Preference Reversal and Information Aggregation in Elections

Sourav Bhattacharya*

Department of Economics, University of Pittsburgh

Email: sourav@pitt.edu

November 7, 2007

Abstract

I analyse informational efficiency of two-alternative elections where the utility of the voters depends on the realisation of an uncertain, binary state variable about which voters receive an independent, noisy signal. I show that large elections aggregate information efficiently for any voting rule in the unique equilibrium if and only if the set of voters who favour an alternative in one state includes the set which favours the same alternative in the other state. I call this the *preference monotonicity* condition. If preference monotonicity fails, we have two groups of voters such that a change in state induces their rankings to change in opposite ways; and I term this phenomenon as *preference reversal*. Under preference reversal, for large classes of voting rules we have equilibria with certain outcomes different from the full information outcome. Preference reversal is the generic condition when voter preferences are multidimensional.

*While writing this paper, I have benefited a lot from discussions with David Austen-Smith and Tim Feddersen. Roger Myerson helped me immensely by suggesting a more efficient proof of Theorem 3. I also thank Steve Callander, Marciano Siniscalchi, Sean Gailmard, Jaehoon Kim, Navin Kartik, Alexandre Debs, Siddarth Madhav and participants in the Voting and Information panel at the Econometric Society Summer Conference (2006), USC Marshall School, University of Pittsburgh and Sabanci University . All responsibility for any errors remaining in the paper is mine.

1 Introduction

A deep problem with group decision making under uncertainty is that decision relevant information is often dispersed throughout the group. Traditionally, elections have been thought to solve this problem by ensuring that all the individuals' private information is incorporated in the election outcome. A strong argument for using elections for making important political decisions is that the society, by aggregating everyone's information, would be better off compared to any individual making the decision on behalf of the society.

Starting from Condorcet (1785), several mathematical models have been formulated to express the idea that voting efficiently aggregates information dispersed in the electorate. In most models, there are two alternatives, a state variable that captures the decision relevant uncertainty and every individual receives a noisy independent signal conditional on the state. These models exhibit that, in a large electorate, the voting outcome is almost surely the same as the one that would have occurred if the state variable were common knowledge: thus the individual uncertainty vanishes in the aggregate. All these models assume diversity in information but either complete homogeneity (e.g. Myerson (1998a, 1998b, 2000), Wit(1998), Meirowitz (2002))or a very limited heterogeneity in preferences (e.g. Feddersen-Pesendorfer (1996, 1997, 1999)) among individuals in the society. The current paper explores the limits of the possibility of information aggregation through voting with plurality rules[1]. It develops a condition on the relationship between preferences and uncertainty in the electorate that is not only sufficient, but also necessary for elections to be full information equivalent. While the condition admits existing results on information aggregation like that in Feddersen-Pesendorfer (1997), henceforth F-P, as a special case, I show that the condition itself holds true only in a non-generic set of situations. In particular, it is hard to aggregate information when voters in an election care about more than one issue.

Although the earlier proofs of what is today called the Condorcet Jury theorem[2] assumed that each individual voted according to their private signal, recent work

[1]In this paper I focus only on elections with two given alternatives. The voting rules considered are plurality rules or q-rules, according to which the candidate getting more than q share of the votes wins the elections, where $q \in (0, 1)$. I, however, denote a voting rule in this paper by θ.

[2]For earlier proofs of this theorem using statistical arguments, see Berg (1992), Ladha (1992, 1993), Nitzan and Paroush (1985).

(Austen-Smith and Banks (1996)) has suggested that such behavior may not be optimal for rational voters[3]. Much like a bidder in an auction, each voter conditions his decision on the event that the others are tied and his vote actually decides the outcome[4]. Conditioning on being pivotal, a voter may learn more about the state from the equilibrium strategies of the other voters, and consequently may vote against his signal. F-P has shown that in equilibrium, even though only a small fraction of voters vote informatively, in a large election the actual number of voters using their information is large enough for the outcome to be almost surely the full information outcome. For this result, their paper assumes what they call "common values"[5]: for each individual, the relative valuation of an alternative is a strictly monotone function of the state variable. Thus, a change in state makes *all* voters more inclined in favour of exactly one alternative.

In this paper I study a spatial model of electoral competition, allow for a more general correlation between preferences and the state variable, and examine a set-up that is generalisable to the case of multidimensional space of voter preferences. For the sake of tractability, unlike F-P which considers a continuum of states, I consider only two states (say L and R). Suppose the alternatives to be chosen from are denoted by P and Q. I consider the limit of a sequence of equilibria of the voting game as the number of voters becomes large. The condition for information aggregation presented here is that under full information, the set of voters who prefer alternative P under state L should include (be included by) the set of voters who prefer P under state R. Since this is a monotonicity condition on sets determined by voter preference, I call this condition *preference monotonicity.* Note that as the state changes from R to L, for *all* those voters whose ranking is sensitive to the state, the preferred alternative changes from Q to P (from P to Q). This monotonicity condition is implied by the common values condition in F-P (and obviously by the setting where all voters have the same preference), but the reverse implication is not true. This makes the settings

[3]See McLennan (1998) for a sophisticated enunciation of Condorcet's Jury Theorem and formal proof showing that if there exists an outcome that aggregates information with sincere voting, there exists a Nash equilibrium that does the same too.

[4]See Battaglini, Morton and Palfrey (2006) and Goeree and Yariv (2006) for experimental evidence that voters condition their decision on information about the state learnt from the event of being pivotal.

[5]F-P use the term "common values" in a sense slightly weaker than the way the term is defined for the first time in the context of auctions in Milgrom and Weber (1982). In the auction context, if the ranking of alternatives is the same for all individuals given a state, then we have common values. However, in this paper, we shall refer to the term in the sense used by F-P.

studied so far in the literature special cases of preference monotonicity.

Moreover, this paper shows that preference monotonicity is also necessary for information aggregation in the sense that if the condition is not satisfied, we always get equilibria which give a "wrong" outcome with a high probability in one state. As long as there are any two groups of voters with positive measure such that one group prefers alternative P in state L and alternative Q in state R while the other group prefers Q in L and P in R, there are multiple equilibria for large classes of voting rules, at most one of which aggregates information. It is worth noting that this failure does not depend on the precision of signals or on the size of the conflicting groups. Therefore, the source of informational failure of the election system is the existence of state-contingent conflict among voter groups, a phenomenon which I term as *preference reversal*. Additionally, if individual preferences are defined over multiple dimensions, then in a sense that will be defined later, preference reversal happens generically. Therefore, the claim that elections aggregate information appears tenuous.

1.1 Examples

1. Members of a jury are trying to determine whether the defendant is guilty or innocent. Each member of the jury wants to convict if the defendant is guilty and acquit him if he is innocent, but they have only information about the state. Here, since everyone has the same ranking over alternatives given the state, both common values and preference monotonicity are satisfied. This is the case of complete homogeneity of preference.

2. Continuing with the jury example, suppose, instead of being guilty and innocent, the state is the level of guilt. Individuals have similar preferences in the sense that everyone wants to acquit for low levels of guilt and convict for high levels, but they vary with the precise level at which they switch from acquittal to conviction. So, as the level of guilt increases, more and more members favour the guilty verdict - this is the classic common values situation. By implication, preference monotonicity is also satisfied.

In the situations depicted by examples 1 and 2, we know that information is fully aggregated. However, we can have situations where monotonicity of preference does not hold, and others in which, even if common values holds, preference monotonicity does not.

3. Suppose a country has so far been isolated and now is voting on whether to allow free trade by joining the WTO. Because of its isolation, it has developed both an industrial sector and an agricultural sector to suit its own consumption needs. If the country allows free trade, the sector in which it has comparative advantage will grow and the other will shrink. Assume that the voters do not know where the comparative advantage of the country lies. If the advantage lay in industry and this was commonly known, those engaged in the industrial sector would vote in favour of joining WTO while those in the agricultural sector would vote against, and conversely if the advantage lay in agriculture. Thus there are two interest groups who have exactly opposite ranking over the alternatives in each state, and preference reversal obtains, leading to a failure of information aggregation.

The next example introduces the basic structure in which the model studied.

4. Consider an election with an incumbent candidate and a challenger. Assume that a candidate can only commit to his own most preferred point on the policy space, and the policy space is the left-right ideological space, which can be represented by an interval of real numbers. The incumbent's best point is known to be Q, but there is some uncertainty about the location of the challenger on the policy space, which can be one of two known points: either L or R. If L is to the left of Q and R to the right, then the extreme leftists prefer the challenger only when he is located at L (to the left of the incumbent) while the extreme rightists prefer the challenger only when he is at R (to the right of the incumbent). In that case, we are in situation of preference reversal similar to the one described in example 3. However, if L is to the left of R, but both locations are to the left of Q, then, for all practical purposes we have a leftist challenger and a rightist incumbent. The location L is the more extreme location and R is a relatively moderate location of the challenger. As the challenger becomes more extreme (state changes from R to L), he loses support of some of the moderate voters but does not gain any new supporters. In other words, the set of voters that prefer the challenger in the extreme state L is a subset of those who prefer the challenger in the moderate state R: hence preference monotonicity prevails. It should be noted that this is not a common values setting because the effect of a change in state on the utility difference between the two alternatives for an

extreme rightist is the opposite of that for an extreme leftist.

5. To see that preference reversal (and not preference monotonicity) is the more generic situation, consider the same example with an incumbent and challenger, but with the policy space being the square $[-1,1]\times[-1,1]$ as shown in Figure 1. Suppose the incumbent's ideal point is located at $Q=(0,0)$, and the two possible locations of the challenger are $L=(-0.5,0)$ and $R=(0,0.5)$. Now, the policy space is divided into four sections as shown in Figure 1. The voters with ideal points lying in southwest rectangle support the incumbent in state R and the challenger in state L, while the voters with their ideal points lying in the northeast rectangle have the exactly opposite ranking in each state. This is a situation of preference reversal. Preference reversal is generic in the two-dimensional setting in the following sense: suppose the policy space is denoted by $[-x,x]\times[-x,x]\subset\mathbf{R}^2$. Then, for any two locations of L and R that do not lie on a ray emanating from Q, if x is large enough, the policy space is always divided into four segments with two segments having exactly opposite ranking in each state.

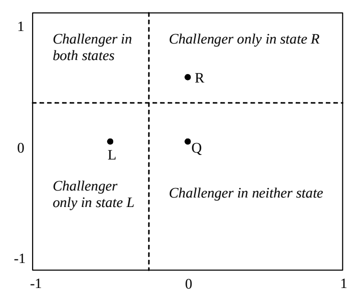

Fig 1: Preference Reversal in a multidimensional policy space

In the main body of the paper, I formally study the setting introduced in example 4. Since the setting has the advantage of incorporating both preference monotonicity and preference reversal depending on the parameters (relative locations of Q, L and R), we can compare equilibria across the two cases and study their aggregation properties. This not only allows us to show when elections may fail to aggregate information, it also sheds light on *why* aggregation fails. Like F-P, we find that in equilibrium only a subset of voters are *responsive* in the sense that changes in information received changes their voting decision. When preferences are monotonic, for *all* responsive voters, the same information induces the same voting decision. But under preference reversal, the same information may induce opposite voting decisions among different groups of responsive voters. What is surprising is that the very possibility of different voters interpreting the same information in different ways leads to outcomes that are "wrong" with a very high probability. The fact that this breakdown does not depend on the size of conflicting groups, the accuracy of signals or on the exact distribution of voter preferences indicates that the source of informational inefficiency in voting is just the existence of two groups of voters that never agree with each other.

We are ultimately interested in the case of example 5, i.e. when voter preferences are defined over many issues. In an extension I show that this analysis and results apply to the case with multidimensional policy space as well, and additionally, we almost always find preference reversal in this case. Therefore, we conclude that information aggregation through voting is an artifact of the assumption of single-dimensional preferences. Whether elections aggregate information or not ultimately depends on whether one believes that elections are fought over one salient issue or over many.

2 The Set-up

In this section 1 discuss the basic set-up and a few definitions that will be used throughout the paper. 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 θ of the votes[6], then $\mathcal{P}$ wins; otherwise the status quo $\mathcal{Q}$ wins. Assume that the policy space is $[-1,1]$. While $\mathcal{Q}$ is known to be located at 0, there is uncertainty about the location of the alternative $\mathcal{P}$. $\mathcal{P}$ is located at $L \in [-1,1]$ or $R \in [-1,1]$ with equal probability. The event that $\mathcal{P}$ is located at S, where $S \in \{L, R\}$ is referred to as state S. To give a natural meaning to the names of the state, I assume that $L < R$[7]. I also assume that the policy never coincides with the status quo, i.e

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

[7] There is some loss of generality - by this assumption, we exclude the case that in the policy location is state invariant, i.e. $L = R$. Thus assuming $L < R$ is tantamount to assuming that there is always some uncertainty about the policy location.

both L and R are non-zero[8]. Each voter receives a private signal $\sigma \in \{l, r\}$ about the state. Signals are independent and identically distributed conditional on the state, with the distribution being:

$$\Pr(l|L) = \Pr(r|R) = q \in (\frac{1}{2}, 1)$$

Voters have single peaked preferences defined on the policy space. Every individual has a privately known bliss point x that is drawn independently from a commonly known distribution $F(\cdot)$ with support $[-1, 1]$ and a density $f(\cdot)$. The utility from the alternative A, when it is located at a, is given by:

$$U(x, A) = -(x-a)^2, \ A \in \{\mathcal{Q}, \mathcal{P}\}$$

Given a draw of x and S, I define $v(x, S)$ as the difference in utility between the policy alternative and the status quo:

$$v\,(x, S) = U(x, \mathcal{P}) - U(x, \mathcal{Q}) = x^2 - (x - S)^2, \ S \in \{L, R\} \tag{1}$$

We shall use $v(x, S)$, the utilitity difference between the two alternatives as given by (1) for all further analysis. If the state S is known, a voter votes for $\mathcal{P}$ if and only if $v(x, S)$ is non-negative. If S is not known, a voter calculates the expected value of this function using the relevant probability distribution over states and votes $\mathcal{P}$ if the expectation is non-negative.

Based on the location of the policy $\mathcal{P}$, the following very important condition distinguishes preference monotonicity from preference reversal.

Definition 1 *Define $\mathbb{P}(S)$ to be the set of types that (weakly) prefer the alternative policy to the status quo if they know that the state is S:*

$$\mathbb{P}(S) = \{x : v(x, S) \geq 0\}$$

Suppose $\mathbb{P}(L) \neq \mathbb{P}(R)$. $\mathbb{P}(S)$ *exhibits* preference monotonicity *if* $\mathbb{P}(L) \subset \mathbb{P}(R)$ *or* $\mathbb{P}(L) \supset \mathbb{P}(R)$, *and* preference reversal *otherwise.*

[8] In other words, we assume that if the state were known, then there will always be a positive interval of types that would strictly prefer to vote for the policy in either state.

