Preference Monotonicity and Information Aggregation in Elections

Sourav Bhattacharya*
Department of Economics, University of Pittsburgh
Email: sourav@pitt.edu

December 12, 2008

Abstract

If voter preferences depend on a noisy state variable, under what conditions do large elections deliver outcomes "as if" the state were common knowledge? While the existing literature models elections using the jury metaphor where a change in information regarding the state induces voters to switch in favour of only one alternative, I allow for more general preferences where a change in information can induce switch in favour of either alternative. I show that information is aggregated for any voting rule if and only if the probability of switch in favour one alternative is strictly greater than the probability of switch away from that alternative for any change in information. In other words, unless preferences closely conform to the jury metaphor, for large classes of voting rules, there are equilibria that produce outcomes different from the full information outcome with high probability. This condition is very fragile and may be easily violated in spatial elections if the policy space is multidimensional. I conclude that state-contingent conflict in voter preferences may often lead to failure of information aggregation.

The lesson from the celebrated Condorcet Jury Theorem (CJT) is that in a two-alternative majoritarian election, as long as voters agree in their ranking over alternatives in each state of the world, the noise in individual information about the state

*I thank Daron Acemoglu, David Austen-Smith, Steven Callander, Joyee Deb, Tim Feddersen, Roger Myerson and many participants in various seminars and conferences for important inputs while writing this paper. All responsibility for any errors remaining in the paper is mine.

does not affect the aggregate outcome. The difficulty of extending this conclusion to real elections is that, unlike the members of a jury, all of whom prefer to convict the defendant in the guilty state and acquit him in the innocent state, preferences of voters may admit substantial diversity. Unfortunately, both the statistical and game theoretic work on information aggregation in elections have incorporated very little or no heterogeneity in voter preferences[1]. In this paper, I allow voters to have diverse, even conflicting preferences, and derive conditions under which information aggregation is guaranteed in all equilibria. Such conditions suggest that other than the stylised situations where elections closely follow the jury metaphor, information aggregation may fail in real elections in an extreme sense: there may be "wrong" outcomes with a very high probability for large classes of voting rules.

In their very important proof of CJT, Feddersen and Pesendorfer (1997), henceforth F-P, admits a limited heterogeneity of preferences by assuming what they call "common values": there is an interval of states and individuals' relative valuation of alternatives is an increasing function of the state variable. Thus, any change in state makes all voters more inclined towards the same alternative. This assumption of unidirectional switching is key to their proof of CJT. Almost all subsequent work in the literature (Feddersen-Pesendorfer (1999), Myerson (1998), Wit (1998), Meirowitz (2002)) makes the implicit assumption that any change in information regarding the state induces switches only in one direction. However, this assumption may be violated in many settings[2]. For example, suppose that voters single peaked preferences on the usual left-right ideological continuum, and the incumbent is to the left of the challenger in state A and to the right in state B. Then a change in state from A to B will induce the leftists to switch in favour of the challenger and the rightists in favour of the incumbent. The same situation may very easily arise when the voters care about many different issues (which arguably they do, in real elections), and there is some issue dimension in which such an uncertainty prevails. I discuss such a case in more detail in section 4.

The model presented in this paper has two alternatives $\mathcal{P}$ and $\mathcal{Q}$, two states A

[1] See Ladha (1992) for a statistical proof of CJT. Austen-Smith and Banks (1996) first pointed out that such sincere voting is not necessarily rational. McLennan (1998) shows that if there exists an outcome that aggregates information with sincere voting, there exists a Nash equilibrium that does the same too.

[2] Kim and Fey (2007) considers a setting with opposed rankings and shows that information aggregation can break down, but since they allow abstention in their model, it is not clear whether the aggregation failure is driven by voter preferences or by the expanded voter strategy space.

and B, and two signals a and b satisfying MLRP. To capture variation in preferences, I define a distribution $F(\mu, x)$ over $\mu \in [0,1]$, the cut-off beliefs on state A at which a voter switches rankings over the alternatives and $x \in \{u, d\}$, the direction of such switch. I derive a condition called *Weak Preference Monotonicity (WPM)* which is necessary and sufficient for information to be aggregated for any consequential rule, i.e. voting rules which induce different outcomes in different states under full information. *WPM* requires that for any non-degenerate belief over states, a change in signal from b to a must induce a strictly higher probability of switch from $\mathcal{Q}$ to $\mathcal{P}$ than from $\mathcal{P}$ to $\mathcal{Q}$. What makes this condition stringent is that it has to be satisfied for every belief over states – if it is violated for some belief, aggregation fails in two extreme ways: one, we get equilibria with a "wrong" outcome with a very high probability in at least one state, and two, the failure occurs for all consequential voting rules.

In general, whether a distribution of preferences F satisfies *WPM* depends on the conditional distribution over signals. I introduce a strengthened condition, *Strong Preference Monotonicity (SPM)* by requiring *WPM* to be satisfied for all information structures satisfying MLRP. *SPM* holds when any increase in belief over state A leads to a strictly higher probability of switch from $\mathcal{Q}$ to $\mathcal{P}$ than from $\mathcal{P}$ to $\mathcal{Q}$. In an essential sense, *SPM* requires that for every belief, the rate of switch in favour of $\mathcal{P}$ be strictly greater than the rate of switch away from $\mathcal{P}$, i.e $\frac{dF(\mu,u)}{d\mu}$ must be greater than $\frac{dF(\mu,d)}{d\mu}$ for all μ, and thus $F(\mu, u) - F(\mu, d)$ be (essentially) strictly increasing. *SPM* is not only sufficient for information being aggregated for any information structure; if *SPM* is violated, there exist information structures for which aggregation fails to obtain. Since SPM is a condition on the derivatives of F, it is extremely fragile: a small local perturbation of F is enough for violation of the condition and consequent aggregation failure. In all the existing literature including F-P, *SPM* is satisfied by the assumption of unidirectional shift. The contribution of this paper lies in pointing out the necessity of this extreme condition for the hitherto familiar property of full information equivalence.

The results in this paper are driven by a single insight: the existence of an equilibrium and characteristics of the outcome depends on the local properties of beliefs over states, in particular, the behaviour of the expected vote shares around a given belief. To ensure that all equilibria generate the full information outcome, one must impose a global property on the vote shares that should hold true for all beliefs. The monotonicity properties in this paper are two such properties, and they are "tight" in the sense that any violation leads to existence of "wrong" equilibrium outcomes.

The rest of the paper is organized as follows. Section 1 describes the set-up and builds the machinery for the results. Section 2 characterises the limiting equilibria as the number of voters becomes large. Section 3 discusses the monotonicity properties that social preferences should have for information to be aggregated in the limit. Section 4 discusses an application to the spatial model and concludes. Most proofs are relegated to the appendix.

