional on the event piv_1 are very close to the beliefs conditional on the event piv_2.

Assume that the critical state is s^1 i.e., $s^1 \in \arg\min_s |F(x(s)) - 1/2|$. (The case where $s^2 = \arg\min_s |F(x(s)) - 1/2|$ is entirely analogous.)

Observe that $\liminf t_{2k}(s^1) \geq 1 - F(x(s^1))$ since every voter with type $x > x(s^1)$ prefers candidate 2 if the state is in the set $s \in \{s^1, s^2\}$ and by Lemma 2 we know that $\beta_k(s^1|piv_j) + \beta_k(s^2|piv_j) \to 1$ as $k \to \infty$. Since candidate 1 wins the election in state s^1 with probability close to one (Proposition 4) it follows that $t_{1k}(s^1) \geq t_{2k}(s^1)$. Since the expected vote shares must be less than or equal to 1 in each state it must be the case that $t_{1k}(s^1) \leq F(x(s^1))$ for sufficiently large k and we get the following inequality:

$$0 \geq \sqrt{t_{2k}(s^1)} - \sqrt{t_{1k}(s^1)} \geq \sqrt{1 - F(x(s^1))} - \sqrt{F(x(s^1))}.$$

Therefore,

$$1 \leq \liminf \sqrt{\rho_k(s^1)} \leq \limsup \sqrt{\rho_k(s^1)} \leq \frac{\sqrt{F(x(s^1))}}{\sqrt{1 - F(x(s^1))}}. \tag{43}$$

Further, observe that since vote shares are monotone it follows that

$$t_2(s^2) \geq 1 - F(x(s^1)).$$

By Lemma 3 $\beta_k(s^1|piv_j) \geq \varepsilon, \beta_k(s^2|piv_j) \geq \varepsilon > 0$ and by Lemma 2 this implies that

$$\left(\sqrt{t_{1k}(s^2)} - \sqrt{t_{2k}(s^2)}\right) - \left(\sqrt{t_{2k}(s^1)} - \sqrt{t_{1k}(s^1)}\right) \to 0$$

As a result we get that

$$\begin{aligned} \liminf \sqrt{t_{1k}(s^2)} &\geq \sqrt{1 - F(x(s^1))} + \sqrt{1 - F(x(s^1)} - \sqrt{F(x(s^1)} \\ &= 2\sqrt{1 - F(x(s^1))} - \sqrt{F(x(s^1)} \end{aligned}$$

and

$$1 \geq \limsup \sqrt{\rho_k(s^2)} \geq \liminf \sqrt{\rho_k(s^2)} \geq \frac{2\sqrt{1-F(x(s^1))} - \sqrt{F(x(s^1)}}{\sqrt{1-F(x(s^1))}}. \tag{44}$$

Equations 43 and 44 imply that for any $\varepsilon > 0$ there is an $\eta > 0$ such that for $F(x(s^1)) - 1/2 \leq \eta$ we have that

$$\begin{aligned} 1 &\leq \liminf \sqrt{\rho_k(s^1)} \leq \limsup \sqrt{\rho_k(s^1)} \leq 1 + \varepsilon \\ 1 &\geq \limsup \sqrt{\rho_k(s^2)} \geq \liminf \sqrt{\rho_k(s^2)} \geq 1 - \varepsilon. \end{aligned} \tag{45}$$

Lemma 1 and the fact that $p(m|s^1)p(m|s^2) > 0$ then implies that for every ε' there is an η such that for $F(x(s^1)) - 1/2 \leq \eta$ implies that

$$\beta_k(s^1|piv_1, m) - \beta_k(s^1|piv_2, m) \leq \varepsilon'.$$

for sufficiently large k and therefore for every ε'' there is an η such that for $F(x(s^1)) - 1/2 \leq \eta$ and for m

$$E\{v(x,s)|piv_1, m\} - E\{v(x,s)|piv_2, m\} \leq \varepsilon'' \tag{46}$$

which yields the result. ■

9 Appendix B

The system of equations for the noisy signals example is given below:

