FULL INFORMATION EQUIVALENCE IN LARGE ELECTIONS

PAULO BARELLI, SOURAV BHATTACHARYA AND LUCAS SIGA

Abstract. We study the problem of aggregating private information in elections with two or more alternatives for a large family of scoring rules. We introduce a feasibility condition, the *linear refinement condition*, that characterizes when information can be aggregated asymptotically as the electorate grows large: there must exist a utility function, linear in distributions over signals, sharing the same top alternative as the primitive utility function. Our results complement the existing work where strong assumptions are imposed on the environment, and caution against potential false positives when too much structure is imposed.

1. Introduction

Elections provide a mechanism to aggregate preferences and information dispersed in the society. To this end, it must be possible for the election outcome to aggregate all the private information that voters possess about the optimal policy. Our aim in this paper is to understand when information aggregation is feasible in large elections.

The existing literature on common-value elections follows an insight by Condorcet (1785): in a large election, when every voter is more likely to vote correctly than incorrectly, the Law of Large Numbers implies that the majority is almost surely correct. This is known as the Condorcet Jury Theorem. Subsequent work formalizes

Date: May 26, 2022.

We thank Nageeb Ali, Mehmet Ekmekci, Timothy Feddersen, Srihari Govindan, Marina Halac, Bård Harstad, Johannes Hörner, Vijay Krishna, Stephan Lauermann, Michael Mandler, Cesar Martinelli, Santiago Oliveros, Wolfgang Pesendorfer, Hamid Sabourian, Larry Samuelson, Joel Sobel, Satoru Takahashi and Roee Teper for helpful comments and suggestions. We also thank Timothy Feddersen for providing us with the first interpretation of Example 1, Anne-Katrin Roesler and Alvaro Sandroni for insightful discussions at conferences, and Max Mihm for numerous comments and suggestions. Bhattacharya thanks Jadavpur University, Kolkata, India, for its hospitality. We also thank the coeditor, Dirk Bergemann, and anonymous referees for their comments and suggestions, which significantly improved the paper.

Condorcet's insight in a game-theoretic framework where voters draw *i.i.d.* signals conditional on an unobservable state. Due to challenges in characterizing equilibria, the literature has restricted attention to settings with a number of simplifying assumptions. The most common approach is to focus on environments with two states and two alternatives (e.g., Austen-Smith and Banks, 1996; Myerson, 1998; Wit, 1998; Duggan and Martinelli, 2001; McMurray, 2013). An alternative approach focuses on environments with many states, two alternatives, and conditionally i.i.d. signals satisfying the Monotone Likelihood Ratio Property (MLRP) (e.g., Feddersen and Pesendorfer, 1997; McMurray, 2017). It is not clear, however, whether the insights developed in this literature extend to more general information environments, or elections with more than two alternatives.

We address this question in a model where a large population of voters chooses an alternative from a finite set of possibilities under plurality rule (which we extend to general scoring rules in Section 5.1). The common preference of the voters depends on the realization of an unobservable state.[1] Conditional on the state, every voter receives an i.i.d. signal and votes. Instead of focusing on limiting properties of a specific equilibrium construction, as in the existing literature, we first focus on the issue of feasibility. We provide necessary and sufficient conditions for the existence of a strategy profile that satisfies *Full Information Equivalence* (FIE): a strategy profile where the alternative chosen by a large electorate is the same alternative that would be chosen if all the information was public. We then draw on an insight from McLennan (1998) on voting with common interests to conclude that if FIE is feasible, it is also achievable in some equilibrium. Thus, we identify the environments where FIE is an equilibrium property of the voting mechanism.

All payoff-relevant aspects of the unobservable state are captured by the distribution of signals, unique to each state. Thus, the utility of an alternative can be viewed as a function of the distributions of signals. Additionally, the law of large numbers

[1]In contrasts with the standard assumption of one state per alternative, a key element in our approach is that an alternative may be best in multiple states, provided that each of these different states lead to distinct distributions of private signals. For instance, a candidate may be best when inequality is the main concern but also in a situation where immigration is the main issue; whereas a different candidate may be best when the main concern is a looming recession or a war.

implies that a strategy profile generates vote shares for each alternative that vary linearly with the distributions of signals. As a consequence, feasibility of FIE boils down to whether the top alternatives from the primitive utility function, u, agree with the top alternatives of a function that is linear in distributions of signals. More formally, we show that FIE can be achieved if and only if the following *linear refinement condition* is satisfied: there exists a utility function v that (i) is linear in distributions[2] and (ii) whose top alternatives are necessarily top alternatives of u.

Our main result identifies the key limitation of the voting mechanism for aggregating information: the linearity requirement is an impediment to FIE. For instance, in the case with two candidates and two signals, if the preference of the voters depends on the political inclination of the candidates on a liberal/conservative one-dimensional scale and a higher probability of one signal indicates that a particular candidate is relatively more conservative, then the environment allows FIE. In such case, the voters' utility function is compatible with a monotone linear utility, so the linear refinement condition holds. On the other hand, if the voters' utility function depends on whether that candidate is relatively more moderate or extreme, while signals only convey information about whether candidates lean left or right, then signals cannot be classified as each favoring one candidate. In this case, there is no linear utility whose top choices match the preference of the voters and, as a consequence, FIE fails (see Example 1 in Section 3.1).

We also provide a geometric characterization of the linear refinement condition in Proposition 1. We show that the linear refinement condition is equivalent to the existence of a family of hyperplanes on the simplex of distributions of signals satisfying three conditions. The first is a separating condition specifying that for every two alternatives there is a hyperplane separating distributions where one alternative is best from distributions where the other alternative is best. The second is a consistency condition between any three alternatives specifying that the three separating hyperplanes intersect, and their upper half spaces point in a direction that prevents violations of transitivity. The third requires that no two hyperplanes coincide with

[2] For each alternative a, $v(a, \mu) = b^a \cdot \mu$, where b^a is a vector of coefficients and μ is the distribution of signals.

one another. The take-away is that FIE is feasible if and only if the simplex can be partitioned into regions with linear boundaries satisfying a certain consistency requirement, with each alternative being the best for states in one and only one of the elements of the partition (see Figures 2 and 3 for illustrations.) This geometric approach provides an additional tool for identifying environments where FIE is not feasible. For instance, consider a standard spatial model of political competition between an alternative and a status quo. Within our framework, the simplex of distributions acts as the multi-dimensional policy space. With standard euclidean preferences with a bliss point and three signals, the simplex is partitioned into two regions: a circle around the bliss point and its complement, where the alternative is best for states inside the circle and the status quo is best for states outside of the circle. Clearly, this environment violates the geometric characterization and hence FIE is not feasible.

