Voting Behavior and Information Aggregation in Elections with Private Information[1]

Timothy Feddersen
and
Wolfgang Pesendorfer
Northwestern University
Evanston, Il 60208

October 1994
revised
August 1996

Abstract

We analyze two-candidate elections in which voters are uncertain about the realization of a state variable that affects the utility of all voters. Each voter has noisy private information about the state variable. We show that the fraction of voters whose vote depends on their private information goes to zero as the size of the electorate goes to infinity. Nevertheless, elections fully aggregate information in the sense that the chosen candidate would not change if all private information were common knowledge. Equilibrium voting behavior is to a large extent determined by the electoral rule, i.e., if a candidate is required to get at least x percent of the vote in order to win the election then in equilibrium this candidate gets very close to x percent of the vote with probability close to one. Finally, if the distribution from which preferences are drawn is uncertain, then elections will generally *not* satisfy full information equivalence and the fraction of voters who take informative action does not converge to zero.

1   Introduction

A strong argument for elections is that society may be collectively better informed about the relative quality of a set of alternatives than any individual. Elections provide a mechanism for aggregating private information ensuring a better collective decision. This idea inspired some of the earliest mathematical models of voting in elections and dates back at least to Condorcet[1]. The set of environments in which elections might usefully aggregate private information about the relative quality of a pair of alternatives goes well beyond the jury setting that was the focus of Condorcet's work. In most elections voters have common values with respect to some characteristic of the alternatives and are privately informed about this characteristic. Consider the following examples:

- An election is held to decide whether or not to increase funding for a local public good. Voters have different valuations for the public good and are uncertain about the cost or the quality of the proposed plan. One particular example is referenda on school funding. While voters' willingness to spend money on schools differs, all agree that better student performance is preferable. There is uncertainty about the degree to which increased spending translates into student performance.

- Voters must decide between an incumbent and a challenger. Voters' preferences have both a private and a common value component. The private value component is voter preferences over the candidates' ideological positions. A common value component is the "character" of the candidates. Candidates with good character can be relied upon to stick closely to their announced positions while those with poor character cannot. Risk averse voters all prefer a candidate with better character. Voters are differentially informed about the record of each candidate and therefore possess private information.[2]

---

[1]For a discussion of Condorcet's Jury Theorem and extensions see Ladha (1992); Miller (1986); and Young (1988,1994).

[2]In the literature on macroeconomic performance and elections competence is frequently introduced as a common value component of voter preferences. See, for example, Alesina, Londregan and Rosenthal (1993), Persson and Tabellini (1990), Rogoff (1990).

- Voters in Presidential primaries are concerned not only about the policy positions of the competing candidates but also about each competitor's probability of winning the general election. All the primary voters prefer any of the candidates running in the primary to any of the candidates from the other party. Voters possess private information about the candidates' electability.[3]

The traditional approach to the question of how well elections aggregate information assumes that voters have identical preferences and behave "naively"[4], i.e., each voter behaves as if her choice alone determines the outcome. However, naive voting is not generally an equilibrium of the corresponding voting game.[5] Voters face a decision problem that is similar to the problem facing bidders in a common value auction. In both elections and auctions an agent's action affects her payoff only in very particular circumstances. As is well known, bidders in a common value auction must condition their belief about the value of the object on the event that their bid is the highest. Similarly, voters must condition their beliefs about the quality of the alternatives on the event that one vote can change the election outcome, i.e., a vote is pivotal. The following example illustrates the problem.

A community must vote on a proposal to increase school funding. There are two equally likely states of the world: the proposal works (w) (e.g., it improves test scores, reduces dropout rates, etc.) or it does not (nw). Everyone in the community favors the proposal in state w and is opposed otherwise. None of the voters knows the state of the world but each voter gets one of two signals: in state w every voter gets the signal w with probability 0.6. In state nw every voter gets the signal nw with probability 0.6. The proposal passes if at least 2/3 vote in favor. Suppose all voters vote "naively", i.e., in favor if they receive signal w and opposed otherwise. Then in a large election, whenever a vote is pivotal (i.e., 2/3 of the voters have voted for the

---

[3]The fact that voters' decisions about which candidate to support are influenced by how they believe others assess the candidates is known as the "bandwagon effect": Candidates who are seen winning early primaries gain support in later primaries (see for example Bartels (1988)). The bandwagon effect is often thought to be a feature of preferences: voters like to support winning candidates just as sports fans enjoy rooting for winning teams. In contrast, we are suggesting that the phenomenon is due to voters learning about the relative merits of the candidates. The bandwagon effect is prima facie evidence that electoral results reveal useful information to voters.

[4]See Ladha (1992); Miller (1986); and Young (1988,1994). See Austen-Smith and Banks 1994; Myerson 1994b; and Klevorick et. al. 1985 for exceptions to the assumption of naive voting.

[5]Austen-Smith and Banks (1995) and Feddersen and Pesendorfer (1995).

proposal), the state is almost certainly w and everyone should vote in favor![6]

As in the above example, we consider a population of voters that uses an election to choose one of two alternatives (labeled Q and A). In contrast to the above example, we allow voters to have different preferences over the two alternatives. Each voter's payoff depends on her preference type, on a state of nature, and on the winning alternative. Preference types are drawn independently from a given distribution whereas the state of nature is common for all voters. Voters know their own preference types but are uncertain about the state of nature. Every voter receives a signal that provides information about the realization of the state of nature. Voting is costless and voters can either vote for Q or for A. Alternative Q wins if the fraction of voters voting for it is at least q. We analyze the *voting equilibria* of this game (symmetric Nash equilibria in which voters do not use weakly dominated strategies).

In a voting equilibrium preference types can be divided into three groups: those types who always vote for Q, those who always vote for A, and those who change their vote depending on their private signal. We say the latter types take *informative action*.

Our first three results analyze voting behavior and information aggregation in relatively simple environments in which voters are uncertain about a one-dimensional state variable.

Theorem 1 demonstrates the inherent tension between information aggregation and informative voting. We show that the fraction of voters who take informative action goes to zero as the size of the electorate goes to infinity. The result that almost no voters take informative action in large elections would seem to put into grave doubt the supposed utility of elections as information aggregation devices. Our next two results show that this is not the case.

Theorem 2 shows that for a wide variety of preference distributions large elections are almost always very close. Theorem 3 shows that elections satisfy *full information equivalence*: with probability arbitrarily close to one, the alternative that would have been chosen if all the private information were common knowledge is selected. This result may appear paradoxical in light of our first result. While the *fraction* of the

---

[6]It should be clear that this example does not depend on the fact that we chose a 2/3 rule rather than simple majority rule. It could easily be modified to show that naive voting is generally not a best response to a population voting naively also in the case of simple majority rule.

electorate's signals revealed in equilibrium goes to zero, the *number* of voters who reveal their signal goes to infinity so that in the limit all information is revealed. Theorem 2 guarantees that the election will be decided by those taking informative action, and thus, large elections effectively aggregate private information.

We use a series of examples to illustrate the effect of relaxing our key assumptions. We also illustrate in section 5.1 that, in contrast to strategic voting, naive voting typically fails to lead to full information equivalence.

Our last result examines the implications of more complicated information environments. We demonstrate that if there is additional uncertainty about the distribution of preferences then elections will generally *not* satisfy full information equivalence and the fraction of voters who take informative action does not converge to zero. The degree to which the election fails to meet the full information equivalence requirement is parameterized by the level of uncertainty about the distribution of preferences. When this uncertainty is small, the election mechanism almost satisfies full information equivalence.

Our approach is related in some respects to the approach taken by Lohmann (1993) and Austen-Smith (1989). Lohmann uses a similar framework to analyze the effects of private information on costly participation in political protest movements while Austen-Smith examines the incentives for strategic voting in small two-alternative elections. Neither Lohmann nor Austen-Smith considered the asymptotic properties of their models. Our results are also related to the literature on information aggregation in auctions: Milgrom (1979); Wilson (1977); and Pesendorfer and Swinkels (1995). In another related paper Palfrey (1985) analyzes information aggregation in a Cournot model.

2   The Model

We analyze a two alternative election. Alternatives are denoted by $j \in \{Q, A\}$. There are $n+1$ voters indexed by $i \in \{1, ..., n+1\}$. A voter's utility depends on a preference parameter $x \in [-1, 1] = X$, the chosen alternative j, and the state $s \in [0, 1]$. We denote by $u(j, s, x)$ the utility function of voters. Let

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

4

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

Each voter knows her preference type but is uncertain about the realization of the state. By $G(s)$ we denote the probability distribution that describes the prior beliefs about the state s. Each voter receives a signal $\sigma \in \{1, ..., M\} \equiv \Sigma$ from an information service $k \in \{1, ..., K\}$. We assume that conditional on state s being realized, the signal that voter i receives is independent of the signal that voter j receives. Thus we can define the function $p_k(\sigma|s)$ which denotes the probability that a voter receives signal σ if $s \in [0, 1]$ is realized and the voter is served by information service k.

A voter's type is characterized by a preference parameter and an information service. Let $T \equiv [-1, 1] \times \{1, ..., K\}$ denote the type space. Let F be a probability distribution over T, where $F(x, k)$ denotes the probability that the type is in the set $[-1, x] \times k$. Let $F_X(x) = \sum_{k=1}^{K} F(x, k)$. The assumption of K information services allows us to introduce correlation between access to information services and preference types.

Nature selects the electorate by choosing $n+1$ voter types independently according to the probability distribution F. Each voter knows her own type but is uncertain about the other voters' types. The distribution F is common knowledge.

A voter can choose Q or A. Let $0 < q < 1$ be a fixed parameter. If the number of voters who choose Q is larger than or equal to $(n + 1) \cdot q$ then Q is the outcome. Otherwise, A is the outcome.

We make the following assumptions:

Assumption 1 $v(x, s)$ *is continuous and increasing with* $|v(x, s) - v(x, s')| \geq \kappa|s - s'|$ *and* $|v(x, s) - v(x', s)| \geq \kappa|x - x'|$ *for some* $\kappa > 0$. *Moreover,* $v(-1, s) < 0, v(1, s) > 0$ *for all* s.

Assumption 2 G *has a density* g *and there is an* $\alpha > 0$ *such that* $1/\alpha > g(s) > \alpha$ *for all* $s \in [0, 1]$.

Assumption 3 $F(x, k)$ *is continuously differentiable in* x *and* $f(x, k)$ *denotes the derivative. There is an* $\alpha > 0$ *such that* $\sum_{k=1}^{K} f(x, k) > \alpha$ *for all* $x \in X$.

Assumption 4 *(Monotone Likelihood Ratio Property).* *If $\sigma > \sigma'$ and $s > s'$ then*
$p_k(\sigma'|s')p_k(\sigma|s) \geq p_k(\sigma|s')p_k(\sigma'|s)$ *for all k.*

Assumption 5 *(Limited Information).* *There is an $\alpha > 0$ such that $p_k(\sigma|s) > \alpha$ for*
all (k, s).

Assumption 6 *nq is an integer.*

Assumption 1 says that the utility difference between alternative A and alternative
Q is continuous and strictly increasing in x and s. Furthermore, voters with preference
parameters at the boundary of X prefer one of two alternatives irrespective of the
state s.

Assumption 2 ensures that every state is in the support of the prior and the relative
likelihood of any pair of states $g(s)/g(s')$ is bounded above and below. Assumption
3 implies that every preference type is in the support of F_X.

Assumption 4 says that the signal satisfies the monotone likelihood ratio prop-
erty (MLRP). One implication is that for $s' > s$, $p_k(\cdot|s')$ first order stochastically
dominates $p_k(\cdot|s)$ (Witt (1980)).[7] In addition, a higher signal indicates to the voter
that a higher state should be expected for *any* prior. More precisely, for $\sigma' > \sigma$
the distribution over states conditional on σ' first order stochastically dominates the
probability distribution over states conditional on σ (Milgrom (1981)).

Assumption 5 says that a voter cannot exclude any state if she receives a particular
signal. Assumption 6 is purely for notational convenience.[8]

3   Strategies and Equilibrium

A pure strategy for voter i, π_i, is a measurable function from her type and her signal
to a vote choice, i.e., $\pi_i : T \times \Sigma \to \{Q, A\}$ and a mixed strategy, $\bar{\pi}_i$, is a measurable
function from a voter's type and her signal to the probability of voting for candidate
Q, i.e., $\bar{\pi}_i : T \times \Sigma \to [0, 1]$.

We define a *voting equilibrium* $\bar{\pi}^*$ to be a symmetric Nash equilibrium in which
no voter uses a weakly dominated strategy.

---

[7]Note that since $p_k(.|s')$ first order stochastically dominates $p_k(.|s)$ for $s' > s$ it follows that
$p_k(1|s)$ is non-increasing in s and $p_k(M|s)$ is non-decreasing in s.

[8]The only change in the analysis when nq is not an integer is that the expression nq must be
replaced with "largest integer that is smaller or equal to nq".

The only time a voter can influence the outcome of the election is if a vote is *pivotal*, i.e., exactly qn of the other n voters voted for Q. A voter will choose Q if conditional on a vote being pivotal the expected payoff of alternative Q is larger than the expected payoff of alternative A.

