A POSSIBILITY THEOREM ON INFORMATION AGGREGATION IN ELECTIONS

PAULO BARELLI, SOURAV BHATTACHARYA AND LUCAS SIGA

Abstract. We study aggregation of private information in large elections where all voters have the same preference. There are many states of the world, and each state is identified with a preference ranking over alternatives and a probability distribution over signals. Each voter draws his private signal independently from the said distribution conditional on the state. When there are two alternatives (say A and B), we obtain a simple condition that is necessary and sufficient for asymptotic aggregation of information: there should be a hyperplane in the simplex over signals that separates the conditional distributions in states where A is preferred from those in states where B is preferred. If this condition is satisfied, information is aggregated in an equilibrium sequence; and If the condition is violated, there exists no feasible strategy profile that aggregates information. While the hyperplane condition is satisfied only in special environments, it holds generically if the state space is discrete and the number of available signals is more than or equal to the number of states. The Condorcet Jury theorem is obtained as a special case of this theorem when there are only two states. In the case of more than two alternatives (say A_1, $A_2, ...A_n$), we find that information is aggregated if, for each pair of alternatives A_i and A_j, there is a hyperplane on the simplex that separates the set of distributions for which A_i is preferred over A_j from those for which A_j is preferred over A_i.

1. Introduction

In large elections, the decision relevant information is often dispersed throughout the electorate. This poses the classic problem of information aggregation in voting: even if voters potentially agree on who the right candidate is, each individual's vote contains only his own *private* information. It is therefore not guaranteed whether the election outcome obtained by aggregating everyone's vote leads to the right outcome, i.e., the outcome that any voter would have preferred if he knew all the information dispersed within the electorate. In this paper, we are interested in the question of information aggregation in large elections: when does voting lead to the same outcome that would have prevailed if all the private information were publicly known? We provide a simple condition on the diversity of information in the electorate that is both necessary and sufficient for information aggregation in equilibrium.

In our benchmark model, there are two alternatives (A and B), and preferences are represented by states: in some states of the world, all voters prefer A while in others, all voters

Date: June 22, 2015.

This paper is partly based on an earlier version with the same title, co-authored with Tim Feddersen and Wolfgang Pesendorfer. We are deeply indebted to both of them for their inputs. We are also thankful to Hari Govindan, Bard Harstad and Roee Teper for helpful suggestions. All errors that remain are our own.

prefer B. Alternative A wins if it obtains more than a threshold $q \in (0,1)$ share of votes. The dispersion of information is captured by the probability distribution over signals conditional on the state. Our central result (Theorem 1) is that there exists some strategy profile that aggregates information efficiently in large voting populations if and only if there is a hyperplane that separates the probability distributions arising from states where A is preferred from those arising from states where B is preferred. This result suggests that information aggregation happens only in special environments. If the hyperplane condition is not satisfied, then there is no *feasible* voting strategy that aggregates information. In such an environment, even if voters could commit to playing any strategy, information would not be aggregated. On the other hand, if the hyperplane condition is satisfied, then full information aggregation obtains as an equilibrium phenomenon (Theorem 2). The set of information aggregating strategies is identified by the hyperplane condition.

To see what our result implies in a very simple setting, suppose that voters' private information is categorized two signals a and b, and voter rankings depend only on the proportion of a-signals in the population. To capture this more formally, suppose the state θ is a number in $[0,1]$, and in state θ, each voter obtains signal a with probability θ. The state then is the same as the expected proportion of a-signals in the electorate, which is also, almost surely, the actual proportion if the population is large. According to our result, information is aggregated for the following type of voter preference: alternative A is preferred for high states (say, larger than θ^*) and alternative B is preferred for low states (i.e., lower than θ^*). On the other hand, information aggregation fails if the voters prefer to elect A for moderate states (say, between θ_1 and θ_2) and B for extreme states (i.e., lower than θ_1 or higher than θ_2). Since a strategy maps individual signals into votes, any strategy that leads lower vote shares for A at low proportions of a-signals and higher vote shares for A at moderate proportions of a-signals must also produce high vote shares for A when the proportion of a-signals is high. Thus, it is impossible to find a strategy that guarantees a win for B both in the very high and very low states, but a win for A in the moderate states. The above example suggests that the possibility of aggregation depends on what information the signals convey: if a higher proportion of signal a (resp. b) indicates higher quality of candidate A (resp. B), then information aggregation holds. On the other hand, if a very high or very low proportion of a-signals conveys than candidate A has a very extreme position on a policy issue, then information aggregation fails.

There is a large literature going back to Condorcet (1786) that argues that information is aggregated in common value environments. In the canonical Condorcet Jury model, there are two states (A and B) and two signals (a and b). State A (resp. B) is simply interpreted as all situations where candidate A (resp. B) is the commonly preferred candidate. In each state, voters draw their signals independently from a distribution: $\Pr(a|A) = p_A > \frac{1}{2}$ and $\Pr(b|B) = p_B > \frac{1}{2}$. Thus, the signal a (resp. b) can be interpreted as an assessment that A (resp. B) is the right candidate: however, the assessment may be mistaken. In this setting, if all voters vote according to their private signal, the majority votes for the correct alternative almost surely by the Law of Large Numbers. In this sense, individual uncertainty does not matter for the aggregate outcome in a large electorate, and information is always aggregated. This result is popularly known as the Condorcet Jury Theorem (CJT).

The earlier statistical work on the theorem have always equated (implicitly or explicitly) the state of the world with a ranking over the two alternatives (see Ladha (1992) and Berg (1993) among others). This strand has also assumed "sincere voting", i.e., that voters vote their signals. The game theoretic literature started with the insight in Austen-Smith and Banks (1996) that the sincere voting profile may not be a Nash equilibrium. Since then, there have been other proofs of CJT showing that information can be aggregated in Nash equilibrium for majority and supermajority rules and for more varied information structure (e.g. Wit (1998), Feddersen and Pesendorfer (1997), Myerson (1998), Duggan and Martinelli (2005)). However, all the papers dealing with common value environments (except Feddersen and Pesendorfer (1997)) have retained the two-state structure, effectively assuming that all situations where one alternative is better for all voters can be lumped into a single state.

Our starting point is that electorates often have a far richer informational diversity than is supposed by the canonical two-state model. Electability of a candidate depends on myriad factors like his policy positions on different issues, his past history, party affiliation, the state of the economy, the geopolitical situation and so forth. We use the state variable to capture all the different factors that affect the preference of the voters. It is also likely that different individuals hold information on different factors, and the overall preference of the electorate is based on the distribution of this dispersed information. Thus, reducing every individual's private signal to a probability assessment of which candidate is better seems to be too narrow a way to describe the dispersion of private information in the society. Our main contribution is to demonstrate how the property of information aggregation depends on the relationship between (common) preference and distribution of information in the electorate.

In the current paper, there is a compact state space Θ, and a compact set of signals X. We allow for both discrete and continuous signals. In each state θ, each voter receives a signal $x \in X$ that is an independent and random draw from the distribution $\eta(\cdot|\theta)$. The state space is partitioned into two sets $\mathcal{A}$ and $\mathcal{B}$: in states lying in the set $\mathcal{A}$, all voters prefer alternative A and in the states lying in the set $\mathcal{B}$, all voters prefer alternative B. Notice that, in a formal sense, voter rankings are simply defined over the space of probability distributions over signals, henceforth denoted by $\Delta(X)$. In large electorates, given a state, the frequency distribution over signals approximates the probability distribution. Therefore, our setup is approximately equivalent to one where there are a large number of voters whose ranking depends on the entire profile of private signals in the electorate (with the added restriction that identity of individuals does not matter for preference).

Formally stated, we have two sets of results. First, we show that there exists some feasible strategy profile that aggregates information if and only if the conditional probability distributions arising from states in $\mathcal{A}$ can be separated from those arising from states in $\mathcal{B}$ by a hyperplane on $\Delta(X)$ (Theorem 1). Our result extends to both symmetric and asymmetric strategies (Corollary 1). The proof of this result simply follows from the fact that the vote share for A is a linear functional of the vectors in $\Delta(X)$. This condition also allows us to identify the class of strategies that do aggregate information for a given voting rule. Moreover, the particular voting threshold is not important - if information aggregation is feasible for a given threshold, it is feasible for every other non-unanimous threshold rule (Corollary 2).

The second result says that if the hyperplane condition is satisfied, then there exists a profile of information-aggregating strategy that is also a Nash equilibrium (Theorem 2). This result borrows the insight from McLennan (1998) that in any voting game, a symmetric strategy profile that maximizes the ex-ante payoff must also be a Nash equilibrium. Combining these two results, the hyperplane condition is both necessary and sufficient for information aggregation in the limit.

We extend these results to the case of multiple alternatives (say 1 through n) in a later section. In this case, we consider preferences directly over probability distributions over signals; and, as a standard for information aggregation, demand that the correct outcome be obtained for almost every distribution on the simplex. In this case, there is a gap between necessary and sufficient conditions on preference for the existence of a strategy that aggregates information. The necessary condition is that the set of distributions for which a given alternative is preferred be convex, but that condition does not guarantee FIE. The existence of an information-aggregating strategy is guaranteed if the hyperplane condition be satisfied for each pair of alternatives i and j : the set of distributions for which i is preferred over j should be separable from the set of distributions for which j is preferred over i by a hyperplane (Theorem 3). An equivalent of Theorem 2 goes through easily in this case.

At this stage, it is important to point out the relationship of our work with McLennan (1998). McLennan points out that if there is some *feasible* symmetric strategy that fully aggregates information, then there exists an equilibrium that aggregates information too. The implication of this result is that aggregation failure is not an equilibrium phenomenon at all. We identify necessary and sufficient conditions for existence of a feasible symmetric strategy that aggregates information, and then use the insight from McLennan to show that in environments where information aggregation is feasible, it is also an equilibrium property.

As implications to our main theorem (Theorem 1), we provide certain sufficient conditions for information aggregation in elections which are new to the literature. In a setting where there are two alternatives, r states and k signals, we show that information is always aggregated if (i) there are at least as many signals as states, i.e., $r \leq k$, and (ii) the conditional probability distribution over signals in any given state cannot be obtained as a convex combination of the conditional distributions in the other states (Corollary 4). This result implies that information is aggregated if the signals are rich enough for the electorate to distinguish between all states. A simple corollary of this result is that whenever there are just two states, information is aggregated as long as the probability distribution over signals is different in the two states. When both states and signals have a natural order, a commonly studied informativeness condition on signals is Monotone Likelihood Ratio Property (MLRP) (Milgrom 1981). We show that the sufficient condition for information aggregation boils down to a weaker version of MLRP in this environment. In fact, these sufficient conditions for information aggregation in large elections have a strong parallel in Siga (2013), which obtains similar results in the context of auctions with multidimensional signals and discrete states.

There is a recent literature showing that information aggregation can fail to obtain in elections, but these papers rely either on preference diversity (Bhattacharya (2013), Acharya (2013) or residual uncertainty, i.e., uncertainty about probability distributions over preferences (Feddersen and Pesendorfer (1997) or information (Mandler 2012). Our paper is the first to show that aggregation can fail in an environment where voters have the same preferences, and there is no residual uncertainty. Mandler (2012) shows that, in a common preference environment, uncertainty over the probability distribution over signals may lead to "wrong" equilibrium assessments about the state due the peculiar logic of pivotality. On the other hand, the logic for aggregation failure that we unveil is not based on any equilibrium reasoning at all.

Feddersen and Pesendorfer (1997) consider a setting with diverse preference along with the restriction that, for every voter, the utility difference between A and B is increasing in the state. In this setting, they show that every sequence of equilibrium aggregates information (provided there is no residual uncertainty over the distribution of preferences). On the other hand, Bhattacharya (2013) shows that if we relax the monotonicity assumption on utility differences, there generically exist equilibrium sequences that do not aggregate information. However, Bhattacharya (2013) is silent about whether information aggregating equilibria do exist in such settings.