Denote $\mathbb{P}(L) \cap \mathbb{P}(R)$ as $\mathbb{P}_{LR}$, and notice that it can be empty. The set of types $\mathbb{P}_{LR}$ always votes for the policy irrespective of the state. They are *committed types,* or *type-$\mathcal{P}$ partisans* according to the nomenclature in Feddersen-Pesendorfer (1996)[9]. Now consider the sets $\mathbb{P}(L) \backslash \mathbb{P}_{LR}$ and $\mathbb{P}(R) \backslash \mathbb{P}_{LR}$. These are the *independent types*, as they change their vote based on the state. Definition (1) says that preferences are monotonic if and only if all independent types switch their votes in the same direction ($\mathcal{P}$ to $\mathcal{Q}$ or $\mathcal{Q}$ to $\mathcal{P}$) when S changes. Under preference reversal, some independents switch from $\mathcal{P}$ to $\mathcal{Q}$ and some from $\mathcal{Q}$ to $\mathcal{P}$ for a change in S. This set theoretic definition is equivalent to the following algebric definition.

Definition 2 *Define $\mathbb{V}(x, S)$ to be the function indicating whether the type x weakly prefers the alternative $\mathcal{P}$ in state S:*

$$\mathbb{V}(x) = \begin{cases} 1 \; if \;\; v(x, S) \geq 0 \\ 0 \; if \;\; v(x, S) < 0 \end{cases}$$

Suppose $\mathbb{V}(x, L) \neq \mathbb{V}(x, R)$ for some x. $\mathbb{V}(x, S)$ exhibits preference monotonicity *if $\mathbb{V}(x, L) \geq \mathbb{V}(x, R)$ for all x or if $\mathbb{V}(x, L) \leq \mathbb{V}(x, R)$ for all x, and* preference reversal *otherwise.*

Definition (2) is a different representation of Definition (1)[10]. For the committed types, we must have $\mathbb{V}(x, L) = \mathbb{V}(x, R)$. For the independent types, we must have $\mathbb{V}(x, L) \neq \mathbb{V}(x, R)$. If for some independents we have $1 = \mathbb{V}(x, L) > \mathbb{V}(x, R) = 0$ and for some others we have $0 = \mathbb{V}(x, L) < \mathbb{V}(x, R) = 1$, the two groups have opposed rankings in each state - and we are in a situation of preference reversal. For preference to be monotonic, we need all independents to have $\mathbb{V}(x, L) > \mathbb{V}(x, R)$ (equivalent to $\mathbb{P}(R) \subset \mathbb{P}(L)$) or $\mathbb{V}(x, L) < \mathbb{V}(x, R)$ (equivalent to $\mathbb{P}(L) \subset \mathbb{P}(R)$). The common values assumption of F-P reduces in the two state framework to $v(x, L) - v(x, L)$ having the same sign for all x. It is easy to show that this assumption implies preference monotonicity.

[9] Given that the location of Q is known and $b_i \neq 0$, there is always an interval of types around 0 that are *Q-partisans.*

[10] A third way to characterise these settings is as follows. Consider any two ideal points $x < x'$ lying in the interior of $[-1, 1]$.Also note that $L < R$. Then, we have preference monotonicity if $v(x, L)v(x', R) > v(x, R)v(x', L)$ and preference reversal if $v(x, L)v(x', R) < v(x, R)v(x', L)$. Notice the similarity with the affiliation property in Milgrom and Weber (1982). However, in this case, the preference monotonicity condition is not equivalent to log supermodularity since the $v(\cdot, \cdot)$ can be negative.

Remark 1 *If L and R have the same sign, we have preference monotonicity, and if they have different signs, we have preference reversal.*

Proof. In Appendix. ■

The intuition behind this remark is illustrated by example 4 in Section 1.

The equilibrium concept we employ is symmetric Bayesian Nash equilibrium in undominated strategies.

Given an individual's private information (bliss point x and signal σ), the strategy specifies a probability of voting for $\mathcal{P}$:

$$\pi(x, \sigma) : [-1, 1] \times \{r, l\} \to [0, 1]$$

Thus, under state S, the expected share of votes is:

$$t(S, \pi) = \int_{-1}^{1} \Pr(l|S)\pi(x, l)dF(x) + \int_{-1}^{1} \Pr(r|S)\pi(x, r)dF(x), \;\; S = L, R \qquad (2)$$

Exapanding (2) we can write

$$\begin{aligned} t(L, \pi) &= q \textstyle\int_{-1}^{1} \pi(x, l)dF(x) + (1 - q) \int_{-1}^{1} \pi(x, r)dF(x) \\ t(R, \pi) &= (1 - q) \textstyle\int_{-1}^{1} \pi(x, l)dF(x) + q \int_{-1}^{1} \pi(x, r)dF(x) \end{aligned}$$

Under a rule θ a voter is pivotal if $n\theta$ votes are cast for the policy $\mathcal{P}$ from among the remaining n voters. So, the probability of being pivotal under state S is given by[11]:

$$\Pr(piv|\pi, S) = \binom{n}{n\theta} (t(S, \pi))^{n\theta} (1 - t(S, \pi))^{n - n\theta}, \;\; S = L, R \qquad (3)$$

Note that (3) actually denotes a pair of equations, one for each state. Call these the *pivot equations*. Note that if $t(S, \pi) \in (0, 1)$ then $\Pr(piv|\pi, S) > 0$. I show later that in any equilibrium of the model, we must have $t(S, \pi) \in (0, 1)$. Assuming the belief on the state conditional on being pivotal is well defined, it is given by:

$$\beta (S|piv, \pi) = \frac{\Pr(piv|\pi, S)}{\Pr(piv|\pi, L) + \Pr(piv|\pi, R)}, \;\; S = L, R \qquad (4)$$

Since $\Pr(piv|\pi, S) > 0$ for both states, we have $\beta (S|piv, \pi) \in (0, 1)$. The strategies played in equilibrium determine the pivot probabilities in each state through (2) and

[11]For technical convenience, we assume that $n\theta$ is an integer.

(3). In return,the probability of state L conditional on being pivotal is determined by Bayes rule by (4). I call $\beta(L|piv,\pi)$ the *induced prior* and denote it as β_L. The posterior beliefs given a signal are:

$$\left.\begin{array}{l}\beta(L|piv,\pi,l)=\frac{q\beta_L}{q\beta_L+(1-q)(1-\beta_L)}\\ \beta(L|piv,\pi,r)=\frac{(1-q)\beta_L}{(1-q)\beta_L+q(1-\beta_L)}\\ \beta(R|piv,\pi,l)=\frac{(1-q)(1-\beta_L)}{q\beta_L+(1-q)(1-\beta_L)}\\ \beta(R|piv,\pi,r)=\frac{q(1-\beta_L)}{(1-q)\beta_L+q(1-\beta_L)}\end{array}\right\} \tag{5}$$

I refer to $\beta(L|piv,\pi,l)$ as p_l and to $\beta(L|piv,\pi,r)$ as p_r. Note that while both p_l and p_r are increasing functions of the induced prior, p_l is concave and p_r is convex throughout. This, coupled with their equality at the extreme values of β_L, i.e. $p_l = p_r = \beta_L$ at $\beta_L = 0$ and $\beta_L = 1$, implies that $p_l > p_r$ for all other values of β_L. Figure 2 graphs the posteriors as functions of the induced prior β_L.

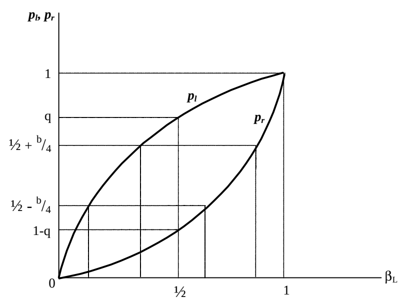

Fig 2: Posteriors as functions of the induced prior belief

The following is an important definition that will serve to distinguish between the nature of equilibria under monotonic and non-monotonic preferences.

Definition 3 *A voting strategy is a* cut-off *strategy if, given a signal σ and an induced belief β_L, the type space $[-1,1]$ can be partitioned into exactly two intervals (one possibly empty) such that every type votes for $\mathcal{Q}$ in one interval and for $\mathcal{P}$ in the other. Suppose, given a signal σ and some induced belief β_L, there is a cut-off type $x_\sigma(\beta_L)$ such that the types to the right (left) of the cut-off vote for $\mathcal{P}$. The cut-off strategy is said to be* ordered[12] *if, for all values of the induced belief $\beta_L \in [0,1]$, the types to the right (left) of the cut-off $x_\sigma(\beta_L)$ vote for $\mathcal{P}$.*

[12]Note that the definition of ordering of cut-offs is different here from the one in F-P (page 1035) where ordering is defined based on whether cut-offs are monotonic in signals. Here, for any location

In other words, a voting strategy is a cut-off strategy if given $\sigma \in \{l, r\}$ and $\beta_L \in (0,1)$, there is some $x_\sigma(\beta_L)$ such that for any $x_1 < x_\sigma(\beta_L)$ and $x_2 > x_\sigma(\beta_L)$, we have $\pi(x_1, \sigma) \in \{0,1\}$.and $\pi(x_2, \sigma) = 1 - \pi(x_1, \sigma)$. If $x_\sigma(\beta_L) \in \{0,1\}$, a cut off strategy requires that $\pi(x, \sigma)$ be 0 or 1 for all x. A cut-off strategy is said to be ordered if, given some β_L and any two types x and x', if we have $\pi(x, \sigma) > \pi(x', \sigma)$,then for all values of β_L, we have $\pi(x, \sigma) \geq \pi(x', \sigma)$.We shall see that while we always have cut-off strategies, cut-offs are ordered when preferences are monotonic and unordered when there is preference reversal.

2.1 Conditions for Information Aggregation

Before identifying equilibrium strategies, I classify the voting rules and lay down the conditions that need to be satisfied for information to be aggregated in equilibrium given a voting rule.

We shall call a voting rule *consequential* if under that rule, we get different outcomes under different states if the states were common knowledge. On the contrary, if the voting threshold is such that under full information, the same outcome obtains in each state, we call the it a *trivial* rule. I shall discuss this classification in more detail in section 3.3.

If the vote of an individual with type x changes with the signal, i.e. if $\pi(x, l) \neq \pi(x, r)$, then type x is said to be *responsive*. Given a voting rule, the characteristics of the responsive set of voters determines whether information will be successfully aggregated.

Definition 4 *Suppose that given a consequential rule, under full information, $\mathcal{P}$ wins under state S and $\mathcal{Q}$ in the other state, i.e. the state $\{L, R\} \backslash S$. A type x is said to be* aligned *with the society if he prefers $\mathcal{P}$ in state S and $\mathcal{Q}$ in the other state. If on the other hand, a type x prefers $\mathcal{Q}$ in state S and $\mathcal{P}$ in the other state, then we call the type* mis-aligned*.*

Note that the responsive set of voters contains only independent types, and independents are either aligned or misaligned. If the majority of types in a set of voters is aligned, the set itself is said to be aligned.

of $\mathcal{P}$, we always have cut-offs monotonic in the signals. However, it is possible that for some values of the induced belief, those to the left of the cut-off vote for $\mathcal{P}$ while for other values, those to the right of the cut-off vote for $\mathcal{P}$. We distinguish those situations as unordered.

For voting with a consequential rule θ, we need the following conditions to be satisfied in equilibrium for the outcomes to be full information equivalent:

1. The responsive set should be *influential*, i.e. the overall voting outcome should change as the responsive types vote differently in the different states. In other words, the vote share for $\mathcal{P}$ should be higher than the threshold θ in one state and lower in the other.

2. The responsive set should be *aligned* with the society, and thus contribute more votes for the "correct" alternative in each state.

Both conditions are satisfied under equilibrium in the common values situation, but in the non-common values situation, each can individually fail in the limiting equilibrium.

On the other hand, for voting with trivial rules, we need the responsive types *not* to be influential for information to be aggregated.

In what follows, I find the equilibria and their aggregation properties under monotonic preferences in Section 3 and under preference reversal in Section 4. I show that while information is aggregated in the unique equilibrium for all voting rules when preferences are monotonic, non-aggregating equilibria occur for consequential rules and some trivial rules when there is preference reversal. In Section 5, I discuss how the model extends to the case of multidimensional policy space. Section 6 discusses the relation to existing literature and concludes.

3 Preference Monotonicity

I start by looking at the benchmark case of preference monotonicity. I show that both the conditions for information aggregation as mentioned above are satisfied for every voting rule.

3.1 Strategies and equilibria

Lemma 1 *Under preference monotonicity, all equilibrium strategies are ordered cut-off strategies.*

Proof. A voter with signal $\sigma, (\sigma \in \{l, r\})$ evaluates the state using the distribution $\beta(S|piv, \pi, \sigma)$ and votes for the policy if and only if the expected value of the function $v(\cdot, \cdot)$ is non-negative. Assume for now that $\beta(S|piv, \pi, \sigma)$ is well-defined. Define $x(p_\sigma)$ as the solution of the equation $E(v(x, S)|piv, \pi, \sigma) = 0$. Solving,

$$x(p_\sigma) = \frac{1}{2}\left(\frac{(L)^2 p_\sigma + (R)^2 (1 - p_\sigma)}{Lp_\sigma + R(1 - p_\sigma)}\right) \in \left[\frac{L}{2}, \frac{R}{2}\right]$$

Thus, $x(p_\sigma)$ always exists uniquely. Also, since $\frac{\partial Ev(x,S)}{\partial x} = 2\left(Lp_\sigma + R(1 - p_\sigma)\right)$, $R > L > 0 \Rightarrow \frac{\partial Ev(x,S)}{\partial x} > 0 \Rightarrow Ev(x, S) > 0$ iff $x > x(p_\sigma)$. Similarly, if $L < R < 0$, $Ev(x, S) > 0$ iff $x < x(p_\sigma)$. This establishes the cut-off nature of strategies. If $L < R < 0$, types to the left of the cut-off $x(p_\sigma)$ vote for $\mathcal{P}$, while if $0 < L < R$, types to the right of the cut-off $x(p_\sigma)$ vote for $\mathcal{P}$. Thus the cut-off strategies are ordered too, and the lemma is proved, under the assumption that $\beta(S|piv, \pi, \sigma)$ is well-defined. ■

Denote $x(p_l)$ as x_l and $x(p_r)$ as x_r. The the cut-off strategies are given by (6) and (7):

$$\left\{\begin{array}{l} \pi(x, l) = \begin{cases} 1 \; if \; x \geq x_l \\ 0 \; otherwise \end{cases}, \\ \pi(x, r) = \begin{cases} 1 \; if \; x \geq x_r \\ 0 \; otherwise \end{cases} \end{array}\right\} \text{when } R > L > 0 \tag{6}$$

$$\left\{\begin{array}{l} \pi(x, l) = \begin{cases} 1 \; if \; x \leq x_l \\ 0 \; otherwise \end{cases}, \\ \pi(x, r) = \begin{cases} 1 \; if \; x \leq x_r \\ 0 \; otherwise \end{cases} \end{array}\right\} \text{when } L < R < 0 \tag{7}$$

Remark 2 *For* $\beta_L = 1$, $x_r = x_l = \frac{L}{2}$, *and likewise for* $\beta_L = 0$, $x_r = x_l = \frac{R}{2}$. *Since* $\frac{dx(p)}{dp} = -\frac{1}{2}\left(\frac{(R-L)LR}{(Lp+R(1-p))^2}\right) < 0$, *and since* $p_l > p_r$ *for* $\beta_L \in (0, 1)$, $x_r > x_l$ *for non-degenerate values of* β_L.