Given a symmetric strategy profile $\bar{\pi}$ we can compute the probability that a vote is pivotal as a function of the state s. Let

$$t(s, \bar{\pi}) = \sum_{k=1}^{K} \sum_{\sigma=1}^{M} p_k(\sigma|s) \int_X \bar{\pi}(x, k, \sigma) f(x, k) dx \qquad (2)$$

denote the probability that a randomly selected voter votes for Q in state s. Let *piv* denote the event that a vote is pivotal. The probability that a vote is pivotal in state s is given by:

$$\Pr(piv|s, \bar{\pi}) = \binom{n}{qn} \cdot t(s, \bar{\pi})^{qn} \cdot (1 - t(s, \bar{\pi}))^{n-qn} \qquad (3)$$

When $1 > t(s, \bar{\pi}) > 0$ for all s then $\Pr(piv|s, \bar{\pi}) > 0$ for all s, and therefore, the probability distribution over states conditional on being pivotal is given by

$$\beta(s|piv, \bar{\pi}) = \frac{\Pr(piv|s, \bar{\pi})g(s)}{\int_0^1 \Pr(piv|w, \bar{\pi})g(w)dw}. \qquad (4)$$

Similarly, the probability distribution over states conditional on being pivotal and observing signal σ from service k is given by:

$$\beta(s|piv, \bar{\pi}, \sigma, k) = \frac{\Pr(piv|s, \bar{\pi})p_k(\sigma|s)g(s)}{\int_0^1 \Pr(piv|w, \bar{\pi})p_k(\sigma|w)g(w)dw} = \frac{\beta(s|piv, \bar{\pi})p_k(\sigma|s)}{\int_0^1 \beta(w|piv, \bar{\pi})p_k(\sigma|w)dw} \qquad (5)$$

Let $E(v(x, s)|piv, \bar{\pi}, \sigma, k)$ denote the expectation of $v(x, s)$ with respect to $\beta(\cdot|piv, \bar{\pi}, \sigma, k)$. Since the signal satisfies the MLRP, $\beta(\cdot|piv, \bar{\pi}, \sigma, k)$ first order stochastically dominates $\beta(\cdot|piv, \bar{\pi}, \sigma', k)$ for $\sigma > \sigma'$ (see Milgrom (1981)), and hence $E(v(x, s)|piv, \bar{\pi}, \sigma, k)$ is non-decreasing in σ.

A strategy is characterized by cutpoints if for every information service and every signal there is a cutpoint x_σ^k such that the voter chooses Q whenever the preference type is smaller than x_σ^k and A otherwise. If the cutpoints x_σ^k are non-increasing in σ then we say that the strategy can be characterized by ordered cutpoints.

Definition 1 *A strategy $\bar{\pi}$ is characterized by* ordered cutpoints *if for every information service k there are cutpoints $(x_\sigma^k)_{\sigma=1,\dots M}$ with the property that $1 > x_1^k \geq \dots \geq x_M^k > -1$ and $\bar{\pi}(x, k, \sigma) = 1$ for $x < x_\sigma^k$, $\bar{\pi}(x, k, \sigma) = 0$ for $x > x_\sigma^k$.*

7

Observe that if a strategy is characterized by ordered cutpoints then voters of type (x, k) with $x < x_M^k$ vote for candidate Q *irrespective of their private signal.* Similarly, voter types (x, k) with $x > x_1^k$ vote for candidate A irrespective of their private signal. Voter types (x, k) with $x \in (x_M^k, x_1^k)$ change their vote depending on the signal they receive. We say such types take *informative action.*

Proposition 1 *Suppose Assumptions 1-6 hold. Then there exists a voting equilibrium $\bar{\pi}^*$. Every voting equilibrium $\bar{\pi}^*$ is characterized by ordered cutpoints (x_σ^k) such that $E(v(x_\sigma^k, s)|piv, \bar{\pi}^*, \sigma, k) = 0$ for all (σ, k). Moreover, $t(s, \bar{\pi}^*)$ is non-increasing in s with $0 < t(s, \bar{\pi}^*) < 1$ for all s.*

Proof. see Appendix. ∎

The fact that voting equilibrium can be characterized by ordered cutpoints follows from the fact that $E(v(x, s)|piv, \bar{\pi}, \sigma, k)$ is strictly increasing in x (Assumption 1) and non-decreasing in σ (Assumption 4). The cutpoints allow us to simplify (2) to

$$t(s, \bar{\pi}^*) = \sum_{k=1}^{K} \sum_{\sigma=1}^{M} p_k(\sigma|s) F(x_\sigma^k, k). \tag{6}$$

The final part of the Proposition 1 now follows since $p_k(\cdot|s')$ first order stochastically dominates $p_k(\cdot|s)$ for $s' > s$ and $F(x_\sigma^k, k)$ is non-increasing in σ.

4 Voting Equilibria in Large Elections

In this section we analyze the limiting properties of a sequence of elections with n voters, where $n \to \infty$. Along any such sequence only the number of voters changes while the information structure, the payoffs, and all other parameters stay fixed. In the following we superscript our notation with n to indicate that we are working with elements of a sequence. We assume that for each element of the sequence qn is an integer (Assumption 6). As before, this assumption is made for convenience only.

4.1 Large Elections and Informative Action

In this section we show that informative action by the electorate creates an incentive for individual voters not to vote informatively. This leads to the central result in this section: in a voting equilibrium with a large number of voters only a vanishing fraction of the electorate takes informative action.

We assume that the probability of receiving signal σ in state s is a continuous function of s.

Assumption 7 $p_k(\sigma|s)$ *is continuous in s for all k and for all* σ.

Assumption 7 implies that for any symmetric strategy profile, $\bar{\pi}^n$, $t(s, \bar{\pi}^n)$ is continuous. For the remainder of the paper we will assume that Assumption 7 holds. In Example 2 we demonstrate how the following results (in particular Theorem 1) fail if Assumption 7 is violated.

As we argued above, voters must evaluate candidates in the event a vote is pivotal. In the following we characterize the probability distribution over states conditional on the event that a vote is pivotal. We define $S_\eta(\bar{\pi}^n)$ as the set of states for which the expected vote share of alternative Q is within η of the vote share of the state that minimizes $|t(s, \bar{\pi}^n) - q|$. More precisely,

$$S_\eta(\bar{\pi}^n) = \left\{ s \in [0,1] : |t(s, \bar{\pi}^n) - q| \leq \min_s |t(s, \bar{\pi}^n) - q| + \eta \right\} \tag{7}$$

If there is a state for which $t(s, \bar{\pi}^n) = q$ then $S_\eta(\bar{\pi}^n)$ simply denotes the set of states for which the expected vote share of alternative Q is within η of q. Lemma 1 demonstrates that for large n, conditional on a vote being pivotal, the probability distribution over states must be concentrated on those states which generate an expected vote share closest to q.

Lemma 1 *Suppose Assumptions 1-7 hold. Consider a sequence of strategy profiles,* $(\bar{\pi}^n)$ *such that* $t(s, \bar{\pi}^n)$ *is continuous, non-increasing and* $0 < t(s, \bar{\pi}^n) < 1$. *For any* $\eta > 0$, $S_\eta(\bar{\pi}^n)$ *is an interval of states with the property that* $\int_{S_\eta(\bar{\pi}^n)} \beta(s|piv, \bar{\pi}^n) \to 1$.

Proof. see Appendix. ∎

To get an intuition for Lemma 1 observe that the number of votes for Q follows a binomial distribution with parameter $t(s, \bar{\pi}^n)$, where $t(s, \pi^n)$ is non increasing in the state. If a vote is pivotal then qn voters out of a population of size n have voted for Q. Thus, if $q \in [t(1, \bar{\pi}^n), t(0, \bar{\pi}^n)]$ and n is large, then the beliefs about the parameter $t(s, \bar{\pi}^n)$ conditional on a vote being pivotal must be concentrated around q. This implies that the beliefs about the state conditional on a vote being pivotal must be concentrated around those states that produce a value for $t(s, \bar{\pi}^n)$ closest to q. If there

is no state such that $t(s, \bar{\pi}^n) = q$ the beliefs must be concentrated around those states where $t(s, \bar{\pi}^n) - q$ is minimized.

Fix a symmetric strategy profile $\bar{\pi}$. Consider a subset of states $S \subset [0, 1]$ with the property that conditional on a vote being pivotal the state is in S with probability 1. If $p_k(\sigma|s)$ is constant on the set S for all signals σ then the information service k does not discriminate between the states in S. Since the state is in S whenever a vote is pivotal information service k is *redundant*. Now consider a sequence of symmetric strategy profiles. We say service k is *asymptotically redundant* if it is redundant in the limit. The following definition makes this precise.

Definition 2 *Fix a sequence of symmetric strategy profiles* $(\bar{\pi}^n)$*. Information service k is* asymptotically redundant *if for every $\epsilon > 0$ there is a sequence of sets (S_ϵ^n), with $S_\epsilon^n \subset [0, 1]$ for all n, such that $\int_{S_\epsilon^n} \beta(s|piv, \bar{\pi}^n) ds \to 1$ and $|p_k(\sigma|s) - p_k(\sigma|s')| < \epsilon$ for any $s, s' \in S_\epsilon^n$ and for all σ .*

In Lemma 2 we assume that the expected fraction of voters who receive their signal from service k and vote informatively is bounded away from zero. We demonstrate that this implies that information service k is asymptotically redundant.

Lemma 2 *Suppose Assumptions 1-7 hold. Consider a sequence of symmetric strategy profiles, $\bar{\pi}^n$, that can be characterized by ordered cutpoints with the property that for some k and some $\delta > 0$, $F(x_1^{k,n}, k) - F(x_M^{k,n}, k) > \delta$ for all n. Then k is asymptotically redundant. In particular, there is a constant $c < \infty$ such that $|p_k(\sigma|s) - p_k(\sigma|s')| < \eta c$ for any $s, s' \in S_\eta(\bar{\pi}^n)$ and for all σ.*

Proof. see Appendix. ∎

To provide an intuition for Lemma 2, first note that by Assumption 7 $t(s|\bar{\pi}^n)$ is continuous. Since $\bar{\pi}^n$ is characterized by ordered cutpoints $t(s|\bar{\pi}^n)$ non-increasing in s and $0 < t(s|\bar{\pi}^n) < 1$. Therefore, we can apply Lemma 1 to conclude that $\int_{S_\eta(\bar{\pi}^n)} \beta(s|piv, \bar{\pi}^n) \to 1$ for every $\eta > 0$. It is therefore sufficient to show that there is a constant $c < \infty$ such that $|p_k(\sigma|s) - p_k(\sigma|s')| < \eta c$ for any $s, s' \in S_\eta(\bar{\pi}^n)$ and for all signals σ.

Suppose there are only two signals, $\sigma = 1, 2$. Now consider a pair of states, $s, s' \in S_\eta(\bar{\pi}^n)$, with $s' > s$. By the MLRP $p_k(1|s) \geq p_k(1|s')$. Recall that all the

10

voters with preference types in the interval $(x_1^{k,n}, x_2^{k,n})$ choose Q if they receive signal 1 and A if they receive signal 2. By assumption $F(x_1^{k,n}, k) - F(x_2^{k,n}, k) > \delta$, and thus, the expected fraction of voters who choose Q decreases by at least $\delta \cdot (p_k(1|s) - p_k(1|s'))$ between s and s'. Since the decrease in the expected vote share must be less than 2η it follows that $p_k(1|s) - p_k(1|s')$ must be less than $2\eta/\delta$ which establishes the Lemma.

In Lemma 3 we consider a sequence of voting equilibria that have the property that information service k is asymptotically redundant. Under this hypothesis we show that the strategy of voters who receive their information from service k is almost independent of the signal they receive. More precisely, we show that $x_1^{k,n} - x_M^{k,n} \to 0$.

Lemma 3 *Suppose Assumption 1 holds. Consider a sequence of voting equilibria $(\bar{\pi}^{*n})$ and assume that information service k is asymptotically redundant. Then, the cutpoints corresponding to $\bar{\pi}^{*n}$ satisfy $x_1^{k,n} - x_M^{k,n} \to 0$.*

Proof. see Appendix. ∎

To get an intuition for Lemma 3, observe that by the definition of asymptotic redundance, we find a sequence of subsets of states with the property that $\Pr(s \in S_\epsilon^n | piv, \bar{\pi}^n) \to 1$ and that the signals from service k discriminate very little between the states in S_ϵ^n if ϵ is small. Therefore, the expected payoff difference between voting for Q and A, conditional on a vote being pivotal, is almost independent of the signal from service k. By Assumption 1, $v(x, s)$ is strictly increasing in x at a rate larger than κ. As a consequence, there is at most a small interval of preference types x with the feature that the voter prefers alternative Q for one signal and alternative A for another signal. Therefore, the range of preference parameters for which a voter takes informative action must be small if ϵ is small, and the Lemma follows.

Theorem 1 says that the expected fraction of voters who take informative action in equilibrium must converge to zero. Furthermore, because every preference type is served by some information service (Assumption 3), the cutpoints of at least one information service must converge.