1 The Set-up

Suppose there is an electorate composed of a finite number $(n+1)$ of people who are voting for or against a policy $\mathcal{P}$. If the policy gets more than a proportion $\theta \in (0,1)$ of the votes, then $\mathcal{P}$ wins; otherwise the status quo $\mathcal{Q}$ wins[3]. I consider all plurality rules other than unanimity. There are two states of nature $S \in \{A, B\}$, and the commonly known prior probability of state A is $\Pr(A) = \alpha \in (0,1)$. The voter receives a private signal $s \in \{a, b\}$ which is drawn randomly from a conditionally independent distribution given by $\Pr(a|A) = q_a$ and $\Pr(a|B) = q_b$. I make the usual assumption on informativeness of signals, i.e. $1 > q_a > q_b > 0$. A specific pair of conditional probabilities $\{q_a, q_b\}$ will be called an *information structure*.

1.1 Preferences

The voter's utility difference $v(S)$ between the policy alternative and the status quo in state S is an independent, random draw from some distribution over real numbers. The voter prefers the alternative $\mathcal{P}$ in state S if the realised value $v(S)$ is positive and the status quo if $v(S)$ is negative. If $v(S)$ has the same sign in both states, then the vote does not depend on information about the state. If $v(S)$ has different signs in the two states, then voter's decision depends on her belief over states based on private information and strategies of other voters. For each such voter, there is a cut-off probability of state A, say μ, which depends solely on $v(A)$ and $v(B)$, such that for all beliefs less than μ, she votes for one alternative and for all beliefs greater

[3]To simplify the analysis, assume the tie breaking rule that if the policy receives exactly θ proportion of votes, the status quo wins.

than μ, she votes for the other. It is easy to see that the distribution of $v(S)$ in the society can be equivalently described in terms of a distribution over the cut-off belief and the direction of switch. Thus, with probability γ_P, a voter is committed to $\mathcal{P}$ in both states, with probability γ_Q, she is committed to $\mathcal{Q}$, and with the remaining probability γ_I, she is independent. Conditional on being independent, her preferences are drawn from the joint distribution $F(\mu, x)$, with $\mu \in [0, 1]$ describing her cut off belief over states where she switches her ranking, and $x \in \{u, d\}$ which is the direction of shift. If $x = u$ and $\mu = \mu_0$, she votes for the alternative $\mathcal{P}$ if her belief (assessment of $\Pr(A)$) is above μ_0 and against $\mathcal{P}$ if her belief is below μ_0. If $x = d$ and $\mu = \mu_0$, she votes for the alternative if her belief is below μ_0 and against the alternative with belief above μ_0. Preferences are private information. Since the committed voters play very little role in the results of the paper, I shall focus on F and call it the *distribution of social preferences.*

Assumption 1: Each of γ_P, γ_Q and γ_I is strictly positive.

Assumption 2: For all $\mu \in (0, 1)$, the joint distribution $F(\mu, x)$ is non-atomic and differentiable in μ for each $x \in \{u, d\}$. I denote $\frac{\partial F(\mu,x)}{\partial \mu}$ by $f_x(\mu)$,and refer to these derivatives as rates of switch, in favour of or away from $\mathcal{P}$, as the case may be. The switch rates $f_x(\mu)$ are continuous and bounded. Also, $f_u(\mu) \neq f_d(\mu)$ at $\mu \in \{0, 1\}$.

Assumption 1 ensures that each alternative gets a positive share of votes, and pivot probabilities are well defined in each state. Assumption 2 is made for technical convenience. I shall later make an assumption that implies a stronger restriction on the rates of switch: that there is no interval on the support of F for which f_u and f_d are equal.

1.2 Strategy

The equilibrium concept employed here is symmetric Bayesian Nash equilibrium in weakly undominated strategies, as has been used in F-P. Weakly undominated strategies imply that equilibria where everyone votes for or against the policy alternative irrespective of private information are ruled out. Symmetry requires all voters with the same (μ, x) to use the same strategy.

The strategy for committed voters is trivial, and indeed they will not play any role except ensuring positive expected vote shares for both alternatives. A pure strategy for independent voters is a function $\pi : [0,1] \times \{u,d\} \times \{a,b\} \to \{0,1\}$, which chooses whether to vote for or against $\mathcal{P}$ given the preference type $\{\mu, x\}$ and a signal s. It must be mentioned here that instead of looking for equilibrium strategies directly, I solve for what beliefs over states will be induced in equilibrium given one's private signal. Based on their equilibrium assessment of $\Pr(A)$ and their cut-offs, the voters vote either for or against $\mathcal{P}$ unless the cut-off is exactly equal to the belief held, in which case, the voter is indifferent. I concentrate on pure strategies, since indifference occurs with zero probability and mixing by the indifferent voters does not change the vote shares and therefore the outcome of the election[4].

If a randomly chosen independent voter holds belief $p \in [0,1]$, her probability of voting for the alternative $\mathcal{P}$ is:

$$V(p) = \gamma_I \left[\int_0^p f_u(t)dt + \int_p^1 f_d(t)dt \right] + \gamma_P = \gamma_I [F(p,u) + F(1,d) - F(p,d)] + \gamma_P \quad (1)$$

Non-atomicity of F in μ guarantees that $V(p)$ is continuous. Note also that $0 < \gamma_P \leq V(p) \leq 1 - \gamma_Q < 1$ for all $p \in [0,1]$.

Notice that we have a Condorcet set up if $F(\mu, x) = 0$ for all μ for either $x = u$ or $x = d$, i.e. if all independents switch in one direction. Then the current setting is similar to that in F-P - while they have a continuum of states, in this model, preferences are described as rankings on a continuum of beliefs over two states.

1.3 Equilibrium

In this set-up, I solve for equilibrium not in terms of strategies but in terms of the common belief over states held by the voters. Suppose we start with a common initial belief $\Pr(A) = \beta \in [0,1]$ Given this belief β, a voter updates based on her signal $s \in \{a,b\}$, and forms a posterior $\beta_s \equiv \Pr(A|\beta, s)$, which can be calculated by Bayes rule.

[4]Non-atomicity of F guarantees that the probability of a voter being indifferent is zero, since only a finite number of beliefs can be held in the society in the symmetric equilibrium, with one belief for each signal.

A random voter holding belief β_s votes for the alternative with a probability $V(\beta_s)$. Since the two states have different distribution over signals, each state produces, in general, a different probability that a random voter votes for $\mathcal{P}$. Note that the probability of a random voter voting for the alternative in state S, in other words, the expected vote share for the alternative in state S is a function of the initial belief β and is denoted by $t(S, \beta)$. Given these vote shares, the voting rule θ and n, we get a probability of a tie in each state. By Bayes rule, we get a belief over states conditioning on the event of a tie. Call this probability $\Pr(A|piv, \beta, n)$. In equilibrium, this belief over states conditioning on pivotality should be the same as the belief β held initially. Thus, the equilibrium condition is

$$\Pr(A|piv, \beta, n) = \beta \tag{2}$$

Since the initial belief β is like a common prior except that it is induced in equilibrium, I call it the *induced prior* belief, and solve the model for this induced prior.