Overall, our results identify a broad class of environments where FIE fails, but also show that FIE is satisfied under more general conditions than those of the prior literature. We relate our results to the literature in several directions. First, generalizing previous findings to multiple alternatives, we show that FIE is always achievable when there is a one-to-one mapping between alternatives and states (Proposition 2). However, such result may well lead to false positives when the modeler assigns one state for each alternative (by coalescing multiple states where the alternative is best into a single aggregate state): with more states than alternatives, many environments do not allow FIE. Second, we connect our linear refinement condition to the MLRP, by establishing that the *betweenness property* identified in Siga and Mihm (2021) for an auction, is sufficient, but not necessary, for the linear refinement condition. As Siga and Mihm establish that the betweenness property is strictly weaker than the MLRP, our linear refinement generalizes the conditions in the literature. Moreover, since the betweenness property holds generically with more signals than states, the linear refinement condition does as well (Proposition 3(a)). On the other end, we establish conditions for robust failure of FIE in Proposition 3(b) for sufficiently large number of states relative to the signals and alternatives. Finally, Proposition 3(c) establishes that FIE fails robustly when we do not restrict the number of states.

Lastly, we show that our results generalize in several directions. We introduce a large class of voting rules (including supermajority, approval voting, storable votes, etc.) and show that if there exists a profile of strategies that satisfies FIE in any voting rule of this class, there exists also a profile of strategies that satisfies FIE under plurality, and conversely. As such, the restrictions we identify on FIE for plurality rule also apply to many other popular voting mechanisms. Second, we show that our feasibility results also extend to non common-value environments, where voters have heterogeneous preferences. While feasibility of FIE is not generally sufficient for existence of an equilibrium strategy satisfying FIE with heterogeneous preferences, the restrictions on information aggregation in common-value environments apply well beyond the common-value setting. Finally, we extend our analysis to environments with infinitely many signals.

1.1. **Related Literature.**

Our paper extends the existing body of work on information aggregation in several ways. An important literature has shown that information is aggregated in all equilibrium sequences by assuming binary states (e.g., Austen-Smith and Banks, 1996; Myerson, 1998; Wit, 1998; Duggan and Martinelli, 2001), or information structures satisfying MLRP (e.g., Feddersen and Pesendorfer, 1997; McMurray, 2017). Chakraborty and Ghosh (2003) show that with binary states and finite signals, any voting threshold rule satisfies FIE in some equilibrium. An implication of our findings is that such coalescing of states and monotonicity assumptions are not innocuous.

Another strand of literature identifies sources of aggregation failures. This includes unanimity rules (Feddersen and Pesendorfer, 1998), alternative voters motivations (e.g., Razin, 2003; Callander, 2008), information acquisition costs (Persico, 2004; Martinelli, 2006), costs of voting (Krishna and Morgan, 2012), aggregate uncertainty (Feddersen and Pesendorfer, 1997), state-dependent size of electorate Ekmekci and Lauermann (2020), and so forth. Our work highlights that complexity of the information structure may itself be a barrier to information aggregation.

A number of papers show existence of equilibrium strategies where information aggregation fail in non-common value environments (e.g., Acharya, 2016; Ali et al., 2019; Bhattacharya, 2013, 2018; Kim and Fey, 2007). The typical approach is to exploit local properties of the vote-share function to construct inefficient equilibria. Ekmekci and Lauermann (2022) show that both efficient and inefficient equilibria coexist if the electorate size varies across the two states. While these papers have an analogous message, we provide a stronger result of failure of FIE: the failure is that of feasibility rather than that of voters coordination due to pivotal inference and/or the presence of aggregate uncertainty. Indeed, there is no advantage to abandoning game-theoretic models with hyper-rational voters to aggregate information: in particular, in environments that fail FIE, aggregation would not be possible even if voters could commit in advance to arbitrary strategy profiles. In contrast, ours is a feasibility result: in environments that allow FIE, there may well exist equilibria that do not aggregate information.

Mandler (2012) constructs inefficient equilibria in a common-value election by introducing aggregate uncertainty. In particular, conditional on a distribution of signals, there is an unobservable random variable that determines the best alternative which does not vanish as the population grows (see also Feddersen and Pesendorfer (1997), Section 6). There is no such aggregate uncertainty in our model and, in addition, by integrating out the aggregate uncertainty, the inefficient equilibria in Mandler's correspond to environments where FIE is not feasible in our setting. We elaborate further after Example 1 in Section 3.1.

A related question is whether communication is useful to produce the correct outcomes. In common-value environments, there is a clear incentive to share information. The common-value environments where we show that FIE fails are precisely those environments where deliberation is necessary to improve outcomes. One interpretation of Coughlan (2000) is that deliberation has a role only when there is preference diversity among voters. We, on the other hand, establish that it can also have a role under common preferences.

There is a growing literature on informational efficiency for different scoring rules. Martinelli (2002) provides necessary and sufficient conditions under unanimity for no incorrect convictions in a large population. Another strand of literature focuses on approval voting with three alternatives: Goertz and Maniquet (2011) and Bouton and Castanheira (2012) focus on non-common-values components, Ahn and Oliveros (2016) present a common-value environment and show that approval voting performs weakly better than other scoring rules. We show that plurality performs as well as all other scoring rules in a large population.