In a separate section, we allow voters to have diverse preferences in addition to diverse information. The feasibility result (Theorem 1) generalizes to a case with diverse preferences, with the only modification that $\mathcal{A}$ (resp. $\mathcal{B}$) is defined as the set of states in which the alternative A (resp. B) would win under full information. Our environment allows both Feddersen and Pesendorfer (1997) and Bhattacharya (2013) as special cases. We can show that in each of these cases, there exist strategies that do aggregate information. However, our proof of Theorem 2 does not directly extend to a setting with diverse preferences. Therefore, the existence of an information aggregating strategy does not automatically imply information aggregation in equilibrium. We are currently working on conditions that guarantee the existence of some equilibrium sequence that aggregates information in the diverse preference case.

The rest of the paper is organized as follows. Section 2 lays out the model with common voter preferences. Section 3 provides the main theorem that establishes conditions under which an environment allows full information aggregation, and discusses the implications for some specific environments. Section 4 shows that existence of a feasible strategy profile that fully aggregates information implies the existence of an equilibrium sequence of profiles that does the same. Section 5 discusses three extensions: the first one generalizes the results to more than two alternatives, the second one shows that our main results do not change if we consider continuous signal spaces, and the other one derives the conditions for existence of information-aggregating strategy profile in the case where voters may have preference heterogeneity. Section 6 concludes.

2. Model

In the model, there are n voters choosing between two alternatives A and B. Alternative A wins if it receives more than $q \in (0,1)$ share of votes and loses if it receives less than q share If A receives exactly nq votes, then we assume that tie is broken randomly.

In this section, we treat every voter as having the same preferences. In a later section, we show that our feasibility result can be extended to a setting where voters in an electorate may have different preferences. The utility of a voter from an alternative depends on an unobserved state variable $\theta \in \Theta$, where Θ is a compact, separable metric space. The utility of each voter is given by a bounded and continuous function $u : \Theta \times \{A, B\} \to \mathbb{R}$. Let the three sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{I}$ denote the respective regions in Θ where A is preferred to B, B is preferred to A and the voters are indifferent.

$$\begin{aligned}
\mathcal{A} &= \{\theta \in \Theta : u(\theta, A) > u(\theta, B)\} \\
\mathcal{B} &= \{\theta \in \Theta : u(\theta, A) < u(\theta, B)\} \\
\mathcal{I} &= \{\theta \in \Theta : u(\theta, A) = u(\theta, B)\}
\end{aligned}$$

Each voter i receives a private signal $x \in X$, where X is a compact, separable, metric space. Profiles of signals are denoted by $x^n \in X^n$. The information structure is captured by a probability measure η on $\Theta \times X$, as follows. When a voter gets a signal $x \in X$, he makes inferences about the true state θ using the conditional $\eta(\cdot\,|x) \in \Delta(\Theta)$, where $\Delta(\cdot)$ denotes the space of probability measures over "$\cdot$", endowed with the weak star topology. Likewise, for a given θ, $\eta(\cdot\,|\theta) \in \Delta(X)$ is the conditional on the signal received by an individual voter. Hence we assume that voters' signals are independent and identically distributed conditional on θ. We assume of $\theta \mapsto \eta(\cdot|\theta)$ is surjective and strongly continuous: for each Borel measurable $E \subset X$, $\eta(E|\theta_k) \to \eta(E|\theta)$ as $\theta_k \to \theta$. Notice that this continuity assumption is trivially satisfied when Θ is discrete. Let $m \in \Delta(\Theta)$ denote the marginal of η on Θ. That is, m is the prior on Θ. We also assume that while both sets $\mathcal{A}$ and $\mathcal{B}$ have positive probability ex-ante, indifference occurs with zero probability, i.e. $m(\mathcal{A}) > 0$, $m(\mathcal{B}) > 0$ and $m(\mathcal{I}) = 0$. This assumption essentially implies that rankings are strict when Θ is discrete.

A tuple $\{u, \Theta, X, \eta, q\}$ is defined as an environment. An environment in addition to an electorate size n defines a game. In a game, a strategy for voter i is a measurable function $s_i : X \to \{0, 1\}$, with $s_i(x) = 1$ meaning that i votes for A at signal x. A mixed strategy is a measurable function $\sigma_i : X \to [0, 1]$ with $\sigma_i(x)$ being the probability of a vote for A at signal x. Unless mentioned otherwise, we consider only symmetric strategies, i.e. voters with the same signal play the same strategy. Hence we can drop the index i and use $s(\cdot)$, $\sigma(\cdot)$ to denote individual strategies. We sometimes abuse terminology and refer to s or σ as a profile of strategies, with the understanding that every player uses the same s or σ, as the case may be.

In the main body of the paper, we shall consider the case where the space of signals X is countable (possibly finite). However, all our results holds in the case were X is an infinite. We deal with the infinite case in a separate section.

In what follows, we define the standard for information aggregation for a given strategy profile.

2.1. **Full Information Equivalence.** We say that the election leads to a wrong outcome if, for $\theta \in \mathcal{A}$, the alternative A fails to win, or for $\theta \in \mathcal{B}$, the alternative B fails to win. Fixing a state θ, strategy σ and electorate size n, we denote by the random variable $z_n^\sigma(\theta)$ as the actual proportion of votes obtained by alternative A. The realization of $z_n^\sigma(\theta)$ depends on the realized signal profile x^n (drawn according to $\eta(\cdot|\theta)$) and the realized vote tally given x^n (drawn according to $\sigma(x^n)$). Clearly, in states $\theta \in \mathcal{A}$, there is a wrong outcome if the realized value of $z_n^\sigma(\theta)$ is weakly less than q and in states $\theta \in \mathcal{A}$, a wrong outcome occurs if the realized value is weakly greater than q. For a given environment, we can find the ex-ante likelihood of error induced strategy σ by integrating the probability of error at each state with respect to the prior distribution over states. Thus, we denote the expected likelihood of a wrong outcome by

$$(1)\quad W_n^\sigma = \int_{\mathcal{A}} \mathbb{1}\{\theta : z_n^\sigma(\theta) \leq q\} m(d\theta) + \int_{\mathcal{B}} \mathbb{1}\{\theta : z_n^\sigma(\theta) \geq q\} m(d\theta) + \int_{\mathcal{I}} \mathbb{1}\{\theta : z_n^\sigma(\theta) = q\} m(d\theta)$$

We say that in an environment $\{u, \Theta, X, \eta, q\}$, the strategy $\sigma(\cdot)$ achieves Full Information Equivalence (FIE) if the ex-ante likelihood of error induced by $\sigma(\cdot)$ in a sequence of games converges to 1 as the number of voters increases unboundedly. More loosely, we say that $\sigma(\cdot)$ aggregates information in this environment. Formally, a strategy profile σ achieves FIE if for every $\delta > 0$, there exists some n such that $W_n^\sigma < \delta$. Notice that there is no guarantee that such a strategy will exist for every environment. In fact, if there exists some strategy that aggregates information in an environment, we say that the environment allows FIE. Next, we provide an equivalent way to define FIE.

Define the expected share of votes for A in state θ under symmetric strategy σ as

$$z^\sigma(\theta) \equiv \sum_{x \in X} \sigma(x) \eta(x|\theta) \tag{2}$$

For an asymmetric strategy profile $\sigma = (\sigma_i)_{i \geq 1}$, we define

$$z^\sigma(\theta) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i^\theta,$$

where $\mu_i^\theta = \sum_{x \in X} \sigma_i(x) \eta(x|\theta)$. Observe that $z^\sigma(\theta)$ is a linear functional of the conditional probability vector $\eta(\cdot|\theta)$ and is continuous in θ.

By the Strong Law of Large numbers,

$$z_n^\sigma(\theta) \to z^\sigma(\theta)$$

$\eta(\cdot|\theta)$-almost surely as $n \to \infty$. In other words, as n becomes large, the realized share of votes for A is very close to the expected share $z^\sigma(\theta)$ with a high probability.

Given σ, let

$$\begin{aligned} \mathcal{A}^\sigma &= \{\theta \in \Theta : z^\sigma(\theta) > q\} \\ \mathcal{B}^\sigma &= \{\theta \in \Theta : z^\sigma(\theta) < q\} \\ \mathcal{I}^\sigma &= \{\theta \in \Theta : z^\sigma(\theta) = q\} \end{aligned}$$

denote the regions in Θ where the expected share of votes for A is higher than, lower than, or equal to q respectively. From continuity arguments, it is easy to see that if the electorate is sufficiently large and every voter uses the same strategy σ, alternative A (resp. B) wins with an arbitrarily high probability in states in $\mathcal{A}^\sigma$ (resp. $\mathcal{B}^\sigma$). The outcome can go either way in the set $\mathcal{I}^\sigma$.

We say that σ achieves Full Information Equivalence (FIE) if the set of states where the preferred alternative fails win almost surely is of m-measure zero.

Definition 1 (Full Information Equivalence)**.** *In a given environment* (u, Θ, X, η, q)*, a strategy* σ *achieves Full Information Equivalence (FIE) if*

$$m(\mathcal{A} \backslash \mathcal{A}^\sigma) = m(\mathcal{B} \backslash \mathcal{B}^\sigma) = 0.$$

We say that an environment (u, Θ, X, η, q) *allows FIE when there exists a profile* σ *that achieves FIE.*

It is easy to check that this definition is equivalent to the expected probability of error $W_n^\sigma \to 0$. For somewhat technical purposes, we also need another definition. For some $\epsilon > 0$, we say that a strategy σ achieves ε-FIE if

$$m(\mathcal{A} \backslash \mathcal{A}^\sigma) = m(\mathcal{B} \backslash \mathcal{B}^\sigma) < \varepsilon.$$

If all voters follow a strategy that achieves ε-FIE, for a large enough electorate, the probability of error is arbitrarily close to ε. Clearly, a strategy that achieves FIE also achieves ε-FIE for any $\varepsilon > 0$.

The next section discusses properties of the environment that allows FIE. Notice that an environment allowing FIE is necessary but not sufficient for information aggregation in equilibrium.

3. Feasibility of Information Aggregation

To demonstrate the condition that determines whether an environment allows FIE or not, we start with a series of examples. In the first example, we show that when a third state is added to a very standard Condorcet Jury setup, we may fail to get information aggregation

Example 1. *Suppose there are two alternatives A and B and three states L, M and R, with A preferred in L and R and B preferred in M. Also, suppose that, $X = \{x, y\}$, and for some $p > \frac{1}{2}$,*

$$\Pr(x|L) = p,\ \Pr(x|M) = \frac{1}{2},\ \textit{and}\ \Pr(x|R) = 1 - p$$

The winner is decided by majority rule. This environment does not allow FIE.

Proof. Any strategy in the above environment is a pair of probabilities $(\sigma(x), \sigma(y))$. Now, in order for A to win almost surely in both L and R, we require the strategy $\sigma(\cdot)$ to satisfy the following two conditions respectively

$$\begin{aligned} z^\sigma(L) &= p\sigma(x) + (1-p)\sigma(y) > \frac{1}{2} \\ z^\sigma(R) &= (1-p)\sigma(x) + p\sigma(y) > \frac{1}{2} \end{aligned}$$

Taken together, we must have $\sigma(x) + \sigma(y) > 1$, which violates the condition for B winning almost surely in state M, which is

$$z^\sigma(M) = \frac{1}{2}\sigma(x) + \frac{1}{2}\sigma(y) < \frac{1}{2}$$

□

The problem with information aggregation in this example is the following: in order for A to win in state L where x is the more frequent signal, $\sigma(x)$ has to be high. Similarly, in order for A to win in state R where y is the more frequent signal, $\sigma(y)$ has to be high. But then, in state M where both these signals occur with moderate probability, the vote share for A is already high and B cannot win.