The expected vote share in each state as a function of induced prior $\beta \in [0, 1]$ is:

$$\begin{aligned} t(A, \beta) &= q_a V(\beta_a) + (1 - q_a)V(\beta_b) \\ t(B, \beta) &= q_b V(\beta_a) + (1 - q_b)V(\beta_b) \end{aligned} \tag{3}$$

Continuity of $V(p)$ guarantees the continuity of $t(S, \beta)$. Also, Since $V(p) \in (0, 1)$ for all $p \in [0, 1]$ we must also have $t(S, \beta) \in (0, 1)$ for $S \in \{A, B\}$ and all $\beta \in [0, 1]$.With a slight abuse of terminology, I shall simply refer to $t(S, \beta)$ as the "vote share function".

Given that all the n other voters hold belief β, the probability of the remaining voter's vote being decisive in state S is given by:

$$\Pr(piv|S, \beta, n) = \binom{n}{[n\theta]} t(S, \beta)^{[n\theta]} \left(1 - t(S, \beta)\right)^{n-[n\theta]}, \ S \in \{A, B\} \tag{4}$$

where $[n\theta]$ is the largest integer weakly smaller than $n\theta$. From the equilibrium condition (2), equation (4) and Bayes Rule, we can write

$$\frac{\beta}{1-\beta} = \frac{\Pr(A|Piv, \beta, n)}{\Pr(B|Piv, \beta, n)} = \left(\frac{\alpha}{1-\alpha}\right) \frac{\Pr(piv|A, \beta, n)}{\Pr(piv|B, \beta, n)}$$

From (4), therefore, we can rewrite the equilibrum condition as

$$\frac{\beta}{1-\beta} = \frac{\alpha}{1-\alpha}\left[\frac{t(A,\beta)^{[n\theta]}\left(1-t(A,\beta)\right)^{n-[n\theta]}}{t(B,\beta)^{[n\theta]}\left(1-t(B,\beta)\right)^{n-[n\theta]}}\right] \equiv H(n,\beta,\theta) \tag{5}$$

Continuity and boundedness of $t(S,\beta)$ guarantees that $H(n,\beta,\theta)$ is strictly positive, continuous and bounded both above and below for any given n and θ. The left hand side of (5), on the other hand, varies monotonically in $(0,\infty)$ as β changes in $(0,1)$. Therefore a solution to equation (5) exists, and it characterises the equilibrium for (n,θ). Note that since the vote shares are strictly between 0 and 1, the equilibrium induced prior must lie in $(0,1)$.

2 Limit properties of equilibria

The analysis in the previous section is for a given, finite electorate. To look at the voting outcome for a large electorate, I hold the preference distribution, information structure and voting rule fixed, and examine the limit of the equilibrium outcome as the group size becomes larger. I denote the equilibrium induced prior (solution to (5)) by β^n, and examine the sequence of equilibria as $n \to \infty$. First note than since the sequence belongs to a compact interval, a limit always exists. Call it β_0. From condition (5), we obtain the *limiting equilibrium condition*:

$$\beta_0 = \lim_{n\to\infty}\left[\frac{1}{1+H(n,\beta^n,\theta)}\right] \tag{6}$$

While the ultimate objective is to find the limiting equilibrium induced prior beliefs β_0 (one for each equilibrium) for each voting rule θ, I consider a given belief β_0 and examine the voting rules which may support an equilibrium with beliefs converging to β_0 for some distribution of preferences. The following Lemma tells us that in the limiting equilibrium, vote shares should bear a particular relationship with the voting rule.

Lemma 1 *For the equilibrium condition to be satisfied at any value of* $\beta^0 \in [0,1]$ *and any* $\theta \in (0,1)$, *we must have*

$$\left(t(A,\beta^0)\right)^{\theta}\left(1-t(A,\beta^0)\right)^{1-\theta} = \left(t(B,\beta^0)\right)^{\theta}\left(1-t(B,\beta^0)\right)^{1-\theta}$$

For beliefs that produce equal expected vote shares in the two states, the above lemma is immedate. For beliefs that produce different expected vote shares in the two states, i.e. $t(A,\beta) \neq t(B,\beta)$,an inspection of equation (5) shows that $H(n,\beta^n,\theta)$ either explodes to infinity or goes to zero if the above condition in the Lemma is not satisfied. In this latter case, there is a unique voting rule that satisfies the condition given the vote shares in the two states, given by[5]

$$\theta^*(\beta) = \frac{\log \frac{1-t(B,\beta)}{1-t(A,\beta)}}{\log \frac{t(A,\beta)(1-t(B,\beta))}{t(B,\beta)(1-t(A,\beta))}} \tag{7}$$

For a belief β, such that $t(A,\beta) \neq t(B,\beta)$, the only voting rule that may support the belief in limiting equilibrium is given by (7). Notice also that $\theta^*(\beta)$ lies strictly between the two vote shares in the two states.

Lemma 2 *Define by $\Theta(\beta)$ the following correspondence:*

> *(i) If $t(A,\beta) \neq t(B,\beta)$, then $\Theta(\beta)$ is the unique value $\theta^*(\beta)$*
> *(ii) If $t(A,\beta) = t(B,\beta) = t$ and $\beta \in (0,1)$, then $\Theta(\beta) = (0,1)$*
> *(iii) $\Theta(0)$ is $(0,V_B]$ if $f_u(0) > f_d(0)$ and $[V_B,1)$ if $f_u(0) < f_d(0)$*
> *(iv) $\Theta(1)$ is $[V_A,1)$ if $f_u(1) > f_d(1)$ and $(0,V_A]$ if $f_u(1) < f_d(1)$*

For a given $\beta \in [0,1]$, consider any sequence $\beta_n \to \beta$. Now, if $\theta \notin \Theta(\beta)$, then $H(n,\beta^n,\theta)$ is bounded away from $\frac{\beta}{1-\beta}$.

Lemma 2 demarcates the set of voting rules which *may* support a given induced prior as a limit of the voting equilibria. For a voting rule $\theta \notin \Theta(\beta)$, no sequence $\beta_n \to \beta$ satisfies the limiting equilibrium condition (6). Perhaps surprisingly, the converse is also true: for almost all β, with each voting rule in $\Theta(\beta)$ there exists a sequence of equilibria with the induced prior going to β in the limit for any distribution of preference once I make two more assumptions on the preference distributions which basically rule out pathological cases.

Assumption 3: The vote share functions for the two states $t(A,\beta)$ and $t(B,\beta)$ are equal only at a countable number of beliefs, and their functions "cross" at those beliefs.

[5]In other words, in equilibrium, suppose $r(\mathcal{A}) = \frac{\text{Vote share for alternative } \mathcal{A} \text{ in state } B}{\text{Vote share for alternative } \mathcal{A} \text{ in state } A}$ for alternative $\mathcal{A} \in \{\mathcal{P},\mathcal{Q}\}$. Then we must either have $r(\mathcal{A}) = 1$ for each alternative, or the voting rule and vote shares should satisfy the property that $\theta^* = \frac{\log r(\mathcal{Q})}{\log r(\mathcal{Q}) - \log r(\mathcal{P})}$

Assumption 4: There is no interval of beliefs for which function $\theta^*(\beta)$ is constant.