Thus, for any induced prior, the strategies in the benchmark case are characterised by cutpoints x_l and x_r, with $x_l \leq x_r$. If $R > L > 0$, types $x < x_l$ always vote for $\mathcal{Q}$, types $x \in [x_l, x_r]$ vote for $\mathcal{P}$ if they get signal l and $\mathcal{Q}$ if they get signal r, and the types $x > x_r$ vote for $\mathcal{P}$ regardless of the signal. If $L < R < 0$, types left of x_l always vote for $\mathcal{P}$ and those right of x_r always vote for $\mathcal{Q}$ while types in $[x_l, x_r]$ vote informatively. In either case, $[x_l, x_r]$ is the responsive set, while the other types vote according to their bias. Henceforth, I shall deal only with the case $L < R < 0$, noting that the other case is completely symmetric.

Note that the ordered cut-off nature of the strategies ensures that there will always be one and only one interval of responsive types. Also, irrespective of the location of the cutoffs, the responsive set is always *aligned* with the society. Thus, for consequential rules, all we need to show for information aggregation is that in any limiting equilibrium, the responsive set is influential. For this, we need vote shares to be monotonic in the induced priors under each state, which is again ensured by the ordered nature of the cut off strategies. I define the probability of an individual voting for the alternative $\mathcal{P}$ given σ as z_σ, *i.e.* $z_\sigma \equiv \int_{-1}^{1} \pi(x, \sigma) dF$. For $L < R < 0$, we have from (7),

$$z_\sigma = F(x_\sigma),\ \sigma = \{l, r\}$$

Therefore, using (2) we write[13]:

$$\left.\begin{aligned} t(L,\pi) &= qz_l + (1-q)z_r = qF(x_l) + (1-q)F(x_r) \\ t(R,\pi) &= (1-q)z_l + qz_r = (1-q)F(x_l) + qF(x_r) \end{aligned}\right\} \tag{8}$$

Note that since the cut-offs x_l and x_r are functions of the induced prior, the vote shares $t(L, \pi)$ and $t(R, \pi)$ are also functions of β_L. The following lemma examines how the vote share in each state changes as a function of the induced prior.

Lemma 2 *If $L < R < 0$, the expected share of votes $t(S, \pi)$ in state $S \in \{L, R\}$ decreases strictly with the induced prior β_L from $F(\frac{R}{2})$ at $\beta_L = 0$ to $F(\frac{L}{2})$ at $\beta_L = 1$. Also, for all interior values of the induced prior, i.e. for all $\beta_L \in (0, 1)$, the vote share $t(L, \pi)$ in state L is stricty less than the share $t(R, \pi)$ in state R*[14].

Proof. By Remark 2, at $\beta_L = 0$, $z_l = z_r = F(\frac{R}{2}) \;\Rightarrow t(S, \pi) = F(\frac{R}{2})$ for $S \in \{L, R\}$. Similarly, at $\beta_L = 1$, $t(S, \pi) = F(\frac{L}{2})$ for $S \in \{L, R\}$. Also, since p_σ is a strictly increasing function of β_L, x_σ is decreasing in β_L by Remark 2. The full support assumption guarantees that $F(\cdot)$ is strictly increasing. Hence, $t(S, \pi)$ is

[13] For $0 < L < R$, we have $z_\sigma = G(x_\sigma)$, where $G(y) \equiv 1 - F(y)$, $y \in [-1, 1]$

[14] If $0 < L < R$, then both $t(L, \pi)$ and $t(R, \pi)$ increase strictly with the induced prior β_L from $F(\frac{R}{2})$ at $\beta_L = 0$ to $F(\frac{L}{2})$ at $\beta_L = 1$. Also, for all $\beta_L \in (0, 1)$, $t(L, \pi) > t(R, \pi)$.

strictly decreasing in β_L. For the second part of the lemma, note that

$$t(L,\pi) - t(R,\pi) = (2q-1)\left(F(x_l) - F(x_r)\right)$$

By remark 2 again, for $\beta \in (0,1)$, $F(x_l) - F(x_r) < 0$, and since $q > \frac{1}{2}$, we have $t(L,\pi) < t(R,\pi)$. ■

Lemma 2 states that as the induced prior probability of the state being L (conditional on being pivotal) increases, the expected share of votes for $\mathcal{P}$ decreases under either state since the state L is deemed to be more "extreme". Informative voting by the responsive set ensures that for any induced prior, the policy receives more votes in the "moderate" state (R) unless the prior is degenerate. Note also that at any induced prior, the difference in expected vote shares is increasing in the informativeness of the signal. The expected vote shares in the two states are plotted against the induced prior in figure 3.

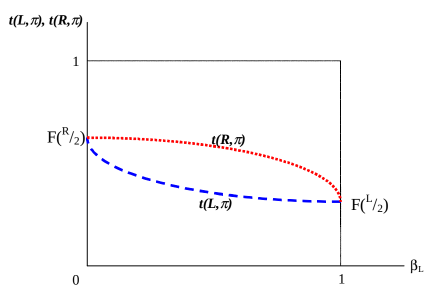

Fig 3: Vote shares in each state under preference monotonicity $(L < R < 0)$

Lemma 2 also ensures that since $t(S,\pi)$ lies strictly between 0 and 1, and $\beta(S|piv,\pi,\sigma)$ is always well-defined. Intuitively, since the types left of $\frac{L}{2}$ are $\mathcal{P}$-partisans and those to the right of $\frac{R}{2}$ are $\mathcal{Q}$-partisans, there is always a positive probability that any given type is pivotal. This finally proves Lemma 1.

The following proposition guarantees the existence of an equilibrium of the preference monotonicity voting game $(F(\cdot), q, L, R, n, \theta)$.

Proposition 1 *In the voting game with monotonic preferences, there exists a voting equilibrium π^* for every population size n and every voting rule $\theta \in (0,1)$, characterized by ordered cut-off strategies x_σ which is given by the solution of the equation $E(v(x_\sigma, s)|piv, \pi^*, \sigma) = 0$ for $\sigma = (l, r)$.*

Proof. For the proof of this proposition, we first note that the strategy for each voter can be denoted by two numbers x_l and x_r, both lying between $\frac{L}{2}$ and $\frac{R}{2}$. Thus the strategy space is a compact, convex and non-empty set $\left[\frac{L}{2}, \frac{R}{2}\right] \times \left[\frac{L}{2}, \frac{R}{2}\right]$. The rest of the proof follows from the proof of Proposition 1 in F-P. ■

To find the equilibrium of the model, I find a fixed point on the belief space $\beta_L \in [0,1]$. Suppose everyone else holds some belief β_L.This β_L determines two distributions $p_\sigma(\beta_L)$ according to (5) and correspondingly, the cut-off strategies $x_\sigma(\beta_L)$ according to (7). From the cut-off strategies, the expected shares of votes for the policies $t(S,\pi)$ in the two states $S = L, R$ is determined by (8). Given these shares, the number of players n and the voting rule θ, a player forms $\Pr(piv|\pi, S)$: probabilities of being pivotal in each state according to the pivot equations (3). These probabilities define belief $\widetilde{\beta_L}$ by (4) : in equilibrium, this $\widetilde{\beta_L}$ should be equal to the initial belief β_L. Thus the induced prior should have the rational expectations property in equilibrium. Note that

$$\frac{\Pr(piv|\pi, L)}{\Pr(piv|\pi, R)} = \frac{\beta(L|piv, \pi)}{\beta(R|piv, \pi)} = \frac{\beta_L}{1-\beta_L}$$

Thus, using the above and the pivot equations, the *equilibrium condition* can be simply stated as:

$$\frac{\beta_L}{1-\beta_L} = \frac{\Pr(piv|\pi, L)}{\Pr(piv|\pi, R)} = \left[\frac{(t(L,\pi^n))^\theta (1-t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^\theta (1-t(R,\pi^n))^{1-\theta}}\right]^n \tag{9}$$

3.2 Limiting Equilibria under Preference Monotonicity

Now, I consider the properties of the voting equilibria as the electorate grows in size arbitrarily, keeping all other parameters of the model constant. Therefore, every quantity is superscripted by the number of voters n. The superscript will be suppressed when there is no ambiguity Suppose, given L, R and θ for some n, the equilibrium is π^n, and the cutoffs are x_σ^n. As long as preference monotonicity assumption is satisfied, existence of equilibrium for any n implies the existence of a convergent subsequence with an accumulation point as $n \to \infty$. If a limit of this sequence exists, I call it π^0. By continuity arguments, as $x_\sigma^n \to x_\sigma^0$, $t(S,\pi^n)$, β_L^n, p_l^n,and p_r^n all converge to finite limits $t(S,\pi^0)$, β_L^0, p_l^0, and p_r^0 respectively along the sequence.

Rewriting the equilibrium condition:

$$\frac{\beta_L^n}{1-\beta_L^n} = \left[\frac{(t(L,\pi^n))^{\theta}(1-t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^{\theta}(1-t(R,\pi^n))^{1-\theta}}\right]^n \quad for\ all\ n \tag{10}$$

By Proposition 1, a solution to (10) exists for every n. From continuity, if a limit exists, we can also say that the above relation has to hold in the limit; call this the *limiting equilibrium condition.*

$$\frac{\beta_L^0}{1-\beta_L^0} = \lim_{n\to\infty}\left[\frac{(t(L,\pi^0))^{\theta}(1-t(L,\pi^0))^{1-\theta}}{(t(R,\pi^0))^{\theta}(1-t(R,\pi^0))^{1-\theta}}\right]^n \tag{11}$$

To avoid writing complicated expressions, I define:

$$\alpha_n = \frac{(t(L,\pi^n))^{\theta}(1-t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^{\theta}(1-t(R,\pi^n))^{1-\theta}} \text{ and } \alpha_0 = \frac{(t(L,\pi^0))^{\theta}(1-t(L,\pi^0))^{1-\theta}}{(t(R,\pi^0))^{\theta}(1-t(R,\pi^0))^{1-\theta}}$$

Note that the vote shares $t(S,\pi^n)$ are functions of β_L^n. Next, I look at the properties of the limit, assuming existence for the time being. I later show that in the preference monotonicity setting, for any voting rule, there is only one accumulation point of π^n which must be the limit.

Lemma 3 *If* $\beta_L^0 \in (0,1),\ \alpha_0 = \lim_{n\to\infty}\alpha_n = 1$

Proof. See Appendix. Note that this lemma does not use the preference monotonicity condition, so it is true of preference reversal too. ■

Lemma 4 *If* $\beta_L^0 = 1$, *then* $x_\sigma^n \to \frac{R}{2}$ *from the left for* $\sigma = l, r$. *Similarly, if* $\beta_L^0 = 0$, *then* $x_\sigma^n \to \frac{L}{2}$ *from the right for* $\sigma = l, r$

Proof. Follows from continuity of x_σ^n in p_σ^n and of p_σ^n is β_L^n, along with Remark 2. ■

As an aside to Lemma 4, note that although under both signals the cutoffs converge to $\frac{R}{2}$ or $\frac{L}{2}$ as the induced prior converges to 1 or 0 respectively, by remark 2, we always have $x_l^n < x_r^n$. Thus, in the responsive set, the voters always vote for $\mathcal{Q}$ if they get moderate signal r and $\mathcal{P}$ if they get the extreme signal l. The responsive interval is vanishingly small as the induced prior distribution converges to state R , grows for intermediate values of the prior, and again shrinks to a vanishing size as the distribution converges to a degenerate distribution at state L. Thus, given a level of precision q of the signals, the difference between expected shares in the two states is low for extreme values of the induced prior and high for intermediate values.

Lemma 3 and Lemma 4 together imply that for any limiting induced prior, given a voting rule θ, under any equilibrium, the vote shares in each state must be related in a certain way. This is stated in Proposition 2 below. According to Lemma 3, if α_n is bounded away from 1, then β_L^0 must be either 0 or1. Under conditions of Lemma 4, if β_L^n is indeed 0 (or 1), then the voters are almost sure of the state in which they are pivotal and vote as if under full information. Every type except those in a vanishing set votes uninformatively, and the vote shares under either state are the same in the limit. Thus, in equilibrium, we have $\alpha_0 = 1$ for all values of the induced prior.

Proposition 2 *In all limiting equilibria, we must have* $\alpha_0 = 1$, *i.e.*

$$\left(t(L,\pi^0)\right)^\theta \left(1-t(L,\pi^0)\right)^{1-\theta} = \left(t(R,\pi^0)\right)^\theta \left(1-t(R,\pi^0)\right)^{1-\theta}, i.e.\ \alpha_0 = 1$$

Proof. For any equilibrium with $\beta_L^0 \in (0,1)$, the proposition follows straightforwardly from Lemma 3. If $\beta_L^0 = 1$, the first part of Lemma 4 implies that

$$\begin{aligned} t(L,\pi^n) &= qF(x_l^n) + (1-q)F(x_r^n) \to qF(\frac{R}{2}) + (1-q)\,F(\frac{R}{2}) \to F(\frac{R}{2}) \\ t(R,\pi^n) &= qF(x_r^n) + (1-q)F(x_l^n) \to qF(\frac{R}{2}) + (1-q)\,F(\frac{R}{2}) \to F(\frac{R}{2}) \end{aligned}$$

$$\therefore \alpha_0 = \lim_{n\to\infty} \frac{(t(L,\pi^n))^\theta\,(1-t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^\theta\,(1-t(R,\pi^n))^{1-\theta}} = \frac{\left(F(\frac{R}{2})\right)^\theta \left(1-F(\frac{R}{2})\right)^{1-\theta}}{\left(F(\frac{R}{2})\right)^\theta \left(1-F(\frac{R}{2})\right)^{1-\theta}} = 1 \left(\because F(\frac{R}{2}) \in (0,1)\right)$$

If $\beta_L^0 = 0$, the proof follows in exactly the same way since the second part of Lemma 4 implies that then $t(S,\pi^n) \to F(\frac{L}{2})$ for $S \in \{L,R\}$. ■

Note that Proposition 2 is based on a necessary condition that must be true for a β_L^0 to which the induced belief converges in the limiting equilibrium. It helps exclude certain voting rules that cannot support a given value of β_L in the limit. To say this formally, define $\Theta(\beta_L)$ as the set of voting rules that can support β_L as an induced belief in the *limiting equilibrium condition* (11) for *some* distribution of preferences in the cut-off equilibrium. To emphasize that $t(S,\pi)$ is a function of β_L, we write $t(S,\pi)$ as $t_S(\beta_L)$ for $S \in \{L,R\}$.