In parallel to the literature in elections, a large literature in auctions draws related insights. In particular, much like in Feddersen and Pesendorfer (1997), Pesendorfer and Swinkels (1997) observe that expected values (in pivotal events) are monotone in signals when signals satisfy the MLRP and strategies are monotone in signals. The similarities are somehow surprising given obvious differences in the mechanisms, incentives, and outcomes. Closest to ours is Siga and Mihm (2021) (SM), who characterize information structures that permit fully informative equilibria in large common-value auctions and allows us to contrast the consequences of the differences in the models. In the auction mechanism, individual actions (i.e., bids) are aggregated into a final outcome—the price—which is given by a specific *quantile* of the bids, and information is aggregated if the price reveals the value in every state. By contrast, in the voting mechanism, individual actions (i.e., votes) are aggregated into a final outcome—the chosen alternative—which is given by the *mode* of the votes, and information is aggregated if the chosen alternative reveals the best alternative in every state. These differences, especially the difference in the relevant statistic (quantile vs. mode), lead to a different characterization of information aggregation in the two models.[3] Under an appropriate mapping between the voting and auction models, our linear refinement condition nests the betweenness condition in the auction environment. The betweenness condition implies the existence of a linear refinement, whereas it is simple to construct examples (e.g. Figures 2(A) and 3(B)) of environments that violate betweenness but satisfy the linear refinement condition, and thus allow FIE.

[3]There is also a significant difference in the incentives in the voting and the auction environment. Even with common values, the bidders in an auction always have private interests. As a result, feasibility is not sufficient for equilibrium in a common value auction. By contrast, with common values, voting is a problem of common interests. As a result, as we show in Section 4, any voting outcome that is feasible can be realized as an equilibrium.

2. Model

We consider an electorate with $n > 1$ voters choosing from a finite set A of alternatives, under plurality voting: each voter casts a vote for one of the alternatives, and the alternative with most votes wins the election. Ties are broken randomly.

Voters have the same utility over alternatives, which depends on a common and unknown state. Before voting, each voter independently receives a private signal from a finite set S.[4] Signals are drawn from a state-dependent distribution. The state therefore impacts voters in two ways: (i) it determines the utility from each of the alternatives and, (ii) it determines the distribution over signals. We assume that each state generates a different distribution over signals and, to simplify notation, we directly identify a state with the distribution over signals that it generates. As such, a state is described by a distribution $\mu \in M \subset \Delta(S)$, where $\Delta(S)$ is the simplex over S. Voters have a common prior λ with supp $\lambda = M$.[5] The common utility is represented by $u \in \mathcal{U}$ where $\mathcal{U}$ is the set of all bounded and measurable functions from $A \times M$ to $\mathbb{R}$. For a function $u \in \mathcal{U}$, let $\alpha_u(\mu) \equiv \arg\max_{a \in A} u(a, \mu)$. Throughout the paper, we use *a.s.* for a property that holds on a set of states of λ-measure one, and for a given set X, we use $|X|$ to denote the cardinality of X. For a fixed number of signals and alternatives, the pair $\{u, \lambda\}$ defines an environment.

To simplify exposition, we focus on symmetric strategies for now. In Section 4 we show how our analysis extends to asymmetric strategies. A mixed strategy $\sigma : S \times A \to [0, 1]$, with $\sum_{a \in A} \sigma(s, a) = 1$ for all $s \in S$, specifies, for each signal, a probability distribution over alternatives. That is, $\sigma(s, a)$ is the probability of voting for alternative a upon obtaining signal s. For $a \in A$ we also use the notation $\sigma^a(s) \equiv \sigma(s, a)$, and $\sigma^a \in [0, 1]^{|S|}$ in its vector form, whenever convenient.

Our interest is whether one of the best alternatives is chosen by a sufficiently large electorate. Given a state μ and strategy profile σ, the strong law of large numbers

---

[4]We extend our analysis to allow S infinite in Section 5.3.

[5]As S is finite, the simplex $\Delta(S)$ belongs to a $|S|-1$ dimensional affine subspace of $\mathbb{R}^{|S|}$. It inherits the subspace topology of $\mathbb{R}^{|S|}$ with its standard Euclidean topology. The prior λ is a probability measure on the resulting Borel sigma-algebra on $\Delta(S)$, and its support, supp λ, is the smallest closed set $F \subset \Delta(S)$ such that $\lambda(F) = 1$.

guarantees that the realized proportion of votes for alternative a converges almost surely to its expected proportion of votes, $E_\mu(\sigma^a) \equiv \sum_{s\in S} \mu_s \sigma^a(s) = \mu \cdot \sigma^a$. As a consequence, we formalize our notion of information aggregation as follows.

Definition 1. *A strategy profile σ satisfies Full Information Equivalence (FIE) if*

$$\arg\max_{a\in A} E_\mu(\sigma^a) \subset \alpha_u(\mu) \ \textit{a.s.}$$

When such strategy exists, we say that the environment allows FIE, and we say that FIE fails otherwise.

That is, a strategy σ satisfies FIE when the alternatives with the largest expected proportion of votes maximize utility in almost all states. We assume that every alternative is among the optimal alternatives in a set of positive measure. That is, for all $a \in A$, $\lambda(\{\mu \in M : a \in \alpha_u(\mu)\}) > 0$.

3. Main Results

Our main result characterizes the environments that allow FIE and it is tied to the following two notions:

Definition 2. *For $v, \hat{v} \in \mathcal{U}$, $\hat{v}$* refines *v if $\alpha_{\hat{v}}(\mu) \subset \alpha_v(\mu)$ a.s.*

Let us use $\hat{v} \succsim v$ to denote that $\hat{v}$ refines v. In words, $\hat{v} \succsim v$ if for almost every state the top alternatives for $\hat{v}$ are necessarily top alternatives for v as well.

Definition 3. *A function $v \in \mathcal{U}$ is* linear *if for all $a \in A$ there exists $b^a \in \mathbb{R}^{|S|}$ such that $v(a, \mu) = \mu \cdot b^a$.*

Definition 4. *An environment $\{u, \lambda\}$ satisfies the* linear refinement condition *if there exists a linear utility function v such that $v \succsim u$.*

Hence, the linear refinement condition is satisfied if the utility function u is refined by a function v that is linear in distributions.