Assumption 3 rules out situations for which the two vote share functions are equal on an interval of beliefs. It also rules out the vote share function for one state being tangent to that for the other for some belief. Assumption 4 rules out two kinds of situations. First, it rules out a situation where both vote share functions are constant over a range of beliefs - this implies the assumption that both f_u and f_d cannot be equal over any open interval. Second, it rules out the more pathological situation where both vote shares change in such a way as to keep $\theta^*(\beta)$ constant over an interval. Define the set of beliefs where $\theta^*(\beta)$ is either locally increasing or decreasing as M. Formally, M includes those beliefs β such that there exists $\epsilon > 0$ such that $\theta^*(\cdot)$ is strictly monotonic in $(\beta - \epsilon, \beta + \epsilon)$. Assumption 3 states that the complement of M is a countable set composed of local maxima and local minima of $\theta^*(\beta)$.

Theorem 1 *For any distribution of preferences satisfying assumptions* 1 *through* 3, *any information structure and any voting rule* $\theta \in (0, 1)$, *as the number of voters becomes large, there is a sequence of equilibria of the voting game with the induced prior belief converging to* β *if and only if* $\beta \in M$ *and* $\theta \in \Theta(\beta)$.

The "only if" direction is clear from Lemma 2, and the proof of the "if" direction is in the appendix. Theorem 1 characterises the limiting equilibria of a voting game as the electorate grows large. Given voting rule θ, if in some equilibrium the limiting induced prior is β^0, then by the Law of Large numbers, for any $\epsilon > 0$ there is some m such that the alternative $\mathcal{P}$ wins in state S with a probability higher than $1 - \epsilon$ for all population sizes $n > m$ if and only if $t(S, \beta^0) > \theta$. Figure 1 helps us "read" the different limiting equilibria given a voting rule and determine the outcomes. For the voting rule θ, there are five equilibria (induced priors β_1 through β_5). In equilibria with induced priors β_1 and β_5, the alternative $\mathcal{P}$ wins only in state A while in the equilibrium with induced prior β_3, it wins only in state B. The alternative wins in both states if the equilibrium belief is β_4 and loses in both states if it is β_2. Also note that since $\theta^*(\beta)$ is "turning" at β_0, there is no equilibrium with β_0 as the limiting belief. Notice, as an aside, that unlike F-P, elections are not necessarily "close", i.e. the expected margin of victory may be strictly positive in the limit.

In the next section, I study the information aggregation properties of the limiting outcome. In particular, I examine conditions on the social preference distribution F that are necessary and sufficient for information to be aggregated in every equilibrium.

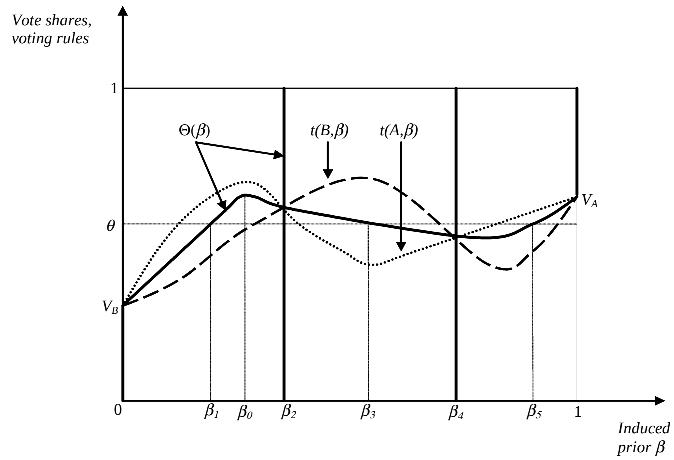

Fig 1: Voting equilibria

3 Information Aggregation Property

To study conditions under which voting as an incomplete information game delivers the same outcome as is obtained under full information, I first introduce a classification of voting rules according to full information outcomes. Denote by V_S the share of voters who prefer the alternative $\mathcal{P}$ over status quo $\mathcal{Q}$ in state S. From equation (1), we can see that

$$V_A = \gamma_I F(1,u) + \gamma_P \quad \text{and} \quad V_B = \gamma_I F(1,d) + \gamma_P$$

WLOG, I assume that $F(1,u) > F(1,d)$, which also implies that $V_A > V_B$[6]. Now we can have a classification of voting rules according to the outcome induced under common knowledge of the state. If we have $\theta \in (V_B, V_A)$, the alternative $\mathcal{P}$ would win in state A and lose in state B. Since these voting rules lead to different outcomes in different states, we call these the *Consequential* rules. On the other hand, voting rules $\theta \in (0, V_B)$ are called *P-trivial* rules, since $\mathcal{P}$ wins in both states under common knowledge. Similarly, voting rules $\theta > V_A$ are called *Q-trivial* rules, since under these rules, the status quo prevails in both states.

[6]The only loss of generality in making this assumption is that it rules out equal full information vote shares for the alternative under each state.

The yardstick of information aggregation used here is an adapted version of the full information equivalence (FIE) criterion used in F-P. Formally, an election is said to satisfy *full information equivalence* if, *every* equilibrium of the voting game satisfies following condition: for any $\epsilon > 0$, there is some m such that, the equilibrium outcome in each state is the same as the outcome under full information with a probability larger than $1 - \epsilon$ if the electorate size is greater than m. By that standard, aggregation fails in the example presented in figure 1, since we get the "right" outcomes in equilibria with beliefs β_1 and β_5, and "wrong" outcomes with high probability in the rest.

3.1 Weak Preference Monotonicity

First, it is easy to see that for $\mathcal{P}$ to win in state A and lose in state B under consequential rules, the vote share for $\mathcal{P}$ in state A has to be strictly greater than that in state B. Also, there cannot be any non-degenerate belief where the vote shares are equal, because then we will get an equilibrium with the same alternative wins in both states for almost all voting rules, including consequential rules.

Lemma 3 *Elections with any consequential voting rule is full information equivalent if and only if $t(A,\beta) > t(B,\beta)$ for any $\beta \in (0,1)$.*

From (1), (3), and $q_a > q_b > 0$, we can say that the condition that $t(A,\beta) - t(B,\beta) > 0$ is equivalent to:

$$\int_{\beta_b}^{\beta_a} f_u(z)dz > \int_{\beta_b}^{\beta_a} f_d(z)dz \quad \text{or} \quad F(\beta_a, u) - F(\beta_b, u) > F(\beta_a, d) - F(\beta_b, d) \tag{8}$$

I call this condition Weak Preference Monotonicity.