The next example develops similar ideas in a richer state space.

Example 2. *Suppose there are two candidates A and B and the state θ is distributed unformly over $[0, 1]$. The signal space $X = \{x, y\}$, and $\Pr(x|\theta) = \theta$. The winner is decided by majority rule. Consider two different preference environments. In the first case, suppose all voters prefer A if $\theta > t$ and B for $\theta < t$, for some $t \in (0, 1)$. In the second case, for some $0 < t_1 < t_2 < 1$, A is preferred whenever $\theta \in (t_1, t_2)$ and B is preferred when $\theta < t_1$ or $\theta > t_2$. The first environment allows FIE but the second environment does not.*

Proof. For any strategy $\sigma = (\sigma(x), \sigma(y))$, the vote share function is given by

$$z^\sigma(\theta) = \theta\sigma(x) + (1-\theta)\sigma(y)$$

Notice that the vote share function is linear in θ.

For σ to satisfy FIE in the first environment, we must have $z^\sigma(\theta) > \frac{1}{2}$ for $\theta > t$ and $z^\sigma(\theta) < \frac{1}{2}$ for $\theta < t$. By continuity, we must have $z^\sigma(t) = \frac{1}{2}$. Any σ that satisfies (i) $z^\sigma(t) = \frac{1}{2}$ and (ii) $\sigma(x) > \sigma(y)$ leads to FIE. It is easy to check that we can always find some σ with these properties.

For FIE in the second environment, we must have $z^\sigma(\theta) > \frac{1}{2}$ for $\theta \in (t_1, t_2)$ and $z^\sigma(\theta) < \frac{1}{2}$ for $\theta \in [0, t_1) \cup (t_2, 1]$. By the same argument as above, we must have $z^\sigma(t_1) = z^\sigma(t_2) = \frac{1}{2}$.

Since $z^\sigma(\theta)$ is linear, this is only possible if $\sigma(x) = \sigma(y) = \frac{1}{2}$. But in this case, we get large probabilities of errors in almost all states. Thus, there is no symmetric strategy profile that achieves FIE in this case. □

The main idea in the paper is contained in example 2. Notice that what drives information aggregation in this example is the convexity of the set of states for which a given alternative is preferred: in the first environment both the sets $\mathcal{A}$ and $\mathcal{B}$ are convex, while in the second environment the set $\mathcal{B}$ is non-convex.

To see why convexity matters, denote as "pivotal" those states that lie on the boundary of $\mathcal{A}$ and $\mathcal{B}$. In order to induce the correct outcome in the neighborhood of a pivotal state, the strategy must produce vote share equal to the threshold q at the pivotal state. In the unidimensional case, if there are multiple pivotal states, only a constant vote share function can produce vote shares equal to q at all pivotal states. In the above example, failure of convexity produces multiple pivotal states, leading to failure of FIE.

The idea contained in example 2 can be actually generalized to multidimensional state and signal spaces. There are two conceptual issues involved in this generalization. First, notice that in example 2, the state space is isomorphic to the simplex over signals: in fact, the preferences are actually described over probability distributions over signals. Second, as equation (2) indicates, the vote share function is linear in these conditional probability vectors (which, in the example, boils down to linearity in the state variable). Therefore, convexity of the sets $\mathcal{A}$ and $\mathcal{B}$ are incidental to the example: in general, the condition for an environment allowing FIE is convexity of the probability distributions for which the same alternative is preferred. This can explain the failure in example 1: $\Pr(x|M)$ is a convex combination of $\Pr(x|L)$ and $\Pr(x|R)$, but while alternative A is preferred in states L and R, alternative B is preferred in state M.

By a simple application of the separating hyperplane theorem, the above condition is equivalent to the existence of a hyperplane on the simplex over signals that separates the set of distributions arising in states for which $\mathcal{A}$ is preferred from the set of distributions arising in states for which $\mathcal{B}$ is preferred. If there are states that correspond to conditional probability distributions lying on this hyperplane, these states must be pivotal states - and the information aggregating strategy must produce a vote share exactly equal to q in these states.

.In what follows, we provide a formal description of this idea.

3.1. **Main Result.** Before stating the main result, we need some more definitions.

3.1.1. *Definitions: Pivotal states and hyperplanes.* We say that a state $\theta \in \Theta$ is *pivotal* if, for each $\varepsilon > 0$, $m(\mathcal{A} \cap B_\varepsilon(\theta))$ and $m(\mathcal{B} \cap B_\varepsilon(\theta))$ are positive, where $B_\varepsilon(\theta)$ is the ε open ball around θ. Let $M^{piv} \subset \Theta$ denote the set of pivotal states. By continuity of the utility function, at all pivotal states, the voters must be indifferent between A and B.

A hyperplane in $\Delta(X)$ is denoted by

$$H = \{\mu \in \Delta(X) : \sum_{x\in X} h(x)\mu(x) = \ell\},$$

for a given measurable function $h : X \to \mathbb{R}$, and a number $\ell \in \mathbb{R}$. Given a hyperplane H, we use

$$H^+ = \{\mu \in \Delta(X) : \sum_{x\in X} h(x)\mu(x) > \ell\} \text{ and}$$
$$H^- = \{\mu \in \Delta(X) : \sum_{x\in X} h(x)\mu(x) < \ell\}$$

to denote the two associated half-spaces.

The following theorem states that information is aggregated by some strategy if and only if there is a hyperplane on the simplex over signals that separates the conditional probability vectors arising from states in the interior of $\mathcal{A}$ from those arising from states in the interior of $\mathcal{B}$; and contains all such vectors arising from pivotal states. The proof relies heavily on the linearity of the vote share function in the conditional probabilities of signals.

Theorem 1. *An environment (u, Θ, X, η, q) allows FIE if and only if there exists a hyperplane H in $\Delta(X)$ such that $\eta(\cdot|\theta) \in H^+$ for $\theta \in \mathcal{A}$, $\eta(\cdot|\theta) \in H^-$ for $\theta \in \mathcal{B}$, and, if $M^{piv} \neq \emptyset$, $\eta(\cdot|\theta) \in H$ for $\theta \in M^{piv}$.*

Proof. Let σ be a profile that achieves FIE. We first show that, if $M^{piv} \neq \emptyset$, then $M^{piv} \subset \mathcal{I}^\sigma$. Assume to the contrary and pick $\theta \in M^{piv} \cap \mathcal{A}^\sigma$. As $z^\sigma(\cdot)$ is continuous, there is $\varepsilon > 0$ with $B_\varepsilon(\theta) \subset \mathcal{A}^\sigma$. Because θ is pivotal, we have $m(B_\varepsilon(\theta) \cap \mathcal{B}) > 0$. Hence $m(\mathcal{B} \cap \mathcal{A}^\sigma) \geq m((B_\varepsilon(\theta) \cap \mathcal{B}) \cap \mathcal{A}^\sigma) = m(B_\varepsilon(\theta) \cap \mathcal{B}) > 0$. But this means that $m(\mathcal{B}\backslash\mathcal{B}^\sigma) > 0$, contradicting FIE.

Next, we claim that $\mathcal{A} \subset \mathcal{A}^\sigma$. Again, by contradiction, pick $\theta \in \mathcal{A}$ and suppose $\theta \notin \mathcal{A}^\sigma$. Observe that because σ achieves FIE, there must exist at least one pair x, x' with $\sigma(x) < q$ and $\sigma(x') > q$. For each integer k, let $\mu_k \in \Delta(X)$ be given by $\mu_k(E) = (1-1/k)\eta(E|\theta)+(1/k)\mu(E)$, for each measurable $E \subset X$, where μ satisfies $\sum_{x\in X} \sigma(x)\mu(x) < q$. Then $\mu_k \to \eta(\cdot|\theta)$ strongly as $k \to \infty$, and $\sum_{x\in X} \sigma(x)\mu_k(x) < q$ for every k. Let $\varepsilon > 0$ satisfy $m(B_\varepsilon(\theta) \cap \mathcal{B}) = 0$, which must exist because $\theta \in \mathcal{A}$. Because the range of $\theta \mapsto \eta(\cdot|\theta)$ is $\Delta(X)$ and $\theta \mapsto \eta(\cdot|\theta)$ is strongly continuous, there exists $\theta' \in B_\varepsilon(\theta)$ and k large such that $\eta(\cdot|\theta') = \mu_k(\cdot)$. Observe that $z^\sigma(\theta') < q$. By continuity of z^σ, we can find $\varepsilon' > 0$ with $B_{\varepsilon'}(\theta') \subset B_\varepsilon(\theta)$ and $z^\sigma(\theta'') < q$ for all $\theta'' \in B_{\varepsilon'}(\theta')$. Hence $m(B_\varepsilon(\theta) \cap \mathcal{B}) = 0$ while $m(B_\varepsilon(\theta) \cap \mathcal{B}^\sigma) > 0$, meaning that $m(\mathcal{A}\backslash\mathcal{A}^\sigma) > 0$, contradicting FIE.

Similarly, we obtain $\mathcal{B} \subset \mathcal{B}^\sigma$. Setting $H = \{\mu \in \Delta(X) : \sum_{x\in X} \sigma(x)\mu(x) = q\}$, the "only if" direction is verified.

For the "if" part, let $H = \{\mu \in \Delta(X) : \sum_{x\in X} h(x)\mu(x) = \ell\}$ denote the required hyperplane. Let σ be a mixed strategy such that $\sigma(x) = q + \varepsilon(h(x) - \ell)$, where $\varepsilon > 0$ ensures that $\sigma(x) \in [0, 1]$ (observe that, by re-scaling, it is without loss to have $|h(x)| \leq 1$ for every $x \in X$, and hence $|h(x)-\ell| \leq 2$ for every $x \in X$.) If $M^{piv} \neq \emptyset$, pick $\theta \in M^{piv}$. As $\eta(\cdot|\theta) \in H$, we have $z^\sigma(\theta) = \sum_{x\in X}(q + \varepsilon(h(x) - \ell)\eta(x|\theta) = q$, so $M^{piv} \subset \mathcal{I}^\sigma$. Likewise, for $\theta \in \mathcal{A}$ (resp. $\theta \in \mathcal{B}$) we have $\eta(\cdot|\theta) \in H^+$ (resp. $\eta(\cdot|\theta) \in H^-$), readily showing that $\mathcal{A} \subset \mathcal{A}^\sigma$ and $\mathcal{B} \subset \mathcal{B}^\sigma$for m-almost all θ (whether M^{piv} is empty or not), and FIE is verified. □

While the analysis above was restricted to symmetric strategies, The exact same results hold true when we consider FIE under asymmetric strategies. If there exist an asymmetric profile that achieves FIE in an environment (u, θ, X, η, q), then exactly as in the symmetric case there will exist a hyperplane H in $\Delta(X)$ as in the statement of Theorem 1. Again setting $\sigma(x) = q+\varepsilon(h(x)-\ell)$ and following the steps of the "if" part above shows that the symmetric profile σ achieves FIE. Hence we have the following corollary, which says that it is without loss of generality to restrict our attention to symmetric strategies.

Corollary 1. *If an environment (u, θ, X, η, q) allows FIE with an asymmetric profile $\hat{\sigma} = (\hat{\sigma}_i)_{i\geq 1}$ then there exists a strategy σ such that the symmetric strategy profile where every voters uses σ also achieves FIE.*

Next, we show that the existence of an information aggregating strategy depends only on how preference interacts with distribution of information and not on the voting rule in use. If there is a strategy that achieves FIE for a given voting rule, then, for each non-unanimous threshold voting rule, there exists some strategy that achieves FIE.