Theorem 1 *Suppose Assumptions (1)-(7) hold. Let $(\bar{\pi}^{*n})$ be a sequence of voting equilibria, and let $(x_\sigma^{k,n})$ be the corresponding cutpoints. Then for all k, $F(x_1^{k,n}, k) - F(x_M^{k,n}, k) \to 0$ and for some k, $x_1^{k,n} - x_M^{k,n} \to 0$.*

11

The proof of Theorem 1 is straightforward. Suppose, contrary to Theorem 1, that the expected fraction of voters who receive their information from service k and take informative action is bounded away from zero. Then, Lemma 2 implies that k is asymptotically redundant and so Lemma 3 implies that the cutpoints for service k must converge. But then the expected fraction of voters who receive their information from service k and take informative action converges to zero, establishing a contradiction.

Proof. By Proposition 1, in any voting equilibrium the cutpoints are ordered for all k, and hence, the first hypothesis of Lemma 3 is satisfied. Also note that F does not have any mass points. Lemmas 2 and 3 imply that in any voting equilibrium $F(x_1^{k,n}, k) - F(x_M^{k,n}, k) \to 0$. This follows since by Lemma 2, if $F(x_1^{k,n}, k) - F(x_M^{k,n}, k)$ stays bounded away from zero along some subsequence then information service k is redundant. Lemma 3 then implies that $x_1^{k,n} - x_M^{k,n} \to 0$ which in turn implies that $F(x_1^{k,n}, k) - F(x_M^{k,n}, k) \to 0$ resulting in a contradiction.

To prove the final part of Theorem 1 let x^{*n} satisfy

$$E(v(x^{*n}, s)|piv) = 0$$

Note that $x_M^{k,n} \leq x^{*n} \leq x_1^{k,n}$ for all k. By Assumption 3 there is a k' such that $f(x^{*n}, k') \geq \alpha/K$. We will show that $x_1^{k',n} - x_M^{k',n} \to 0$. Suppose $x_1^{k',n} - x_M^{k',n} \geq \delta > 0$ for all n. Continuity of $f(., k')$ then implies that $F(x_1^{k',n}, k') - F(x_M^{k',n}, k') \geq \eta > 0$ for some $\eta > 0$ which yields the desired contradiction. ∎

4.2   Voting Behavior and Full Information Equivalence

In this section we show the following results: Theorem 2 demonstrates that in equilibrium large elections must be very close, i.e., the fraction of the electorate that supports alternative Q must be very close to the critical fraction q. Theorem 3 demonstrates that elections effectively aggregate information. More precisely, we show that large elections almost always choose the alternative that would have been chosen if the state variable were common knowledge. In order to show these results, we require two preliminary Lemmas.

Lemma 4 provides the converse of Lemma 3. It says that if the cutpoints of an information service converge then the information service must be asymptotically redundant.

Lemma 4 *Suppose Assumptions 1-7 hold. Consider a sequence of voting equilibria $(\bar{\pi}^{*n})$ and the corresponding cutpoints $(x_\sigma^{k,n})$. If $x_1^{k,n} - x_M^{k,n} \to 0$ for some k then k is asymptotically redundant.*

Proof. see Appendix. ∎

As an intuition, observe that if the cutpoints for information service k converge to one point, it must be that the expected utility difference between the alternatives, conditional on a vote being pivotal, changes very little as the voter's signal changes. This can only be the case if the signal adds very little information once a voter conditions on being pivotal. Hence, information service k is redundant.

The following results use the strict monotone likelihood ratio property (SMLRP). The SMLRP implies that sampling many signals from *any* information service makes it possible to determine the state with great accuracy.

Assumption 8 *(Strict Monotone Likelihood Ratio Property). For all k,*

$$\frac{p_k(M|s)}{p_k(1|s)}$$

is strictly increasing in s and $p_k(\sigma|s)$ satisfies MLRP.

Assumption 8 implies that every information service discriminates between any pair of states. This assumption excludes a situation (as in Example 3 below) in which some preference types have access to an informative information service that satisfies the Strict Monotone Likelihood Ratio Property (SMLRP) while others do not.

Lemma 5 states that if the SMLRP holds then in a large election voters can predict the state with great accuracy if a vote is pivotal. More precisely, the distribution over states, conditional on a vote being pivotal, converges to a distribution that is arbitrarily concentrated around some state s^n, and s^n solves

$$\max_{s \in S} \Pr(piv|s, \bar{\pi}^{*n}). \tag{8}$$

Lemma 5 *Consider a sequence of voting equilibria $(\bar{\pi}^{*n})$ and suppose Assumptions 1-8 hold. Then there is a unique state s^n that solves $\max_{s \in S} \Pr(piv|s, \bar{\pi}^{*n})$. For every $\delta > 0$, $\int_{\{s:|s-s^n| \leq \delta\}} \beta(s|piv, \bar{\pi}^{*n})ds \to 1$.*

Proof. see Appendix. ∎

To get an intuition for Lemma 5, recall that by Lemma 4 at least one information service k must be asymptotically redundant. Thus, for every $\epsilon > 0$ we find a sequence S_ϵ^n with the property that (1) $\int_{S_\epsilon^n} \beta(s|piv, \bar{\pi}^{*n})ds \to 1$ and (2) the probability of receiving any signal from service k varies by less than ϵ on S_ϵ^n. If Assumption 8 holds then every information service discriminates between every pair of states. Therefore, (2) can only hold if the maximum distance between any pair of states in S_ϵ^n is small. Hence, the probability distribution over states, conditional on a vote being pivotal, must be arbitrarily concentrated around one state for n large enough. Since s^n maximizes the probability that a vote is pivotal, it follows that the conditional probability distribution must be concentrated around s^n.

Theorem 2 says that in large elections the expected vote share of alternative Q will be very close to q. Let x_Q denote the preference type who is indifferent between Q and A in state $s = 1$ and let x_A denote the preference type who is indifferent between A and Q in state $s = 0$. Then, by the assumption that voters never play weakly dominated strategies in a voting equilibrium, all preference types below x_Q always vote for Q, and all types above x_A always vote for A. Therefore, if $\bar{\pi}^{*n}$ is a sequence of voting equilibria, then $F_X(x_A) \geq t^n(s, \bar{\pi}^{*n}) \geq F_X(x_Q)$.

Theorem 2 *Suppose Assumptions (1)-(8) hold and suppose that $F_X(x_A) > q > F_X(x_Q)$. Consider a sequence of symmetric voting equilibria $(\bar{\pi}^{*n})$. Then for all $\eta > 0$ there is an $\bar{n}$ such that for $n > \bar{n}$, $|q - t(s, \bar{\pi}^{*n})| < \eta$ for all s.*

In a large election the actual vote shares are close to the expected vote shares with high probability. Theorem 2 therefore implies that large elections will be close with probability close to one in every state. Note that Theorem 2 holds for a wide variety of preference distributions.

To give an intuition for Theorem 2, suppose there is a state such that the expected vote share of Q is less than $q - \eta$ for all n. By Theorem 1 vote shares change very little as a function of s if the electorate is large. Therefore, for large enough n the vote share of Q is less than $q - \eta/2$ for *all* states. Since the expected vote share of Q is decreasing in s (Proposition 1), it must be that $s = 0$ minimizes the difference between the expected vote share and q. But then (by Lemma 5), conditional on a vote being pivotal, the state is close to 0 with probability close to one. Since the fraction

of voters who prefer Q in state $s = 0$ is larger than q by assumption, the expected vote share of Q must be larger than q. This establishes the desired contradiction.

Proof. Suppose that $q \geq t(s, \bar{\pi}^{*n}) + \eta$ for some s and for all n along some subsequence. From Equation 6 we get that

$$t(0, \bar{\pi}^{*n}) - t(1, \bar{\pi}^{*n}) \leq \max_k F_X(x_1^{k,n}) - F_X(x_M^{k,n})$$

Since the right hand side of the above inequality tends to zero, by Theorem 1 we can find an n' such that for $n > n'$

$$t^n(0, \bar{\pi}^{*n}) - t^n(1, \bar{\pi}^{*n}) \leq \eta/2$$

and hence $q > t^n(s) + \eta/2, \forall s$. Since $t^n(s, \bar{\pi}^{*n})$ is decreasing in s, it follows that $\Pr(piv|s, \bar{\pi}^{*n})$ is maximized at $s = 0$ and hence $s^n = 0$. By Lemma 5 this implies that for every $\epsilon' > 0$

$$\int_0^{\epsilon'} \beta\left(s \,|piv, \bar{\pi}^{*n}\right) \to 1$$

But then, for every $\epsilon > 0$ there is an n' such that for $n > n'$ all voters with preference parameters $x > x_Q + \epsilon$ must vote for A, and therefore $F_X(x_Q + \epsilon) \geq t(s, \bar{\pi}^{*n})$ for all $\epsilon > 0$. Since $F_X(x_Q) < q$ we obtain a contradiction to the hypothesis that $t(s, \bar{\pi}^{*n}) \geq q + \eta$ for all n. (For $q \leq t^n(s) - \eta$ an analogous contradiction can be obtained.) ∎

The probability with which large elections choose the alternative that would have been chosen if the state variable were common knowledge serves as a natural benchmark for the performance of elections as information aggregation mechanisms.[9] We say large elections satisfy *full information equivalence* if the alternative that wins a large election is almost certainly the same as the alternative that would have been chosen if the electorate were fully informed about the state variable.

In order to formally define full information equivalence, let

$$x^* = F_X^{-1}(q). \tag{9}$$

If $q = 1/2$ then x^* is the expected preference parameter of the median voter. For arbitrary q we call the voter type with preference parameter x^* the expected q-median.

---

[9]Alternatively, we could use as a benchmark the situation in which all the private signals are common knowledge among voters. Note, however, that Assumption 8 and the law of large numbers imply that in a large electorate knowing all signals is almost equivalent to actually knowing the true state of nature.

In an election in which the state is known the actual q-median's preferred alternative wins. In a large election the actual q-median's preference parameter is very close to x^* with probability close to one. Therefore, full information equivalence is satisfied in a large election if the expected q-median's preferred alternative wins with probability close to one.

Clearly, the alternative preferred by the expected q-median depends on the state. Let

$$s^* = \arg\min_{s \in S} |v(x^*, s)|. \tag{10}$$

If $v(x^*, s^*) = 0$, then s^* is the state in which the expected $q-$median voter is indifferent between the two alternatives. If $v(x^*, s^*) > 0$, then there is no state in which type x^* prefers Q to A, and hence, $s^* = 0$. Similarly, if $v(x^*, s^*) < 0$, then $s^* = 1$. Informally, full information equivalence will be satisfied if Q is almost certainly the winner when $s < s^*$ and A almost certainly wins otherwise. We now formally define full information equivalence as follows:

Definition 3 *We say that a sequence of strategy profiles satisfies* full information equivalence *if for all $\epsilon > 0$, there is an n such that for $n' > n$, the following holds: if $s < s^* - \epsilon$ then Q is elected with probability greater than $1 - \epsilon$; if $s > s^* + \epsilon$ then A is elected with probability greater than $1 - \epsilon$.*

We now prove that full information equivalence holds for any sequence of voting equilibria.

Theorem 3 *If Assumptions 1-8 hold then every sequence of voting equilibria satisfies full information equivalence.*

To give an intuition, consider the case in which there is a state that makes the expected $q-$median voter indifferent between the two alternatives, i.e., $v(x^*, s^*) = 0$. Lemma 5 implies that, conditional on a vote being pivotal, the distribution over states puts almost all the weight on the neighborhood of one state s^n. Thus, voters essentially behave as if state s^n has occurred. First we show that $\lim s^n = s^*$. To see this, note that if, e.g., $v(x^*, s^n) > \epsilon > 0$ then the fraction of voters who prefer Q in state s^n is smaller than and bounded away from q. But then, the fraction of voters who vote for Q must be smaller than and bounded away from q which contradicts

Theorem 2. From Lemma 5 we know that the election is tied only if the state is very close to s^*. Since the vote share of alternative Q is strictly decreasing in s this can only be the case if for $s < s^* - \epsilon$ alternative Q wins with probability close to one, and for $s > s^* + \epsilon$ alternative A wins with probability close to one.

Proof. *Case 1:* If $F_X(x_Q) > q$ then, since all voters with $x < x_Q$ will vote for Q, alternative Q will be chosen with probability close to one for large n. Moreover, this choice satisfies full information equivalence because voters with $x < x_Q$ prefer alternative Q in every state s. A similar argument shows that the Theorem is satisfied if $F_X(x_A) < q$.