Definition 1 *Consider a given information structure $\{q_a, q_b\}$. A distribution of social preferences F is said to satisfy Weak Preference Monotonicity (*WPM*) if a change in signal from b to a induces a strictly larger probability of switch from $\mathcal{Q}$ to $\mathcal{P}$ than from $\mathcal{P}$ to $\mathcal{Q}$ for any non-degenerate prior belief over states.*

Given a prior belief β, we get posteriors $\beta_a > \beta_b$. As the signal changes from b to a, $F(\beta_a, u) - F(\beta_b, u)$ mass of voters switch from $\mathcal{Q}$ to $\mathcal{P}$, while $F(\beta_a, d) - F(\beta_b, d)$ switch from $\mathcal{P}$ to $\mathcal{Q}$. Notice that posteriors are a function of the prior β and the precision of signals q_a and q_b. WPM requires that for *every* prior belief $\beta \in (0,1)$, the shift in favour of $\mathcal{P}$ is larger than the shift away from $\mathcal{P}$. The next proposition follows from Lemma 3 and definition of *WPM.*

Proposition 1 *Given assumptions* 1 *through* 4 *and an information structure, voting with any consequential voting rule is full information equivalent if and only if social preference distribution F satisfies WPM.*

It is important to note that if *WPM* fails to obtain, aggregation fails in two extreme ways. First, in at least one state, an outcome different from the full information outcome obtains with a probability arbitrarily close to 1. Second, if aggregation fails for one consequential rule, it fails for all consequential rules (and either all P-trivial rules, or all Q-trivial rules).

There are two weaknesses of *WPM* as a condition for information aggregation. First, whether or not a distribution F satisfies *WPM* depends on the distribution of signals. Suppose $q_a = 1 - q_b = q > \frac{1}{2}$. It is easy to construct examples where a distribution satisfies *WPM* for high values of q but not for low values. Second, *WPM* does not guarantee information aggregation for trivial rules. Unless we can guarantee that $\theta^*(\beta)$ lies strictly between V_B and V_A, we will have $\mathcal{P}$ winning in one state but losing in another even for trivial rules, which is not the full-information outcome. This is stated the following proposition.

Proposition 2 *Given assumptions* 1 *through* 4 *and an information structure, voting with any non-unanimous plurality rule is full information equivalent if and only if both of the following conditions hold:*

(i) *The distribution of preferences F satisfies WPM*
(ii) $\theta^*(\beta) \in [V_B, V_A]$ *for all* $\beta \in (0,1)$

3.2 Strong Preference Monotonicity

WPM is a joint condition on the information structure and the preference distribution. Now, I introduce a condition on preferences alone that guarantees FIE for any information structure. Notice that *WPM* is simply a requirement that the mass of voters switching from $\mathcal{P}$ to Q be strictly greater than those switching in the opposite direction for certain intervals of posterior beliefs over the two states. These intervals are defined by the prior belief and the conditional distribution of signals over states. *Strong Preference Monotonicity (SPM)* holds when the requirement is satisfied over *any* interval of posterior belief. In other words, a distribution satisfying *SPM* satisfies *WPM* for every distribution of signals over states.

Definition 2 *A distribution of social preferences F is satisfies Strong Preference Monotonicity (*SPM*) if the probability of a switch from $\mathcal{P}$ to $\mathcal{Q}$ is strictly greater than that of a switch from $\mathcal{Q}$ to $\mathcal{P}$ for any increase in belief over state A.*

SPM requires that $f_u(\mu) > f_d(\mu)$ for all $\mu \in [0,1]$ except possibly for a finite number of values μ where $f_u(\mu)$ and $f_d(\mu)$ may be equal. While it is clear that any distribution F satisfying *SPM* also satisfies *WPM* for any allowable value of q_a and q_b, the converse is also true, i.e. any distribution that satisfies *WPM* for all allowable values of q_a and q_b also satisfies *SPM*. To see that, suppose $f_u(z_0) < f_d(z_0)$ for some value of $z_0 \in (0,1)$. Since f_u and f_d are continuous, it follows that $f_u(z) < f_d(z)$ over some interval (β_1, β_2) containing z_0. Set $\widehat{\beta} = \frac{1}{2}((\beta_1, \beta_2)$. Now, choosing the values of q_a and q_b appropriately, we can get psoteriors $\beta_a(\widehat{\beta}, q_a) = \beta_2$ and $\beta_b(\widehat{\beta}, q_b) = \beta_1$. It is easy to see now that *WPM* will not hold for such a pair (q_a, q_b).

The next lemma tells us that *SPM* implies condition (ii) of Proposition 2. This follows from the fact that *SPM* guarantees strict monotonicity of the vote share functions.

Lemma 4 *If F satisfies SPM, $\theta^*(\beta)$ must lie in the interval* (V_B, V_A) for all $\beta \in (0,1)$.

The main theorem of the paper now follows from Lemma 4, the definition of *SPM* and Proposition 2.

Theorem 2 *Under assumptions* 1 *through* 4, *the following statements are equivalent:*

(i) *The distribution of social preferences satisfies Strong Preference Monotonicity.*
(ii) *Voting is full information equivalent for any information structure and any non-unanimous voting rule.*

This theorem tells us that if we want to guarantee information aggregation for *every* information structure, the preferences have to satisfy a very strong condition, the *SPM*. The condition basically says that the *net* rate of switch in favour of the alternative $\mathcal{P}$ must be strictly positive for all beliefs over the states except possibly at countable number of beliefs where the rate can be zero. In other words, switchers in one direction should always be dominated by the switchers in the other direction. Put differently, SPM requires that $F(\mu, u) - F(\mu, d)$ be strictly increasing in every interval of μ. The condition is very fragile in the sense that we can always change the function $F(\cdot, d)$ over any small open interval $\zeta \subset (0, 1)$ of the support and set $f_d(\mu) > f_u(\mu)$ over ζ, and the condition is violated.

The bulk of the existing literature on CJT has assumed unidirectional switching, i.e. $f_d(\mu) = 0$ for all μ. This assumption automatically implies *SPM*, and it is not surprising that the literature has broadly agreed that elections efficiently aggregate information. The point of this paper is to demonstrate that modelling elections using the jury metaphor may lead to erroneous conclusions. The very elegant property of full information equivalence cannot be extended much beyond the jury setting. The specific distribution of social preference in an election depends on the particular case under consideration, but in many situations, monotonicities may be hard to obtain.