Theorem 1. *An environment $\{u, \lambda\}$ allows FIE if and only if it satisfies the linear refinement condition.*

Proof. Suppose σ satisfies FIE. Consider a linear utility function $v(a,\mu) \equiv E_\mu(\sigma^a)$. From the definition of FIE, it follows directly that $\alpha_v(\mu) \subset \alpha_u(\mu)$ a.s. and hence $v \succsim u$. Conversely, suppose that there is a linear utility function $v(a,\mu) = \mu \cdot b^a$ that refines u. Consider a linear $\hat{v} \in \mathcal{U}$ given by $\hat{v}(a,\mu) = \mu \cdot \hat{b}^a$, where $\hat{b}^a = \gamma b^a + \frac{1}{|A|}(1 - \sum_a \gamma b^a)$ with $\gamma > 0$ chosen to ensure that $\hat{b}^a > 0$ for all $a \in A$. Since $\hat{v}$ is affine to v, $\alpha_{\hat{v}}(\mu) = \alpha_v(\mu)$ for all μ, and hence $\alpha_{\hat{v}}(\mu) \subset \alpha_u(\mu)$ a.s. Let $\sigma^a = \hat{b}^a$ for all $a \in A$. By construction $\sigma^a \in [0,1]^{|S|}$ and $\sum_a \sigma^a(s) = 1$ for every s. Thus, σ is a well-defined strategy. Since $\arg\max_a E_\mu(\sigma^a) = \alpha_{\hat{v}}(\mu)$, σ satisfies FIE. $\square$

Theorem 1 establishes that the utility functions compatible with information aggregation belong to a restricted class. Even though the utility function need not be linear, there must exist some linear utility with the same top alternatives as the voters' original utility function. This this is a substantive restriction on primitives which is circumvented by a large body of literature where feasibility of information aggregation is built in.

3.1. FIE and Information Structures.

In this section we present a geometric characterization of information structures that allow FIE, and illustrate it with examples with two, three, or more alternatives. Throughout this section, we focus on environments where best alternatives are generically unique, as captured by Assumption 1:

Assumption 1. $|\alpha_u(\mu)| = 1$ *a.s.*.

For all $a \in A$, let $M_a^u = \{\mu \in M : \alpha_u(\mu) = \{a\}\}$ be the set of states where a is the unique best alternative. Thus, under Assumption 1, $M = \bigcup_{a \in A} M_a^u$ a.s..

For $b \in \mathbb{R}^{|S|}$, let $h_b = \{x \in \mathbb{R}^{|S|} : x \cdot b = 0\}$ be a hyperplane in $\mathbb{R}^{|S|}$ with b as its normal vector, and $h_b^+ = \{x \in \mathbb{R}^{|S|} : x \cdot b > 0\}$ be its strict upper half-space. Existence of a linear refinement, and hence FIE, is characterized by the existence of hyperplanes with the following conditions:

Proposition 1. *An environment $\{u, \lambda\}$ satisfies the linear refinement condition if and only if there exists $b : A^2 \to \mathbb{R}^{|S|}$ with $b(a,a') = -b(a',a)$ such that: (i) for each*

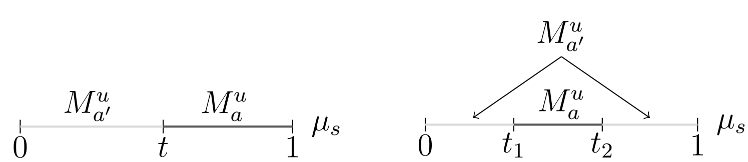
a and all $a' \neq a$, $M^u_a \subset h^+_{b(a,a')}$ a.s.; and, for any three distinct alternatives a, a', a'', (ii) $h^+_{b(a,a')} \cap h^+_{b(a',a'')} \subset h^+_{b(a,a'')}$ and (iii) $h_{b(a,a')} \neq h_{b(a,a'')}$.

Condition (i) is a separating condition that specifies that, for every two alternatives, there exists a hyperplane separating the set of states where one alternative is best from the set where the other alternative is best. Condition (ii) is a consistency condition between any three alternatives specifying that the three hyperplanes intersect, and their upper half spaces point in a direction that prevents violations of transitivity. Finally, if (iii) was not met, then one of the three alternatives would never be optimal.[6]

In the case with two alternatives, a and a', only the separating condition applies, so FIE is characterized by the existence of a separating hyperplane such that M^u_a is a.s. on one side of the hyperplane and $M^u_{a'}$ is a.s. on the other. Example 1 illustrates.

Example 1. *Consider two alternatives, $A = \{a, a'\}$, and two signals, $S = \{s, s'\}$, and let λ be the uniform distribution on $M = \Delta(S)$. Consider two environments summarized by the information structures illustrated in Figure 1. In Figure 1(A), all voters prefer a if $\mu_s > t$ and a' if $\mu_s < t$, for some threshold $t \in (\frac{1}{2}, 1)$. In Figure 1(B), for some $0 < t_1 < t_2 < 1$, a is preferred whenever $\mu_s \in (t_1, t_2)$ and a' is preferred otherwise. Our separation condition (i) in Proposition 1 implies that (A) allows FIE but (B) does not.*

FIGURE 1



(A) (B)

Example 1 illustrates that the substantive meaning of each signal matters for FIE. For instance, if the preference of the voters depends on the political inclination of the candidates on a liberal/conservative one-dimensional scale and a higher probability of

[6]It is possible to present a similar result without Assumption 1 at the cost of additional notation and a restriction to control for non-zero measures of states lying on separating hyperplanes.

the s signal indicates that candidate a is relatively more conservative than candidate a', then the environment allows FIE, as in Figure 1(A). Conversely, if voter preferences depend on whether candidate a is relatively more moderate or extreme, while signals only convey information about whether she leans relatively more left or right, then signals cannot be classified as each favoring one candidate. This interpretation applies to Figure 1(B), where FIE fails.[7]