Case 2: Suppose that $F_X(x_Q) < q < F_X(x_A)$. Lemma 5 implies that there is an s^n such that for all $\delta > 0$ and for all $\sigma \in \{1, ..., M\}$, $k \in \{1, ..., K\}$

$$\int_{\{s:|s-s^n|\leq\delta\}} \beta\left(s|piv, \bar{\pi}^{*n}, \sigma, k\right) \to 1$$

We must show that $v(x^*, s^n) \to 0$. Suppose that along some subsequence

$$v(x^*, s^n) \to \epsilon > 0$$

Let $\hat{x} < x^*$ be such that $v(\hat{x}, s^n) \to 0$ along that subsequence. Then, since $\beta\left(s|piv, \bar{\pi}^{*n}\right)$ is arbitrarily concentrated around s^n for large n, it follows that all voters with preference type $x > (x^* + \hat{x})/2$ strictly prefer to vote for A. But this implies that $t^n(s) \leq F_X((x^* + \hat{x})/2) < q - \epsilon'$ for some $\epsilon' > 0$ for all n which contradicts Theorem 2. Thus, we have established that $v(x^*, s^n) \to 0$ and hence $s^n \to s^*$.

It remains to be shown that for large n whenever $s > s^* + \epsilon$, the probability that A is chosen is larger than $1 - \epsilon$, and whenever $s < s^* - \epsilon$, the probability that A is chosen is smaller than ϵ. Let $w(m|s, \pi^{*n})$ denote the probability that m voters choose alternative Q if the state is s and the strategy profile is π^{*n}. Recall that $\Pr(piv|s, \pi^{*n})$ is a single peaked function of s. Thus $\beta\left(s|piv, \bar{\pi}^{*n}\right)$ can only be concentrated around s^* if for $s > s^* + \epsilon$, $\frac{\Pr(piv|s,\pi^{*n})}{\Pr(piv|s^*,\pi^{*n})} \to 0$. Therefore, for every $\epsilon > 0$ there is an n' such that for $n > n'$

$$\frac{w(qn|s, \pi^{*n})}{w(qn|s^*, \pi^{*n})} = \frac{\Pr(piv|s, \pi^{*n})}{\Pr(piv|s^*, \pi^{*n})} < \epsilon.$$

Since $t(s, \bar{\pi}^{*n})/t(s^*, \bar{\pi}^{*n}) < 1$ for $s > s^* + \epsilon$ it follows that for $m > qn$

$$\begin{aligned}
\frac{w(m|s, \pi^{*n})}{w(m|s^*, \pi^{*n})} &= \frac{t(s, \bar{\pi}^{*n})^m \cdot (1 - t(s, \bar{\pi}^{*n}))^{n-m}}{t(s^*, \pi^{*n})^m \cdot (1 - t(s^*, \pi^{*n}))^{n-m}} \\
&\leq \frac{t(s, \bar{\pi}^{*n})^{qn} \cdot (1 - t(s, \bar{\pi}^{*n}))^{n-qn}}{t(s^*, \pi^{*n})^{qn} \cdot (1 - t(s^*, \pi^{*n}))^{n-qn}} < \epsilon
\end{aligned}$$

for n sufficiently large. And hence for all $s > s^* + \epsilon$

$$\sum_{m>qn} w(m|s, \pi^{*n}) < \epsilon \sum_{m>qn} w(m|s^*, \pi^{*n}) \leq \epsilon$$

This implies that A will be chosen with probability larger than $1 - \epsilon$. An analogous argument shows that for $s < s^* - \epsilon$ the probability that A is chosen is smaller than ϵ. ∎

5 Examples

For the following examples we assume that

$$v(x, s) = -1 + 2x + 2s. \tag{11}$$

Voter preference parameters are distributed uniformly.[10] Further, we assume that for each information service k there are two signals $\sigma \in \{1, 2\}$.

5.1 Strategic Versus Naive Voting

Suppose that voters *naively* act as assumed in the literature on the Condorcet Jury Theorem: each voter behaves as if her choice alone determines the outcome. Thus, a voter of type (x, k) with signal σ votes for Q if $-1 + 2x + 2E[s|\sigma, k] < 0$ and for A if $-1 + 2x + 2E[s|\sigma, k] > 0$. In this case, a larger fraction of voters vote informatively than in a voting equilibrium, and hence, more information is revealed by the vote share. However, in contrast to voting equilibria, naive voting does not imply full information equivalence.

Suppose $g(s) = 2s$. Suppose, further, that there is one information service and $p(2|s) = s$ and that $q = .5$. A simple calculation shows that $E[s|\sigma = 2] = 3/4$ and $E[s|\sigma = 1] = 1/2$, and hence, under naive voting all preference types $x > 0$ vote for A independent of their private signal. But this implies that in a large electorate, A will be elected with probability close to one for all $s > 0$. Full information equivalence requires that A is elected only if $s > 1/2$.[11]

---

[10]The model setup used in this example is nearly identical to the setup used in Lohmann (1993) with the key differences that we assume no costs to participate and uncertainty about the location of voter ideal points. As mentioned above Lohmann does not analyze the assymptotic properties of her model.

[11]In Feddersen and Pesendorfer (1994) we show that for the preferences and the information

5.2 Example 2

In this example we demonstrate how a failure of Assumption 7 (continuity of $p_k(\sigma|s)$) may lead to a voting equilibrium in which the fraction of voters who take informative action does not converge to zero, and hence, Theorems 1 and 2 do not hold. However, voting equilibria still satisfy full information equivalence.

Suppose that $g(s) = 1$, $q = 1/2$, and there is one information service that is described by

$$p(1|s) = \left\{ \begin{array}{ll} 1 - \alpha & \text{if} \quad s < 1/2 \\ \alpha & \text{if} \quad s > 1/2 \end{array} \right.$$

where $\alpha < 1/2$. The unique voting equilibrium is given by the cutpoints $x_1 = 1/4 - \alpha/2$ and $x_2 = -1/4 + \alpha/2$. The equilibrium strategies in this example are independent of n.

To see why the prescribed strategies are an equilibrium, note that $|t(s, \bar{\pi}^n) - q| = \frac{1}{8}(2\alpha - 1)^2$ for every s, and therefore, conditioning on the event that a vote is pivotal provides no information. As a consequence, the signal is informative, conditional on a vote being pivotal, and private information remains valuable for all n.

5.3 Example 3

We now give an example that demonstrates how Theorem 3 depends on the SMLRP. What is critical in this example is that voters with preference types around the expected $q-$median voter do not have access to an information service that discriminates between states as precisely as voters on the extremes. Suppose there are two information services, $k \in \{1, 2\}$ and $p_1(1|s) = 1 - s$ and $p_2(1|s) = 1/2$, i.e., information service 2 is not informative. Further, let $q = 1/2$. The distribution F is such that all voters with preference parameters $x \in [-1, -.2] \cup [.2, 1]$ have access to information service 1 with probability 1, while all voters with preference parameters $x \in [-1/6, 1/6]$ have access to information service 2 with probability 1.

We assume that $g(s) = 1$ for all s. Consider the cutpoints $x_\sigma^2 = 0$ for all $\sigma = 1, 2$ and $x_1^1 = 1/6, x_2^1 = -1/6$. Since all voters who receive their information from service

---

service given in this example naive voting does not lead to full information equivalence whenever $E(s) \neq 1/2$. See also Austen-Smith and Banks (1995) and Myerson (1995) for a discussion of strategic voting and Jury theorems.

1 have preference types outside the interval $(-1/6, 1/6)$, no voter takes informative action in this strategy profile.

To see that this is an equilibrium, note that $E(s|piv, \bar{\pi}, 1, 1) = 1/3$ and $E(s|piv, \bar{\pi}, 2, 1) = 2/3$. Hence, $-1 + 2x + 2E(s|piv, \bar{\pi}, 1, 1) < 0$ for $x < -1/6$ and $-1 + 2x + 2E(s|piv, \bar{\pi}, 2, 1) > 0$ for $x > 1/6$.

Therefore, irrespective of the state, each alternative has a 50% chance of winning the election and full information equivalence is not satisfied.

6 Uncertainty about the Distribution of Voters' Preferences

Up to now we have assumed that voters know the distribution from which preferences are drawn. In this section, we show how introducing uncertainty about this distribution upsets the results. To simplify the analysis we assume that voters are uncertain about the expected fraction of partisans, i.e., voters who choose either alternative Q or alternative A irrespective of the state. Let F be a probability distribution that satisfies Assumption 3. In this section we assume that the distribution function according to which nature selects the electorate depends on the parameter $\lambda \in [0, 1]$ and is given by:

$$H_\lambda(x) = \begin{cases} (1 - \phi)F(x) + \phi(1 - \lambda) & \text{if } -1 \leq x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

Thus, H_λ has $\phi(1 - \lambda)$ mass at -1 and $\phi\lambda$ mass at $+1$. We assume that for all $\lambda \in [0, 1]$,

$$H_\lambda(x_Q) < q < H_\lambda(x_A) \tag{12}$$

which implies that the expected fraction of voters who prefer one alternative irrespective of the state is always smaller than the fraction necessary to elect that alternative.

In the first stage of the game, nature chooses both s and λ independently. By $\ell(\lambda)$ we denote the density that describes the prior beliefs about the state λ. We assume that there is an $\alpha > 0$ such that $1/\alpha > \ell(\lambda) > \alpha$ for all $\lambda \in [0, 1]$. After choosing the state (s, λ), nature selects an electorate by taking n independent draws from the distribution H_λ.[12]

---

[12]In Feddersen and Pesendorfer (1994) voters, in addition to the signal σ also get a signal that provides noisy information about λ. All the following results also hold in this case and hence we omit the second signal.

For the remainder of this section we assume that Assumptions 1,2,4-8 hold and that there is one information service described by $p(\sigma|s)$. It is straightforward to verify that, in this modified environment, Proposition 1 still holds.[13] Thus, there exists a voting equilibrium, and every voting equilibrium can be described by ordered cutpoints.

Let

$$x(\lambda) = H_\lambda^{-1}(q) \tag{13}$$

denote the expected $q-$ median voter if λ is realized. Further, let $s(\lambda)$ be such that $v(s(\lambda), x(\lambda)) = 0$. Thus, $s(\lambda)$ is the state at which the expected $q-$ median voter is indifferent between the two alternatives if λ is realized. Note that (12) implies that $s(\lambda)$ is well-defined. Moreover, $s(\lambda)$ is a strictly decreasing function of λ.

A sequence of voting equilibria, therefore, satisfies *full information equivalence* if for all $\epsilon > 0$, there is an n such that if $n' > n$ then the following holds for every λ: if $s < s(\lambda) - \epsilon$ then Q is elected with probability greater than $1 - \epsilon$; if $s > s(\lambda) + \epsilon$ then A is elected with probability greater than $1 - \epsilon$;

The first part of Theorem 4 says that the set of voters who use their private signal σ stays bounded away from zero in measure when the distribution of the electorate is uncertain. The second part says that full information equivalence does not hold. This latter result will be shown to hold for a typical utility function $v(x, s)$. To make this precise, denote the set of utility functions that satisfy Assumption 1 by P. Endow P with the topology of uniform convergence. We say that a property holds for a *generic* utility function if it holds for all $v \in O \subset P$ where O is open and dense.

The third part of the theorem says that if the uncertainty about the distribution of preferences is small, as compared to the uncertainty about s, (if ϕ is small) then full information equivalence nearly holds. This should be seen as a continuity check. As the uncertainty about the distribution of preference types becomes small, the information aggregation results of the previous section are a good approximation of a situation where voters are uncertain also about the distribution of preference types.

---

[13]Note that if a voter learns her preference parameter x and if $x \in (-1, 1)$ she does not learn anything about the realization of λ since the likelihood of observing $x \in (-1, 1)$ is independent of λ. The only voters who get information about the realization of λ by observing their preference parameter are voters with $x \in \{-1, +1\}$. However, these voters are partisans and will always vote for A (in the case of $x = 1$) or Q (in the case of -1) by Assumption 1. See Alesina and Rosenthal (1995) for a similar formulation.

Theorem 4 *Suppose $k = 1$, Assumptions 1,2,4-8 hold, and the preference types are drawn according to the procedure described in this section. Consider a sequence of voting equilibria $(\bar{\pi}^{*n})$. Then*

(i) *There is an $\eta > 0$ such that $x_1^n - x_M^n > \eta$ for all n.*

(ii) *There exists an $O \subset P$, where O is open and dense, such that for every $v \in O$ the election does not satisfy full information equivalence.*

(iii) *For every $\epsilon > 0$ there is a $\phi' > 0$ and an n' such that if $\phi < \phi', n > n'$ then there are (s_0, s_1) with the following properties: (1) if $s < s_0$ then Q is elected with probability greater than $1 - \epsilon$; if $s > s_1$ then A is elected with probability greater than $1 - \epsilon$ and (2) $|s_0 - s_1| \leq \epsilon$ and $s_0 \leq s(\lambda) \leq s_1$ for all λ.*

Proof. see Appendix.

To provide an intuition for the proof of part (i) of Theorem 4, suppose for (s, λ) the expected vote share of alternative Q is q. Since the vote share for Q is a strictly decreasing function of λ, if the vote share for Q is responsive to changes in s, we can decrease s and simultaneously increase λ so that the expected vote share stays unchanged. Conditional on being pivotal, a voter believes that one of the states has occurred for which the expected vote share of alternative Q is q. Thus, whether or not the vote share is responsive to changes in s, it is impossible to invert the map from states to vote counts. There is now a whole interval of states such that the expected vote share of alternative Q is q. Therefore, the beliefs over states, conditional on being pivotal, do not converge to a degenerate distribution. But then the private information of voters provides useful information, and hence, the set of voters who take informative action does not converge to zero in measure.