$$
\begin{aligned}
A &= 1 - t_1(1) - t_1(2) \\
A_i &= .5(x_1^2 - x_1^1) \\
A_u &= -x_\emptyset^1 \\
x_\emptyset^1 &= \frac{2}{1 + \sqrt{\frac{t_1(1)}{t_1(2)}}} - 1 \\
x_1^1 &= \frac{2}{1 + \frac{1-p}{p}\sqrt{\frac{t_1(1)}{t_1(2)}}} - 1 \\
x_1^2 &= \frac{2}{1 + \frac{1-p}{p}\sqrt{\frac{t_1(2)}{t_1(1)}}} - 1 \\
t_1(1) &= .5q\left((1 + x_1^1)p + (1 - x_1^2)(1-p)\right) + .5(1-q)(1 + x_\emptyset^1) \\
t_1(2) &= .5q\left((1 + x_1^1)(1-p) + (1 - x_1^2)p\right) + .5(1-q)(1 + x_\emptyset^1)
\end{aligned}
$$

The system of equations for the biased distribution of information example is given below:

$$
\begin{aligned}
A_L &= .5(x_\emptyset^2 - x_\emptyset^1) \\
x_\emptyset^1 &= \frac{2}{1 + \gamma_1} - 1 \\
x_\emptyset^2 &= \frac{2}{1 + \gamma_2} - 1 \\
\gamma_1 &= \sqrt{\frac{t_1(1)t_2(2)}{t_1(2)t_2(1)}}\gamma_2 \\
\sqrt{t_1(1)} - \sqrt{t_2(1)} &= \sqrt{t_2(2)} - \sqrt{t_1(2)} \\
t_1(1) &= .5 + .5(1 + x_\emptyset^1) \\
t_1(2) &= .5(1 + x_\emptyset^1) \\
t_2(1) &= .5(1 - x_\emptyset^2) \\
t_2(2) &= .5 + .5(1 - x_\emptyset^2)
\end{aligned}
$$

References

[1] Abramson, Paul R., John H. Aldrich and David W. Rohde. 1991. *Change and Continuity in the 1988 Elections.* Revised edition. Washington D.C.: CQ Press.

[2] Austen-Smith, David and Jeff Banks. 1996. "Information Aggregation, Rationality and the Condorcet Jury Theorem" *American Political Science Review*, Vol. 90, 34-45.

[3] Feddersen, Timothy J. 1992. "A Voting Model Implying Duverger's Law and Positive Turnout." *American Journal of Political Science.* Vol. 36. No. 4, 938-62.

[4] Feddersen, Timothy J., and Wolfgang Pesendorfer 1996. "The Swing Voter's Curse". *American Economic Review*, Vol. 86, No. 3, 408-424..

[5] Feddersen Timothy J., and Wolfgang Pesendorfer. 1997. "Voting Behavior and Information Aggregation in Elections with Private Information", *Econometrica,* Vol. 65, No. 5 1029-1058.

[6] Grofman, Bernard Ed.. 1993. *Information, Participation and Choice: An Economic Theory of Democracy in Perspective.* Ann Arbor: University of Michigan Press.

[7] Ghiradato, Paolo and Jonathan Katz. 1997. "Indecision Theory: An Informational Theory of Roll-Off." Mimeograph. Caltech.

[8] Klevorick, Alvin K., Michael Rothschild and Christopher Winship. 1984. "Information Processing and Jury Decision making." *Journal of Public Economics* Vol. 23, 245-78.

[9] Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review.* Vol. 91, No. 1, 1-14.

[10] Milgrom, Paul. 1981. "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, Vol. 12, 380-391.

[11] Myerson, Roger (1997a). "Population Uncertainty and Poisson Games". Mimeograph. Northwestern University.

[12] Myerson, Roger (1997b). "Extended Poisson Games and the Condorcet Jury Theorem". Forthcoming in *Games and Economic Behavior.*

[13] Niemi, Richard G., and Herbert F. Weisberg. 1984. *Controversies in Voting Behavior.* Second edition. Washington D.C.: CQ Press.

[14] Palfrey, Thomas R., and Howard Rosenthal. 1983. "A Strategic Calculus of Voting." *Public Choice,* Vol. 41, No. 1, 7-53.

[15] Palfrey, Thomas R., and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *American Political Science Review,* Vol. 79, No. 1, 62-78.

[16] Riker, William H., and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review,* Vol. 62, No. 1, 25-42.

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