A related issue arises in Mandler (2012), who constructs equilibria that fail FIE in a model with aggregate uncertainty. In his formulation, a given distribution of signals (our states) does not pin down which alternative is best. Distributions are drawn from one of two densities, with known prior probabilities depending on which alternative is best. As a consequence, a distribution cannot be used as a state and the information available in the electorate is not sufficient to determine the best alternative, regardless of the electorate size. Since voters are expected utility maximizers, we can use the expected utility conditional on a distribution of signals as our primitive utility. In this sense, his model can be cast in our framework, with the aggregate uncertainty integrated out. When the two densities that generate distribution of signals cross multiple times, as in Figures 1 and 2 in his paper, the corresponding sets of distributions associated with each alternative constitute information structures analogous to our Figure 1(B), and hence to failure of FIE. In other words, the environments where Mandler constructs inefficient equilibria translate to our framework as environments where FIE is not feasible.

An alternative interpretation for the environment in Figure 1(B) is that it is a special case (with only one dimension) of a spatial model where the simplex is the policy space and voters are to decide between a status quo a and an alternative a' whose desirability depends on the state. Allowing for more dimensions, let us consider a general finite set of signals S and the simplex $\Delta(S)$ as the policy space. Assume the utilities are Euclidean over $\Delta(S)$: the voters' common bliss point is μ^*, the "location" of the status quo a is μ_q and the utility from a' in state μ is $-\|\mu - \mu^*\|$. Therefore, the states where a' is preferred is a ball with center μ^* and radius $\|\mu_q - \mu^*\|$, and the rest

[7]There is nothing special about using intervals of states; already with three states in a line, with a preferred in the "middle state" and a' in the two "extreme states", FIE fails.

of the simplex is where a is preferred. In other words, voters prefer the alternative a' when the state is close enough to the bliss point. Since there is no way to separate the two sets with a hyperplane, FIE cannot be satisfied in this environment.

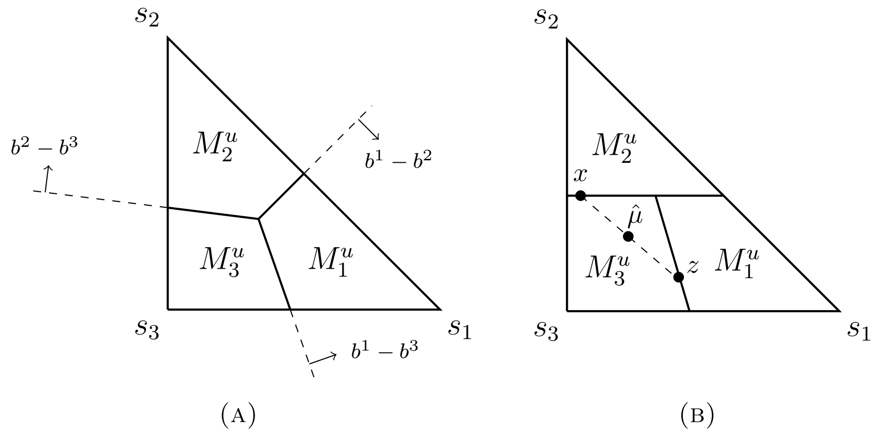
With three alternatives, all conditions in Proposition 1 are binding and they imply that the separating hyperplanes must form a convex partition of $\Delta(S)$ with three components, with the interior of each component associated with a specific alternative a.s.. Condition (iii) imposes the additional mild restriction that no two boundaries can lie on the same hyperplane. This illustrated in Example 2 below.

Example 2. *Let $A = \{a_1, a_2, a_3\}$, $S = \{s_1, s_2, s_3\}$, and let λ be the uniform distribution on $M = \Delta(S)$. Consider two environments with preferences as in Figure 2, with M_i^u denoting $M_{a_i}^u$ for ease of notation. FIE is possible in environment (A) that satisfies (i), (ii), and (iii) in Proposition 1, and not possible in environment (B) that fails condition (iii) in Proposition 1. We can construct a refining linear utility v for environment (A) as follows. For each of the three lines depicted, consider the affine hyperplanes in $\mathbb{R}^2$ that they belong to, and have the normals of such hyperplanes be equal to $b^i - b^j$ for the hyperplane containing the line separating $M_i^u - M_j^u$, for $i, j = 1, 2, 3$ and suitable vectors b^1, b^2, b^3 in $\mathbb{R}^2$. This is illustrated in Figure 2(A). Let $\tilde{b}^i = (b_1^i, b_2^i, 0)$ extend b^i to $\mathbb{R}^3$ and define a linear utility $v(\mu, a_i) = \mu \cdot \tilde{b}^i$. It is simple to verify that $\alpha_v(\mu) = \{a_i\}$ for every state $\mu \in M_i^u$, $i = 1, 2, 3$: as we move inside M_i^u from its boundary with M_j^u, the corresponding affine hyperplane is shifted in the direction of increased v-utility for alternative a_i. Thus, $v \succsim u$.*

As for the environment in Figure 2(B), for a contradiction, suppose there is a refining linear utility $v(a, \mu) = \mu \cdot b^a$ for $a = 1, 2, 3$. Consider the points x, z in Figure 2(B). As x is on the separating line between 1 and 2, as well as between 2 and 3 we have $x \cdot (b^2 - b^1) = x \cdot (b^2 - b^3) = 0$. Then, it follows that $x \cdot (b^1 - b^3) = 0$. But since $z \cdot (b^1 - b^3) = 0$, any state $\hat{\mu}$ in the line connecting x to z must satisfy $\hat{\mu} \cdot (b^1 - b^3) = 0$ so $v(1, \hat{\mu}) = v(3, \hat{\mu})$. But $\hat{\mu} \in M_3^u$, a contradiction.

Example 2 suggests that FIE is characterized by the existence of convex partitions. This is generally not true for more than 3 alternatives since the consistency requirement in Proposition 1 imposes additional restrictions. We illustrate this in a richer environment with 4 alternatives in Example 3.