To provide an intuition for part (ii), note that full information equivalence requires that for states $(s(\lambda), \lambda)$, the expected vote share of Q must be close to q for a large electorate, since otherwise, close to $(s(\lambda), \lambda)$, the wrong candidate is chosen with high probability. This follows from the fact that the derivative of the expected vote share of A with respect to s is uniformly bounded above for all n. We show that for a generic choice of v, equilibrium strategies allow too few degrees of freedom to have the expected vote share equal to q for all states $(s(\lambda), \lambda)$.

Underlying Theorem 4 is the fact that there are two random variables both of which are correlated with the votes for each alternative. This makes it impossible for voters to invert back from votes to the payoff relevant state variable. As an alternative to the introduction of uncertainty about the distribution of voter preferences, we could allow s to be a two dimensional variable and get similar results.[14]

7   Conclusion

Taken together, our results demonstrate the importance of the information environment in determining the effectiveness of elections as information aggregation mechanisms. If voters are uncertain about a one-dimensional state variable, strategic voting results in effective information aggregation. If, for example, there is additional uncertainty about the distribution of preferences or if the payoff relevant uncertainty is of higher dimension, then electoral mechanisms do not perform so well. The importance of the dimensionality of uncertainty for the performance of elections suggests that future research should focus on the events that precede elections–nominating procedures, campaigns, polls, etc.,–as such events determine the information environment.

8   Appendix

Proof of Proposition 1: First we demonstrate that any best response to a weakly undominated strategy can be characterized by ordered cutpoints.

Note that by Assumption 1 there is an $\epsilon > 0$ such that $v(x, s) < 0$ for all s if $x \in [-1, -1 + \epsilon]$ and $v(x, s) > 0$ for all s if $x \in [1 - \epsilon, 1]$. Therefore, for any strategy $\bar{\pi}$ that is not weakly dominated, types with $x \in [-1, -1 + \epsilon]$ vote for Q irrespective of the signal and types with $x \in [1 - \epsilon, 1]$ vote for A irrespective of their signal. This in turn implies for any strategy that is not weakly dominated $\Pr(piv|s) > 0$ for all s and hence $\beta(s|piv, \bar{\pi}, \sigma, k)$ is well defined.

By Assumption 1 v is strictly increasing and continuous in x. In addition $v(-1, s) < 0, v(1, s) > 0$ for all s. Thus, it follows that there is a unique cutpoint $x_\sigma^k \in [-1, 1]$

---

[14]A result similar to the one given in Theorem 4 will hold as long as the two dimensions are not perfectly correlated, i.e., there does not exist a function $a(s_1, s_2)$ such that $v(x, s_1, s_2) = v'(x, a(s_1, s_2))$ for some v'.

such that

$$E[v(x_\sigma^k, s)|piv, \bar{\pi}, \sigma, k] = 0 \tag{14}$$

(the expectation is taken with respect to $\beta(s|piv, \bar{\pi}, \sigma, k)$). Clearly, $1-\epsilon \geq x_\sigma^k \geq -1+\epsilon$. If $x < x_\sigma^k$ then a voter type (x, k) who receives signal σ strictly prefers to vote for Q, and if $x > x_\sigma^k$ then a voter type (x, k) who receives signal σ strictly prefers to vote for A.

By the MLRP, it follows that $\beta(s|piv, \bar{\pi}, \sigma', k)$ first order stochastically dominates $\beta(s|piv, \bar{\pi}, \sigma, k)$ whenever $\sigma' > \sigma$. Since $v(x, s)$ is increasing in s, it follows that

$$E[v(x, s)|piv, \bar{\pi}, \sigma', k] \geq E[v(x, s)|piv, \bar{\pi}, \sigma, k]$$

and $x_\sigma^k \geq x_{\sigma'}^k$. Thus, any best response to a weakly undominated strategy can be characterized by ordered cutpoints.

Using the above characterization of best responses we now demonstrate existence of a voting equilibrium. By the argument above, the cutpoints corresponding to any best response to a weakly undominated strategy profile are in the interval $[-1+\epsilon, 1-\epsilon]$. Thus, to demonstrate existence, consider the following function

$$\psi : [-1 + \epsilon, 1 - \epsilon]^{KM} \rightarrow [-1 + \epsilon, 1 - \epsilon]^{KM}$$

To any KM-tuple $a = (a_{11}, ..., a_{1M}, ..., a_{K1}, ..., a_{KM})$, let $\psi(a)$ be the (unique) set of cutpoints associated with the best responses to the strategy characterized by the cutpoints $(a_{k\sigma})$. Note that substituting the cutpoints into Equation (6) we get

$$t(s, a) = \sum_{k=1}^{K} \sum_{\sigma=1}^{M} p_k(\sigma|s) F(a_{k\sigma}, k) \tag{15}$$

which is continuous in a since F is continuous in x. Therefore, $\beta(s|piv, a, \sigma, k)$ is continuous in a, and hence, continuity of v implies ψ is continuous (a straightforward application of the Theorem of the Maximum). Thus, by Kakutani's fixed point theorem, the map ψ has a fixed point, and hence, the game has a voting equilibrium.

The proof that $t(s, \bar{\pi}^*)$ is non-increasing in s is in the text. ∎

Proof of Lemma 1: By assumption, $t(s, \bar{\pi}^n)$ is non-increasing and hence $S_\eta(\bar{\pi}^n)$ must be an interval which proves the first part of the lemma. Also, observe that since $0 < t(s, \bar{\pi}^n) < 1$, $\beta(s|piv, \bar{\pi}^n)$ is well defined for all n and all s.

Observe that if $S_\eta(\bar{\pi}^n) = [0,1]$ then $\int_{S_\eta(\bar{\pi}^n)} \beta(s|piv, \bar{\pi}^n) = 1$, and hence, if we prove the lemma for the sequence of those n for which $S_\eta(\bar{\pi}^n) \neq [0,1]$ holds then we are done. Thus, we assume in the following that the complement of $S_\eta(\bar{\pi}^n)$ is non-empty.

For $t \in [t_1, t_2] \subseteq [0,1]$ define

$$L(t) = t^q(1-t)^{1-q}$$

and note that this is a concave, single peaked function which reaches a maximum at the t that solves $\min_{t \in [t_1, t_2]} |t - q|$.

Let $t^{*n} = \arg\min_{t \in [t(1, \bar{\pi}^n), t(0, \bar{\pi}^n)]} |t - q|$. Note that by the continuity of $t(s, \bar{\pi}^n)$, there is an $s \in [0,1]$ such that $t^{*n} = t(s, \bar{\pi}^n)$. Since for $s \notin S_\eta(\bar{\pi}^n)$, $|t(s, \bar{\pi}^n) - t^{*n}| \geq \eta$, single-peakedness and continuity of L implies that there is a δ_η such that for all n

$$L(t^{*n}) - \sup_{s \notin S_\eta(\bar{\pi}^n)} L(t(s, \bar{\pi}^n)) \geq \delta_\eta$$

We define the set of states $P_\eta(\bar{\pi}^n) \subset S_\eta(\bar{\pi}^n)$ by

$$P_\eta(\bar{\pi}^n) = \{s : L(t(s, \bar{\pi}^n)) \geq L(t^{*n}) - \delta_\eta/2\}.$$

Since $t(s, \bar{\pi}^n)$ is non-increasing and continuous, it follows that $P_\eta(\bar{\pi}^n)$ is a non-empty interval. In addition, there is a $\gamma > 0$ such that the length of $P_\eta(\bar{\pi}^n)$ is larger than γ for all n. To see this note that there is an $\epsilon > 0$ such that for all n

$$\max_{P_\eta(\bar{\pi}^n)} t(s, \bar{\pi}^n) - \min_{P_\eta(\bar{\pi}^n)} t(s, \bar{\pi}^n) > \epsilon.$$

(This follows since $\max_{P_\eta(\bar{\pi}^n)} L(t(s, \bar{\pi}^n)) - \min_{P_\eta(\bar{\pi}^n)} L(t(s, \bar{\pi}^n)) = \delta_\eta/2$ and since L is continuous on $[0,1]$). By the definition of $t(s, \bar{\pi}^n)$, (see Equation (2))

$$|t(s, \bar{\pi}^n) - t(s', \bar{\pi}^n)| \leq \max_{\sigma, k} |p_k(\sigma|s) - p_k(\sigma|s')|.$$

The (uniform[15]) continuity of $p_k(\sigma|s)$ in s implies that there is an $\gamma > 0$ (independent of n) such that

$$\max_{\sigma, k} |p_k(\sigma|s) - p_k(\sigma|s')| < \epsilon$$

whenever $|s - s'| < \gamma$. Therefore, $P_\eta(\bar{\pi}^n)$ is an interval of size at least γ.

---

[15]Recall that the domain of $p_k(\sigma|s)$ is compact and hence continuity of p_k implies uniform continuity.

To prove the Lemma we now show that $\int_{s\notin S_\eta(\bar\pi^n)}\beta(s|piv,\bar\pi^n)\to 0$.

$$\int_{s\notin S_\eta(\bar\pi^n)}\beta(s|piv,\bar\pi^n) = \frac{\int_{s\notin S_\eta(\bar\pi^n)}\Pr(piv|s)g(s)}{\int_0^1\Pr(piv|w)g(w)dw} \leq \frac{\int_{s\notin S_\eta(\bar\pi^n)}\Pr(piv|s)g(s)ds}{\int_{s\in P_\eta(\bar\pi^n)}\Pr(piv|w)g(w)dw}$$

$$< \frac{\sup_{s\notin S_\eta(\bar\pi^n)}\Pr(piv|s)\int_{s\notin S_\eta(\bar\pi^n)}g(s)ds}{\inf_{s\in P_\eta(\bar\pi^n)}\Pr(piv|s)\int_{s\in P_\eta(\bar\pi^n)}g(s)ds}$$

$$\leq \left(\frac{1-\delta_\eta}{1-\delta_\eta/2}\right)^n\frac{1}{\alpha\gamma} \tag{16}$$

To see the last inequality, note that $g(s)\geq\alpha>0$ implies that $\int_{s\in P_\eta(\bar\pi^n)}g(s)ds\geq\alpha\gamma$, and using Equation (3), we get that for $s'\notin S_\eta(\bar\pi^n)$ and for $s\in P_\eta(\bar\pi^n)$

$$\frac{\Pr(piv|s')}{\Pr(piv|s)} = \left(\frac{L(t(s',\bar\pi^n))}{L(t(s,\bar\pi^n))}\right)^n \leq \left(\frac{L(t^{*n})-\delta_\eta}{L(t^{*n})-\delta_\eta/2}\right)^n \leq \left(\frac{1-\delta_\eta}{1-\delta_\eta/2}\right)^n. \tag{17}$$

Since $\delta_\eta>0$, inequality (16) implies that $\int_{s\notin S_\eta(\bar\pi^n)}g(s)ds$ converges to zero as $n\to\infty$ thus proving the Lemma. ∎

Proof of Lemma 2: By the assumption of ordered cutpoints it follows that $0<t(s,\bar\pi^n)<1$ and that $t(s,\bar\pi^n)$ is non-increasing in s. By Assumption 7 it follows that $t(s,\bar\pi^n)$ is continuous and hence the hypothesis of Lemma 1 is satisfied. Therefore, it suffices to demonstrate that there is a constant $c<\infty$ such that $|p_k(\sigma|s_1)-p_k(\sigma|s_2)|<\eta c$ for any $s_1,s_2\in S_\eta(\bar\pi^n)$ with $s_1<s_2$ and for all σ.

Let $t_k(s,\bar\pi^n)=\sum_\sigma F(x_\sigma^{k,n},k)p_k(\sigma|s)$ be the probability that a voter receives a signal from service k and votes for Q in state s. We first demonstrate that

$$2\eta \geq t(s_1,\bar\pi^n)-t(s_2,\bar\pi^n)\geq t_k(s_1,\bar\pi^n)-t_k(s_2,\bar\pi^n) \tag{18}$$

The first inequality in (18) follows from the definition of $S_\eta(\bar\pi^n)$. Note that $F(x_\sigma^{k,n},k)$ is non-increasing in σ by the assumption of ordered cutpoints, and therefore, by the MLRP, $t_k(s_1,\bar\pi^n)-t_k(s_2,\bar\pi^n)\geq 0$ for all k. Since

$$t(s_1,\bar\pi^n)-t(s_2,\bar\pi^n)=\sum_k\left(t_k(s_1,\bar\pi^n)-t_k(s_2,\bar\pi^n)\right)$$

the second inequality in (18) follows.

Next we show that when $F(x_1^{k,n},k)-F(x_M^{k,n},k)\geq\delta$ either

$$t_k(s_1,\bar\pi^n)-t_k(s_2,\bar\pi^n)\geq\left(p_k(1|s_1)-p_k(1|s_2)\right)\delta \tag{19}$$

or

$$t_k(s_1,\bar\pi^n)-t_k(s_2,\bar\pi^n)\geq\left(p_k(M|s_2)-p_k(M|s_1)\right)\delta \tag{20}$$

26

must hold.