Lemma 5 *Under preference monotonicity, (i) If* $\beta_L \in (0,1)$, *then* $\Theta(\beta_L)$ *is a strictly* *decreasing function $\theta^*(\beta_L)$, with $t_L(\beta_L) < \theta^*(\beta_L) < t_R(\beta_L)$. (ii) Otherwise,* $\Theta(1) = \{\theta : \theta < F(\frac{L}{2})\}$, *and* $\Theta(0) = \{\theta : \theta > F(\frac{R}{2})\}$

Proof. In Appendix. ■

The first part of the lemma is almost a corollary of Proposition 2. For each interior value β_L of the induced prior, it identifies a unique θ as the only possible voting rule to support β_L in the limiting equilibrium. As long as the expected vote shares in the two states are different, the only voting rule that can satisfy Proposition 2 is one that lies strictly between the two shares. This has the implication that under one state the status quo wins, while in the other, the policy wins. If there are any equilibria with beliefs that place positive probability on each state, then the responsive set of types for these equilibria are always influential. The second part of the lemma says that the extreme beliefs can be supported only by extreme values of the voting rules. The main inplication of the Lemma is that while the responsive set is influential for any possible equilibria with consequential rules, it is never infleuntial for trivial rules.

Note that since $\theta^*(\beta_L)$ is strictly decreasing, its inverse function $\beta_L^{-1}(\theta)$ exists for $\theta \in \left(F(\frac{L}{2}), F(\frac{R}{2})\right)$ and is strictly increasing. Thus, according to Lemma 5, for every θ, there is a unique β_L that can be supported as an equilibrium induced prior in the limit, *for any distribution of types.* Call it $\beta(\theta)$. We can write:

$$\beta(\theta) = \begin{cases} 1 \; if \; \theta < F(\frac{L}{2}) \\ \quad \beta_L^{-1}(\theta) \;\; if \; \theta \in \left(F(\frac{L}{2}), F(\frac{R}{2})\right) \\ 0 \; if \; \theta > F(\frac{R}{2}) \end{cases} \tag{12}$$

I plot the correspondence $\Theta(\beta_L)$ along with the expected vote shares in each state against the induced prior in Figure 4.

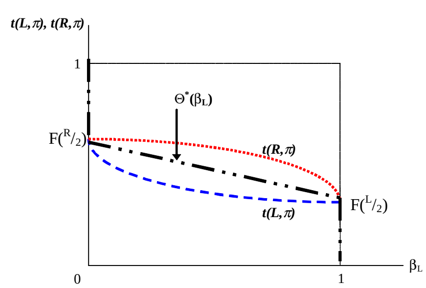

Fig 4: Correspondence $\Theta(\beta_L)$ under preference monotonicity

The next theorem gives a characterization of cut-off equilibria in large populations for different voting rules under preference monotonicity.

Theorem 1 *Assume* $L_, R$ *satisfy the preference monotonicity condition,* $F(\cdot)$ *satisfies full support, and* $q \in (\frac{1}{2}, 1)$. *Fix a voting rule* $\theta \in (0, 1)$. *Then there is a unique limiting equilibrium* π^0 *with ordered cut-off strategies and with the induced prior converging to* β_L *if and only if* $\theta \in \Theta(\beta_L)$, *or alternatively, if and only if* $\beta_L = \beta(\theta)$.

Proof. According Proposition 2, a voting equilibrium π^n with ordered cut-off strategies exists for a given θ, for any n. Since β_L lies in a compact set, there is an accumulation point π^a, given θ. I show in the appendix that this π^a is the limiting equilibrium π^0 given θ. Lemma 5 states that for any distribution of types, if a limit exists, there is a unique number $\beta(\theta)$ to which the induced prior converges in the limit along the sequence of equilibria under voting rule θ. ■

Note that once the limiting value of the induced prior β_L is established, the limiting posterior distributions p_σ, the limit cut-offs x_σ etc. are all determined from β_L. Thus this theorem describes all relevant information about strategies, vote shares and statewise outcomes in equilibria with a voting rule when the population size becomes large. Also, by the Law of Large numbers, the actual vote shares are arbitrarily close to the expected vote shares[15]. From here onwards, I do not distinguish between the expected and actual, and just call it "vote share".

3.3 Outcomes and Information Aggregation

In Section 2, I informally discussed a classification of voting rules according to the outcomes produced under full information. Here I formalise the discussion, and then examine the information aggregation properties of each class of voting rules.

For the purposes of this paper define a social choice rule H as a function that maps a state to an outcome, i.e.

$$H : \{L, R\} \to \{\mathcal{P}, \mathcal{Q}\}$$

When the function $H(\cdot)$ maps different states to different outcomes, i.e. $H(L) \neq H(R)$, I call it a *consequential choice rule.* When $H(\cdot)$ is a constant function, i.e.

[15] More specifically, given any $\epsilon > 0$ and $\delta > 0$, we can find some number N such that as long as the polupation size is larger than N, the actual vote share is within ϵ of the exoected share with a probability higher than $1 - \delta$.

when the planner wants the same outcome in both states, I call it a *trivial choice rule*. There are two trivial rules - one where the planner always wants the status quo to prevail ($H(L) = H(R) = \mathcal{Q}$), and the one that maps both states to the policy ($H(L) = H(R) = \mathcal{P}$). I call the first one $\mathcal{Q}$*-trivial* and the second one $\mathcal{P}$*-trivial choice rule*. If .

A voting rule is identified by the particular social choice rule it implements when the state is common knowledge. With full information, under state L, the policy would get $F(\frac{L}{2})$ share of votes; and similarly under state R, the policy would get $F(\frac{R}{2})$ share of votes. Therefore:

- Any voting rule $\theta < F(\frac{L}{2})$ is a $\mathcal{P}$-trivial rule, i.e. $\mathcal{P}$ wins under both states.
- Any voting rule $F(\frac{L}{2}) < \theta < F\left(\frac{R}{2}\right)$ is a consequential rule, i.e. $\mathcal{P}$ wins in state R and $\mathcal{Q}$ in state L[16].
- Any voting rule $\theta > F\left(\frac{R}{2}\right)$ is a $\mathcal{Q}$-trivial rule.

A voting rule is said to satisfy *full information equivalence*[17] if, for any $\epsilon > 0$, we can find a number N such that when the population size is larger than N, in either state the outcome of the voting game *under incomplete information* is the same as the outcome under full information with a probability larger than $1-\epsilon$. A full information equivalent voting rule implements the corresponding social choice rule.

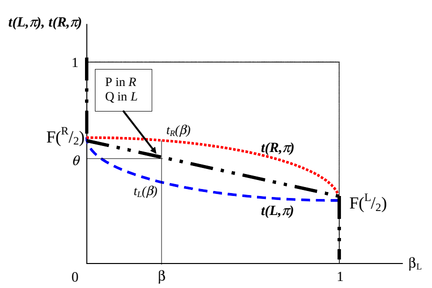

Figure 5(a): Outcome under a Consequential rule θ under Preference monotonicity *(L < R < 0)*

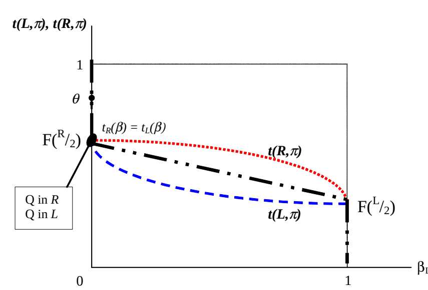

Figure 5(b): Outcome under a Q-trivial rule θ under preference monotonicity *(L < R < 0)*

Fig 5(a), 5(b)

[16] Note that the other consequential rule, i.e. $\{G(L) = \mathcal{P}, G(R) = \mathcal{Q}\}$ cannot be implemented under full information by the plurality rule with the common values case we are considering, i.e. $L < R < 0$.

[17] The concept of full information equivalence was formalised by F-P, and I adapt their definition to my setting.

Theorem 2 *Under preference monotonicity, any plurality rule $\theta \in (0,1)$ satisfies full information equivalence for any distribution of types.*

Proof. In appendix. ■

According to the theorem, under preference monotonicity, any voting rule aggregates information. Since the vote shares in each state is between $F(\frac{L}{2})$ and $F\left(\frac{R}{2}\right)$, the responsive set is never influential and any trivial rule aggregates information. With $\mathcal{P}$-trivial rules, being pivotal at state L (when $\mathcal{P}$ receives least votes) is infinitely more probable than being pivotal at state R, and everyone is virtually sure that conditional on being pivotal, the state is L. Similarly, with any $\mathcal{Q}$-trivial rule, conditioning on being pivotal, the state is almost surely R (when $\mathcal{P}$ receives most votes). I depict the outcome in the limiting equilibrium with a $\mathcal{Q}$-trivial rule in figure 5(a). On the other hand, for any consequential rule, the induced prior places positive probability on both states in the limit, and the responsive set is influential. Since the responsive types are aligned too, we have outcome $\mathcal{P}$ in state R and $\mathcal{Q}$ in state L almost surely, and hence we have information aggregation. The limiting equilibrium outcome with a consequential rule is depicted in figure 5(b).

4 Preference Reversal

Recall that preference reversal occurs if $L < 0 < R$. I now look at the strategies and equilibria in this situation and compare and contrast their properties with that of the benchmark model with monotonic preferences. Specifically, I show how voting can fail to aggregate information in the presence of groups with competing interests. Since the voters with $x \leq \frac{L}{2}$ support the alternative policy only in state L and those with $x \geq \frac{R}{2}$ do so only in state R, I will call these two groups of voters the L-group and the R-group respectively. These are groups of voters with state-dependent rankings such that the two groups have exactly opposite ranking over alternatives in each state. Note that within a group while rankings are the same, there is some heterogeneity in terms of strength of preferences for each alternative.

I shall simplify the model a bit and consider a slightly special case with $L = -b$ and $R = b > 0$. Note that this is not too strong an assumption as I consider all possible distributions of voter ideal points. However, we need to make an additional assumption on the informativeness of the signals.

Assumption I (Informativeness): $\Pr(l|L) = \Pr(r|R) = q > \frac{1}{2} + \frac{b}{4}$

The full support assumption is heneceforth referred to as Assumption F. A preference reversal setting is denoted by the collection $(F(\cdot), q, b)$. In this section, I shall use the same methodology I used in the previous section to examine the preference reversal situation.

4.1 Strategies and equilibria

A voter with signal σ, ($\sigma \in \{l, r\}$) evaluates the state using the distribution $\beta(S|piv, \pi, \sigma)$ and votes for $\mathcal{P}$ if and only if the expected value is non-negative. So, the condition for voting for the policy after having received σ is:

$$Ev(x, \sigma) \geq 0 \Rightarrow 2x(1 - 2p_\sigma) \geq b$$

Hence, the voter votes for $\mathcal{P}$ iff

$$1 \geq |x| \ \geq \frac{b}{2(1 - 2p_\sigma)} \tag{13}$$

Using (13), we can determine the cut-offs:

$$x_\sigma = \begin{cases} \min(1, \frac{b}{2(1-2p_\sigma)}), & 0 \leq p_\sigma < \frac{1}{2} \\ \max(-1, \frac{b}{2(1-2p_\sigma)}), & \frac{1}{2} \leq p_\sigma \leq 1 \end{cases} \tag{14}$$

Now, according to the above definitions of the cut-off, we get:

$$\pi(x, \sigma) = \begin{cases} \left.\begin{array}{c} 1 \ for \ x \leq x_\sigma \\ 0 \ for \ x > x_\sigma \end{array}\right\} if \ \frac{1}{2} \leq p_\sigma \leq 1 \\ \left.\begin{array}{c} 1 \ for \ x \geq x_\sigma \\ 0 \ for \ x < x_\sigma \end{array}\right\} if \ \ 0 \leq p_\sigma < \frac{1}{2} \end{cases} \tag{15}$$

Or alternatively, combining (14) and (15), we define the strategies in terms of p_σ as follows:

$$\pi(x,\sigma)=\left\{\begin{array}{l}\left.\begin{array}{l}1\ for\ x\leq\frac{b}{2(1-2p_\sigma)}\\0\ for\ x>\frac{b}{2(1-2p_\sigma)}\end{array}\right\}\ if\ p_\sigma\geq\frac{1}{2}+\frac{b}{4}\\0\ for\ all\ x\quad if\ p_\sigma\in\left(\frac{1}{2}-\frac{b}{4},\frac{1}{2}+\frac{b}{4}\right)\\\left.\begin{array}{l}1\ for\ x\geq\frac{b}{2(1-2p_\sigma)}\\0\ for\ x<\frac{b}{2(1-2p_\sigma)}\end{array}\right\}\ if\ p_\sigma\leq\frac{1}{2}-\frac{b}{4}\end{array}\right\}$$

Any equilibria must have strategies of the above form. Note that $p_\sigma\in[0,1]\Rightarrow -1\leq 1-2p_\sigma\leq 1$ and so $x_\sigma\in[-1,-\frac{b}{2}]\cup[\frac{b}{2},1]$. Also, for all values of p_σ, $\pi(x,\sigma)=0$ in the range $(-\frac{b}{2},\frac{b}{2})$. Thus a voter with his bliss point in this range always votes for the status quo irrespective of the signal.

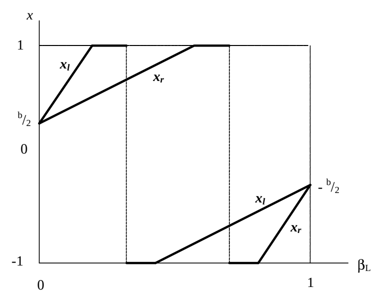

Fig 6: Cut-offs under preference reversal as functions of induced prior

Thus, although all equilibria must have cut-off strategies, the cut-offs are not ordered. The cutoffs as functions of the induced prior are plotted in Figure 6. When a cut-off is in $[-1,-\frac{b}{2}]$ (the L-group), the types to the left of the cut-off vote for $\mathcal{P}$, and when the cut-off lies in $[\frac{b}{2},1]$ (the R-group), types to the right of the cut-off vote for $\mathcal{P}$. This has several implications. First, the responsive types lying in these two groups would vote in opposite ways based on the same information since one of the groups is aligned with the society and the other is not. Second, in each state, the vote share is a non-monotonic function of the induced belief. Note that the monotonicity in vote shares was crucial for information aggregation with consequential rules in the preference monotonicity case. Third, with unordered cut-offs, the existence of a well-defined induced prior is no longer trivial, and we need the informativeness assumption I on signals to guarantee that. Lastly, with a loss of the ordering property, uniqueness of the responsive set is no longer assured. This can give rise to a certain kind of equilibria that is not seen in the preference monotonicity case, as we shall see in

Proposition 4.