4 Conclusion

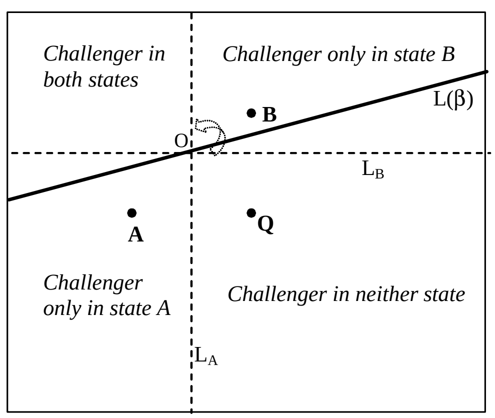

Fig 2: Spatial Voting on multiple dimensions

To see what Strong Monotonicity entails in an applied setting, consider a spatial election with a two-dimensional policy space, as shown in figure 2. Voters have their ideal points distributed over the policy space, a voter prefers the candidate with location nearest to his ideal point. The incumbent's location is Q, but there is uncertainty regarding the challenger who could be located either at A (state A) or at B (state B). Under full information, in state A, those to the left of the vertical line L_A vote for the challenger, and in state B, those above the horizontal line L_B vote for the challenger. L_A and L_B intersect at O, dividing the policy space into four quadrants. Observe that the group of voters with ideal points in the southwest quadrant and those in the northeast quadrant have exactly opposite rankings over candidates in each state. Now, for each belief β over states, we will have a line $L(\beta)$ passing through O which partitions the policy space into two elements - those to the northwest of $L(\beta)$ vote for the challenger and those to the southwest vote for the incumbent. Define $V(\beta)$ as the probability mass of voters to the northwest of $L(\beta)$. *SPM* would require that as β changes from 0 to 1 and consequently $L(\beta)$ rotates counterclockwise about O from L_B to L_A, the function $V(\beta)$ should be strictly increasing. Such an arbitrary restriction on distribution of voter ideal points is very unnatural and may be hard to obtain in reality. Weak Preferenve Monotonicity will also entail a very similar requirement that is equally unnatural and restrictive.

Before concluding, it is customary to discuss the limitations of the model. First, I have not included "realistic features" of elections such as the possibility of abstention, signaling motivation of voters, unanimity rules, costly information and so on[7]. A strand of the literature finds precisely these as possible reasons for breakdown of information aggregation. The reason these features have not been included is to demonstrate that the central source of aggregation failure is preference conflict, and to indicate the possibility that such conflicts may easily occur in real elections. Admittedly, this paper uses a particularly strict standard for evaluating whether elections aggregate information: aggregation is said to fail as long as there is *some* equilibrium in which there are "wrong" outcomes with high probability. The difficulty with the multiplicity of equilibria in this model is that there is no standard way to refine away one equilibrium or another, and therefore a sharper prediction cannot be made. This leaves open a research question which might be considered the "dual" of the issue addressed in this paper: what characteristics of voter preferences determine whether there exists *any* equilibrium that aggregates information? Hopefully, future research will throw more light on the issue.

[7] For these other sources of aggregation failure, see Feddersen-Pesendorfer (1998), Razin (2004), Persico (2004) and Martinelli (2006).

5 References

1. Austen-Smith, D., and J. Banks (1996): Information Aggregation, Rationality and the Condorcet Jury Theorem, *American Political Science Review*, 90(1), 34-45

2. Condorcet, Marquis de. [1785] (1976): Essay on the Application of Mathematics to the Theory of Decision-Making. Reprinted in *Condorcet: Selected Writings*, Keith Michael Baker, ed., 33-70. Indianapolis: Bobbs-Marrill Co.

3. Feddersen, T., and W. Pesendorfer (1997):Voting Behavior and Information Aggregation in Large Elections with Private Information, *Econometrica*, 65(5), 1029-1058

4. Feddersen, T., and W. Pesendorfer (1998):Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting, *The American Political Science Review*, 92(1), 23-35

5. Feddersen, T., and W. Pesendorfer (1999):Abstention in Elections with Asymmetric Information and Diverse Preferences, *The American Political Science Review*, 93(2), 381-98

6. Kim, J., and M. Fey (2007): The Swing Voter's Curse with Adversarial Preferences, *Journal of Economic Theory,* 135, 236-252

7. Ladha, K (1992): The Condorcet Jury Theorem, Free Speech and Correlated Votes, *AmerICAN Journal of Political Science,* 36, 617-34

8. Martinelli, C. (2005): Would Rational Voters Acquire Costly Information?, *Journal of Economic Theory,* 129, 225-51

9. Mclennan, A. (1998): Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents, *The American Political Science Review*, 92(2), 413-18

10. Meirowitz, A (2002):Informative Voting and Condorcet Jury Theorems with a Continuum of Types, *Social Choice and Welfare*, 19, 219-236

11. Milgrom, P. and R.J. Weber (1982): A Theory of Auctions and Competitive Bidding, *Econometrica*, 50, 1089-1122

12. Myerson, R.B. (1998): Extended Poisson Games and the Condorcet Jury Theorem, *Games and Economic Behaviour*, 25, 111-131

13. Myerson, R.B. (2000): Large Poisson Games, *Journal of Economic Theory,* 94 (1), 7-45

14. Persico, N. (2004): Committee Design with Endogenous Information, *Review of Economic Studies*, 71(1), 165-91

15. Razin, R. (2003): Signaling and Electing Motivations in a Voting Model with Common Values and Responsive Candidates, *Econometrica* 71(4), 1083-1120.

16. Wit, J. (1998): Rational Choice and the Condorcet Jury Theorem, *Games and Economic Behavior*, 22, 364-376

6 Appendix - Omitted Proofs

Proof of Lemma 1

Note first, that by the usual continuity arguments, as $\beta^n \to \beta_0$, $t(S, \beta^n) \to t(S, \beta_0)$. If $t(A, \beta_0) = t(B, \beta_0)$, and the Lemma holds trivially, which is the case if (but not necessarily only if) $\beta_0 \in \{0, 1\}$. Now, consider $\beta_0 \in (0, 1)$ and rewrite $H(n, \beta^n, \theta)$ as $\frac{\alpha}{1-\alpha}\left[\frac{1-t(A,\beta^n)}{1-t(B,\beta^n)}\right]^{n-m}\left[\frac{t(A,\beta^n)^\theta(1-t(A,\beta^n))^{1-\theta}}{t(B,\beta^n)^\theta(1-t(B,\beta^n))^{1-\theta}}\right]^m$ where $m = \frac{[n\theta]}{\theta}$. Since $m \geq n - \frac{1}{\theta}$, $m \to \infty$ as $n \to \infty$. Also, there is some $0 < \underline{t} < \bar{t}$ such that $\underline{t} \leq \left[\frac{1-t(A,\beta^n)}{1-t(B,\beta^n)}\right]^{n-m} \leq \bar{t}$ for all m and n, since $t(S, \beta) \in (\gamma_P, 1 - \gamma_Q)$ and $m - n \in \left[0, \frac{1}{\theta}\right]$. If there is some $\epsilon > 0$ such that $\frac{(t(A,\beta^n))^\theta(1-t(A,\beta^n))^{1-\theta}}{(t(B,\beta^n))^\theta(1-t(B,\beta^n))^{1-\theta}} > 1 + \epsilon$ for all n large enough, then $\lim_{n\to\infty} H(n, \beta^n, \theta) > \lim_{n\to\infty} \left(\frac{\alpha}{1-\alpha}\right) \underline{t} \left[\frac{t(A,\beta^n)^\theta(1-t(A,\beta^n))^{1-\theta}}{t(B,\beta^n)^\theta(1-t(B,\beta^n))^{1-\theta}}\right]^{\frac{[n\theta]}{\theta}} > \left(\frac{\alpha}{1-\alpha}\right) \underline{t} \left[\lim_{m\to\infty} (1 + \epsilon)^m\right] = \infty$. Hence the RHS of equation (6) is not bounded away from zero, which is a contradiction. Similarly, if there is some $\epsilon > 0$ such that $\frac{(t(A,\beta^n))^\theta(1-t(A,\beta^n))^{1-\theta}}{(t(B,\beta^n))^\theta(1-t(B,\beta^n))^{1-\theta}} < 1 - \epsilon$ for all n large enough, then $\lim_{n\to\infty} H(n, \beta^n, \theta) = 0$, and the RHS of equation (6) is not bounded away from 1.