Corollary 2. *If an environment (u, θ, X, η, q) allows FIE then for any $\hat{q} \in (0, 1)$, $(u, \theta, X, \eta, \hat{q})$ allows FIE.*

Proof. Let H be the hyperplane in $\Delta(X)$ associated with the environment (u, θ, X, η, q), and pick $\hat{q} \in (0, 1)$. As in the "if" part above, set $\hat{\sigma}(x) = \hat{q} + \varepsilon(h(x) - \ell)$, where $\varepsilon > 0$ again ensures that $\hat{\sigma}(x) \in [0, 1]$, and follow the same steps to establish that $\hat{\sigma}$ achieves FIE. □

It must be noted that while we cannot uniquely identify the strategies that aggregate information, all strategy profiles that aggregate information are characterized by a specific property: in pivotal states, they lead to a vote share exactly equal to the threshold necessary for the alternative A to win. If there are multiple pivotal states, the conditional distributions arising in such states must lie on a hyperplane on the simplex. As one moves in the simplex along the normal to that hyperplane, the vote share for A produced by any information-aggregating strategy changes monotonically. Later, we show that in environments that allow FIE, there also exists an equilibrium strategy profile with the same properties. McMurray (2014) finds a result with a similar flavor and interprets it as the endogenous emergence of a single dimension of political conflict in a multidimensional world..

3.2. **Some Special Environments.** In the previous section we have shown that in general environments, FIE can obtain only in rather special cases. Now, we turn to some specific environments that are of interest to us.

3.2.1. *State space isomorphic to the simplex over signals.* Suppose that the state space is convex and the conditional probability distribution is a linear function of the state. In other words, $\eta(\cdot|\alpha\theta + (1-\alpha)\theta') = \alpha\eta(\cdot|\theta) + (1-\alpha)\eta(\cdot|\theta')$ for all pairs θ, θ' and $\alpha \in [0,1]$. Additionally, assume that there exist some pivotal state(s). In this setting, the condition for existence of FIE strategies simply boils down to all probability distributions corresponding to pivotal states lying on the same hyperplane in the simplex. By the linearity assumption, this condition is equivalent to there being a hyperplane in the state space such that all pivotal states lie on that hyperplane.

A special case of the above is when X is finite, θ is continuous and $\Theta = \Delta(X)$. Additionally, for any θ, $\eta(\cdot|\theta) = \theta$. The situation considered in examples 2 and 3 (which comes later) are special instances of this case. In this situation, the voter preferences are simply defined over the probability distribution over signals rather than a different state variable. In other words, in a large electorate, the preferences depends only on the profile of signals, i.e., the distribution of information in the electorate. In this case, the condition for information aggregation is simply that the set of states in which a particular alternative is preferred is convex.

Corollary 3. *If Θ is a convex subset of a vector space, $\eta(\cdot|\alpha\theta + (1-\alpha)\theta') = \alpha\eta(\cdot|\theta) + (1-\alpha)\eta(\cdot|\theta')$ for all pairs θ, θ' and $\alpha \in [0,1]$, and $M^{piv} \neq \emptyset$, then an environment allows FIE if and only if there exists a hyperplane $H \in \Delta(X)$ such that $\eta(\cdot|\theta) \in H$ for $\theta \in M^{piv}$.*

Proof. There are two steps to this proof. First, we show that under the conditions, an environment allows FIE if and only if there exists a hyperplane H in $\Delta(X)$ such that $\eta(\cdot|\theta) \in H$ for $\theta \in M^{piv}$.

As in the proof above, we have $M^{piv} \subset \mathcal{I}^\sigma$ for the profile σ defined by $\sigma(x) = q + \varepsilon(h(x) - \ell)$ for all x. Say that there are $\theta^A \in \mathcal{A}$ and $\theta^B \in \mathcal{B}$ with θ^A, θ^B in $\mathcal{A}^\sigma$. There must exist $\alpha \in (0,1)$ and $\bar{\theta} \in M^{piv}$ with $\bar{\theta} = \alpha\theta^A + (1-\alpha)\theta^B$. By linearity of $\theta \mapsto \eta(\cdot|\theta)$, we would have $z^\sigma(\bar{\theta}) > q$, contradicting $M^{piv} \subset \mathcal{I}^\sigma$. Hence the sets $\mathcal{A}$ and $\mathcal{B}$ are each mapped to one of the halfspaces determined by the hyperplane associated with $\mathcal{I}^\sigma$. Changing signs if necessary, we have $M^A \subset \mathcal{A}^\sigma$ and $M^B \subset \mathcal{B}^\sigma$, and the rest follows as in the proof above. Observe that the existence of a hyperplane H in $\Delta(X)$ necessarily implies the existence of a hyperplane H' in Θ satisfying: $\theta \in H'$ for $\theta \in M^{piv}$. Indeed, by linearity of η, the mapping $\theta \mapsto \sum_{x \in X} h(x)\eta(x|\theta)$ defines the required hyperplane in Θ. □

3.2.2. *Monotone Likelihood Ratio Property.* Suppose that both signals and states have a natural order. A standard informativeness assumption on signals in this setting is the Monotone Likelihood Ratio Property (MLRP), which ensures that a signal is a "sufficient statistic" of the state (Milgrom 1981) in the sense that higher signals indicate higher states. Feddersen and Pesendorfer (1997) assumes strict MLRP condition on signals and shows (albeit in a model of diverse preferences) that information is aggregated in equilibrium. We obtain a sufficient condition for an environment to allow FIE which entertains MLRP as a specific case.

Definition 2 (Monotone Likelihood Ratio Property)**.** *Suppose* $\Theta = [0,1]$ *and* $X = \{x_1, ..., x_k\} \in [0,1]^k$*, with* $x_1 < x_2 < \cdots < x_k$*. The signals are said to satisfy strict MLRP if, for any two signals* $x < x'$*, the likelihood ratio* $\frac{\eta(x|\theta)}{\eta(x'|\theta)}$ *is a decreasing function of* θ*.*

We obtain a sufficient condition for the existence of a strategy that achieves FIE in this environment that is weaker than MLRP. Assume that the prior m is non-atomic and has full support over $[0,1]$. Moreover, suppose that for some $\theta^* \in (0,1)$, A is preferred for $\theta > \theta^*$ and B is preferred for $\theta < \theta^*$. In other words, $\mathcal{A} = (\theta^*, 1]$ and $\mathcal{B} = [0, \theta^*)$.

Let $F(x|\theta) = \sum_{x_j \leq x} \mu_j(x|\theta)$ denote the cumulative distribution function of $\eta(\cdot|\theta)$. Strict MLRP implies that for every x, the cumulative distribution $F(x|\cdot)$ is a decreasing function. Now consider the following property: For each $\theta^a \in \mathcal{A}$ and each $\theta^b \in \mathcal{B}$, we have for all $x \in X$

$$F(x|\theta^a) < F(x|\theta^{piv}) < F(x|\theta^b) \tag{3}$$

As long as the property (3) is satisfied, there exists a strategy that achieves FIE. To see that, let x^* be the smallest $x \in X$ such that $1 - F(x|\theta^*) \geq q$. Now, set $\sigma(x) = 0$ for $x \leq x^*$ and $\sigma(x) = 1$ for $x > x^*$. It is easy to verify that the strategy profile σ achieves FIE.[1]

Note that the property (3) is weaker than strict MLRP. While strict MLRP implies that $F(x|\cdot)$ is decreasing over the entire interval $[0,1]$, property (3) does not require $F(x|\cdot)$ to be decreasing within $\mathcal{A}$ or within $\mathcal{B}$.

3.2.3. *Finite State and Signal Spaces.* Another case that is of interest to us is when both the state and signals are finite. Suppose that there are r states and k signals. By our assumption that $m(\mathcal{I}) = 0$, the ranking over alternatives is strict in every state. In this setting, any environment allows FIE if the signal space is sufficiently rich vis-a-vis the state space. As long as there are at least as many signals as there are states, and none of the conditional probability vectors on the simplex can be expressed as an affine combination of the conditional probability distributions in the other states, there is a strategy that achieves FIE. As a particular case, when there are only two states, as long as the conditional probability distributions are not the same in the two states, there exists some strategy that achieves FIE. Notice that under this richness condition on signals, the utility function assigning states to rankings is immaterial for FIE.

Corollary 4. *Suppose there are* r *states and* k *signals, i.e.,* $\Theta = \{\theta_1, ..., \theta_r\}$ *and* $X = \{x_1, ..., x_k\}$ *with* $r \leq k$*. Moreover, assume that there exists no state* $\theta_t \in \Theta$ *such that* $\eta(\cdot|\theta)$ *is an affine combination of members of the set* $\{\eta(\cdot|\theta_k)\}_{k \neq t}$*. In such an environment, there exists a strategy that achieves FIE.*

Proof. Consider any partition of Θ into two nonempty sets $\mathcal{A}$ and $\mathcal{B}$. Denote by A_μ the set $\{\mu \in \Delta(X) : \theta \in \mathcal{A} \text{ and } \eta(\cdot|\theta) = \mu\}$, i.e., the set of conditional probability vectors arising in states in $\mathcal{A}$. Similarly, denote by B_μ the set $\{\mu \in \Delta(X) : \theta \in \mathcal{B} \text{ and } \eta(\cdot|\theta) = \mu\}$. Denote the

[1] When the state space is not $[0,1]$, a sufficient condition for obtaining an FIE strategy with signals $x_1 < ... < x_k$ is that (1) each $\theta^{piv} \in M^{piv}$ leads to the same cumulative distribution $F(x|\theta^{piv})$ for all $x \in X$, and (2) for any $\theta^a \in \mathcal{A}$ and $\theta^b \in \mathcal{B}$, we obtain $F(x|\theta^a) < F(x|\theta^{piv}) < F(x|\theta^b)$ for all $x \in X$

respective convex hulls by coA_μ and coB_μ. If coA_μ and coB_μ are disjoint, there exists some hyperplane H in $\Delta(X)$ that separates A_μ and B_μ, which is sufficient for the existence of a strategy that achieves FIE from our main Theorem and the subsequent remark. Suppose now that coA_μ and coB_μ are not disjoint. Then there must exist convex weights (α_i) and (β_j) associated only with elements of $\mathcal{A}$ and $\mathcal{B}$, respectively, such that $\sum_i \alpha_i \eta(\cdot|\theta_i) = \sum_j \beta_j \eta(\cdot|\theta_j)$. Without loss of generality, let α_1 be the largest of such weights. Then

$$\eta(\cdot|\theta_1) = \frac{1}{\alpha_1}\left(\sum_j \beta_j \eta(\cdot|\theta_j) - \sum_{i\neq 1} \alpha_i \eta(\cdot|\theta_i)\right)$$

is an affine combination of members of the set $\{\eta(\cdot|\theta_k)\}_{k\neq 1}$, contrary to our assumption. Hence, the intersection must be empty. □

Observe that the result above is not true if we weakened the requirement from affine to convex independence. Consider the following example with $r = k = 4$ to illustrate. The conditionals $\eta(\cdot|\theta_1) = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, $\eta(\cdot|\theta_2) = (\frac{1}{2}, 0, 0, \frac{1}{2})$, $\eta(\cdot|\theta_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$, and $\eta(\cdot|\theta_4) = (0, 0, 0, 1)$ satisfy convex independence; assume that $\mathcal{A} = \{\theta_1, \theta_2\}$ and $\mathcal{B} = \{\theta_3, \theta_4\}$, so that $A_\mu = \{\eta(\cdot|\theta_1), \eta(\cdot|\theta_2)\}$ and $B_\mu = \{\eta(\cdot|\theta_3), \eta(\cdot|\theta_4)\}$. It is clear that these sets cannot be separated by a hyperplane. In fact, the vector $(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5})$ is equal to $\frac{3}{5}\eta(\cdot|\theta_1) + \frac{2}{5}\eta(\cdot|\theta_2)$ and also to $\frac{3}{5}\eta(\cdot|\theta_3) + \frac{2}{5}\eta(\cdot|\theta_4)$, so it lies in co$A_\mu \cap$ coB_μ, and we cannot separate these two convex sets.