Recall that the probability of an individual voting for the alternative $\mathcal{P}$ given σ is $z_\sigma, i.e.\ z_\sigma \equiv \int_{-1}^{1} \pi(x,\sigma)dF$. In any equilibrium, we have:

$$z_\sigma = \begin{cases} F(x_\sigma)\ if\ x_\sigma \leq -\frac{b}{2} \\ 1 - F(x_\sigma)\ if\ x_\sigma \geq \frac{b}{2} \\ 0\ otherwise \end{cases} \tag{16}$$

Although the definition of z_σ is different in the preference reversal case, the vote shares in the two states in terms of z_σ are still given by equation (8):

$$\begin{aligned} t(L,\pi) &= qz_l + (1-q)z_r \\ t(R,\pi) &= (1-q)z_l + qz_r \end{aligned}$$

Lemma 6 *In any equilibrium in the preference reversal setting, the expected share of votes in any state lies strictly between* 0 *and* 1, *i.e.* $t(S,\pi) \in (0,1)$ *for* $S \in \{L, R\}$.

Proof. See Appendix. ■

Lemma 6 guarantees that the induced prior is indeed always well-defined. The expected share of people voting is less than unity because there is always a set of types close enough to 0 (between $-\frac{b}{2}$ and $\frac{b}{2}$) who vote for the $\mathcal{Q}$. On the other hand, the signal being informative enough (Assumption I) guarantees that the cut-offs are sufficiently distant for moderate values of the induced prior. This is needed to ensure that if for one signal, no type votes for $\mathcal{P}$, there is an interior cut-off for the other signal. To see that from Figure 2, note that the range of β_L for which p_l lies between $\frac{1}{2} - \frac{b}{4}$ and $\frac{1}{2} + \frac{b}{4}$ lies entirely to the left of $\frac{1}{2}$, while the range of β_L for which p_r lies between $\frac{1}{2} - \frac{b}{4}$ and $\frac{1}{2} + \frac{b}{4}$ lies entirely to the right of $\frac{1}{2}$. This guarantees that, for any induced prior, at least one signal always leads to an interior cut-off - leading to positive expected share for $\mathcal{P}$.

Next, the existence of an equilibrium for the preference reversal game $(F(\cdot), q, b, n, \theta)$.is proved. This is the analogous result to Proposition 1. Although the strategy set is non-convex and we cannot use a fixed point theorem to prove existence the way we did in the preference monotonicity setting, we can still show the existence of a solution to equation (10), which is the *equilibrium condition.*

Remark 3 *In the preference reversal game, there exists a voting equilibrium* π^* *for* *every population size n and every voting rule $\theta \in (0,1)$. The equilibrium is characterized by unordered cut-off strategies x_σ given by the solution of $E(v(x_\sigma, s)|piv, \pi^*, \sigma) = 0$ for $\sigma = (l, r)$.*

Proof. From Lemma 5, we know that $t(S, \pi)$, is bounded by positive numbers both above and below. This implies that for any n, the right hand side of equation (10) is bounded above and below. However, as β_L goes from 0 to 1, the left hand side continuously increases from 0 to ∞. This guarantees the existence of a solution β_L^n to the equation, and hence existence. ■

We can immediately identify one particular equilibrium for the case with a distribution of types with density $f(\cdot)$ that is symmetric about 0.

Proposition 3 *For any $F(\cdot)$ for which the density $f(\cdot)$ is symmetric about 0, there is an equilibrium with $x_l^* = -\frac{b}{2(2q-1)}$ and $x_r^* = -x_l^*$. This is an equilibrium for all values of $\theta \in (0,1)$ and is independent of the number of voters n.*

Proof. Consider the situation where everyone else plays $x_\sigma = x_\sigma^*$, and $\sigma \in \{l, r\}$. Note that $x_l^* < -\frac{b}{2}$ and $x_r^* > \frac{b}{2}$. So, $z_l^* = F(x_l^*)$ and $z_r^* = 1 - F(x_r^*) = 1 - F(-x_l^*) = F(x_l^*) = z_l^*$, by symmetry of $f(\cdot)$. Therefore, $t(L, \pi) = t(R, \pi) = F(x_l^*)$ for each n, which implies that $\beta_L = \frac{1}{2}$ for every θ and n. Thus, the signals are fully informative, and we have $p_l = q$ and $p_r = 1 - q$. These, coupled with the Assumption I, imply that the best response to x_σ^* is indeed x_σ^*, which establishes the claim. ■

The proposition says that if the commonly held induced priors are uninformative, then sufficiently extreme types vote for the alternative $\mathcal{P}$ if and only if they get favourable signals, and everyone else votes uninformatively, disregarding their signal. There are a few things to be noted about the above equilibrium. First, this is the only "stable" equilibrium sequence in the sense that the strategies do not change with the number of players. Second, in this equilibrium, the expected vote share does not change with the state or the voting rule. If the required plurality for the policy to pass is higher than $F(x_\sigma^*)$, then the status quo always passes, and if the required share is lower than $F(x_\sigma^*)$, then the status quo always loses. If $\theta = F(x_\sigma^*)$, then we get either alternative (policy or status quo) with equal probability. As we shall see later in Section 4.3, this constitutes a failure of information aggregation. Note here that we do not even require the full force of symmetry of $f(\cdot)$ here. As long as we have $F\left(-\frac{b}{2(2q-1)}\right) = 1 - F\left(\frac{b}{2(2q-1)}\right)$, we shall have this equilibrium. I later establish that even if the distribution of ideal points is not symmetric, there is always an equilibrium at some belief β_L^* (not necessarily equal to $\frac{1}{2}$) that has the same vote share for each state and is independent of the voting rule.

Next, let us examine the vote share as a function of the induced prior in the preference reversal set-up.

Lemma 7 *Under preference reversal, the expected share of votes $t(S,\pi)$ in state $S \in \{L,R\}$ is a U-shaped function of the induced prior β_L. There exists some number β_L^* satisfying $0 < \beta_L^* < 1$ such that $\beta_L < \beta_L^*$, $t(R,\pi) > t(L,\pi)$, for $\beta_L > \beta_L^*$, $t(R,\pi) > t(L,\pi)$ and for $\beta_L = \beta_L^*$, $t(R,\pi) = t(L,\pi)$.*

Proof. See Appendix. ■

This lemma says that there is a critical value β_L^* of the induced belief of the state being L below (above) which the expected vote share in favour of the policy alternative in state L is higher (lower) than that in state R. Also, given a state, the expected share of the votes in favour of the policy alternative increases with more extreme beliefs. As the voters get more unsure about the state, only the very extreme types vote for the policy. Note that at β_L^*, we have $F(x_l) + F(x_r) = 1$, and under a symmetric distribution of types, $\beta_L^* = \frac{1}{2}$.. The expected share of votes under the two states in the preference reversal situation (according to Lemma 7) as functions of the induced prior are shown in figure 7. To illustrate how the shares are constructed according to (8), we also show the functions z_l and z_r (i.e. the probability of voting for $\mathcal{P}$ on getting the signal l and r respectively) in the figure.

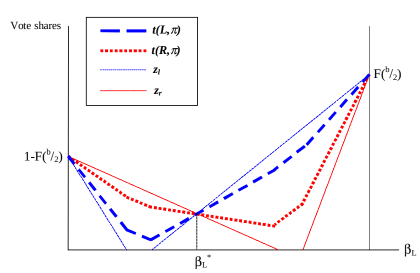

Fig 7: Construction of vote shares as functions of induced prior

4.2 Limiting equilibria in large elections

Given Proposition 3, equilibrium exists for every n. Therefore, I use the same notation as in Section 3.2. Since the cutoffs are bounded within a compact set, any sequence of x_σ^n will have a convergent subsequence. We look at such convergent subsequences x_σ^n as $n \to \infty$. We call an accumulation point of such a sequence of cutoffs as x_σ^0, and the resulting equilibrium as π^0. By the continuity arguments , as $x_\sigma^n \to x_\sigma^0$, $t(S,\pi^n)$, β_L^n, p_l^n,and p_r^n all converge to $t(S,\pi^0)$, β_L^0, p_l^0, and p_r^0 respectively along the subsequence. In this section I examine which outcomes can be supported in the limit.

The necessary conditions for the limit, the *limiting equilibrium condition* as identified in equation (11) remains exactly the same. Lemma 3 goes through without any change. Lemma 4 goes through too, with the slight modification that it is no longer true of all n, but it holds for large enough n. I state this in Lemma 8. For a suffieicntly large electorate, if the induced prior converges to 0 (1), both cut-offs are in the L-group (R-group).

Lemma 8 *If $\beta_L^0 = 1$, (i) $\exists$ some m such that $x_l^n > x_r^n$ for all $n > m$; and (ii) $x_\sigma^n \to -\frac{b}{2}$ from the left for $\sigma = l, r$. Similarly, if $\beta_L^0 = 0$, (i) $\exists$ some m_1 such that $x_l^n > x_r^n$ for all $n > m_1$; and (ii) $x_\sigma^n \to \frac{b}{2}$ from the right for $\sigma = l, r$*

Proof. See Appendix. ■

Proposition 2 now goes through in exactly the same form. The proof follows from Lemma 3 and Lemma 8 analogously. In other words, the local properties of the limiting equilibria are the same in the case of preference monotonicity and that of preference reversal. Next, I examine which voting rules can be supported by a given value of the induced prior in the limit, for which an equivalent of Lemma 5 is necessary.

Lemma 9 *Under preference reversal, (i)for $\beta_L \in (0, \beta_L^*) \cup (\beta_L^*, 1)$, $\Theta(\beta_L)$ is a continous function $\theta^*(\beta_L)$, with $t_L(\beta_L) < \theta^*(\beta_L) < t_R(\beta_L)$ for $\beta_L < \beta_L^*$, and $t_L(\beta_L) > \theta^*(\beta_L) > t_R(\beta_L)$ for $\beta_L > \beta_L^*$, (ii)Otherwise, $\Theta(1) = \{\theta : \theta > F(-\frac{b}{2})\}$, $\Theta(0) = \{\theta : \theta > 1 - F(\frac{b}{2})\}$ and $\Theta(\beta_L^*) = \{\theta : \theta \in (0,1)\}$.*

In Appendix.

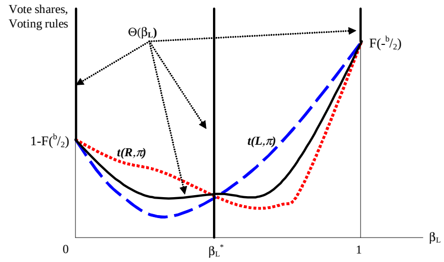

Fig 8: Correspondence $\Theta(\beta_L)$ under preference reversal

The correspondence $\Theta(\beta_L)$ for the preference reversal case, as inferred in Lemma 9, is depicted in figure 8. For any value of induced belief β_L for which the vote share in the two states are different, there is a unique voting rule $\theta^*(\beta_L)$ that can support such a belief in the limiting equilibrium. The extreme (degenerate) priors can be supported by "large enough" voting rules. For the belief β_L^* where the vote shares are equal in the two states, we cannot rule out any voting rule. Proposition 4 is an example where any voting rule leads to a limiting equilibrium at $\beta_L^* = \frac{1}{2}$. Note that in this case, if we invert the correspondence to get the supporting induced belief β_L for each voting rule θ, we no longer get a function $\beta(\theta)$ as defined in (12) in the preference monotonicity case, but rather a correspondence.

The above lemma only says that it is possible that some voting rule $\theta \in \Theta(\beta_L)$ *may* support an equilibirum with belief convering to β_L for *some* distribution of ideal points. The next theorem states that given an induced prior β_L, any voting rule in $\Theta(\beta_L)$ indeed supports a limiting equilibrium with beliefs converging to β_L for *any* distribution of ideal points.

Theorem 3 *For any $b > 0$ satisfying the preference reversal condition, and for any q satisfying Assumption I and any distribution of preferences $F(\cdot)$ satisfying assumption F, given a voting rule θ, there is a limiting equilibrium π^0 with cut-off strategies and with the induced prior converging to β_L if $\theta \in \Theta(\beta_L)$*[18].

Proof. In appendix. ■

Denote $t(L, \beta_L^*) = t(R, \beta_L^*))$ by z. Because of the non-monotonic vote share functions, for any voting rule $\theta > z$, there can be three different limiting equilibria: one

[18]This theorem requires an assumption that $\theta^*(\beta_L)$ is not constant over any range. We ignore that as a non-generic case.

with equilibrium limiting belief less than, one more than and one exactly equal to β_L^*.The first equilibrium has responsive set of types entirely (or mostly) in the R-group and the second one has the responsive set mostly or entirely in the L-group. The third equilibrium is similar to the one identified in Proposition 4 with the responsive set equally divided in both groups.

4.3 Voting rules and Information Aggregation

From Lemma 9, we can deduce possible outcomes for each value of the induced prior. All these outcomes occur almost surely, in the same way as in the preference monotonicity case.

- For $\beta_L = 0$, the only possible outcome is $\mathcal{Q}$ under both states. Here, the responsive set is in the R-group but is not influential.
- For $\beta_L \in (\beta_L^*, 1)$, the only possible outcome is $\mathcal{Q}$ under state L and $\mathcal{P}$ under state R. Here, the responsive set is in the R-group and is influential.
- For $\beta_L = \beta_L^*$, the vote share in each state is fixed at z and the outcome depends on whether the voting rule is greater or less than z.
- For $\beta_L \in (0, \beta_L^*)$, the only possible outcome is $\mathcal{P}$ under state L and $\mathcal{Q}$ under state R. Here, the responsive set is in the L-group and is influential.
- For $\beta_L = 0$, the only possible outcome is $\mathcal{Q}$ under both states. Here, the responsive set is in the L-group but is not influential.

From here onwards, I assume with a slight loss of generality that $F(-\frac{b}{2}) > 1 - F(\frac{b}{2})$[19]. In other words, I assume that the L-group is the larger interest group, and hence the group that is aligned with the society. Therefore,

- Any voting rule $\theta < 1 - F\left(\frac{b}{2}\right)$ is $\mathcal{P}$-trivial
- Any voting rule $1 - F\left(-\frac{b}{2}\right) \leq \theta < F\left(\frac{b}{2}\right)$ is a consequential rule[20], i.e. the policy wins in state L and the status quo in state R.

[19] If $F(-\frac{b}{2}) = 1 - F(\frac{b}{2})$,then there are no consequential rules. $\theta = F(-\frac{b}{2})$ would implement a random social choice rule under full information if the L-group is the larger interest group.

[20] Note that the other consequential rule, i.e. $\{G(L) = Q, G(R) = P\}$ cannot be implemented under full information by the plurality rule

- Any voting rule $\theta \geq F\left(-\frac{b}{2}\right)$ is a $\mathcal{Q}$-trivial rule.

For all $\mathcal{Q}$-trivial rules, the beliefs that can be supported in equilibrium are $\beta = \{0, \beta_L^*, 1\}$. Since the maximum share of received by the alternative $\mathcal{P}$ in any state is $F\left(-\frac{b}{2}\right)$, $\mathcal{Q}$-trivial rules always aggregate information. Figure 9(a) depicts the limiting equilibria for a $\mathcal{Q}$-trivial rule.