Proof of Lemma 2

Consider the proof separately for each case. The proof for case (i) follows from Lemma 1. If β belongs to case (ii), the statement is vacuous. Next, consider case (iii) with the subcase that $f_u(0) > f_d(0)$. This implies that $t(A, \beta^n) > t(B, \beta^n)$ as $\beta^n \to 0$. Notice now that the the function $z^\theta(1 - z)^{1-\theta}$ is single peaked in z and attains its maximum at $z = \theta$. Now, if $\theta > V_B$, then for all large enough n, we have $\theta > t(A, \beta^n) > t(B, \beta^n)$, which would imply that $(t(A,\beta^n))^\theta (1 - t(A,\beta^n))^{1-\theta} > (t(B,\beta^n))^\theta (1 - t(B,\beta^n))^{1-\theta}$ for all n large enough. Therefore, we must have $H(n, \beta^n, \theta) > \left(\frac{\alpha}{1-\alpha}\right) \underline{t} \left[\frac{t(A,\beta^n)^\theta(1-t(A,\beta^n))^{1-\theta}}{t(B,\beta^n)^\theta(1-t(B,\beta^n))^{1-\theta}}\right]^{\frac{[n\theta]}{\theta}} > \left(\frac{\alpha}{1-\alpha}\right) \underline{t}$, which is bounded away from 0 in the limit. The subcase with $f_u(0) < f_d(0)$ and case (iv) follow similar logic.

Proof of Theorem 1 ("if" direction)

This is a proof by construction. Define the function $f_n(\beta, \theta) = \frac{1}{1+H(n,\beta,\theta)}$. If we can show that, given (n, θ), the function $f_n(\beta, \theta)$ has a fixed point β_n, then that β_n is the solution to the equilibrium condition (5). We show that for any $\theta \in \Theta(\beta^0)$, there is a sequence of fixed points β_n of $f_n(\beta, \theta)$ such that $\beta_n \to \beta^0$ as $n \to \infty$. We prove this separately for different values of β^0.

Suppose $g(x, y, \theta) = \frac{x^\theta(1-x)^{1-\theta}}{y^\theta(1-y)^{1-\theta}}$. for some $1 > x > y > 0$. It is then easy to show that $\frac{\partial g(x,y,\theta)}{\partial\theta} > 0$. This result is repeatedly used in the proof.

First consider some β^0 such that $t(A, \beta^0) \neq t(B, \beta^0)$. WLOG, assume $t(A, \beta^0) > t(B, \beta^0)$. For such a β^0, $\Theta(\beta^0) = \theta^*(\beta^0)$, and notice that $\theta^*(\cdot)$ has countinuous and bounded derivatives since $f_x(\mu)$ are continuous and bounded. Since we have ruled out the cases where $\theta^*(\beta^0)$ is neither increasing nor decreasing, there must be a neighbourhood $(\beta^0 - \epsilon, \beta^0 + \epsilon)$ where $\theta^*(\beta)$ is either only increasing or only decreasing, and because f_u and f_d are bounded, $t(A, \beta) > t(B, \beta)$. Suppose first that $\theta^*(\cdot)$ is decreasing in $(\beta^0 - \epsilon, \beta^0 + \epsilon)$. Write $H(n, \beta, \theta) = \frac{\alpha}{1-\alpha}\left[\frac{1-t(A,\beta)}{1-t(B,\beta)}\right]^{n-m}\left[\frac{t(A,\beta)^\theta(1-t(A,\beta))^{1-\theta}}{t(B,\beta)^\theta(1-t(B,\beta))^{1-\theta}}\right]^m$ as $B(\beta)\left[g(x, y, \theta)\right]^m$ where $B(\beta) = \frac{\alpha}{1-\alpha}\left[\frac{1-t(A,\beta)}{1-t(B,\beta)}\right]^{n-m}$ is bounded above and below and $g(x, y, \theta) = \frac{t(A,\beta)^\theta(1-t(A,\beta))^{1-\theta}}{t(B,\beta)^\theta(1-t(B,\beta))^{1-\theta}}$ with $x = t(A, \beta)$ and $y = t(B, \beta)$. Now, for $\beta \in (\beta^0, \beta^0 + \epsilon)$, $g(x, y, \theta^*(\beta^0)) > 1$, since $\theta^*(\beta^0) > \theta^*(\beta)$ as $\theta^*(\cdot)$ is decreasing, and $g(x, y, \theta^*(\beta)) = 1$ by definition. As $n \to \infty$, m must also go to ∞, and then, $\left[g(x, y, \theta^*(\beta^0))\right]^m \to \infty$, implying that $H(n, \beta, \theta^*(\beta^0)) \to \infty$, i.e $f_n(\beta, \theta^*(\beta^0)) \to 0$.

Hence, for $\beta \in (\beta^0, \beta^0 + \epsilon)$, we must have $f_n(\beta, \theta^*(\beta^0)) \to 0$ as $n \to \infty$. On the other hand, for $\beta \in (\beta^0 - \epsilon, \beta^0)$, we must have $f_n(\beta, \theta^*(\beta^0)) \to 1$ as $n \to \infty$. Consider the (continuous) function $f_n(\beta, \theta^*(\beta^0)) - \beta$ in the range $\beta \in (\beta^0 - \epsilon, \beta^0 + \epsilon)$. Given ϵ, for large enough n, it is positive for $\beta = \beta^0 - \epsilon$, and negative for $\beta = \beta^0 + \epsilon$. Thus, there must exist some $\beta_n \in (\beta^0 - \epsilon, \beta^0 + \epsilon)$ such that $f_n(\beta_n, \theta^*(\beta^0)) - \beta_n = 0$ for all n large enough. Thus, there exists a sequence β_n such that for any $\epsilon > 0$ small enough, there is some m such that for all $n > m$, $f_n(\beta_n, \theta^*(\beta^0)) = \beta_n$ and $\left|\beta_n - \beta^0\right| < \epsilon$. If $\theta^*(\beta)$ is increasing in $(\beta^0 - \epsilon, \beta^0 + \epsilon)$, then we can prove the theorem in an analogous way.