As far as genericity is concerned, affine independence is already quite weak: in general, when $r \leq k$, the vectors $\{\eta(\cdot|\theta_i)\}_{i=1}^r$ will be linearly (and hence affinely) independent, so the typical finite case is that FIE is possible regardless of the specification of preferences. Thus, we expect FIE to obtain generically when the state and signal spaces are discrete, and the signal space is sufficiently rich compared to the state space.

3.2.4. *Examples - rich state space.* The previous section suggests that FIE obtains generically in case where the signals are sufficiently numerous compared to the states. In particular, as long as a mild condition is satisfied, whether an environment allows FIE does not depend at all on the details of the preference. On the other hand, the following two example illustrate that when, the state space is richer than the signal space, FIE may obtain only in special environments. In particular, given the information structure, information is aggregated only for specific utility functions.

Example 3: Aggregation of private opinion

Suppose that a policy proposal (alternative A) is being voted on against a status quo (alternative B). The signal $x \in \{1, 0, -1\}$ stands for whether a voter's private assessment of the proposal's merits over the status quo. A voter with $x = 1$ has private information that favors the proposal, one with $x = -1$ has private information that is against the proposal and one with $x = 0$ is indifferent between the two. Assume that the state space is isomorphic to the simplex over signals, i.e., preferences are defined directly over the simplex. Denote

$\mu = (\mu_1, \mu_0, \mu_{-1}) \in \Delta(X)$ as a generic state, with μ_x being the proportion of signal x in the population. Suppose $v(\mu)$ denotes the utility difference between A and B for $\mu \in \Delta(X)$, and it is continuous. Our interpretation of signals imposes the following restriction on the utility function v: (i) $v(\mu_1, \mu_{-1})$ increases in μ_1 and decreases in μ_2. Therefore, along the locus of indifference, as the proportion of signal $x = 1$ increases, the proportion of signal $x = -1$ must also increase, i.e., $\left.\frac{d\mu_1}{d\mu_{-1}}\right|_{v(\mu)=0} > 0$. However, the content of Theorem 1 as applied to this setting, is the following: in environments that satisfy FIE, all indifferent states must lie on a hyperplane. In other words, there must be some $k > 0$ such that $\left.\frac{d\mu_1}{d\mu_{-1}}\right|_{v(\mu)=0} = k$. Thus, FIE is satisfied only for very specific preferences in this setting.

Example 4: Spatial Model

Our final example is an application to a very standard case of spatial model of political competition between two alternatives. We continue the metaphor of a policy proposal (alternative A) being voted on against a status quo (alternative B). There is a policy space $Y = [0, 1]^2$, in which both alternatives are located. Voter utility for policy y is given by $u(|y - y^*|)$, $u' < 0$. Thus, $y^* \in Y$ is the voter ideal policy and voters prefer policies closer to y^* than further from it. This type of utility functions is very commonly assumed in the applied literature. The status quo is known to be located at $y_B \neq y^*$ on the policy space. On the other hand, there is uncertainty about the location of the proposed policy: we denote the location of the proposed alternative A on the policy space by $\theta = (\theta_1, \theta_2) \in [0, 1]^2$ Now, the voters' preference over the policy proposal vs the status quo depends on whether the state is closer to y^* than is the status quo y_B, i.e.,

$$\begin{aligned}
\mathcal{A} &= \{\theta : |\theta - y^*| < |y_B - y^*|\\
\mathcal{B} &= \{\theta : |\theta - y^*| > |y_B - y^*|\\
\mathcal{I} &= \{\theta : |\theta - y^*| = |y_B - y^*|
\end{aligned} \tag{4}$$

Notice that $\mathcal{I}$ is denoted by the circumference of a circle (or the portion of the circumference that intersects with the state space). This is entirely a property of the utility function.

Now assume that the signal $x = (x_1, x_2)$ is also two dimensional, and uppose that for $i = 1, 2$, $x_i \in \{0, 1\}$, with $\Pr(x_i = 1|\theta_i) = \theta_i$. Thus, x_1 provides information on θ_1 and x_2 on θ_2 independently of each other. Alternatively, the state θ_i can simply be thought of as the proportion of 1-signals in dimension i. Thus, the space of signals is $X = \{(0, 0),\ (0, 1),\ (1, 0),\ (1, 1)\}$ and $\eta(\cdot|\cdot)$ is given by

$$\begin{aligned}
\eta((1,1)|\theta) &= \theta_1\theta_2\\
\eta((1,0)|\theta) &= \theta_1(1-\theta_2)\\
\eta((0,1)|\theta) &= (1-\theta_1)\theta_2\\
\eta((0,0)|\theta) &= (1-\theta_1)(1-\theta_2)
\end{aligned}$$

Theorem 1 dictates that the necessary condition for FIE in this setting is that the conditionals arising from the indifferent states $\theta \in \mathcal{I}$ should lie on a hyperplane on $\Delta(X)$. This implies that there should be numbers $a',\ b',\ c',\ d'$ and e' such that for all states $\theta \in \mathcal{I}$,

$$a'\theta_1\theta_2 + b'\theta_1(1-\theta_2) + c'(1-\theta_1)\theta_2 + d'(1-\theta_1)(1-\theta_2) = e'$$

which simplifies to the following statement: the set of indifferent states should be described by four numbers $a,\ b,\ c$ and d, not all zero such that

$$a\theta_1 + b\theta_2 + c\theta_1\theta_2 = d \tag{5}$$

There can be no vector of numbers (a, b, c, d) for which equation (5) describes the circumference of a circle, as required by the utility function (specified in equation (4)). Therefore, it is impossible to obtain FIE in this simple and commonly used setting. This example, along with example 3, shows that the requitement of FIE imposes strong conditions on preferences when the state space is sufficiently rich compared to the signal space.

4. Equilibrium analysis

From the previous section, it is clear that only special environments allow full information equivalence in the sense that there exist strategies that achieve FIE. However, it is not clear whether, even in such environments, voters have an incentive to use such strategies. In order to check whether voters find it in their interest to use such strategies, we consider voting as a game played in such environments. A game is defined as an environment (u, θ, X, η, q) along with a number of players n. We fix an environment and consider a sequence of games by letting the number of voters grow. Following the logic in McLennan (1998), we show that under common preferences, any environment that allows FIE also has a sequence of Nash equilibrium profiles that achieves FIE.

First, we define the game derived from the environment (u, θ, X, η, q) along with a number of players n more formally. We provide the analysis for the case where θ is a continuous random variable, the case where θ is discrete is exactly analogous.

It will be necessary to distinguish finite electorates, so let us use the notation $x^n = (x_1, ..., x_n)$, $s^n = (s_1, ..., s_n)$, and $\sigma^n = (\sigma_1, ..., \sigma_n)$ for profiles of signals and strategies in a finite electorate $\{1, ..., n\}$. Given η on $\Theta \times X$, construct ν on $\Theta \times X^n$ as $d\nu = dm \otimes (\otimes_{i=1}^n d\eta(\cdot|\theta))$ and let η_x be the marginal of ν on X^n. Denote by $\eta(\cdot|x^n)$ the conditional of ν on Θ given a profile x^n.[2] For a given x^n, let

$$u^n(a, x^n) = \int_\Theta u(a, \theta)\eta(d\theta|x^n),$$

for $a \in \{A, B\}$. Given a profile of pure strategies s^n, let $u(s^n(x^n), x^n)$ denote the utility at profile x^n at the outcome induced by $s^n(x^n)$ (it will be A (resp. B) if $\frac{1}{n}\sum_{1=1}^n s_i(x_i) > q$ (resp. $\frac{1}{n}\sum_{1=1}^n s_i(x_i) < q$) with ties broken by a coin flip.) The multilinear extension at a

[2] The construction is presented in general form so that it also covers the case of uncountable X that we will deal later.

profile of behavioral strategies is denoted $u(\sigma^n(x^n), x^n)$. Finally, the ex ante expected utility for a given profile σ^n is

$$u^n(\sigma^n) = \sum_{x^n \in X^n} u(\sigma^n(x^n), x^n)\eta_x(x^n).$$

The Bayesian game G^n played by the n voters is the game where each voter has the same space of behavioral strategies, $\Sigma_i = \{\sigma_i : X \to [0,1]\}$, endowed with the narrow topology that makes it a compact, convex LCTVS, and the payoffs are the ones we just derived.

Suppose that σ_n^* is a maximizer of $u^n(\sigma^n)$. The existence of such a maximizer follows from compactness of the domain and continuity of u on σ^n.[3] Following McLennan (1998), σ_n^* is a Bayesian Nash equilibrium of the game G^n. It is straightforward to restrict to profiles of symmetric strategies and ensure existence of a symmetric BNE. The next theorem tells us that the sequence σ_n^* achieves FIE as long as the environment (u, θ, X, η, q) allows FIE.

Theorem 2. *If the environment (u, θ, X, η, q) allows FIE, there exists a sequence σ^n of Nash equilibria of the game G^n that achieves FIE., i.e., $W_n^{\sigma^n} \to 0$.*

Proof. Observe that $u(\sigma^n)$ can be written as

$$\int_\Theta \sum_a u(a,\theta) \sum_{x^n \in X^n} \varphi_a^{\sigma^n}(x^n)\eta(x^n|\theta)m(d\theta)$$

where $a \in \{A, B, D\}$, "D" standing for a draw (and the associated coin flip to decide the winner), and $\varphi_a^{\sigma^n}(x^n)$ is the probability that a is the outcome of the election at the profile x^n. For each size n of electorate, consider a symmetric profile of strategies $\sigma^n = (\sigma, ..., \sigma)$, so that $\sigma^\infty = (\sigma, \sigma, ...)$. For each $\theta \in \Theta$, the proportion of votes for A converges to $z^\sigma(\theta)$ $\eta(\cdot|\theta)$-almost surely as $n \to \infty$. Hence $\sum_{x^n \in X^n} \varphi_a^{\sigma^n}(x^n)\eta(x^n|\theta)$ converges, so Lebesgue Dominated Convergence implies that $u(\sigma^\infty) = \lim_{n\to\infty} u(\sigma^n)$ is well defined.

Observe that if the symmetric profile $\hat{\sigma}^\infty$ achieves FIE, then $u(\hat{\sigma}^\infty)$ is the maximum attainable value: for m-almost every $\theta \in \mathcal{A}$, A wins, and for m-almost every $\theta \in \mathcal{B}$, B wins. States in $\mathcal{I}$ are irrelevant for the evaluation above because they are of m-measure zero. So, given that $u(\sigma^n)$ is linear in $u(a,\theta)$, the claim is verified.

For each finite electorate $\{1, ..., n\}$, choose σ^n as a maximizer of $u(\sigma^n)$. We know such profile is an equilibrium of the corresponding game G^n. We also know that $u(\hat{\sigma}^\infty)$ is the maximum feasible value of the ex ante utility. Hence

$$u(\hat{\sigma}^\infty) \geq u(\sigma^\infty) = \lim_n u(\sigma^n) \geq \lim_n u(\hat{\sigma}^n) = u(\hat{\sigma}^\infty),$$

establishing the result. In fact, if $W_n^{\sigma^n}$ were not to converge to zero, then we would have to have, say, $m(\mathcal{A} \backslash \mathcal{A}^{\sigma^\infty}) > 0$. That is, a set of positive measure in $\mathcal{A}$ where B wins under σ^∞, whereas we know that no such set exists for $\hat{\sigma}^\infty$. But then $u(\hat{\sigma}^\infty) > u(\sigma^\infty)$, contradicting what we just established. □

[3]By construction, for each given θ, the probability measure on profiles is the product $\otimes_{i=1}^n \eta(\cdot|\theta)$; hence the integral of the utility function is a continuous function of the profile σ^n; integrating out θ then recovers the ex ante utility, which must then be a continuous function of σ^n.