For any subset of signals $O \subset \Sigma$ let $\Pr(O|s) = \sum_{\sigma \in O} p_k(\sigma|s)$. The probability that a voter votes for Q in state s if he receives a signal in O from service k is

$$E(F(x_\sigma^{k,n})|\sigma \in O, s) = \sum_{\sigma \in O} F(x_\sigma^{k,n}, k) \frac{p_k(\sigma|s)}{\Pr(O|s)}. \tag{21}$$

This probability is non-increasing in s. To see this note that by the MLRP of $p_k(s|\sigma)$ it follows that the random variables s and σ are affiliated (see Milgrom and Weber (1982), Theorem 1). Since $F(x_\sigma^{k,n}, k)$ is a non-increasing function of σ we can apply Theorem 5 of Milgrom and Weber (1982) to show that (21) is non-increasing in s. Thus $E(F(x_\sigma^{k,n}, k)|O, s_1) \geq E(F(x_\sigma^{k,n}, k)|O, s_2)$.

Let $O = \{2, ..., M-1\}$ denote the event that service k produces a signal σ such that $2 \leq \sigma \leq M - 1$, then

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) =$$
$$(p_k(1|s_1) - p_k(1|s_2)) F(x_1^{k,n}, k) + (p_k(M|s_1) - p_k(M|s_2)) F(x_M^{k,n}, k)$$
$$+ \Pr(O|s_1) E(F(x_\sigma^{k,n}, k)|O, s_1) - \Pr(O|s_2) E(F(x_\sigma^{k,n}, k)|O, s_2).$$

Since $E(F(x_\sigma^{k,n}, k)|O, s_1) \geq E(F(x_\sigma^{k,n}, k)|O, s_2)$ it follows that

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) \geq (p_k(1|s_1) - p_k(1|s_2)) F(x_1^{k,n}, k) +$$
$$(p_k(M|s_1) - p_k(M|s_2)) F(x_M^{k,n}, k) + (\Pr(O|s_1) - \Pr(O|s_2)) E(F(x_\sigma^{k,n}, k)|O, s_1).$$

Now either $\Pr(O|s_1) - \Pr(O|s_2) > 0$ or $\Pr(O|s_1) - \Pr(O|s_2) \leq 0$. Note that by construction $F(x_1^{k,n}, k) \geq E(F(x_\sigma^{k,n}, k)|O, s) \geq F(x_M^{k,n}, k)$ for any s. Suppose $\Pr(O|s_1) - \Pr(O|s_2) > 0$. Then

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) \geq (p_k(1|s_1) - p_k(1|s_2)) F(x_1^{k,n}, k) +$$
$$(p_k(M|s_1) - p_k(M|s_2)) F(x_M^{k,n}, k) + (\Pr(O|s_1) - \Pr(O|s_2)) F(x_M^{k,n}, k).$$

Since, by definition, $\Pr(O|s) = 1 - p_k(1|s) - p_k(M|s)$ we can simplify the right hand side of the last inequality to obtain 19 as follows:

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) \geq (p_k(1|s_1) - p_k(1|s_2)) (F(x_1^{k,n}, k) - F(x_M^{k,n}, k))$$
$$\geq (p_k(1|s_1) - p_k(1|s_2)) \delta.$$

On the other hand, if $\Pr(O|s_1) - \Pr(O|s_2) < 0$ then

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) \geq (p_k(1|s_1) - p_k(1|s_2)) F(x_1^{k,n}) +$$

$$(p_k(M|s_1) - p_k(M|s_2)) F(x_M^{k,n}, k) + (\Pr(O|s_1) - \Pr(O|s_2)) F(x_1^{k,n}, k)$$

which can again be simplified to yield 20 as follows:

$$t_k(s_1, \bar{\pi}^n) - t_k(s_2, \bar{\pi}^n) \;\geq\; (p_k(M|s_2) - p_k(M|s_1)) (F(x_1^{k,n}, k) - F(x_M^{k,n}, k))$$

$$\geq\; (p_k(M|s_2) - p_k(M|s_1)) \delta.$$

We complete the proof by showing that 18 and 19 imply there exists a constant $c < \infty$ such that

$$\eta c \geq |p_k(\sigma|s_1) - p_k(\sigma|s_2)| \tag{22}$$

for all σ. The argument for the case where 18 and 20 holds is entirely analogous and is therefore omitted.

If 18 and 19 hold then $2\eta/\delta \geq p_k(1|s_1) - p_k(1|s_2)$. Now the MLRP implies $\frac{p_k(1|s_1)}{p_k(1|s_2)} \geq \frac{p_k(\sigma|s_1)}{p_k(\sigma|s_2)}$. Recall that $p_k(\sigma|s) \geq \alpha$ by Assumption 5, and therefore, $\frac{\alpha + 2\eta/\delta}{\alpha} \geq \frac{p_k(1|s_1)}{p_k(1|s_2)}$. Thus, it must be the case that $p_k(\sigma|s_1) - p_k(\sigma|s_2) \leq 2\eta/(\delta\alpha)$ for all σ. The equality $\sum_{\sigma=1}^{M} p_k(\sigma|s_1) = \sum_{\sigma=1}^{M} p_k(\sigma|s_2)$ then implies that $p_k(\sigma|s_2) - p_k(\sigma|s_1) \leq 2\eta \cdot (M-1)/(\delta\alpha)$. Therefore $|p_k(\sigma|s_1) - p_k(\sigma|s_2)| \leq 2\eta \cdot (M-1)/(\delta\alpha)$ for any σ. ∎

Proof of Lemma 3: Let $S_\epsilon^n \subset [0,1]$ be such that $\Pr\{s \in S_\epsilon^n | piv, \bar{\pi}^{*n}\} \to 1$ and

$$|p_k(\sigma|s) - p_k(\sigma|s')| < \epsilon \tag{23}$$

for any $s, s' \in S_\epsilon^n$ and for all σ. Thus, we can choose n' such that for $n > n'$

$$\left| \int_{s \notin S_\epsilon^n} v(x, s) \beta(s|piv, \bar{\pi}^{*n}, \sigma, k) ds \right| < \epsilon \max_{s,x} |v(x,s)| \tag{24}$$

By (23) for any $s \in S_\epsilon^n$ we can write $p_k(\sigma|s) = a_\sigma + \delta(s)$ where $\alpha < a_\sigma < 1$ and $|\delta(s)| < \epsilon$. Note that $v(1,1) > 0, v(-1,0) < 0$, and $v(1,1) \geq v(x,s) \geq v(-1,0)$ by Assumption 1. Thus it follows that

$$E[v(x,s)|piv, \bar{\pi}^{*n}, M, k, S_\epsilon^n] - E[v(x,s)|piv, \bar{\pi}^{*n}, 1, k, S_\epsilon^n] \tag{25}$$

$$\leq v(1,1) \int_{S_\epsilon^n} \left(\frac{\beta(s|piv, \bar{\pi}^{*n})(a_1 + \epsilon)}{\int_{S_\epsilon^n} \beta(w|piv, \bar{\pi}^{*n})(a_1 - \epsilon) dw} - \frac{\beta(s|piv, \bar{\pi}^{*n})(a_M - \epsilon)}{\int_{S_\epsilon^n} \beta(w|piv, \bar{\pi}^{*n})(a_M + \epsilon) dw} \right) ds +$$

$$v(-1,0) \int_{S_\epsilon^n} \left(\frac{\beta(s|piv, \bar{\pi}^{*n})(a_1 - \epsilon)}{\int_{S_\epsilon^n} \beta(w|piv, \bar{\pi}^{*n})(a_1 + \epsilon) dw} - \frac{\beta(s|piv, \bar{\pi}^{*n})(a_M + \epsilon)}{\int_{S_\epsilon^n} \beta(w|piv, \bar{\pi}^{*n})(a_M - \epsilon) dw} \right) ds$$

$$= v(1,1) \left(\frac{a_1 + \epsilon}{a_1 - \epsilon} - \frac{a_M - \epsilon}{a_M + \epsilon} \right) + v(-1,0) \left(\frac{a_1 - \epsilon}{a_1 + \epsilon} - \frac{a_M + \epsilon}{a_M - \epsilon} \right) = o_1(\epsilon)$$

28

where $o_1(\epsilon)$ can be made arbitrarily small for small ϵ. Inequality (24) together with (25) then imply that

$$
\begin{aligned}
E[v(x,s)|piv, M, k] - E[v(x,s)|piv, 1, k] &< 2\epsilon \max_{s,x} |v(x,s)| + o_1(\epsilon) \\
&= o(\epsilon) \qquad (26)
\end{aligned}
$$

Recall that for all σ, $E[v(x_\sigma^{kn}, s)|piv, \bar{\pi}^{*n}, \sigma, k] = 0$. By Assumption 1, for $x \geq x'$, $v(x,s) - v(x',s) \geq \kappa(x - x')$ for some $\kappa > 0$ and therefore

$$
0 = E[v(x_1^{kn}, s)|piv, M, k] - E[v(x_M^{kn}, s)|piv, 1, k] \geq \kappa(x_1^{kn} - x_M^{kn}) - o(\epsilon) \qquad (27)
$$

Thus it follows that

$$
x_1^{kn} - x_M^{kn} \leq o(\epsilon)/\kappa
$$

which proves the Lemma since ϵ can be chosen arbitrarily close to zero. ∎

Proof of Lemma 4: By Proposition 1, we have that

$$
E(v(s, x_\sigma^{k,n})|piv, \bar{\pi}^{*n}, \sigma, k) = E(v(s, x_{\sigma'}^{k,n})|piv, \bar{\pi}^{*n}, \sigma', k) = 0
$$

Since $x_\sigma^{k,n} - x_{\sigma'}^{k,n} \to 0$, it follows from the continuity of $v(x,s)$ in x that for any $x \in [x_\sigma^{k,n}, x_{\sigma'}^{k,n}]$,

$$
E(v(s, x)|piv, \bar{\pi}^{*n}, \sigma', k) - E(v(s, x)|piv, \bar{\pi}^{*n}, \sigma, k) \to 0
$$

Note that $v(x, s)$ is strictly increasing in s with $|v(x, s) - v(x, s')| \geq \kappa|s - s'|$ for some $\kappa > 0$ by Assumption 1. Thus

$$
\begin{aligned}
&E(v(s, x)|piv, \bar{\pi}^{*n}, M, k) - E(v(x, s)|piv, \bar{\pi}^{*n}, 1, k) \\
&\geq \kappa \left(E(s|piv, \bar{\pi}^{*n}, M, k) - E(s|piv, \bar{\pi}^{*n}, 1, k) \right)
\end{aligned}
$$

and hence it is sufficient to show that $E(s|piv, \bar{\pi}^{*n}, M, k) - E(s|piv, \bar{\pi}^{*n}, 1, k) > \delta$ for some $\delta > 0$ whenever asymptotic redundancy is violated.

Suppose asymptotic redundancy is violated for some k. Then, along some subsequence, there must exist a $\delta > 0$ an $\epsilon > 0$ and a sequence s_1^n, s_2^n, such that

$$
\Pr([0, s_1^n]|piv, \bar{\pi}^{*n}) > \delta, \Pr([s_2^n, 1]|piv, \bar{\pi}^{*n}) > \delta \ \forall n
$$

and

$$
|p_k(\sigma|s) - p_k(\sigma|s')| > \epsilon
$$

for some $s_1^n \leq s < s' \leq s_2^n$ and some σ. This must hold since otherwise for every $\epsilon > 0$ and every $\delta > 0$ there exists a sequence of intervals I^n such that $\Pr(I^n|piv, \bar{\pi}^{*n}) \geq 1 - 2\delta$ for sufficiently large n, and $|p_k(\sigma|s) - p_k(\sigma|s')| \leq \epsilon$ for all $s, s' \in I^n$ and every signal σ which implies asymptotic redundance. It follows from the continuity of $p_k(\sigma|s)$ that we can choose ϵ such that $s_2^n - s_1^n > \epsilon$.