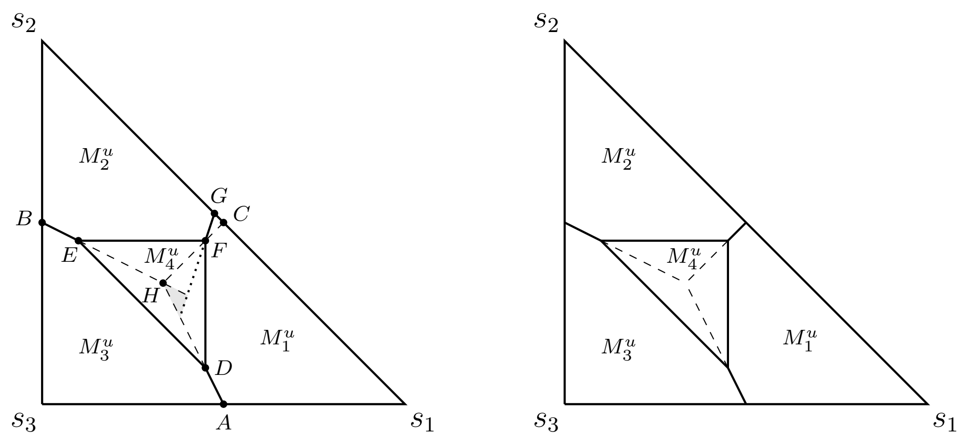
FIGURE 2



Example 3. *Let $A = \{a_1, a_2, a_3, a_4\}$, $S = \{s_1, s_2, s_3\}$, and λ be the uniform distribution on $M = \Delta(S)$. Consider the environment illustrated in Figure 3(A). While each M_i^u is convex and no two boundaries lie on the same hyperplane, FIE fails. To see why, suppose a strategy σ satisfies FIE, and let $E_\mu(\sigma^{ij}) = \mu \cdot (\sigma^{a_i} - \sigma^{a_j})$ be the expected difference in proportion of votes between alternative a_i and alternative a_j in state μ. Now, it must be the case that for all μ on the line passing through AD, $E_\mu(\sigma^{13}) = 0$. Similarly, on the line passing through BE, $E_\mu(\sigma^{23}) = 0$. By linearity of $E_\mu(\sigma^{ij})$ in μ if the lines passing through BE and AD intersect at H then at H, $E_\mu(\sigma^{13}) = E_\mu(\sigma^{23}) = 0$ so that also $E_\mu(\sigma^{12}) = 0$. We also know that along the points GF, $E_\mu(\sigma^{12}) = 0$. However, it is impossible for a line $E_\mu(\sigma^{12}) = 0$ to pass through the line GF and FH simultaneously, so no strategy can generate this partition. Observe that using h_{ij} to denote the hyperplane $h_{b(a_i,a_j)}$, since the line GF must lie on h_{23}, the shaded area below point H will belong to $h_{32}^+ \cap h_{21}^+$. But since it does not belong to h_{31}^+, condition (ii) in Proposition 1 fails.*

On the other end, the environment illustrated in Figure 3(B) allows FIE. In particular, the following strategy satisfies FIE: for $i = 1, 2, 3$, a voter with signal s_i votes for alternative a_i with probability $1 - t$ and votes for alternative a_4 with probability t, where $t \in (1/4, 1/3)$ depending on the size of M_4^u (as drawn, $t = 5/12$). This strategy leads to winning regions as in Figure 3(B). It is simple to verify that all conditions from Proposition 1 hold.

Figure 3



(a) Non-FIE

(b) FIE

Example 3 is analogous to the example in Green and Osband (1991). This turns out not to be a coincidence since the conditions in Green and Osband's work such that an agent's choices can be rationalized by an expected utility are mathematically equivalent to our linear refinement condition. In their environment, our alternatives would correspond to acts chosen by the decision maker, our set of signals would correspond to the set of states identifying the payoffs of all degenerate lotteries, and finally, our states would correspond to the set of probability assessments from which decision makers make choices. As a consequence, the linear refinement condition is satisfied if and only if there exists a utility function that rationalizes an optimal behavior. Green and Osband show that the existence of a rationalizable behavior is tied to an integrability condition that provides an alternative tool for evaluating FIE. Likewise, our Proposition 1 provides an alternative characterization in their setting. We make this connection precise in Appendix 7.5.

3.2. **Sufficient Conditions for FIE.**

It is standard in the voting literature to assume that either (i) there is one state per alternative, or (ii) states are ordered and signals are ordered by the MLRP. In this section we connect our results with these assumptions, caution against potential false positives when too much structure is imposed, and identify genericity or robust failure of FIE depending on the relative sizes of the state and signal spaces.

3.2.1. *One state per alternative.* The first standard assumption of one state per alternative contrasts with a key element in our approach, which is that an alternative may be best in multiple states with distinct distributions of signals. This distinction has important consequences on FIE, as we now establish.

Proposition 2. *An environment with $|M| = |A|$ and $|S| \geq 2$ allows FIE.*

That is, when for each alternative there is a unique state where the alternative is top-ranked, FIE is guaranteed *regardless of the number of alternatives and the number of signals.* Hence, the common practice of coalescing all states associated with one alternative into one aggregate state, in effect, trivializes the question of FIE. If there are more states than alternatives, FIE can fail as illustrated in previous examples.

3.2.2. *Relationship with alternative properties.* The second standard assumption is the MLRP which is substantially more demanding than the linear refinement condition. In particular, we next introduce a substantial weakening of the MLRP, the *betweenness condition* analogous to the *betweenness property* identified in the auction environment in Siga and Mihm (2021) (henceforth SM), that is strictly more demanding than our linear refinement condition.

Much like the MLRP, the betweenness property in SM is associated with the natural order of values in the auction. In our setting, this is equivalent to having an order on alternatives.