5. Extensions

In this section, we study three extensions to the setup in which the main results are presented. First, we consider the case of more than two alternatives. In this case, while the basic ideas go through, the convexity of sets of conditional probability vectors for which a given alternative is top ranked is necessary for FIE. However, we need some extra conditions on preferences to ensure that FIE holds. Second, we revert back to the two-alternative case and show that all our results go through if we consider a signal space X that is uncountable rather than finite or uncountable. Finally, we study a relaxation of the common preference assumption. In this setting, we obtain a generalization of Theorem 1 for two alternatives with diverse preference.

5.1. **Multiple alternatives.** Let $\mathcal{A} = \{1, 2, ..., n\}$ be the set of alternatives, and Θ be the space of states. The utility function of each voter is $u : \mathcal{A} \times \Theta \to \mathbb{R}$. Let L be the number of signals, and let $\eta \in \Delta(\{x_1, x_2, ...x_L\} \times \Theta)$ denote the information structure. It is convenient to work directly with preferences over the distributions of signals, so let us assume that $\theta \mapsto \eta(\cdot|\theta)$ is one-to one, so that, for $i \in \mathcal{A}$ and $\mu \in \Delta(\{x_1, x_2, ...x_L\})$, the utility function $u(i, \mu) = u(i, \theta)$, where $\eta(\cdot|\theta) = \mu$, is well-defined. We assume the utility function to be continuous. For ease of notation, let us use $\Delta \equiv \Delta(\{x_1, x_2, ...x_L\})$. Assume that m has full support over Δ and is non-atomic. Whichever alternative gets the most votes wins. While we do not consider abstention in this paper, none of our arguments change if we allow abstention. In the case of multiple alternatives, a strategy σ is a set of n functions $\sigma_i : X \to [0, 1]$, $i = 1, 2, ...n$, where $\sigma_i(x_l)$ is the probability of voting for alternative i on obtaining signal x_l. Therefore, the expected vote share for alternative i at state μ is simply $\sigma_i \cdot \mu$.

We start with a simple case of a linear utility function.

Remark 1. *Suppose that the utility function $u(i, \mu)$ is linear in μ, that is, $u(i, \mu) = u_i \cdot \mu$, for all $i \in \mathcal{A}$ and $\mu \in \Delta$, where u_i is a vector in $\mathbb{R}^L$. There exists a strategy with the FIE property.*

Proof. Without loss of generality, assume that u_i are nonnegative, and $\sum_{\ell=1}^{L} u_{i\ell} = 1$, that is, the vector u_i belongs to Δ. (First, we add a positive constant d_i to each coordinate of u_i so that $u_i \cdot \mathbf{e} + Ld_i = M > 0$. Finally, we divide each coordinate $u_{i\ell} + d_i$ by M). Let σ be the matrix $\{u_1, u_2, ..., u_n\}$ with typical element $\sigma_{i\ell}$ denoting the probability of voting for alternative i when observing signal ℓ. As the number of votes obtained by alternative i is equal to $\sigma_i \cdot \mu$, the alternative i that generates the greatest payoff also generates the greatest number of votes under the symmetric strategy where every voter plays according to σ. $\square$

Moreover, consider any ordinal preference over $\mu \in \Delta$ that admits a strategy σ which delivers FIE. As a corollary to the above remark, there exists a utility function, linear in μ, which agrees with the ordinal preference for the top-ranked alternative in each state x. To see this, we just need to let $u_i(\mu) = \sigma_i(\mu)$ for every alternative i and every state $\mu \in \Delta$.

Next, we turn to the general question of conditions on preference over distributions of signals for which FIE obtains. Denote $\mathcal{A}_i$ as the set of states μ such that i is the best alternative, i.e.

$$\mathcal{A}_i \equiv \{\mu \in \Delta : u(i,\mu) \geq u(j,\mu) \text{ for all } j \neq i\}$$

Remark 2. *Assume that for any two states i, j, $m(\mathcal{A}_i \cap \mathcal{A}_j) = 0$. A necessary condition for FIE is that $\mathcal{A}_i$ is convex for all $i = 1, 2, ..n$.*

Proof. Denote by σ that strategy that produces FIE. If i is preferred over other alternatives for every state, then $\mathcal{A}_i$ is the simplex and we are done. Suppose from now on there are at least two alternatives such each is strictly preferred over every other altermative for some positive measure of states. Now, consider the case of two alternatives i and j such that $\mathcal{A}_i \cap \mathcal{A}_j \neq \phi$ and denote $\mathcal{A}_i \cap \mathcal{A}_j = M_{ij}$. By assumption, $m(M_{ij}) = 0$. By continuity, we must have $(\sigma_i. - \sigma_j) \cdot \mu = 0$ for all $\mu \in M_{ij}$. Since $(\sigma_i. - \sigma_j) \cdot \mu = 0$ describes a hyperplane on the simplex, M_{ij} must lie on a hyperplane. Thus, $\mathcal{A}_i$ and $\mathcal{A}_j$ must be separated by a hyperplane. □

From the previous section, we also know that convexity is sufficient for FIE when $n = 2$. To see that convexity may not be sufficient for FIE when there are more than two alternatives, suppose that $\mathcal{A}_i$ is convex for all i, and thus M_{ij} as defined above lies on the hyperplane H_{ij}. Denote by the vector h_{ij} the norm of H_{ij}. For σ to produce FIE, $\sigma_i - \sigma_j$ must be a scalar multiple of h_{ij} for every pair (i, j). This implies that potentially, there are ${}_nC_2$ equations with n unknowns, and it is not obvious that the system has a solution. However, under an extra condition in addition to convexity, we can guarantee FIE in the general case. In order to state it, it is necessary to first extend the domain of u to $\mathcal{A} \times \Sigma$, where $\Sigma = \{\mu \in \mathbb{R}^L : \mu \cdot \mathbf{e} = 1\}$. In particular, we assume

Assumption E There exists $\hat{u} : \mathcal{A} \times \Sigma \to \mathbb{R}$ such that $u : \mathcal{A} \times \Delta \to \mathbb{R}$ is the restriction of $\hat{u}$ to $\mathcal{A} \times \Delta$.

Let

$$H_{ij}^+ := \{x \in \Sigma : \hat{u}(i,x) \geq \hat{u}(j,x)\}$$
$$H_{ij} := H_{ij}^+ \cap H_{ji}^+$$

Under Assumption E, we can state our main condition, which requires that the indifference between any pair of alternatives can be represented by a hyperplane.

Assumption H For all $i, j \in \mathcal{A}$, there exists $h_{ij} \in \mathbb{R}$ such that $H_{ij}^+ = \{x \in \Sigma : x \cdot h_{ij} \geq 0\}$.

Assumption H strengthens convexity of $\mathcal{A}_i$ in the following sense. Convexity of $\mathcal{A}_i$ imposes conditions only on the most preferred alternative in each state. Assumption H requires that for each pair of alternatives i and j, the simplex be separated into two convex sets - one where i is preferred and the other where j is preferred.

We shall refer to the vector h_{ij} as the normal to the hyperplane H_{ij} denoting the region of indifference between alternatives i and j. It is evident that $-h_{ij} = h_{ji}$ by construction so the order of the subindexes matter. Notice that under assumption H, we can simply describe any given preference in terms of the set of normals $< h_{ij} >,\ i \neq j$ induced by it.

Before stating the main theorem, we provide an important Lemma that says that, for any preference, these normals are linearly dependent. The Lemma follows simply from transitivity of preference and its content is as follows: For any three alternatives i, j, k, if $\mu \in H_{ij} \cap H_{jk}$, then $\mu \in H_{ik}$. In words, for every three alternatives, the three hyperplanes that describe indifference between each pair are either parallel to each other or have a common intersection. There is no state that belongs to two of these hyperplanes but not on the third one. In absence of assumption E, the hyperplanes could intersect pairwise outside the simplex. Thus, assumption E simply imposes a restriction on the preferences over distributions within the simplex, i.e., on preferences described by $u(\cdot)$.

Lemma 1. *If assumptions* ***E*** *and* ***H*** *hold, for all* $i, j, k \in \mathcal{A}$*, there exists positive constants* α_{ij}, α_{jk}*, and* α_{ik} *such that*

$$\alpha_{ij} h_{ij} + \alpha_{jk} h_{jk} = \alpha_{ik} h_{ik}. \tag{6}$$

Proof. Let $M = H_{ij} \cap H_{jk}$. Suppose M is non-empty. By assumption on preferences, if $\mu \in M$ then $\mu \in H_{ik}$ (this is true because $\hat{u}(i, \mu) = \hat{u}(j, \mu) = \hat{u}(k, \mu)$). By assumption **H**, $\hat{X}$ is a vector space of dimension $k - 2$. It's orthogonal complement therefore has dimension 2. In this orthogonal complement, h_{ij} and h_{jk} lie on a two dimensional linear space, they are independent (because their corresponding hyperplanes intersect), and therefore, they span the space. Since h_{ik} also lies in this space it must be true that there exists constants α_{ij}, α_{jk} such that $h_{ik} = \alpha_{ij} h_{ij} + \alpha_{jk} h_{jk}$. Suppose instead $\hat{X}$ is empty. This is true whenever H_{ij} and H_{jk} are parallel. It is clear that there is always constants satisfying equation (6). Therefore, we establish that whether $\hat{X}$ is empty or not, there will always exist constants satisfying equation (6).

Next we show by contradiction that the constants need to be positive.

Case 1. Suppose $\alpha_{ij} < 0$, and $\alpha_{jk} < 0$. There are two possible scenarios: either there exists μ such that $\mu \cdot h_{ij} > 0$ and $\mu \cdot h_{jk} > 0$, or there exists μ such that $\mu \cdot h_{ij} < 0$ and $\mu \cdot h_{jk} < 0$. Otherwise, one of the alternatives is never the best choice. We will focus on the former and note that the analogous argument holds inverting the inequalities in the following argument. Thus, $\hat{u}(i, \mu) > \hat{u}(j, \mu)$ and $\hat{u}(j, \mu) > \hat{u}(k, \mu)$. By transitivity, $\hat{u}(i, \mu) > \hat{u}(k, \mu)$. On the other end, $\mu \cdot (\alpha_{ij} h_{ij} + \alpha_{jk} h_{jk}) < 0$. Thus $\mu \cdot h_{ik} < 0$, and thus, $\hat{u}(i, \mu) < \hat{u}(k, \mu)$, a contradiction.

Case 2. Suppose $\alpha_{ij} > 0$, and $\alpha_{jk} < 0$. Consider μ such that $\mu \cdot h_{ik} = 0$ and $\mu \cdot h_{ij} \neq 0$. $\mu \cdot h_{ik} = 0$ implies $\mu \cdot (\alpha_{ij} h_{ij} + \alpha_{jk} h_{jk}) = 0$. Then $\alpha_{ij} \mu \cdot h_{ij} = -\alpha_{jk} \mu \cdot h_{jk}$. There are two possibilities: (i) $\mu \cdot h_{ij} > 0$. This implies $\mu \cdot h_{jk} > 0$. Then $\hat{u}(i, \mu) > \hat{u}(j, \mu) > \hat{u}(k, \mu)$. (ii) $\mu \cdot h_{ij} < 0$. This implies $\mu \cdot h_{jk} < 0$. Then $\hat{u}(i, \mu) < \hat{u}(j, \mu) < \hat{u}(k, \mu)$. In either case this is a contradiction since $\mu \cdot h_{ik} = 0$ implies that $\hat{u}(i, \mu) = \hat{u}(k, \mu)$.

Case 3. Suppose $\alpha_{ij} < 0$, and $\alpha_{jk} > 0$. This is symmetric to the case 2.