Next, consider $\beta^0 \in (0,1)\backslash\{1-\alpha\}$ such that $t(A,\beta^0) = t(B,\beta^0) = t$. By assumption 3, since the graphs of $t(A,\beta)$ and $t(B,\beta)$ "cross" at β^0, WLOG, consider a small interval $(\beta^0-\epsilon,\beta^0+\epsilon)$ such that $t(A,\beta) > t(B,\beta)$ for $\beta \in (\beta^0-\epsilon,\beta^0)$ and $t(A,\beta) < t(B,\beta)$ for $\beta \in (\beta^0,\beta^0+\epsilon)$. Since $t(A,\beta^0) = t(B,\beta^0)$, $f_n(\beta^0,\theta) = 1-\alpha$ for all (n,θ). Suppose now that $1-\alpha-\beta^0 < 0$. Then, consider any $\theta > t$. Since $t(A,\beta) > \theta^*(\beta) > t(B,\beta)$ for all $\beta \in (\beta^0-\epsilon,\beta^0)$, given θ we can choose ϵ small enough such that $\theta > \theta^*(\beta)$ for all $\beta \in (\beta^0-\epsilon,\beta^0)$. Therefore $f_n(\beta,\theta) \to 1$ in this interval. Now, consider the continuous function $f_n(\beta,\theta)-\beta$ in this interval. For large enough n, it is positive at $\beta^0-\epsilon$ and negative at β^0. Therefore, $f_n(\beta,\theta)$ must have a fixed point β_n in this interval. Thus, there exists a sequence β_n such that for any $\epsilon > 0$ small enough, there is some m such that for all $n > m$, $f_n(\beta_n,\theta) = \beta_n$ and $\left|\beta_n-\beta^0\right| < \epsilon$ for any $\theta > t$. By continuity of $f_n(\beta,\theta)$ in both β and θ, there is also a sequence of fixed points in beliefs arbitrarily close to β^0 for $\theta = t$. On the other hand, if $1-\alpha-\beta^0 > 0$, choose the interval $(\beta^0,\beta^0+\epsilon)$ where we have $f_n(\beta,\theta) \to 0$. The sequence of fixed points in beliefs will lie then in this interval for large enough n. To show the existence of a sequence of beliefs converging to β^0 for voting rules $\theta < t$, follow an analogous method.

Next, consider the cases with $\beta^0 \in \{0,1\}$. If $\beta^0 = 0$ and $f_u(0) > f_d(0)$, then we must have $t(A,\beta) > t(B,\beta)$ in some interval $(0,\epsilon)$. We also have $1-\alpha-\beta^0 = 1-\alpha > 0$. By the above method, we can show that for any $\epsilon > 0$ small enough, there exists a sequence of fixed points of $f_n(\beta,\theta)$ for any $\theta \in (0,V_B] = t(A,0) = t(B,0)$ in the interval $(0,\epsilon)$ that converges to $\beta^0 = 0$.We are done, since $\Theta(\beta^0) = (0,V_B]$. Other cases are similar.

Lastly, if $\beta^0 = 1-\alpha$ and $t(A,\beta^0) = t(B,\beta^0)$, consider a sequence $\beta_n = \beta_0$ for all n. We are done, since $f_n(\beta^0,\theta) = 1-\alpha = \beta^0$ for all n.

Proof of Lemma 3

First, suppose $t(A,\beta) > t(B,\beta)$ for all $\beta \in (0,1)$. From Lemma 2, $\Theta(\beta) = \theta^*(\beta)$ for all $\beta \in (0,1)$. Note that $\theta^*(\beta)$ is a continuous function that goes from V_B to V_A as β goes from 0 to 1. For any voting rule $\theta_c \in (V_B,V_A)$, by the Intermediate Value Theorem there must be some $\beta_c \in (0,1)$ such that $\theta^*(\beta_c) = \theta_c$. Now, there are two possibilities: $\beta_c \in M$ and $\beta_c \notin M$. If $\beta_c \in M$, by Theorem 1, there is an equilibrium for voting rule θ_c with belief β_c. If $\beta_c \notin M$, then, by assumption 4, it is possible to show that there must some other solution $\beta'_c \in M$ where $\theta^*(\beta'_c) = \theta_c$. In this case, rename β'_c as

β_c. Since $t(A, \beta_c) > \theta^*(\beta_c) > t(B, \beta_c)$, information is aggregated in this equilibrium in the limit. It remains to show that there is no equilibrium with $\beta = 0$ or 1 with $\theta_c \in (V_B, V_A)$. It suffices to show that $\Theta(0)$ and $\Theta(1)$ does not include consequential rules. Note that $t(A, \beta) > t(B, \beta)$ implies that $\int_{\beta_b}^{\beta_a} f_u(z)dz > \int_{\beta_b}^{\beta_a} f_d(z)dz$. This has to be true as $\beta \to 0$, which implies that $f_u(0) > f_d(0)$. It follows that $\Theta(0) < V_B$ and $\Theta(1) > V_A$, and we are done showing that there are no equilibria with $\theta_c \in (V_B, V_A)$ that do not aggregate information.

To see the other direction, note that $t(S, \beta)$ is continuous in β. So, if there is some β such that $t(A, \beta) < t(B, \beta)$, there must be some β' such that we have $t(A, \beta') = t(B, \beta') = t'$ then $\Theta(\beta') = (0, 1)$. It follows from Theorem 1 that there is an equilibrium at β' which does not aggregate information.

Proof of Proposition 2

Proof of Necessity: If condition (i) does not hold, FIE fails for consequential rules, according to Proposition 1. If condition (ii),does not hold, suppose there is some β' such that $\theta^*(\beta') > V_A$, then there is an equilibrium for $\theta = \theta^*(\beta')$ with limiting belief β', and $\mathcal{P}$ wins in one state and loses in another. Since $\theta > V_A$, it is a $Q-$trivial rule and FIE fails in this equilibrium. Similarly, if there is some β'' such that $\theta^*(\beta'') < V_B$, then FIE fails for some $P-$trivial rule.

Proof of Sufficiency: FIE holds for consequential rules as long as condition (i) holds, again by Proposition 1. To see sufficiency for P-trivial rules, notice that WPM implies that we must have $f_u(0) > f_d(0)$, and therefore, $\Theta(0) = (0, V_B]$. Thus, for all P-trivial rules, the equilibrium induced prior is 0, when the vote share is V_B in each state. Information is aggregated in this equilibrium. Condition (ii) implies that there is no other limiting equilibrium for P-trivial rules. The proof for Q-trivial rules is similar.

Proof of Lemma 4

Notice that $\frac{dt(S,\beta)}{d\beta} = q_s [f_u(\beta_a) - f_d(\beta_a)] + (1 - q_s) [f_u(\beta_a) - f_d(\beta_a)] > 0$ by SPM. Since $t(S, \beta)$ is strictly monotonic and $t(S, 0) = V_B$ and $t(S, 1) = V_A$, we must $t(S, \beta) \in (V_B, V_A)$ for all $\beta \in (0, 1)$ and $S \in \{A, B\}$. Since $\theta^*(\beta)$ must lie strictly between $t(B, \beta)$ and $t(A, \beta)$, $\theta^*(\beta)$ must also lie in the interval (V_B, V_A) for all $\beta \in (0, 1)$.