Definition 5. *An environment $\{u, \lambda\}$ satisfies* the betweenness condition *if there exists a linear refinement v and an ordering of alternatives (where we let $A = \{1, ..., k\}$ with the natural order after reindexing), such that if $v(a, \mu) > v(a + 1, \mu)$, then $v(a + 1, \mu) > v(a + 2, \mu)$ for all $\mu \in \Delta(S)$ and all $a = 1, ..., k - 2$.*[8]

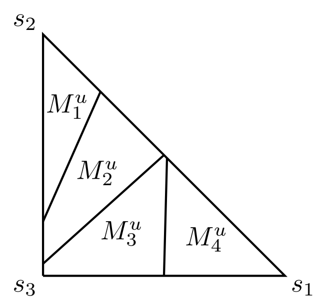
That is, on top of refining u, $v(a, \mu) = b^a \cdot \mu$ must also satisfy the condition that $\mu \cdot (b^a - b^{a+1}) > 0$ implies $\mu \cdot (b^{a+1} - b^{a+2}) > 0$ for all $a = 1, ..., k - 2$, so the upper

[8]The betweenness condition is weaker than the betweenness property since it allows weak rankings of alternatives whereas the betweenness property has strict separation for every distinct value.

half-spaces of the separating hyperplanes between sets of states with consecutive top-alternatives form a sequence of nested sets *on the simplex*.[9] This is illustrated in Figure 4 with 3 signals and 4 alternatives where a line separating M_a^u and M_{a+1}^u is defined by the hyperplane $h_{b^a - b^{a+1}}$.[10] By definition, the betweenness condition is a strengthening of the linear refinement condition. Figures 2(A) and 3(B) illustrate examples satisfying the linear refinement condition but not the betweenness condition.

The connection of the linear refinement condition to the MLRP follows since from SM it is established that the MLRP implies the betweenness property.[11] Thus, the linear refinement condition contains both the betweenness condition and the MLRP as special cases.

FIGURE 4



3.2.3. *Genericity.* Finally, we move to the discussion of genericity of FIE. In the case of finitely many states, the following remark captures how the relationship between number of states and number of signals influences whether the linear refinement condition is satisfied.

[9]Alternatively, we can interpret the betweenness condition as the linear refinement condition plus a single peakedness condition for some arbitrary order on alternatives as in the median voter theorem in Downs (1957).

[10]It is straightforward to verify that the betweenness condition implies the existence of a betweenness preference (with linear indifference curves in the simplex that are not necessarily parallel) therefore satisfying the betweenness property in SM.

[11]Mihm and Siga (2021) show that there is a precise connection of the MLRP and the betweenness property: the betweenness property is satisfied if states are separated on the simplex by some betweenness order, while the MLRP is satisfied if states are separated for every betweenness order.

Remark 1. *When $|M|$ is finite, FIE is allowed when the states $\mu_1, ..., \mu_{|M|}$, viewed as vectors in $\mathbb{R}^{|S|}$, are linearly independent.*

In order to see this, we can arrange the states $\mu_1, ..., \mu_{|M|}$ on a matrix H where the element H_{ms} is the probability of receiving signal s in state m. By independence, the system $Hb = c$ has a solution for all c. The vector b is, therefore, the normal vector of a family of parallel hyperplanes that separate states according to the arbitrary vector c. Since the order in c is arbitrary, we can choose it so that we can separate alternatives by parallel shifts on the hyperplanes with normal b and, therefore, the betweenness condition is satisfied.

An immediate implication of Remark 1 is that FIE is generic when $|M| \leq |S|$. SM establishes that this is an "if and only if" statement in the auction. As the linear refinement condition is a weakening of betweenness, the "only if" part is not true in our voting environment, as we now establish.

Each set of states is associated with an environment $\{u, \lambda\}$ satisfying supp $\lambda = M$. So we focus on the space of subsets of states (that is, on subsets of $\Delta(S)$) and we say that one such subset $M \subset \Delta(S)$ allows FIE if $\{u, \lambda\}$ allows FIE *for any environment* $\{u, \lambda\}$ *associated with* M. Conversely, we say that FIE fails for M if *there is at least one environment* $\{u, \lambda\}$ *associated with* M such that FIE fails. With $|M| < \infty$, the information structure can be expressed as a $|M| \times |S|$ matrix with rows that belong to $\Delta(S)$, so we can use the Lebesgue measure on $\mathbb{R}^{|M|(|S|-1)}$ to measure the size of sets of information structures where FIE is or is not allowed. When M is allowed to have arbitrary cardinality, we use the space of all information structures with fixed S, $\mathcal{M}(S)$, as the set of all non-empty and closed subsets of $\Delta(S)$ endowed with the topology generated by the Hausdorff distance between closed subsets of $\Delta(S)$.

Proposition 3.

(a) If $|M| \leq |S|$ then FIE is allowed in a set of full Lebesgue measure.[12]

(b) If $|M| \geq |S| + |A| - 1$, then FIE fails in a set of positive Lebesgue measure.

(c) There exists an open set $V \subset \mathcal{M}(S)$ such that FIE fails for each $M \in V$.

[12]Using a similar argument as in the proof of Proposition 2, it is simple to show that FIE is also allowed in a set of full measure when $|M| = |A| + 1$ and $|S| > 2$, under Assumption 1.

Part (b) requires $|M|$ to be sufficiently larger than $|S|$ for us to be able to find a finite-dimensional subspace of information structures where FIE fails in a set of positive measure. Part (c) states that in the space of all possible information structures with fixed number of signals, FIE cannot be generic.[13]

4. Equilibrium

In this section we fix the environment $\{u, \lambda\}$ and let the number of voters n grow to show that if the linear refinement condition is satisfied, then there exists a sequence of equilibrium strategies with limit strategy satisfying FIE. Thus, the linear refinement condition is also necessary and sufficient for FIE in equilibrium. We use an alternative notion of FIE, FIE*, that does not impose symmetric strategies, and we show that FIE is equivalent to FIE*. Hence, if there is a sequence of asymmetric equilibrium strategy profiles where the correct alternative wins almost surely, then there is also a sequence of symmetric equilibrium strategy profiles that achieves the same. Therefore, our earlier focus on symmetric strategies is without loss of generality.