Finally, let $\alpha_{ik} = 1$ and the result follows. □

Now, we are ready to state and prove the main theorem.

Theorem 3. *If Assumptions **E** and **H** hold, then there exists a strategy that achieves FIE.*

Proof. Consider the system,

$$\alpha_{ij}h_{ij} + \alpha_{jk}h_{jk} = \alpha_{ik}h_{ik} \text{ for all } i, j, k \in \mathcal{A} \tag{7}$$

From lemma 1, we know that every single equation is well defined. Since every choice of 3 different alternatives generate a different equation, the number of equations in the system (7) is given by ${}_nC_3$. The norms of the hyperplanes h's are parameters given by the agents preferences.

We will show that this system of equations has a non trivial solution (infinite to be more precise) for the variables α's. Furthermore, if the solution is non trivial, it must be true that all α's are strictly positive. Why? If there were to be some $\alpha_{ij} = 0$, this would force $h_{jk} = ch_{ik}$, for some constant c. However this implies that $H_{jk} = H_{ik} \neq H_{ij}$. This is not possible because the first equality implies that there exist some μ such that $\hat{u}(i, \mu) = \hat{u}(j, \mu) = \hat{u}(k, \mu)$ but the second inequality implies that $\hat{u}(i, \mu) \neq \hat{u}(j, \mu)$.

The number of variables α's is ${}_nC_2$. For $n < 6$, ${}_nC_2 > {}_nC_3$ and therefore the system has a non trivial solution. However, for $n \geq 6$, ${}_nC_2 < {}_nC_3$ so there are more equations than unknowns. We need to show that there are sufficiently many linearly dependent equations so that the system has a solution.

Consider the subsystem of equations in which we fix an alternative, that without loss of generality we will call alternative 1, and we combine with all the other possible combinations of the remaining two alternatives. This is the set of equations containing all equations in which alternative 1 is present. The number of equations in this subsystem is given by ${}_{(n-1)}C_2 < {}_nC_2$ and contains all α's, and therefore it has a non trivial solution. It only remains to show that any equation in the system given by (7) can be generated using this subsystem. For simplicity of exposition, and without loss of generality, consider an equation with alternatives $(2, 3, 4)$:

$$\alpha_{23}h_{23} + \alpha_{34}h_{34} = \alpha_{24}h_{24} \tag{8}$$

This equation will not be contained in our subsystem because alternative 1 is not present. Consider the following three equations from our subsystem:

$$\alpha_{12}h_{12} + \alpha_{23}h_{23} = \alpha_{13}h_{13} \tag{9}$$

$$\alpha_{12}h_{12} + \alpha_{24}h_{24} = \alpha_{14}h_{14} \tag{10}$$

$$\alpha_{13}h_{13} + \alpha_{34}h_{24} = \alpha_{14}h_{14} \tag{11}$$

We do the following operation: equations (9) minus equation (10) plus equation (11) and we note that this is equal to equation (8). Since the choice of alternatives is without loss of generality we convince ourselves that our subsystem generates the full system.

Let $\Upsilon = \{h_{1j}\}_{j>1}$ with $n-1$ elements. For all $h_{1j} \in \Upsilon$ let τ_1, and τ_j be such that $\tau_1 - \tau_j = \alpha_{1j}h_{1j}$. There are n variables τ's and $n-1$ equations so this system has a solution.

Finally, for all i, choose $\sigma_i = (\tau_i + \delta)\epsilon$, where δ is a sufficiently large constant such that $\tau_i + \delta \geq 0$, and $\epsilon = (\mathbf{e} \cdot \sum_i (\tau_i + \delta))^{-1}$. Then, $\sigma_i \geq 0$, and $\sum_i \sigma_i = \mathbf{e}$, and therefore constituting a well defined symmetric mixed strategy profile without the need of abstention.

To show that this strategy aggregates information we need to show that if alternative i is the best, then the strategy selects alternative i over j, for any $j \in \mathcal{A}$.

Consider the simplest case where alternative 1 is the best alternative. By construction, for all $j \neq 1$, $u(1,\mu) > u(j,\mu) \iff \mu \cdot h_{1j} > 0 \iff \mu \cdot (\tau_1 - \tau_j) > 0 \iff \mu \cdot (\sigma_1 - \sigma_j) > 0$. Then, alternative 1 obtains more votes than alternative j. The same relationship holds with weak inequality and equality.

Consider the case where alternative $i \neq 1$ is the best alternative. Then, $u(i,\mu) > u(j,\mu) \iff \mu \cdot h_{ij} > 0 \iff \mu \cdot (\alpha_{1j} h_{1j} - \alpha_{1i} h_{1i}) > 0 \iff \mu \cdot (\tau_1 - \tau_j - \tau_1 + \tau_i) = \mu \cdot (\tau_i - \tau_j) \iff \mu \cdot (\sigma_i - \sigma_j) > 0$.

Therefore, the strategy chooses the right candidate always in the limit and this concludes the proof. □

It is straightforward to generalize Theorem 2 to the case of multiple alternatives, so we skip it.

5.2. **Uncountable signal space.** The analysis above focused on the case that X was either finite or countable. In other words, voters have discrete pieces of information. On the other hand, we may also be interested in situations where the signal space is continuous, i.e., it is very unlikely that any two voters have the same private information. We show that while we have to use a different set of tools for this case, all our results in the previous section go through.

The setting where the signal space is uncountable has some technical disadvantages. In particular, we can no longer use the useful property that, for a given behavioral strategy $\sigma : X \to [0,1]$ and distribution $\eta(\cdot|\theta) \in \Delta(X)$, the asymptotic relative frequency of votes for A is equal to the expectation of σ with respect to $\eta(\cdot|\theta)$, for almost all sample paths. In the uncountable case, the property certainly goes through with pure strategies. Therefore, we use purification ideas to tackle this case. To distinguish between pure and behavioral strategy profiles, we will use notations $\sigma(\cdot)$ and $s(\cdot)$ respectively. Unless otherwise stated, when we refer to a pure strategy profile $s(\cdot)$, we will imply that all voters use the pure strategy $s(\cdot)$. Also, a hyperplane in $\Delta(X)$ now is defined as $H = \{\mu \in \Delta(X) : \int h(x)\mu(dx) = \ell\}$.

We make the following regularity assumption in addition to the maintained assumptions above.

A1. The conditional distribution $\eta(\cdot|\theta)$ is absolutely continuous with respect to a fixed non-atomic measure $\mu \in \Delta(X)$, with strictly positive density, for every $\theta \in \Theta$.

Next, we define an approximate measure of FIE. Recall that the expected likelihood of a wrong outcome is denoted by W_n^σ, given by equation (1). We say that a strategy profile

σ achieves ε-FIE if there is some $\varepsilon > 0$ such that $W_n^\sigma < \varepsilon$ for all n sufficiently large. This definition is equivalent to

$$m(\mathcal{A}\backslash\mathcal{A}^\sigma) = m(\mathcal{B}\backslash\mathcal{B}^\sigma) < \varepsilon.$$

We show that if there is some behavioral strategy σ that satisfies the hyperplane condition, then, for any $\varepsilon > 0$, there is some profile of pure strategies that achieves ε-FIE. The following proposition extends Theorem 1 in the context of the infinite signal space.

Proposition 1. *Consider an environment (u, Θ, X, η, q) where the signal space X is uncountable, and suppose that assumption A1 holds. For each $\varepsilon > 0$, there exists a profile of pure strategies s^ε that achieves ε-FIE if there is a hyperplane H in $\Delta(X)$ such that $\eta(\cdot|\theta) \in H^+$ for $\theta \in \mathcal{A}$, $\eta(\cdot|\theta) \in H^-$ for $\theta \in \mathcal{B}$, and, if $M^{piv} \neq \emptyset$, $\eta(\cdot|\theta) \in H$ for $\theta \in M^{piv}$. If in addition Θ is a finite set, then there exists a profile of pure strategies s that achieves FIE if the above condition is satisfied. If the above condition fails, then there exists no strategy profile that achieves FIE.*

Proof. A straightforward adaptation of Theorem 1 shows that the existence of a hyperplane H satisfying the conditions listed in the statement of the Proposition are equivalent to the existence of a behavioral strategy σ such that $m(\mathcal{A}\backslash\mathcal{A}^\sigma) = m(\mathcal{B}\backslash\mathcal{B}^\sigma) = 0$. Consider first the case of Θ finite. Applying Lyapunov's theorem establishes the existence of a function $g : X \to \{0, 1\}$ such that

$$\int \sigma(x)\eta(dx|\theta) = \int g(x)\eta(dx|\theta)$$

for every $\theta \in \Theta$. Hence, setting $s(x) = g(x)$, we are able to replace σ with a pure strategy s with the property that $z^s(\theta) = z^\sigma(\theta)$ for every $\theta \in \Theta$. By the SLLN, $\frac{1}{n}\sum_{i=1}^n s \to z^s(\theta)$, $\eta(\cdot|\theta)$-a.e. Hence $m(\mathcal{A}\backslash\mathcal{A}^s) = m(\mathcal{B}\backslash\mathcal{B}^s) = 0$ establishes that s achieves FIE.

The case that Θ is infinite is complicated by the failure of Lyapunov's theorem in infinite dimensions. We will resort to an approximation result to bypass this issue. First, let Σ denote the sigma-algebra of μ-measurable sets of X. Let $Y = \{\eta(E|\cdot)\}_{E\in\Sigma}$, a subset of the separable Banach space space $C(\Theta, [0, 1])$ of continuous functions endowed with the sup norm. Realizing Σ as a complete metric space on its own as a subspace of $L_1(\mu)$ (Aliprantis and Border (2006), Lemma 13.13), Y is also complete: a Cauchy sequence $(\eta(E_n|\cdot))_n$ in Y induces a Cauchy sequence $(E_n)_n$ in Σ, which must converge. In fact, for each $\varepsilon > 0$ and all θ, we have $|\eta(E_n|\theta) - \eta(E_m|\theta)| < \varepsilon$ for n, m large, so $|\mu(E_n) - \mu(E_m)| < \varepsilon$ as well, as the densities are positive. This means that there exist E such that $E_n \to E$. Now take a convergent sequence in Y, so that $\eta(E_n|\cdot)$ is converges uniformly. As it also converges pointwise to $\eta(E|\cdot)$, it must be that $\eta(E_n|\cdot)$ converges to $\eta(E|\cdot)$ uniformly. That is, Y is closed subset of $C(\Theta, [0, 1])$, and hence a separable Banach space with the subspace metric.

Now, by the extension of Lyapunov's theorem to the infinite dimensional case (Aubin and Frankowska (1989), Theorem 8.7.4), the closure of the set $\{\eta(\cdot|\theta)\}_{\theta\in\Theta}$ is convex and compact. Hence, for a given FIE σ, there exists a sequence $g^n : X \to \{0, 1\}$ such that

$$\sup_{\theta\in\Theta}\left|\int g^n(x)\eta(dx|\theta) - \int \sigma(x)\eta(dx|\theta)\right| \to 0.$$

For a given $\delta > 0$, let $V_\delta = \{\theta : z^\sigma(\theta) \in [q-\delta, q+\delta]\}$, a closed neighborhood of $\mathcal{I}^\sigma$ that converges to $\mathcal{I}^\sigma$ as $\delta \to 0$. It follows that $m(V_\delta) \to 0$ as well. That is, there exists $\varepsilon(\delta) > 0$ such that $m(V_\delta) < \varepsilon(\delta)$ and $\varepsilon(\delta) \to$ as $\delta \to 0$. For a given $\varepsilon > 0$, let $\delta > 0$ be such that $\varepsilon(\delta) \leq \varepsilon$. Let g^ε be an element of the sequence g^n such that $\sup_{\theta\in\Theta} \left|\int g^\varepsilon(x)\eta(dx|\theta) - \int \sigma(x)\eta(dx|\theta)\right| < \delta$. Set $s^\varepsilon(\cdot) = g^\varepsilon(\cdot)$ as the common strategy for all voters, and conclude that outside of a set Θ^ε with $m(\Theta^\varepsilon) < \varepsilon$, $\int s^\varepsilon(x)\eta(dx|\theta) > q$ (resp. $\int s^\varepsilon(x)\eta(dx|\theta) < q$) whenever $\int \sigma(x)\eta(dx|\theta) > q$ (resp. $\int \sigma(x)\eta(dx|\theta) < q$).