For informaion to be aggregated under consequential rules, we need the responsive set to be influential and in the L-group. For these rules however, there is always one equilibrium with $\beta_L = 0$ where the responsive set in the R-group and is not influential. Hence we get $\mathcal{Q}$ in both states. In another equilibrium for these rules, $\beta_L = \beta_L^*$, and here too, we get $\mathcal{Q}$ in both states with a very high probability. However, there is a third equilibrium with induced prior converging to some belief in $(0, \beta_L^*)$ with the responsive set entirely in the L-group and influential. This equilibrium aggregates information. Figure 9(b) depicts all the possible limiting equilibria for a consequential rule.

For $\mathcal{P}$-trivial rules greater than z we have two equilibria with opposite outcomes in the different states: one with equilibrium induced prior in the set $(0, \beta_L^*)$ and the other in the set $(\beta_L^*, 1)$. The responsive sets are influential here when information aggregation requires that they not be so. So, for these voting rules we have no information-aggregating equilibrium. The third equilibrium has beliefs converging to β_L^*. Since at this belief, the vote share in both states is z, in this equilibrium we always get the status quo. Figure 9(c) shows the possible equilibria for one such rule. However, information is aggregated almost surely by the very low $\mathcal{P}$-trivial rules[21].

I summarise the inferences about information aggregation for different voting rules in a preference reversal setting in the next theorem. I use the same definition of full information equivalence as in Section 3.3. I define an equilibrium as *non-information aggregating* when in at least one state, voting under incomplete information delivers an outcome different from the full-information outcome with a probability arbitrarily close to 1.

Theorem 4 *All (limiting) voting equilibria with $\mathcal{Q}$-trivial voting rules satisfy the full information equivalence property. For consequential rules, there is one equilibrium*

[21] More speficically, the $\mathcal{P}$-trivial voting rules that aggregate information for sure for any distribution of preferences are those that are below the minimum share of votes received by $\mathcal{P}$ for any belief, i.e. those rules that satisfy $\theta < \min\{\min_{\beta_L} t(L, \pi), \min_{\beta_L} t(L, \pi)\}$. Equilibrium induced prior is β_L^* and equilibrium shares in both states are $z > \theta$ in the limit.

that satisfies full information equivalence and two that are non-information aggregating. For $\mathcal{P}$-trivial rules that are sufficiently large, all equilibria are non-information aggregating. All $\mathcal{P}$-trivial rules below some threshold aggregate onformation.

The above theorem establishes the bias in favour of the status quo. Unless the required vote share for the policy to win is very low, competition between two groups along with risk aversion ensures that the status quo wins in at least one state. Note that the only voting rules for which information is aggregated in any equilibrium are all $\mathcal{Q}$-trivial rules and the very low $\mathcal{P}$-trivial rules.

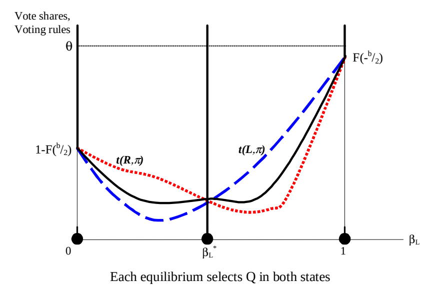

Fig 9(a): Equilibria under a Q-trivial rule

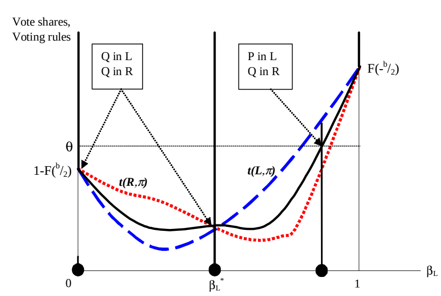

Fig 9(b): Equilibria under a consequential rule

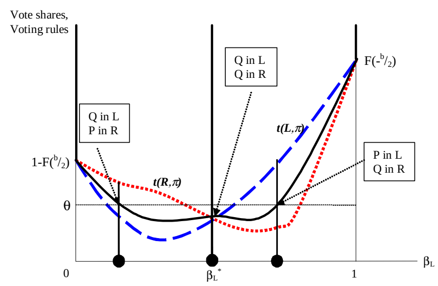

Fig 9(c): Equilibria under a large P-trivial rule

5 Multidimensional extension

In the previous two sections, we have established that at least in the case of convex voter preferences over a unidimensional policy space, the question of information aggregation boils down to the empirical question of whether preference monotonicity or reversal prevails. In the unidimensional model, unless the uncertainty is somewhat extreme, we do not encounter preference reversal. For example, in most elections, it is known whether the challenger is to the left or right of the incumbent. However, in a multidimensional policy space, the monotonic preferences assumption is much harder to justify. The framework developed in the previous section can readily handle the extension to a multidimensional policy space, and the main conclusions carry over. Moreover, convexity of preferences does not seem to be necessary for the results. In this paper, I only provide the intuition for this.

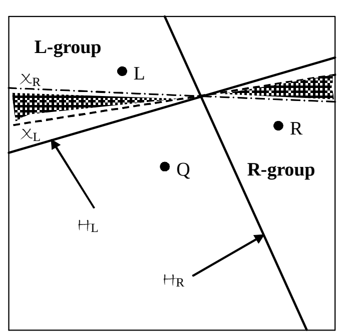

Fig 10: Cut-offs and responsive set in the multidimensional policy space

Think of the policy space as a many-dimensional cube, with each dimension being $[-1, 1]$. Suppose that the status quo $\mathcal{Q}$ is located at the origin, and the policy alternative $\mathcal{P}$ is located at two points L and R under states L and R respectively. Given a state S, a hyperplane $\mathcal{H}_S$ separates the cube into two parts, one composed of types that support $\mathcal{Q}$ (containing the origin) and the other containing types that support $\mathcal{P}$ under full information. Just as described in Section 2, we can define as $\mathbf{P}(S)$ the set of types that prefer $\mathcal{P}$ in state S. The preference monotonicity condition is exactly the same - that $\mathbf{P}(L)$ be included in $\mathbf{P}(R)$ or vice versa. Note that this is harder to satisfy in the multidimensional setting. In particular, for a given location L, as the size of the cube increases, the set of locations R for which $\mathbf{P}(S)$ exhibits preference monotonicity keeps shrinking and approches a ray connecting L with $\mathcal{Q}$ at the origin.

If the hyperplanes $\mathcal{H}_L$ and $\mathcal{H}_R$ are parallel, we are either in a monotonic preference situation or in a situation where there are two disjoint, completely opposed interest groups, much like the unidimensional preference reversal situation. Otherwise, for a large enough policy cube, we have four sets: two of opposed independent types, one type committed to $\mathcal{P}$ under both states and one type committed to $\mathcal{Q}$ under both states. Denote the set of indpendents preferring $\mathcal{P}$ under L and $\mathcal{Q}$ under R by the L-group and the set of independents preferring $\mathcal{P}$ under R and $\mathcal{Q}$ under L as the R-group.

Suppose the hyperplanes $\mathcal{H}_L$ and $\mathcal{H}_R$ meet at a straight line $\mathcal{L}$[22]. Under uncertainty, given a signal σ, the "cut-offs" that separate those who vote for $\mathcal{P}$ from those who vote $\mathcal{Q}$ are hyperplanes $\mathcal{X}_\sigma$. As the induced prior changes from 0 to 1, $\mathcal{X}_\sigma$ rotates about $\mathcal{L}$[23], starting at $\mathcal{H}_R$, and ending at $\mathcal{H}_L$. The strategy of a voter can also be described by the angle that each of the cut-off hyperplanes makes with the line $\mathcal{L}$. This is a compact set, and therefore, an equilibrium exists. If the hyperplanes $\mathcal{H}_L$ and $\mathcal{H}_R$ are parallel, then the cutoffs $\mathcal{X}_\sigma$ do not rotate, but translate from $\mathcal{H}_R$ to $\mathcal{H}_L$.Thus we can trace vote shares t_L and t_R in the two states as a function of the induced prior β_L. Once we have done that, the rest of the analysis is exactly as in the unidimensional model. Note here that the responsive set is always non-convex: it has two subsets, one of which lies in the L-group and the other in the R-group. Figure 10 demonstrates the cut-offs in the multidimensional policy space, with the responsive

[22] If the policy space is two-dimensional, the hyperplanes $\mathcal{H}_L$ and $\mathcal{H}_R$ will be straight lines and $\mathcal{L}$ will be a point.

[23] Using a simple geometric argument, it is easy to show that every type on $\mathcal{L}$ should be indifferent between $\mathcal{P}$ and $\mathcal{Q}$ for any beliefs. Hence the cut-offs should always contain $\mathcal{L}$.

set being the shaded area.

In a monotonic preferences setting, the vote shares in both states are monotonic functions of the induced prior. Thus all results from Section 3 extend to the multi-dimensional case. With preference reversal case though, we do not necessarily have U-shaped share functions. The equilibria depend on the particular shape of the distribution of preferences. This makes generalised equilibrium characteristics and aggregation (or non-aggregation) results difficult to get in a multidimensional set-up. However, given a distribution of preferences we can use the limiting equilibirium conditions developed in this paper to identify all the possible voting equilibria for that particular case and make judgements about information aggregation properties of each voting rule.

6 Discussion

The chief idea of the paper is that the source of informational inefficiency in elections is the existence of groups of voters who always have opposed interests, and such conflict, as the multidimensional model shows, is inherent in the electorate. To demonstrate the results, I develop a methodology of determining all the limiting equilibria in a spatial model of electoral competition for any voting rule short of unanimity. While this method applies to any finite dimensional policy space, for the unidimensional policy space, we can identify all possible equilibria and the aggregation properties of each voting rule except unanimity.

The fact that voting under incomplete information may produce outcomes inferior to those under complete information is hardly a surprise. Several papers point out sources of partial aggregation failure using completely preference homogeneity or only limited heterogeneity in the form of common values: use of unanimity rules (Feddersen and Pesendorfer 1998), voters signaling their preferences through their votes (Razin 2003), information being costly (Persico 2004, Martinelli 2006), abstention (Oliveros 2005) and so on[24]. Since the agenda of the current paper is to pinpoint that the fundamental source of aggregation failure is in competing interests among groups which is endemic to any democracy, I do not allow for these other possible causes of

[24] Another paper that looks at a similar common values context where the members of the jury vary over what counts as "reasonable doubt" for acquittal is Li, Rosen and Suen (2001). While they examine aggregation failure in small committes, F-P shows that such conflict does not affect aggregation if the jury size is large.

partial failure and consider an environment with a more general correlation between the state variable and preferences. There are a few papers (Kim 2006, Kim and Fey 2006, Meirowitz 2006, 2005a) that consider groups with opposed rankings in each state, but in these papers, the voters within the group have exactly same preferences. Kim (2006), which is probably the paper that is closest to the current one, finds that information is fully aggregated for most voting rules as long as voters care enough about mistakes in each state. In the current paper, voters with the same *ranking* over the alternatives under full information need not have the same *intensity* of preference for them, leading to different behaviour under uncertainty. Allowing any intra-group heterogeneity uncovers the deeper problem with inter-group conflict in preferences. The claim made here is that real elections may involve voter preferences which look neither like adversarial committees nor like jury boards. In elections, voters may feel differently about different issues and may in fact care about myriad issues well beyond the left-right ideological space.

It must be noted here that that I have two related, but distinct sets of results: one, the necessity and sufficiency of preference monotonicity for information aggregation; and two, the characterization of all equilibria in the unidimensional model for any voting rule.

In addition, I demonstrate that monotonicity is hard to obtain in a multidimensional policy space. While the assumption of convexity of preferences is necessary for the characterization of equilibria under preference reversal, but is not necessary for the first set of results. With the help of this assumption we can point out, in a substantive way, how exactly the electoral system may fail to produce the outcomes desired by the majority. Since there are different equilibria, there are different reasons why elections can fail to aggregate information.

The multiplicity issue makes the role of beliefs in a political system crucial. The model endogenises the process of formation of beliefs about which types are going to be responsive to information in equilibrium. Aggregation failure for consequential rules can simply be thought of a co-ordination failure because of "wrong" beliefs. For example, while a consequential rule needs the responsive set to be in the larger interest group, voters can believe that almost everyone is voting uninformatively. Independent of information received, the larger interest group votes for the status quo and almost everyone in the smaller interest group votes for the alternative. Voter behaviour in this equilibrium is akin to what we know as block voting. In another "bad" equilibrium, only the extremists at either end of the ideological spectrum are responsive – but aggregation fails because most of the voters vote for the status quo in either state.

In each of these "bad" equilibria, whatever be the mode of failure of aggregation, the failure is of an extreme nature in the sense that the "wrong" outcomes occur with a very high probability in a large electorate. It is worth noting that these results do not depend on the relative size of the conflicting groups or on the extent of noise in the signals. Therefore, any improvement in the accuracy of information that individuals have will fail to produce superior outcomes in the limit.

Aside from the question of information aggregation, this paper addresses an important question of implementation of social choice rules. Can we find a voting rule which delivers two pre-specified outcomes in two states with a very high probability in all equilibria? For example, we might look for a voting rule that delivers the majority preferred outcome in both states. I show that under preference reversal, such a rule does not exist unless the pre-specified outcome is the same in both states.

One interpretation of Condorcet Jury Theorem is that communication among voters is not necessary in large elections for the information problem to be solved. This paper indicates that we are faced with the possibility of multiple equilibria, some or all of which produce informationally inferior outcomes. Thus, voting cannot perform the role of communication among voters. Can democratic deliberation[25] improve election outcomes? Note that since all members within each conflicting interest group have the same state-contingent rankings[26], members each group have an incentive to share information among themselves. However, this needs the voter preferences to be public information. We can think of each set of independent voters with similar rankings as belonging to a political party or a special interest group, and thus this paper highlights the role of political institutions like parties or interest groups as information aggregators in an electorate.

[25] See Coughlan(2000), Austen-Smith and Feddersen (2005, 2006), Gerardi and Yariv (2007), Meirowitz (2005a, 2005b, 2006) for models of deliberation before voting. Goeree and Yariv (2006) demonstrates in an experimental setting that communication can improve outcomes.

[26] In this case, the condition for full revelation of information between any two members of the same group is satisfied according to Baliga and Morris (2002)

7 Bibliography

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. AUSTEN-SMITH, D., AND FEDDERSEN T. (2005): "Deliberation and Voting Rules", In Austen-Smith D. and J. Duggan (eds.) Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks, Berlin: Springer

3. AUSTEN-SMITH, D., AND FEDDERSEN T. (2006): "Deliberation, Preference Uncertainty and Voting Rules", *American Political Science Review*, 100, 209-218

4. BALIGA, S. AND S. MORRIS (2002): "Co-ordination, Spillovers, and Cheap Talk", *Journal of Economic Theory*, 105, 450-468

5. BATTAGLINI, M., R. MORTON AND T. PALFREY (2006): "The Swing Voters' Curse in the Laboratory", Mimeo, Princeton University

6. BERG, S. (1993): "Condorcet's Jury Theorem, Dependency among Jurors", *Social Choice and Welfare*, 10, 87-96

7. 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.