Now let $S_1^n = [0, s_1^n]$, $S_2^n = (s_1^n, s_2^n)$ and $S_3^n = [s_2^n, 1]$. By the MLRP, $E(s|S_i^n, piv, \bar{\pi}^{*n}, M, k) \geq E(s|S_i^n, piv, \bar{\pi}^{*n}, 1, k)$ for $i = 1, 2, 3$, and therefore,

$$E(s|piv, \bar{\pi}^{*n}, M, k) - E(s|piv, \bar{\pi}^{*n}, 1, k) \geq \tag{28}$$
$$\sum_{i=1}^{3} [\Pr(S_i^n|piv, \bar{\pi}^{*n}, M, k) - \Pr(S_i^n|piv, \bar{\pi}^{*n}, 1, k)] E(s|S_i^n, piv, \bar{\pi}^{*n}, M, k).$$

Suppose $p_k(\sigma|s') - p_k(\sigma|s) > \epsilon$ for some $s_1^n \leq s < s' \leq s_2^n$ (an analogous argument can be made for $p_k(\sigma|s) - p_k(\sigma|s') < -\epsilon$) and some signal σ. Recall that by Assumption 5, $p_k(\sigma|s) \geq \alpha$ for all (σ, s). The MLRP then implies that $p_k(M|s') - p_k(M|s) \geq \epsilon\alpha$. Since $p_k(M|s)$ is non-decreasing in s, and $p_k(1|s)$ is non-increasing in s this in turn implies that for any $s_1 \in S_1^n$, $s_3 \in S_3^n$, $p_k(M|s_1) - p_k(M|s_3) \geq \epsilon\alpha$ and $p_k(1|s_1) - p_k(1|s_3) \leq 0$. Since $\Pr(S_i^n|piv, \bar{\pi}^{*n}) > \delta$ for $i = 1, 3$, we may conclude that there is a $\gamma > 0$ such that $\Pr(S_1^n|piv, \bar{\pi}^{*n}, 1, k) - \Pr(S_1^n|piv, \bar{\pi}^{*n}, M, k) > \gamma$ and $\Pr(S_3^n|piv, \bar{\pi}^{*n}, M, k) - \Pr(S_3^n|piv, \bar{\pi}^{*n}, 1, k) > \gamma$. But then (28) and the fact that $E(s|S_i^n, piv, \bar{\pi}^{*n}, M, k)$ is increasing in i with $E(s|S_3^n, piv, \bar{\pi}^{*n}, M, k) - E(s|S_1^n, piv, \bar{\pi}^{*n}, M, k) \geq \epsilon$ imply that

$$E(s|piv, \bar{\pi}^{*n}, M, k) - E(s|piv, \bar{\pi}^{*n}, 1, k) \geq \gamma\epsilon$$

which completes the proof. ∎

Proof of Lemma 5: By Proposition 1, we have that $0 < t(s, \bar{\pi}^{*n}) < 1$, and hence $\beta(s|piv, \bar{\pi}^{*n}, \sigma, k)$ is well defined.

First we show that $t(s, \bar{\pi}^{*n})$ is strictly decreasing in s. To see this, let $(x_\sigma^{k,n})$ denote the cutpoints corresponding to $\bar{\pi}^{*n}$. Also, let $\hat{x}^n$ be defined by

$$E[v(\hat{x}^n, s)|piv, \bar{\pi}^{*n}] = 0$$

Since the signal satisfies the SMLRP it follows that

$$E[v(x, s)|piv, \bar{\pi}^{*n}, 1, k] < E[v(x, s)|piv, \bar{\pi}^{*n}] < E[v(x, s)|piv, \bar{\pi}^{*n}, M, k] \tag{29}$$

This implies that $x_1^{k,n} > \hat{x}^n > x_M^{k,n}$. Since $\hat{x}^n$ is in the support of $F(\cdot, k)$ for at least one k it follows that $F(x_1^{k,n}, k) - F(x_M^{k,n}, k) > 0$ for some k. Since $p_k(\sigma|s)$ satisfies the

SMLRP it follows that $p_k(\cdot|s)$ is ordered by strict first order stochastic dominance. Hence $t(s, \bar{\pi}^{*n}) = \sum_k \sum_\sigma F(x_\sigma^{k,n}, k) p_k(\sigma|s)$ is strictly decreasing in s.

As a consequence, $\Pr(piv|s, \bar{\pi}^{*n})$ is strictly increasing if $t(s, \bar{\pi}^{*n}) > q$ and strictly decreasing if this inequality is reversed. Continuity and monotonicity of $t(s, \bar{\pi}^{*n})$ then imply that there is a unique s^n that solves

$$\max_{s \in S} \Pr(piv|s, \bar{\pi}^{*n}),$$

and the first part of the Lemma follows.

By the preceding argument and the fact that $t(s, \bar{\pi}^{*n})$ is decreasing in s, it follows that $\Pr(piv|s, \bar{\pi}^{*n})$ is monotonically decreasing for $s \leq s^n$ and for $s \geq s^n$. This implies that for any $\delta < 1/2$ either (i) $s^n - \delta \geq 0$ and $\sup_{\{s \in [0,1]:|s-s^n|>\delta\}} \Pr(piv|s', \bar{\pi}^{*n}) = \Pr(piv|s^n-\delta, \bar{\pi}^{*n})$ or (ii) $s^n+\delta \leq 1$ and $\sup_{\{s \in [0,1]:|s-s^n|>\delta\}} \Pr(piv|s', \bar{\pi}^{*n}) = \Pr(piv|s^n+\delta, \bar{\pi}^{*n})$. Suppose that (i) holds. (Case (ii) is entirely analogous with the interval $[s^n, s^n + \delta]$ replacing $[s^n - \delta, s^n]$). Then $\frac{\sup_{\{s:|s-s^n|>\delta\}} \Pr(piv|s, \bar{\pi}^{*n})}{\inf_{[s^n-\delta, s^n]} \Pr(piv|s, \bar{\pi}^{*n})} \geq 1$ and since $g(s) \geq \alpha > 0$ it follows that

$$\int_{\{s:|s-s^n|\leq\delta\}} \beta(s|piv, \bar{\pi}^{*n})ds \;=\; \frac{\int_{\{s:|s-s^n|\leq\delta\}} \Pr(piv|s, \bar{\pi}^{*n})g(s)ds}{\int_0^1 \Pr(piv|s, \bar{\pi}^{*n})g(s)ds} \qquad (30)$$

$$= \frac{1}{1 + \frac{\int_{\{s:|s-s^n|>\delta\}} \Pr(piv|s, \bar{\pi}^{*n})g(s)ds}{\int_{\{s:|s-s^n|\leq\delta\}} \Pr(piv|s, \bar{\pi}^{*n})g(s)ds}}$$

$$\geq \frac{1}{1 + \frac{\int_{\{s:|s-s^n|>\delta\}} \Pr(piv|s, \bar{\pi}^{*n})g(s)ds}{\int_{[s^n-\delta, s^n]} \Pr(piv|s, \bar{\pi}^{*n})g(s)ds}}$$

$$\geq \frac{1}{1 + \frac{1}{\delta\alpha}} > 0.$$

To show that in fact $\int_{\{s:|s-s^n|\leq\delta\}} \beta(s|piv, \bar{\pi}^{*n})ds \to 1$, we first claim that for any $\epsilon > 0$ we can find a sequence of sets I^n with $\sup I^n - \inf I^n < \epsilon$ such that $\Pr\{s \in I^n|piv, \bar{\pi}^{*n}\} \to 1$. Assume for the moment this claim is true. Then, by (30) it must either be the case that $s^n \in I^n$ or that $\inf_{I^n} |s - s^n| \to 0$ since otherwise $\Pr\{s \notin I^n|piv, \bar{\pi}^{*n}\}$ stays bounded away from zero. As a consequence, for every $\delta > 0$ there is a $\epsilon > 0$ and an n' such that for $n > n'$, $I^n \subset \{s : |s - s^n| \leq \delta\}$ and the Lemma follows.

To complete the proof of the Lemma it therefore suffices to show that we can find a sequence of sets with $\Pr\{s \in I^n|piv, \bar{\pi}^{*n}\} \to 1$ and $\sup I^n - \inf I^n < \epsilon$. Assumption 8 implies that $p_k(\cdot|s)$ is strictly stochastically dominated by $p_k(\cdot|s')$ for $s' > s$

(see Witt (1980)). As a consequence, $p_k(1|s)$ is strictly decreasing on $[0,1]$. This, together with continuity of $p_k(1|s)$, implies that for every $\epsilon > 0$ there is a $\epsilon' > 0$ such that $|p_k(1|s) - p_k(1|s')| > \epsilon'$ whenever $|s - s'| > \epsilon$. By Theorem 1 there is an information service k for which $x_1^{k,n} - x_M^{k,n} \to 0$. Now, Lemma 4 implies that this information service is asymptotically redundant, i.e., there is a sequence of subsets of states $S_{\epsilon'}^n$ with $\Pr\{s \in S_{\epsilon'}^n | piv, \bar{\pi}^{*n}\} \to 1$. In addition, for any $s, s' \in S_{\epsilon'}^n$, we have that $|p_k(\sigma|s) - p_k(\sigma|s')| < \epsilon'$, for all σ. Thus it follows that $\sup S_{\epsilon'}^n - \inf S_{\epsilon'}^n < \epsilon$. Hence choosing $I^n = S_{\epsilon'}^n$ completes the proof of the Lemma. ∎

Proof of Theorem 4 Part (i): Suppose that contrary to the Theorem $x_1^n - x_M^n \to 0$. Therefore, $|x_\sigma - x_{\sigma'}| \to 0$ for all σ, σ'. Note that Lemma 4 can be applied to this modified framework without changing the proof and hence we conclude that the information service is asymptotically redundant. The SMLRP implies that $p(1|s)$ is strictly decreasing on $[0,1]$. This, together with continuity of $p(1|s)$, implies that for every $\epsilon > 0$ there is a $\epsilon' > 0$ such that $|p(1|s) - p(1|s')| > \epsilon'$ whenever $|s - s'| > \epsilon$. By asymptotic redundance, there is a sequence of subsets of states $S_{\epsilon'}^n$ with $\Pr\{s \in S_{\epsilon'}^n | piv, \bar{\pi}^{*n}\} \to 1$. In addition, for any $s, s' \in S_{\epsilon'}^n$ we have that $|p(\sigma|s) - p(\sigma|s')| < \epsilon'$ for all σ. Thus it follows that $\sup S_{\epsilon'}^n - \inf S_{\epsilon'}^n < \epsilon$, and hence, the probability distribution over states $s \in [0,1]$ must converge to a probability distribution that has all its mass concentrated at some s^n.

First we show that $\epsilon < s^n < 1 - \epsilon$ for some $\epsilon > 0$. To see this suppose, for example, $s^n \to 0$. For large n, (12) implies that the fraction of voters who prefer Q at $s = 0$ is larger than $q + \eta$, for some $\eta > 0$. Therefore, the vote share of q must be larger than $q + \eta/2$ for large n for all λ and all s. But this in turn implies that $s = 1$ is the state for which voters are most likely to be pivotal which contradicts $s^n \to 0$.

Given the equilibrium cutpoints x_σ^n, let

$$t(s, \lambda, \bar{\pi}^{*n}) = (1 - \phi) \sum_\sigma p(\sigma|s) F(x_\sigma^n) + \phi(1 - \lambda) \tag{31}$$

$$\Pr(piv|s, \lambda, \bar{\pi}^{*n}) = \binom{n}{qn} \cdot t(s, \lambda, \bar{\pi}^{*n})^{qn} \cdot (1 - t(s, \lambda, \bar{\pi}^{*n}))^{n-qn} \tag{32}$$

and

$$\Pr(piv|s, \bar{\pi}^{*n}) = \int_0^1 \Pr(piv|s, \lambda, \bar{\pi}^{*n}) \ell(\lambda) d\lambda$$

Since $\epsilon < s^{*n} < 1 - \epsilon$, the relative likelihood of being pivotal in state s^{*n} and states $s = 0, 1$ must be small. More precisely, it must be the case that

$$\frac{\Pr(piv|0, \bar{\pi}^{*n})}{\Pr(piv|s^{*n}, \bar{\pi}^{*n})} \to 0, \frac{\Pr(piv|1, \bar{\pi}^{*n})}{\Pr(piv|s^{*n}, \bar{\pi}^{*n})} \to 0 \tag{33}$$

Let $\lambda^n(s) = \arg\min_\lambda |t(s, \lambda, \bar{\pi}^{*n}) - q|$ and note that $t(s, \lambda, \bar{\pi}^{*n})^q \cdot (1 - t(s, \lambda, \bar{\pi}^{*n}))^{1-q}$ is a single peaked function that reaches its unique maximum at $t(s, \lambda^n(s), \bar{\pi}^{*n})$.

Suppose $\lambda^n(s^{*n}) \geq \epsilon > 0$ for all n. We will show that in this case $\frac{\Pr(piv|0, \bar{\pi}^{*n})}{\Pr(piv|s^{*n}, \bar{\pi}^{*n})}$ stays bounded away from zero and therefore contradicts 33. First, define $b^n = t(0, \lambda, \bar{\pi}^{*n}) - t(s^{*n}, \lambda, \bar{\pi}^{*n}) > 0$ and hence $t(0, \lambda, \bar{\pi}^{*n}) = t(s^{*n}, \lambda + \frac{b^n}{\phi}, \bar{\pi}^{*n})$. Note that b^n is independent of λ and that $b^n \to 0$ (this follows since $x_1^n - x_M^n \to 0$).

$$\frac{\Pr(piv|0)}{\Pr(piv|s^{*n})} = \frac{\int_0^1 t(0, \lambda, \bar{\pi}^{*n})^{qn} \cdot (1 - t(0, \lambda, \bar{\pi}^{*n}))^{n-qn} \ell(\lambda) d\lambda}{\int_0^1 t(s^{*n}, \lambda, \bar{\pi}^{*n})^{qn} \cdot (1 - t(s^{*n}, \lambda))^{n-qn} \ell(\lambda) d\lambda} \geq$$

$$\frac{\int_{\frac{b^n}{\phi}}^1 t(s^{*n}, \lambda, \bar{\pi}^{*n})^{qn} \cdot (1 - t(s^{*n}, \lambda^n, \bar{\pi}^{*n})^{n-qn} \ell(\lambda) d\lambda}{\int_0^1 t(s^{*n}, \lambda, \bar{\pi}^{*n})^{qn} \cdot (1 - t(s^{*n}, \lambda^n, \bar{\pi}^{*n})^{n-qn} \ell(\lambda) d\lambda} \to 1$$

since $\lambda(s^{*n}) \geq \epsilon > 0$.