Hence, for each $\varepsilon > 0$, outside a set Θ^ε with $m(\Theta^\varepsilon) < \varepsilon$, $\mathcal{A}^{s^\varepsilon} = \mathcal{A}^\sigma$ and $\mathcal{B}^{s^\varepsilon} = \mathcal{B}^\sigma$, so s^ε achieves ε-FIE. □

Next, we turn to equilibrium analysis for this case. First note that the game with an uncountable X is defined in the same way as above, using integrals over X^n rather than summations. While proposition 1 is slightly weaker than theorem 1 in the sense that we only obtain ε-FIE, that is enough to guarantee a sequence of strategies that achieve FIE in the limit.

Proposition 2. *If an environment allows ε-FIE for all $\varepsilon > 0$, then there exists a sequence σ^n of Nash equilibria of the game G^n that achieves FIE, i.e., $W_n^{\sigma^n} \to 0$.*

Proof. For each $\varepsilon > 0$ there exists an ε-FIE, s^ε, which achieves the maximum feasible ex ante utility outside a set Θ^ε with $m(\Theta^\varepsilon) < \varepsilon$. Let u^* be the maximum feasible ex ante utility. As in the arguments establishing Theorem 2 above, let σ^n maximize $u(\cdot)$ in the finite electorate game. Then $u^* \geq \lim_{n\to\infty} u(\sigma^n) \geq u(s^\varepsilon) \geq u^* - \varepsilon$. As $\varepsilon > 0$ is arbitrary, $\lim_n u(\sigma^n) = u^*$, so the same argument as in Theorem 2 establishes that σ^∞ achieves FIE. □

5.3. **Diverse Preferences.** So far we have assumed that all voters have the same preferences. In this section we extend our results to a case where the voters in the electorate may have different preferences. We maintain the assumption that all voters are ex ante identical, and draw their information and preferences from some distribution conditional on the state. To do so, we retain the elements of the set-up in the main section and assume in addition that the private signal x is also payoff relevant. Thus, the private draw of an individual serves two functions: it is a view about the outcomes and it provides information about how others view the outcomes. We may think of $x_i = (s_i, t_i)$, where s_i is the common value component and t_i is the private value component of the preference. Notice that this is a general setting that can encompass many different environments. In particular, it admits the environments studied in Feddersen and Pesendorfer (1997) with continuous state space and Bhattacharya (2013) with just two states.

Consider, therefore, that voters preferences are given by $u : \Theta \times X \times \{A, B\} \to \mathbb{R}$. We deal only with the case that X is at most countable.

We now normalize $u(\theta, x) = 1$ if $u(\theta, x, A) > u(\theta, x, B)$, $u(\theta, x) = 0$ if $u(\theta, A) < u(\theta, x, B)$, $u(\theta, x) = \frac{1}{2}$ if $u(\theta, x, A) = u(\theta, x, B)$. This normalization is innocuous for the feasibility result.

Since there is no commonly preferred candidate under diverse preferences, we have to be careful while defining the standard for information aggregation. We say that a strategy profile aggregates information if, for a large electorate, the outcome is the same as the outcome when the state is commonly known. Notice that in a large electorate with finite signals, knowing the state can be interpreted as knowing the entire profile of private signals.

By the SLLN, asymptotically the proportion of voters that prefer A to B, given a state θ, is

$$u(\theta) = \sum_{x \in X} u(\theta, x)\eta(x|\theta).$$

In a large electorate, A would get a vote share very close to $u(\theta)$ if the state were known to be θ. Therefore, under full information A wins depending on whether $u(\theta)$ is greater or less than the threshold q. Now fix $q \in (0,1)$ and redefine the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{I}$ as

$$\begin{aligned} \mathcal{A}_q &= \{\theta \in \Theta : u(\theta) > q\} \\ \mathcal{B}_q &= \{\theta \in \Theta : u(\theta) < q\} \\ \mathcal{I}_q &= \{\theta \in \Theta : u(\theta) = q.\} \end{aligned}$$

In states $\mathcal{A}_q$ (resp.$\mathcal{B}_q$), alternative A (resp. B) wins under full information and large electorates. We then say that an environment (u, Θ, X, η, q) allows FIE if there exists a strategy σ such that

$$m(\mathcal{A}_q \backslash \mathcal{A}^\sigma) = m(\mathcal{B}_q \backslash \mathcal{B}^\sigma) = 0.$$

It is simple to verify now that the argument in the proof of Theorem 1 follows through line-by-line, so FIE is again characterized by the hyperplane condition. More formally, we have the following result,

Proposition 3. *An environment (u, Θ, X, η, q) and diverse preference allows FIE if and only if there exists a hyperplane H in $\Delta(X)$ such that $\eta(\cdot|\theta) \in H^+$ for $\theta \in \mathcal{A}_q$, $\eta(\cdot|\theta) \in H^-$ for $\theta \in \mathcal{B}_q$, and, if $M^{piv} \neq \emptyset$, $\eta(\cdot|\theta) \in H$ for $\theta \in M_q^{piv}$.*

It is also simple to verify that the arguments for Corollaries 1 and 3 remain valid. Corollaries 2 and 4 on the other hand, do not go through any longer: changing q changes the sets $\mathcal{A}_q$, $\mathcal{B}_q$, and $\mathcal{I}_q$, so the arguments given above do not show that FIE is obtained for the new level $\hat{q}$.

Notice that since Feddersen and Pesendorfer (1997) result already tells us that information is aggregated in equilibrium, the existence of FIE strategies is trivial in their setting. More interestingly, while Bhattacharya (2013) concentrates on showing that, for any consequential rule, there exists an equilibrium that fails to aggregate information, it can be checked that in Bhattacharya's two-state setting, there always exists some feasible strategy that achieves FIE. It would therefore be very interesting to examine the conditions under which, in a general setting with diverse preferences, there exists some equilibrium sequence that aggregates information.

However, the proof of Theorem 2 explicitly utilizes the common value setting, and therefore does not automatically generalize to an environment with diverse preferences. In particular, we do not know yet whether, given preference diversity in the electorate, the existence of a feasible strategy profile guaranteeing FIE also implies that FIE is achieved in equilibrium. Our efforts are currently focussed on analyzing conditions under which the hyperplane result also implies that information is aggregated in equilibrium in presence of preference heterogeneity.

6. Conclusion

The existing literature on information aggregation in large elections has largely focussed on specific preference and information environments. In this paper, we consider general environments in order to analyze conditions under which information is aggregated. Preferences depend on the state of the world, and each state of the world is synonymous with a probability distribution over private signals. Therefore, preferences are simply mappings from allowable probability distributions over private information to rankings over the two alternatives A and B. In a large electorate, the frequency distribution over signals is approximately the same as the probability distribution. Thus, our question is whether the election achieves the outcome that would obtain if the entire profile of private signals were publicly known. If an environment permits a feasible strategy profile that can induce the full information outcome with a high probability in almost all states, we say that the environment allows Full Information Equivalence (FIE). Moreover, we are interested in whether such a strategy profile is incentive compatible, i.e., it constitutes a Nash equilibrium in the underlying game.

We study both environments with and without preference heterogeneity. In both cases, we find that an environment allows FIE if and only if a hyperplane on the simplex over signals separates the probability distributions arising in states where A is preferred from those arising in states where B is preferred. Therefore, if the information environment is sufficiently complex, there is no strategy profile that aggregates information. We like to stress here that the failure of FIE has nothing to do with equilibrium assessments over the states based on the criterion of one's vote being pivotal in deciding the election.

We obtain sharper positive results in the common preference case where all voters would have agreed on their rankings if they had known the profile of signals. In this case, voting aggregates information alone (and not preference). We find that as long as an environment allows FIE, there is a sequence of equilibria that achieves FIE. We must mention here that there may be other equilibrium sequences that do not aggregate information - but ours is only a possibility result. As implications of this result, we provide several examples of common preference environments where information will be aggregated. We show that if there are only a finite number of states and signals, FIE holds in equilibrium for any preference mapping from states to alternatives as long as the signal space is sufficiently rich compared to the space of states. As a special case, we show that whenever there are two states, FIE is generically achieved in equilibrium. We also show that in the common preference environment, the voting rule does not matter for information aggregation: as long as FIE is achieved under some voting rule, FIE is achieved under every other non-unanimous voting rule.

Feasibility of FIE does not automatically imply FIE in equilibrium when voting has the burden of aggregating information and preferences simultaneously, i.e., in the diverse preference case. However, we conjecture that under some conditions, feasibility of FIE indeed implies the existence of an equilibrium sequence that also achieves FIE in the case with diverse preferences. Our current research efforts are focussed on unveiling these conditions.

7. References

(1) Acharya, Avidit. "Equilibrium False Consciousness." mimeo, Stanford University (2014).

(2) Aliprantis, C. D., and K. C. Border. "Infinite Dimensional Analysis", Springer-Verlag. (2006).

(3) Aubin, Jean-Pierre, and Hélène Frankowska. "Set-Valued Analysis", Birkhäuser (2009).

(4) Austen-Smith, David, and Jeffrey S. Banks. "Information aggregation, rationality, and the Condorcet jury theorem." American Political Science Review (1996): 34-45.

(5) Berg, Sven. "Condorcet's jury theorem, dependency among jurors." Social Choice and Welfare 10, no. 1 (1993): 87-95.

(6) Bhattacharya, Sourav. "Preference Monotonicity and Information Aggregation in Elections." Econometrica 81, no. 3 (2013): 1229-1247

(7) Cremer, Jaques, and Richard P. McLean, "Full Extraction of Surplus in Bayesian and Dominant Strategy Auctions", Econometrica 56, no. 6, (1988), 1247-1257.

(8) Duggan, John, and César Martinelli. "A Bayesian model of voting in juries." Games and Economic Behavior 37, no. 2 (2001): 259-294.

(9) Feddersen, Timothy, and Wolfgang Pesendorfer. "Voting behavior and information aggregation in elections with private information." Econometrica: Journal of the Econometric Society (1997): 1029-1058.

(10) Ladha, Krishna K. "The Condorcet jury theorem, free speech, and correlated votes." American Journal of Political Science (1992): 617-634.

(11) Mandler, Michael. "The fragility of information aggregation in large elections." Games and Economic Behavior 74, no. 1 (2012): 257-268.

(12) McLennan, Andrew. "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents." American Political Science Review (1998): 413-418.

(13) McMurray, Joseph: "Why the Political World is Flat: An Endogenous Left-Right Spectrum in Multidimensional Political Conflict", mimeo, Brigham Young University

(14) Milgrom, Paul R. "Good news and bad news: Representation theorems and applications." The Bell Journal of Economics (1981): 380-391.

(15) Myerson, Roger B. "Extended Poisson games and the Condorcet jury theorem." Games and Economic Behavior 25, no. 1 (1998): 111-131.

(16) Siga, Lucas. "Complex Signals and Information Aggregation in Large Markets" mimeo, New York University (2013)

(17) Wit, Jörgen. "Rational choice and the Condorcet jury theorem." Games and Economic Behavior 22, no. 2 (1998): 364-376.

Department of Economics, University of Rochester, Email:paulo.barelli@rochester.edu.

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

Department of Economics, New York University Abu Dhabi, Email: lucas.siga@nyu.edu.