8. COUGHLAN, P.J. (2000): "In Defense of Unanimous Jury Verdicts: Mistrials, Communication and Strategic Voting", *American Political Science Review*, 94(June), 375-93

9. FEDDERSEN, T., AND W. PESENDORFER (1996): "The Swing Voter's Curse", *American Economic Review*, 86(3), 408-424

10. FEDDERSEN, T., AND W. PESENDORFER (1997): "Voting Behavior and Information Aggregation in Large Elections with Private Information", *Econometrica*, 65(5), 1029-1058

11. 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

12. FEDDERSEN, T., AND W. PESENDORFER (1999):"Abstention in Elections with Asymmetric Information and Diverse Preferences", *The American Political Science Review*, 93(2), 381-98

13. GERARDI, D. AND L. YARIV (2007): "Deliberative Voting", *Journal of Economic Theory*, 134(1), 317-338

14. GOEREE, J. AND L. YARIV (2006): "An Experimental Study of Jury Decision Rules", Mimeo, California Institute of Technology

15. GUL, F. AND W. PESENDORFER (2006): "Partisan Politics and Aggregation Failure with Ignorant Voters", Mimeo, Princeton University

16. KIM, J (2005): "A Model of Adversarial Committees", Mimeo, W. Allen Wallis Institue of Political Economy, University of Rochester

17. KIM, J., AND M. FEY (2006): "The Swing Voter's Curse with Adversarial Preferences", *Journal of Economic Theory,* Forthcoming

18. LADHA, K., (1992): "The Condorcet Jury Theorem, Free Speech, and Correlated Votes, *American Journal of Political Science*, 36, 617-34

19. LADHA, K., (1993): "Condorcet's Jury Theorem in the Light of de Finetti's Theorem: Majority Rule with Correlated Votes, *Social Choice and Welfare*, 10, 69-86

20. LI, H., ROSEN, S., AND W. SUEN (2001): "Conflicts and Common Interests in Committees", *American Economic Review,* 91(5), 1478-97

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

22. 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

23. MEIROWITZ, ADAM (2006): "Designing institutions to Aggregate Preferences and Information", *Quarterly Journal of Political Science,* 2006 1(4), 373-92

24. MEIROWITZ, ADAM (2005a):."In Defence of Exclusionary Deliberation: Communication and Voting with Private Beliefs and Values", Mimeo, Princeton University

25. MEIROWITZ, ADAM (2005b):."Communication and Bargaining in the Spatial Model", Mimeo, Princeton University

26. MEIROWITZ, ADAM (2002):"Informative Voting and Condorcet Jury Theorems with a Continuum of Types", *Social Choice and Welfare*, 19, 219-236

27. MILGROM, P. AND R.J. WEBER (1982): "A Theory of Auctions and Competitive Bidding", *Econometrica*, 50, 1089-1122

28. MYERSON, R.B. (1998): "Population Uncertainty and Poisson Games", *International Journal of Game Thoery*, 27(3), 375-392

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

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

31. NITZAN, S., AND J. PAROUSH (1985): *Collective Decision Making*, Cambridge: Cambridge University Press

32. OLIVEROS, S. (2006): "Information acquisition and flexible preferences in committees", Mimeo, University of Wisconsin-Madison

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

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

8 Appendix

8.1 Proof of Remark 1

Let us first look at the situation with $0 < L < R$. Here, $\mathbb{P}(R) = \{x : x \geq \frac{R}{2}\} \subset \{x : x \geq \frac{L}{2}\} = \mathbb{P}(L)$. Similarly, if we have $L < R < 0$, $\mathbb{P}(L) = \{x : x \leq \frac{L}{2}\} \subset \{x : x \leq \frac{R}{2}\} = \mathbb{P}(R)$. On the other hand, if $L < 0 < R$, $\mathbb{P}(L) = \{x : x \leq \frac{L}{2}\}$ and $\mathbb{P}(R) = \{x : x \geq \frac{R}{2}\}$, thus $\mathbb{P}(L) \cap \mathbb{P}(R) = \phi$.

8.2 Proof of Lemma 3

By hypothesis of the lemma, $\lim_{n\to\infty} \frac{\beta_L^n}{1-\beta_L^n} = \frac{\beta_L^0}{1-\beta_L^0}$ is a finite, positive number. Now suppose $\exists$ some $\varepsilon > 0$ such that $\alpha_n > 1 + \varepsilon$ for all n. Then $\frac{\beta_L^n}{1-\beta_L^n} = (\alpha_n)^n > (1+\varepsilon)^n \to \infty \; as \; n \to \infty$ which is a contradiction. On the other hand, suppose $\exists$ some $\varepsilon \in (0, 1)$ such that $\alpha_n < 1 - \varepsilon$ for all n. Then $\frac{\beta_L^n}{1-\beta_L^n} = (\alpha_n)^n < (1-\varepsilon)^n \to 0 \; as \; n \to \infty$, which is again a contradiction.

8.3 Proof of Lemma 5

For part (i) of the lemma, since $\beta_L \in (0, 1)$, Proposition 2 holds. Suppose $0 < y < x < 1$, and $f(z, \theta) = z^\theta (1-z)^{1-\theta}$, with both z and θ lying in $(0, 1)$. Note that if we fix θ, the function $f(z, \theta)$ is continuous and single peaked in z with the peak lying at θ. From the properties of this function, we can show that for any$0 < y < x < 1$, there exists a unique $\theta^* s.t. \; f(x, \theta^*) = f(y, \theta^*)$, and $x < \theta^* < y$. To be specific, $\theta^* = \frac{\log \frac{1-y}{1-x}}{\log \frac{x(1-y)}{y(1-x)}}$. Also, if both x and y increase, θ^* must increase. Since $0 < F(\frac{L}{2}) < t_L(\beta_L) < t_R(\beta_L) < F(\frac{R}{2}) < 1$, taking $t_R(\beta_L) = x$ and $t_L(\beta_L) = y$ and noting that $t_R(\beta_L)$ and $t_L(\beta_L)$ are strictly increasing functions of β_L, part (i) of the Lemma is established.

For part (ii), note that for any n, by Remark 2, we have $x_l^n < x_r^n$. Since $z_\sigma^n = F(x_\sigma^n)$, we have $z_r^n > z_l^n > 0$. Define, for any n, $h^n = z_r^n - z_l^n > 0$. Substituting, we have: $t(R, \pi^n) = z_l^n + qh^n$, and $t(L, \pi^n) = z_l^n + (1-q)h^n$. Therefore:

$$\frac{1-\beta_L^n}{\beta_L^n} = \left[\frac{(t(R,\pi^n))^\theta (1-t(R,\pi^n))^{1-\theta}}{(t(L,\pi^n))^\theta (1-t(L,\pi^n))^{1-\theta}}\right]^n = \left[\frac{(z_l^n + qh^n)^\theta (1 - z_l^n - qh^n)^{1-\theta}}{(z_l^n + (1-q)h^n)^\theta (1 - z_l^n - (1-q)h^n)^{1-\theta}}\right]^n$$

If $\beta_L^0 = 0$ (or 1), the left hand side of the above equation goes to infinity (or 0). This requires the term in the bracket large enough n to be greater (or less) than unity, or its logarithm to be positive (or negative). We can write,

$$\log \frac{\left(z_l^n + qh^n\right)^{\theta}\left(1 - z_l^n - qh^n\right)^{1-\theta}}{\left(z_l^n + (1-q)h^n\right)^{\theta}\left(1 - z_l^n - (1-q)h^n\right)^{1-\theta}} > 0 \Leftrightarrow \theta > \zeta(z_l^n, h^n)\ \forall n$$

where the function $\zeta(z_l^n, h^n)$ is defined as:

$$\zeta(z_l^n, h^n) \equiv \frac{-\log\left[\frac{1-z_l^n-qh^n}{1-z_l^n-(1-q)h^n}\right]}{\log\left[\frac{\left(z_l^n+qh^n\right)\left(1-z_l^n-(1-q)h^n\right)}{\left(z_l^n+(1-q)h^n\right)\left(1-z_l^n-qh^n\right)}\right]}$$

By Lemma 4, we know that for any sequence, with $\beta_L^0 \in \{0, 1\}$, $h^n \to 0^+$. Hence,

$$\lim_{h^n \to 0^+,\ z_l^n = t} \zeta(z_l^n, h^n) = \lim_{h^n \to 0^+,\ z_l^n = t} \left(\frac{-\log\left[\frac{1-z_l^n-qh^n}{1-z_l^n-(1-q)h^n}\right]}{\log\left[\frac{\left(z_l^n+qh^n\right)\left(1-z_l^n-(1-q)h^n\right)}{\left(z_l^n+(1-q)h^n\right)\left(1-z_l^n-qh^n\right)}\right]} \right) = \lim_{z_l^n = t} z_\sigma^n = t$$

By Lemma 4, if $\beta_L^0 = 0, t = F(\frac{R}{2})$, and $\theta > \zeta(z_l^n, h^n)\ \forall n \Rightarrow \theta > \lim_{h^n \to 0^+,\ z_l^n = t} \zeta(z_l^n, h^n) = F(\frac{R}{2})$. Similarly, if $\beta_L^0 = 1, t = F(\frac{L}{2})$, and $\theta < \zeta(z_l^n, h^n)\ \forall n \Rightarrow \theta < \lim_{h^n \to 0^+,\ z_l^n = t} \zeta(z_l^n, h^n) = F(\frac{L}{2})$.

8.4 Proof of Theorem 1

Here we only show that the only accumulation point is also the limit. For this, it is enough to show that given $\theta \in \Theta\left(\beta_L^0\right)$, for any neighbourhood ϵ of β_L^0, there is some large enough N, such that β_L^n in the equilibrium sequence must lie within the nighbourhood for all values of $n > N$.

First consider $\beta_L^0 \in (0, 1)$.Suppose the accumulation point is not the limit, and there is an infinite equilibrium subsequence β_L^m of the sequence β_L^n, such that for any $\epsilon > 0$, there is some M so that for all values of m larger than M, β_L^m lies outside $\left(\beta_L^0 - \epsilon, \beta_L^0 + \epsilon\right)$. Since even this subsequence must have an accumulation point, it must be either 0 or 1. But, by the second part of Lemma 5, since the limiting equilibrium condition must hold for accumulation points too, there cannot be an accumulation point for θ in $\Theta\left(\beta_L^0\right)$ at 0 or 1. Hence there is no such infinite subsequence.

The proof for $\beta_L^0 \in \{0, 1\}$ is similar.

8.5 Proof of Theorem 2

Theorem 1 guarantees existence of limiting equilibrium for all θ. Now consider $\theta < F\left(\frac{L}{2}\right)$. By Lemma 2, $t(S, \pi^n) > F\left(\frac{L}{2}\right) \forall n$ for $S = L, R$. Let $\delta = F\left(\frac{L}{2}\right) - \theta$. By Law of large numbers, given ϵ we can find N such that actual share of votes $\tau(S, \pi^n, \theta)$ under rule θ in any state S is greater than $F\left(\frac{L}{2}\right) - \delta > \theta$ for any $n > N$ with a probability larger than $1 - \epsilon$. Thus, under both states, $\mathcal{P}$ wins with a probability larger than $1 - \epsilon$.

Since $t(S, \pi^n) < F\left(\frac{R}{2}\right) \forall n \forall S$, by the same logic as above, any $\mathcal{Q}$-trivial rule aggregates information too.

Consider a consequential rule θ, for which the only equilibrium induced prior in the limit is $\beta_L^{-1}(\theta)$. By Lemma 5, $t_L\left(\beta_L^{-1}(\theta)\right) < \theta < t_R\left(\beta_L^{-1}(\theta)\right)$.Also, for any such θ, we can find a positive number η such that $F\left(\frac{L}{2}\right) + \eta < \theta < F\left(\frac{R}{2}\right) - \eta$. By Lemma 5, we can find a similar number $\kappa > 0$ such that $\kappa < \beta_L^{-1}(\theta) < 1 - \kappa$. Now, by Lemma 1 and Lemma 2, we can find some $\lambda > 0$ such that $t_R\left(\beta_L^{-1}(\theta)\right) - t_L\left(\beta_L^{-1}(\theta)\right) > \lambda$. Now, from Proposition 2, we can derive θ from $t_R\left(\beta_L^{-1}(\theta)\right)$ and $t_L\left(\beta_L^{-1}(\theta)\right)$ and can find another number $\mu > 0$ such that $t_L\left(\beta_L^{-1}(\theta)\right) + \mu < \theta < t_R\left(\beta_L^{-1}(\theta)\right) - \mu$. Since $t_R,\ t_L$ and θ^* are all continuous functions of β_L, we can find a number $\xi > 0$ such that for a range $\left(\beta_L^{-1}(\theta) - \xi, \beta_L^{-1}(\theta) + \xi\right)$ around $\beta_L^{-1}(\theta)$, $t_L - \frac{\mu}{2} < \theta < t_R + \frac{\mu}{2}$. Given ξ,we can find M_1 such that $\beta_L^n \in \left(\beta_L^{-1}(\theta) - \xi, \beta_L^{-1}(\theta) + \xi\right)$ in any π^n whenever $n > M_1$. Now consider $\delta = \min\left((t_R\left(\beta_L^{-1}(\theta) - \xi\right) + \frac{\mu}{2} - \theta,\ \theta - t_L(\beta_L^{-1}(\theta) + \xi) - \frac{\mu}{2}\right)$. By Law of large numbers, given ϵ we can find M_2 such that actual share of votes under rule θ under state R, $\tau(R, \pi^n, \theta)$ is less than $t_R\left(\beta_L^{-1}(\theta) - \xi\right) + \frac{\mu}{2} - \delta < \theta$ for any $n > M_2$ and the actual share under state L, $\tau(L, \pi^n, \theta)$ is greater than $t_L\left(\beta_L^{-1}(\theta) + \xi\right) + \frac{\mu}{2} - \delta > \theta$ for any $n > M_2$ with a probability larger than $1 - \epsilon$. Set $N = \max(M_1, M_2)$ and we are done.