If $\lambda(s^{*n}) \to 0$ then it follows that $\lambda(s^{*n})$ stays bounded away from 1, and we can make an analogous argument showing that $\frac{\Pr(piv|1, \bar{\pi}^{*n})}{\Pr(piv|s^{*n}, \bar{\pi}^{*n})}$ stays bounded away from zero.

Thus we have again a contradiction to 33, and therefore, we demonstrated that it cannot be the case that $x_1^n - x_M^n \to 0$ which completes the proof of part (i).

Part (ii): Since $p(\sigma|s)$ is continuous in s, it follows that $t(s, \lambda, \bar{\pi}^{*n})$ is uniformly continuous in (s, λ), and therefore, full information equivalence requires that

$$\lim_{n \to \infty} |t(s(\lambda), \lambda, \bar{\pi}^{*n}) - q| = 0$$

for all λ. To prove part (ii) we will demonstrate that for a generic utility function there is a $\lambda \in (0, 1)$ and an $\epsilon > 0$ such that $|t(s(\lambda), \lambda, \bar{\pi}^{*n}) - q| > \epsilon$.

Let $\mathcal{B} = \{B \in [0, 1]^M : B_1 \leq B_2 \leq ... \leq B_M \text{ and } B_1 < B_M\}$ and observe that

$$t(s, \lambda, \bar{\pi}^{*n}) = (1 - \phi) \sum_{\sigma=1}^M p(\sigma|s) B_\sigma + \phi(1 - \lambda) \equiv \tau_B(s, \lambda) \tag{34}$$

for some vector $B \in \mathcal{B}$. For each $B \in \mathcal{B}$ we define

$$\zeta(B, \lambda) = \arg\min_s [\tau_B(s, \lambda) - q] \tag{35}$$

Note that $\zeta(B, \lambda)$ is a continuous function since (35) has a unique solution. (Recall that $p(\sigma|s)$ satisfies the SMLRP).

Let P' denote the set of strictly decreasing continuous functions $\mathbf{s} : [0, 1] \rightarrow (0, 1)$ and endow it with the topology of uniform convergence. We will show that there is an open and dense set $O' \subset P'$ such that for every $\mathbf{s} \in O'$ there is an $\eta > 0$ such that

$$\max_{\lambda} |\mathbf{s}(\lambda) - \zeta(B, \lambda)| > \eta$$

for all $B \in [0, 1]^M$. Suppose for the moment that this claim is true. Since $x(\lambda)$ is a continuous and strictly increasing function it follows that for any $\mathbf{s} \in P'$ there is a $v(x, s)$ that satisfies Assumption 1 and the equation $v(\mathbf{s}(\lambda), x(\lambda)) = 0$ for all λ. If $v(\mathbf{s}(\lambda), x(\lambda)) = 0$ for all λ, we will say in the following that $\mathbf{s}$ is generated by v. Let $O = \{v \in P : v(\mathbf{s}(\lambda), x(\lambda)) = 0, \text{ for some } \mathbf{s} \in O'\}$. Since O' is open O is also an open set. It remains to be shown that O is dense. To this end suppose that $\mathbf{s}$ is generated by v. If $\mathbf{s}'$ satisfies $\|\mathbf{s}' - \mathbf{s}\| < \epsilon$ then we can define $v'(s, x) \equiv v(s + \mathbf{s}(\lambda) - \mathbf{s}'(\lambda), x)$ for all $s \in [\mathbf{s}(\lambda) - \mathbf{s}'(\lambda), 1 - \mathbf{s}(\lambda) - \mathbf{s}'(\lambda)]$. For ϵ small enough $\mathbf{s}'(\lambda) \in [\mathbf{s}(\lambda) - \mathbf{s}'(\lambda), 1 - \mathbf{s}(\lambda) - \mathbf{s}'(\lambda)]$ and hence any extension of v' to all of $[0, 1] \times [-1, 1]$ generates $\mathbf{s}'(\lambda)$. Note that for $s \in [\mathbf{s}(\lambda) - \mathbf{s}'(\lambda), 1 - \mathbf{s}(\lambda) - \mathbf{s}'(\lambda)]$ we have that $|v'(s, x) - v(s, x)| \leq \max_{\epsilon \leq s \leq 1, x} (v(s, x) - v(s - \epsilon, x))$. Therefore, we can extend v' to all of $[0, 1] \times [-1, 1]$ such that $\|v - v'\| \leq \max_{\epsilon \leq s \leq 1, x} (v(s, x) - v(s - \epsilon, x))$. Since v is (uniformly) continuous, it follows that for every $\delta > 0$ there is an $\epsilon > 0$ such that if $\|\mathbf{s}' - \mathbf{s}\| < \epsilon$ then we can find a v' that generates $\mathbf{s}'$ with the property that $\|v' - v\| < \delta$, and hence, O is dense in P.

To prove the claim, consider points $(\lambda_1, ..., \lambda_{M+1})$ and let $0 < \lambda_1 < \lambda_2 < ... < \lambda_{M+1} < 1$ and let

$$S = \{(s_1, ..., s_{M+1}) : s_i = \zeta(B, \lambda_i) \text{ for all } i = 1, .., M + 1 \text{ and some } B \in \mathcal{B}\}.$$

Since ζ is a continuous function of B, it follows that S is contained in an M-dimensional manifold. Let $\bar{S}$ be the closure of S and note that $\bar{S}$ is also contained in an M-dimensional manifold. Consider the set of functions $\mathbf{s}$ that satisfy $(\mathbf{s}(\lambda_1), ..., \mathbf{s}(\lambda_{M+1})) \notin \bar{S}$. Let O' denote this set and note that O' is open since $\bar{S}$ is a closed set. To see that O' is dense suppose that $\mathbf{s} \notin O'$. The set

$$T = \{(s_1, ..., s_{M+1}) : s_i = \mathbf{s}'(\lambda_i) \text{ for all } i = 1, .., M + 1 \text{ and some } \mathbf{s}' \text{ with } \|\mathbf{s}' - \mathbf{s}\| < \epsilon\}$$

34

is an open subset of $\mathbf{R^{M+1}}$. Since $\bar{S}$ is contained in a M-dimensional manifold $T\backslash\bar{S}$ is non-empty, and hence, there exists an $\mathbf{s}'$ with $\|\mathbf{s}-\mathbf{s}'\| < \epsilon$ and $(\mathbf{s}'(\lambda_1),...,\mathbf{s}'(\lambda_{M+1})) \notin \bar{S}$. Therefore, $\mathbf{s}' \in O'$ which proves that O' is dense.

Part (iii) Let s^* be as in Theorem 3 (i.e., corresponding to $\phi = 0$) and suppose that part (iii) of the Theorem does not hold. Then there must exist an $\epsilon > 0$ and a sequence (n, ϕ_n) with $n \to \infty$, $\phi_n \to 0$ and (1) for $s < s^* - \epsilon$ alternative A is elected with probability greater than ϵ or (2) for $s > s^* + \epsilon$ alternative Q is elected with probability greater than ϵ for all n.

We will derive a contradiction. The proof repeats arguments given above and is therefore only sketched. First, we can use the same argument as in the proof of Theorem 1 to demonstrate that $\left|x_1^{n,\phi_n} - x_M^{n,\phi_n}\right| \to 0$. (As before we can show that if $\left|x_1^{n,\phi_n} - x_M^{n,\phi_n}\right| > \nu > 0$ for all n then the information service is asymptotically redundant. The argument is a slight modification of the argument given in Lemmas 2, and therefore, omitted. Lemma 3 demonstrates that asymptotic redundancy of an information service implies that the cutpoints converge. This Lemma can be applied without modification, and hence, we demonstrated that cutpoints converge.) As in the proof of part (i), cutpoint convergence implies that the beliefs conditional on a vote being pivotal must converge to a point mass. I.e., there is a sequence of states s^n such that for every $\delta > 0$, $\Pr\left(\left|s - s^n\right| > \delta\right| piv, \bar{\pi}^n) \to 1$. But then, by the same argument as in the proof of Theorem 2, the probability that any given voter votes for Q must converge to q. Repeating the argument of Theorem 3, this implies that $s^n \to s^*$ and that we can choose $\delta > 0$ such that for $s < s^* - \epsilon$, Q is elected with probability larger than $1 - \epsilon$ whereas for $s > s^* - \epsilon$, A is elected with probability greater than $1 - \epsilon$ which contradicts our initial hypothesis. ∎

References

[1] Alesina, A. and H. Rosenthal. 1995. Partisan Politics, Divided Government, and the Economy. New York; Cambridge University Press.

[2] Alberto Alesina, John Londregan, and Howard Rosenthal. 1993. "A Model of the Political Economy of the United States." *American Political Science Review.* 87 (1), pp. 12-33.

[3] Austen-Smith, David. 1990. "Information Transmission in Debate". *American Journal of Political Science* 4:1, 124-52.

[4] Austen-Smith, David and Jeff Banks. 1995. "Information Aggregation, Rationality and the Condorcet Jury Theorem" *American Political Science Review*, forthcoming.

[5] Calvert, Randall. 1986. Models of Imperfect Information in Politics. New York: Harwood.

[6] Downs, Anthony. 1957. An Economic Theory of Voting. New York: Harper and Row.

[7] Feddersen, Timothy J. and Wolfgang Pesendorfer 1995. "The Swing Voter's Curse". *American Economic Review*, forthcoming.

[8] Feddersen T. and W. Pesendorfer, (1994) "Voting Behavior and Information Aggregation in Elections with Private Information", CMSEMS working paper no.1117.

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

[10] Grofman, Bernard and Julie Withers. 1993. "Information-pooling Models of Electoral Politics." in Information, Participation and Choice: An Economic Theory of Democracy in Perspective. ed. Bernard Grofman. Ann Arbor: The University of Michigan Press.

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

[12] Ladha, Krishna K. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes". *American Journal of Political Science* 36:3 617-34.

[13] Ledyard, John O. 1984. "The Pure Theory of Large Two-Candidate Elections." *Public Choice* 44: 7-41.

[14] Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review* 87: 319-33.

[15] McKelvey, Richard D. and Peter C. Ordeshook. 1985. "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources", *Journal of Economic Theory* 36, 36-55.

[16] Milgrom, Paul. 1979. "A Convergence Theorem for Competitive Bidding with Differential Information," *Econometrica*, Vol. 47, No.3 679-688.

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

[18] Milgrom, Paul, and Robert Weber. 1982. "A Theory of Auctions and Competitive Bidding," *Econometrica* 50: 1089-1122.

[19] Milgrom, Paul and Robert Weber. 1985. "Distributional Strategies for Games with Incomplete Information", *Mathematics of Operations Research*, Vol. 10, No 4, November 1985.

[20] Miller, N. 1986. "Information, Electorates, and Democracy; some Extensions and Interpretations of the Condorcet Jury Theorem," in: B. Grofman and G. Owen. ed. Information Pooling and Group Decision Making. Greenwich Ct. JAI Press.

[21] Myerson, Roger (1994a). "Population Uncertainty and Poisson Games". mimeo. Northwestern University.

[22] Myerson, Roger (1994b). "Extended Poisson Games and the Condorcet Jury Theorem". mimeo. Northwestern University.

[23] Palfrey, Thomas R. (1985). "Uncertainty Resolution, Private Information Aggregation and the Cournot Competitive Limit", *Review of Economic Studies*, 69-83.

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

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

[26] Pesendorfer Wolfgang, and Jeroen Swinkels, 1995. "The Loser's Curse and Information Aggregation in Common Value Auctions", CMSEMS Discussion Paper No. 1147, December 1995.

[27] Persson, Torsten and Tabellini, Guido. Macroeconomic Policy, Credibility and Politics. London: Harwood. 1990.

[28] Riker, William H., and Peter C. Ordeshook . 1968. "A Theory of the Calculus of Voting". *American Political Science Review* 62:1, 25-42.

[29] Rogoff, K. 1990. "Equilibrium Political Budget Cycles", *American Economic Review*, 80, (1), pp. 21-37.

[30] Wilson, Robert. 1977. "A Bidding Model of Perfect Competition," *Review of Economic Studies*, 44, 511-518.

[31] Young, Peyton. 1988. "Condorcet's Theory of Voting", *American Political Science Review* 82(4); 1231-44.

[32] Witt, Ward. 1980. "Uniform Conditional Stochastic Order," *Journal of Applied Probability*, (1980), 112-123.