8.6 Proof of Lemma 6

If $x_\sigma \leq -\frac{b}{2},\ z_\sigma = F(x_\sigma) \leq F(-\frac{b}{2})$ since $F(\cdot)$ is nondecreasing. If on the other hand, $x_\sigma \geq \frac{b}{2},\ z_\sigma = 1 - F(x_\sigma) \leq 1 - F(\frac{b}{2})$. Thus, for $\sigma \in \{l, r\},\ z_\sigma \leq \max\left(F(-\frac{b}{2}), 1 - F(\frac{b}{2})\right)$.Therefore,

$$t(S, \pi) \leq q \max(z_l, z_r) + (1 - q) \max(z_l, z_r) = \max(z_l, z_r) \leq \max\left(F(-\frac{b}{2}), 1 - F(\frac{b}{2})\right) < 1$$

The last inequality in the chain is guaranteed by assumption F. To show $t(S, \pi) > 0$, it is sufficient to show that both z_l and z_r cannot be 0 simultaneously. From assumption

F and the definition of x_σ, $z_\sigma = 0 \Rightarrow p_\sigma \in [\frac{1}{2} - \frac{b}{4}, \frac{1}{2} + \frac{b}{4}]$. We show that both p_l and p_r cannot be simultaneouly in this range. We start by noting that p_l and p_r increase in tandem, since both increase with β_L. When $p_l = q$, $\beta_L = \frac{1}{2}$. So, $p_r = 1 - q$. By the above positive relationship, $p_l < q \Rightarrow p_r < 1 - q$ and $p_r > 1 - q \Rightarrow p_l > q$. Note that by Assumption I, $q > \frac{1}{2} + \frac{b}{4}$ and $1 - q < \frac{1}{2} - \frac{b}{4}$.Hence,

$$p_l \in [\frac{1}{2} - \frac{b}{4}, \frac{1}{2} + \frac{b}{4}] \Rightarrow p_r < 1 - q < \frac{1}{2} - \frac{b}{4} \; and \; p_r \in [\frac{1}{2} - \frac{b}{4}, \frac{1}{2} + \frac{b}{4}] \Rightarrow p_l > q > \frac{1}{2} + \frac{b}{4}$$

8.7 Proof of Lemma 7

At $\beta_L = 0$, $x_l = x_r = \frac{b}{2} \Rightarrow z_l = z_r = 1 - F\left(\frac{b}{2}\right)$. Now, consider the interval of β_L such that p_l lies in $(0, \frac{1}{2} + \frac{b}{4}]$. In this interval, $x_l \in (\frac{b}{2}, 1] \cup \{-1\} \Rightarrow z_l = 1 - F(x_l)$. Also, in this interval of β_L, $p_r < \frac{1}{2} - \frac{b}{4} \Rightarrow x_r \in (\frac{b}{2}, 1) \Rightarrow z_r = 1 - F(x_r) > 0$, by assumptions F and I. For values of β_L such that $x_l \leq 1$, $x_r < x_l \Rightarrow z_l = 1 - F(x_l) < 1 - F(x_r) = z_r$, again by assumption F. For values of β_L such that $x_l = -1$, $z_l = 1 - F(-1) = 0 < z_r$. Thus, over this entire interval $z_r > z_l$. Note also that over this set of values of β_L, z_r is strictly decreasing, while z_l first strictly decreases and then stays at 0. For β_L such that $p_l = \frac{1}{2} + \frac{b}{4}$, $z_r = \overline{z_r}$, say. In the same way, consider the interval of β_L such that p_r lies in $[\frac{1}{2} - \frac{b}{4}, 1]$. Here, by the same token, $z_r < z_l$ except for $\beta_L = 1$ where $z_l = z_r = F\left(-\frac{b}{2}\right)$. z_l increases strictly from $\overline{z_l} > 0$ to $F\left(-\frac{b}{2}\right)$ over this interval, while z_r is initially 0 and then strictly increases.

Now, consider the remaining interval of β_L which is $\left(p_l^{-1}(\frac{1}{2} + \frac{b}{4}), p_r^{-1}(\frac{1}{2} - \frac{b}{4})\right)$. That this is a valid nonempty interval is guaranteed by assumption I. In this interval, $x_r \in (\frac{b}{2}, 1]$, and x_r increases with β_L. Thus, $z_r = 1 - F(x_r)$ is a strictly falling continuous function, going from $\overline{z_r} > 0$ to 0 over this interval. Similarly, z_l strictly and continuously increases from 0 to $\overline{z_l} > 0$. Therefore, there exists a unique β_L^* in this interval where $z_l = z_r$. This implies that at β_L^*, $t(L, \pi) = t(R, \pi)$. For all $\beta_L < \beta_L^*$, $z_l < z_r \Rightarrow t(L, \pi) = qz_l + (1-q)z_r < qz_r + (1-q)z_l = t(R, \pi)$. Similarly, for $\beta_L > \beta_L^*$, where $z_l > z_r$, we have $t(L, \pi) > t(R, \pi)$.

8.8 Proof of Lemma 8

We prove the result for the case $\beta_L^0 = 1$, the other one follows symmetrically. First we look at how $\frac{p_l}{p_r}$ changes with β_L.

$$\frac{p_l}{p_r} = \left(\frac{q}{1-q}\right)\left(\frac{q\beta_R + (1-q)\beta_L}{q\beta_L + (1-q)\beta_R}\right) = \left(\frac{q}{1-q}\right)\left(\frac{q + (1-q)\alpha}{q\alpha + (1-q)}\right),$$

where $\alpha = \frac{\beta_L}{\beta_R}$. Therefore, we have:

$$\frac{d}{d\beta_L}\left(\frac{p_l}{p_r}\right) = \frac{d\alpha}{d\beta_L}\cdot\frac{d}{d\alpha}\left(\frac{p_l}{p_r}\right) = \frac{1}{(1-\beta_L)^2}\left(\frac{q}{1-q}\right)\frac{(1-q)^2 - q^2}{(q\alpha + (1-q))^2} < 0$$

At $\beta_L = 1$, we have $p_l = p_r = 1$. Thus, for $\beta_L \in [0,1)$, we always have $p_l > p_r$ by the above strictly monotonic relationship. Since $\beta_L^0 = 1 \Rightarrow p_r^n \to 1$, by continuity we can find some m large enough such that for all $n > m$, we have $p_r^n > \frac{1}{2} + \frac{b}{4}$. Since $p_l^n > p_r^n$, for all $n > m$, $p_l^n > \frac{1}{2} + \frac{b}{4}$ too. Since we always have $\beta_L^n < 1$, $p_\sigma^n < 1$. Therefore, for all $n > m$, both x_l^n and x_r^n lie in the open interval $(-1, -\frac{b}{2})$. Also, $p_l^n > p_r^n \Rightarrow x_l^n > x_r^n$ for all $n > m$. This proves part (i). Part (ii) follows trivially from $p_\sigma^n \to 1$.

8.9 Proof of Lemma 9

Part (i) follows from Lemma 5 and Lemma 7.

For part (ii), we first consider the case with $\beta_L^0 = 1$. By Lemma 8, we know that for any such sequence, $x_\sigma^n \to \left(-\frac{b}{2}\right)^-$ for $\sigma = \{l, r\}$, and $x_l^n > x_r^n$ for all large enough n. For large enough n, $p_\sigma^n > \frac{1}{2} + \frac{b}{4} \Rightarrow z_\sigma^n = F(x_\sigma^n) \Rightarrow z_l^n > z_r^n > 0$ and $z_\sigma^n \to F(-\frac{b}{2})$. Define $h^n = z_l^n - z_r^n \to 0^+$. Substituting, we have: $t(L, \pi^n) = z_r^n + qh^n$, and $t(R, \pi^n) = z_r^n + (1-q)h^n$. Therefore:

$$\frac{\beta_L^n}{1-\beta_L^n} = \left[\frac{(t(L,\pi^n))^\theta (1 - t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^\theta (1 - t(R,\pi^n))^{1-\theta}}\right]^n = \left[\frac{(z_r^n + qh^n)^\theta (1 - z_r^n - qh^n)^{1-\theta}}{(z_r^n + (1-q)h^n)^\theta (1 - z_r^n - (1-q)h^n)^{1-\theta}}\right]^n$$

If $\beta_L^0 = 1$, the left hand side of the above equation goes to infinity. This requires the term in the bracket large enough n to be greater than unity, or its logarithm to be positive .

For the case with $\beta_L^0 = 0$, we again use Lemma 8 which tells us that $x_\sigma^n \to \left(\frac{b}{2}\right)^+$ for $\sigma = \{l, r\}$, and $x_l^n > x_r^n$ for all large enough n.We also know that for large enough

n, $p^n_\sigma > \frac{1}{2} - \frac{b}{4} \Rightarrow z^n_\sigma = 1 - F(x^n_\sigma) \Rightarrow z^n_r > z^n_l > 0$ and $z^n_\sigma \to 1 - F(-\frac{b}{2})$. Define $h^n = z^n_r - z^n_l \to 0^+$. Substituting, we have: $t(R, \pi^n) = z^n_r + qh^n$, and $t(L, \pi^n) = z^n_r + (1-q)h^n$. Therefore:

$$\frac{\beta^n_L}{1-\beta^n_L} = \left[\frac{(t(L,\pi^n))^\theta (1-t(L,\pi^n))^{1-\theta}}{(t(R,\pi^n))^\theta (1-t(R,\pi^n))^{1-\theta}}\right]^n = \left[\frac{(z^n_r + qh^n)^\theta (1 - z^n_r - qh^n)^{1-\theta}}{(z^n_r + (1-q)h^n)^\theta (1 - z^n_r - (1-q)h^n)^{1-\theta}}\right]^{-n}$$

Since the LHS goes to 0 in the limit, the term within the bracket in the RHS has to be greater than 1. Thus we have the exact same situation as in the proof of Lemma 5, and therefore, we need.

$$\log \frac{(z^n_r + qh^n)^\theta (1 - z^n_r - qh^n)^{1-\theta}}{(z^n_r + (1-q)h^n)^\theta (1 - z^n_r - (1-q)h^n)^{1-\theta}} > 0 \Leftrightarrow \theta > \zeta(z^n_r, h^n)\ \forall n$$

where the function $\zeta(z^n_l, h^n)$ is defined as in the proof of lemma 5.

By Lemma 4, if $\beta^0_L = 0$, $t = 1 - F(\frac{b}{2})$, and $\theta > \zeta(z^n_l, h^n)\ \forall n \Rightarrow \theta > \lim_{h^n \to 0^+,\ z^n_l = t} \zeta(z^n_l, h^n) = 1 - F(\frac{b}{2})$. Similarly, if $\beta^0_L = 1$, $t = F(-\frac{b}{2})$, and $\theta > F(-\frac{b}{2})$.For $\beta^0_L = \beta^*_L$, from Proposition 4, no value of θ can be ruled out.

8.10 Proof of Theorem 3

This is a proof by construction. Consider any preference reversal setting $(F(\cdot), q, b)$. Define the function

$$f_n(\beta, \theta) = \frac{1}{1 + \left[\frac{t_R(\beta)^\theta (1-t_R(\beta))^{1-\theta}}{t_L(\beta)^\theta (1-t_L(\beta))^{1-\theta}}\right]^n}$$

If given (n, θ) we can show that there is some fixed point β_n of the function $f_n(\beta, \theta)$, then that β_n is the solution to the equilibrium condition (10), proving that π^n exists for that θ. We show that for any $\theta \in \Theta(\beta^0)$, there is a sequence of fixed points of beliefs β_n such that $\beta_n \to \beta^0$ as $n \to \infty$. We prove this separately for different values ranges of β^0.

Case 1 $\beta^0 \in (0, \beta^*_L) \cup (\beta^*_L, 1)$.

By Lemma 9, in this range, $\Theta(\beta_L)$ is a continuous function $\theta^*(\beta_L)$. Since F admits a pdf f, $\theta^*(\beta_L)$ is differentiable too. Thus, there exists a neighbourhood $(\beta^0 - \epsilon, \beta^0 + \epsilon)$ where $\theta^*(\beta_L)$ is either only increasing, only decreasing or constant.

Suppose first that $\theta^*(\beta_L)$ is decreasing in $(\beta^0-\epsilon, \beta^0+\epsilon)$. Now, 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 $n \to \infty$. Thus, for δ small enough, there must exist some m such that $f_n(\beta+\epsilon, \theta^*(\beta^0)) < \delta$ and $f_n(\beta-\epsilon, \theta^*(\beta^0)) > 1-\delta$ for all $n > m$. In particular, choose $\delta < \epsilon$. Then, for all $n > m$, if $f_n(\beta, \theta^*(\beta^0))$ is plotted against β, it intersects the 45^0 line for some $\beta \in (\beta^0-\epsilon, \beta^0+\epsilon)$, which is the fixed point of the function. Call it β_n. To be specific, β_n is the solution of $f_n(\beta, \theta^*(\beta^0)) = \beta$, and for all $n > m$, $\beta_n \in (\beta^0-\epsilon, \beta^0+\epsilon)$. 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 $\theta^*(\beta_L)$ is increasing in $(\beta^0-\epsilon, \beta^0+\epsilon)$, then we can prove the theorem in an analogous way. However, if $\theta^*(\beta_L)$ is constant in the range $(\beta^0-\epsilon, \beta^0+\epsilon)$, the theorem may not hold. To be clear, this case requires that $t_L(\beta_L)$ increases (decreases) and $t_L(\beta_L)$ decreases (increases) so as to keep $\frac{t_R(\beta)^\theta(1-t_R(\beta))^{1-\theta}}{t_L(\beta)^\theta(1-t_L(\beta))^{1-\theta}}$ constant over the range. Ignore this case as non-generic.

Case 2 $\beta^0 \in \{0, \beta_L^*, 1\}$

First, consider the case $\beta^0 = 0$. Note that $\Theta(0) = \{\theta : \theta > 1 - F(\frac{R}{2})\}$.Select $\epsilon > 0$ small enough such that $t_R(\beta_L) > t_L(\beta_L)$ in the range $\beta_L \in (0, 2\epsilon)$. Choose $\delta < \epsilon$. By Case 1, for voting rule $\theta^*(\epsilon)$ there exists a sequence of equilibria β_n such that $f_n(\beta_n, \theta^*(\epsilon)) = \beta_n$ and $\beta_n \in (\epsilon-\delta, \epsilon+\delta)$ for n large enough. This implies $\beta_n < 2\epsilon$ for all n large enough. Now consider the sequence β_n such that $f_n(\beta_n, \theta^*(\epsilon)) = \beta_n$. For $\theta > 1 - F(\frac{R}{2}) > \theta^*(\epsilon)$,we must have $f_n(\beta_n, \theta) < \beta_n$. Now, consider the function $f_n(\beta, \theta) - \beta$. At $\beta = \beta_n$, the function is negative while at $\beta = 0$, the function is positive due to the boundedness of the shares. Since $f_n(\beta, \theta)$ is continuous, there is some $0 < \beta_n' < \beta_n < 2\epsilon$ such that $f_n(\beta_n', \theta) = \beta_n'$. Thus, given $\theta \in \Theta(0)$, for any ϵ small enough, there exists a sequence β_n'.such that $f_n(\beta_n', \theta) = \beta_n'$ and $|\beta_n' - 0| < 2\epsilon$ for all n large enough.

In the same way, we can prove the theorem for $\beta^0 = 1$.Next, consider the case with $\beta^0 = \beta_L^*$. Note that $\Theta(\beta_L^*) = (0, 1)$. To show existence of a limiting equilibrium for $\theta < t_L(\beta_L^*)$, use the neighbourhood $(\beta_L^* - \epsilon, \beta_L^*)$ to the left of β_L^*, and to show existence of a limiting equilibrium for $\theta > t_L(\beta_L^*)$, use the neighbourhood $(\beta_L^*, \beta_L^* + \epsilon)$ to the right of β_L^* and apply the same method.