For an electorate of size n, each player's strategy set can be viewed as the set of $|A| \times |S|$ matrices with entries in $[0, 1]$ with each row as a point in $\Delta(S)$. Hence, the common strategy set is a compact subset of $\mathbb{R}^{|A||S|}$. Let us use $a^{(n)} = (a^1, ..., a^n)$, $s^{(n)} = (s^1, ..., s^n)$, and $\sigma^{(n)} = (\sigma_1, ..., \sigma_n)$ to denote profiles of voters' choices, signals, and voters' strategies, respectively. Let A^n and S^n denote the sets of n-fold profiles of choices and signals. Let P^n be the joint distribution over $S^n \times \Delta(S)$, given by $P^n(C \times D) = \int_C \mu^n(D) d\lambda(\mu)$, for Borel sets C and D in $\Delta(S)$ and S^n respectively, where μ^n is the n-fold product of μ's. Let P^n_S denote the marginal of P^n over S^n and $P^n(\cdot|s^{(n)})$ the conditional over $\Delta(S)$ given a realization $s^{(n)} \in S^n$. For each alternative $a \in A$, let $\varphi_a(a^{(n)})$ denote the probability that alternative a is the outcome of the election at the profile $a^{(n)}$. That is, $\varphi_a(a^{(n)})$ is equal to 0 if a is not an alternative

[13] The set of information structures allowing FIE cannot be a dense set, and therefore it cannot be generic, in either a topological or a measure-theoretic sense. Indeed, as $\Delta(S)$ has a complete metric, $\mathcal{M}(S)$ is also a complete metric space and Baire's Theorem implies that FIE cannot be a residual set. Likewise, because it is not dense, FIE cannot be a *prevalent* set (see Ott and Yorke (2005) for a survey of *Prevalence* as the extension of finite-dimensional Lebesgue measure genericity to infinite dimensional spaces.)

with most votes in $a^{(n)}$ and it is equal to $1/q$ when a is one of the q alternatives with most votes in $a^{(n)}$.

The common ex ante utility at the strategy profile $\sigma^{(n)}$ is

$$v(\sigma^{(n)}|u,\lambda) \equiv \sum_{s^{(n)}} \int_M \sum_{a^{(n)}} \sum_a u(a,\mu)\varphi_a(a^{(n)}) \prod_{i=1}^n \sigma_i(s^i, a^i) dP^n(\mu|s^{(n)}) P_S^n(s^{(n)}).$$

Observe that it is continuous in $\sigma^{(n)}$. This defines a voting game, G^n.

Definition 6. *A sequence of strategy profiles $\{\sigma^{(n)}\}_{n>0}$ satisfies FIE* if $v(\sigma^{(n)}|u,\lambda) \to \bar{u}$, where $\bar{u} = \int_M \max_a u(a,\mu) d\lambda(\mu)$ is the common payoff when a correct alternative is chosen a.s.*

When there exists a sequence satisfying FIE* we say that the environment $\{u, \lambda\}$ allows FIE*. Observe that a sequence satisfying FIE* need not be symmetric.

Proposition 4. *An environment allows FIE if and only if it allows FIE*.*

Thus, our linear refinement condition also characterizes FIE*.

Theorem 2. *If the environment allows FIE, then for any sequence of strategy profiles $\{\sigma^{(n)}\}_{n>0}$ such that $\sigma^{(n)} \in \arg\max v(\sigma^{(n)}|u,\lambda)$ for every n, we have that: (i) $\sigma^{(n)}$ is a Nash equilibrium of G^n and (ii) $\{\sigma^{(n)}\}_{n>0}$ satisfies FIE*.*

Thus, in environments where FIE is feasible, FIE can be satisfied by a sequence of Nash equilibria. As such, in a common-value environment, failures of FIE are due to a failure of feasibility rather than issues with incentives.

5. Extensions

Thus far, we have studied environments allowing FIE under the assumptions of plurality voting, common-values, and finitely many signals. In this section, we show how our main insights extend to other voting rules, heterogeneous preferences, and environments with an infinite number of signals.

5.1. **Voting Rules.**

A number of well-known voting rules, including the scoring rules in Myerson (2002), can be captured by expanding the space of actions. The family of scoring rules is characterized by a feasible ballot $V \subset [0,1]^A$ which a voter is allowed to submit. A vector $v \in V$ is a submitted vote, with its coordinate a, v_a, representing the points assigned to alternative $a \in A$. The alternative with most points across all voters wins the election. Within this family of scoring rules, we introduce the following voting rules by imposing restrictions on V.[14]

Definition 7.

(i) Unrestricted scoring: $V^{us} = [0,1]^A$.
(ii) Scoring voting: $V^{sv}(\hat{v}) = \{v \in V^{us} : v_a \in \{\frac{j}{\hat{v}} : j = 0, ..., \hat{v}\}\}$, $\hat{v} \in \mathbb{N}_+$.
(iii) Storable voting: $V^{st}(\hat{v}) = \{v \in V^{sv}(\hat{v}) : \sum_a v_a = 1\}$, $\hat{v} \in \mathbb{N}_+$.
(iv) Approval voting: $V^{ap} = V^{sv}(1)$.
(v) Plurality with abstention: $V^{pa} = \{v \in V^{ap} : \sum_a v_a \in \{0,1\}\}$.
(vi) Plurality: $V^{p} = \{v \in V^{pa} : \sum_a v_a = 1\}$.

Using $\Delta(V)$ as the set of Borel probability measures over V, a mixed strategy $\sigma : S \to \Delta(V)$ specifies, for each signal, a probability measure over ballots. Let $\bar{\sigma}(s) = \int_V v d\sigma(v|s)$, so that $\bar{\sigma}_a(s)$ is the expected vote score to alternative a by a voter with signal s. Then, our notion of FIE is adapted to scoring rules by replacing $E_\mu(\sigma^a)$ by $E_\mu(\bar{\sigma}_a) \equiv \sum_{s\in S} \mu_s \bar{\sigma}_a(s)$ in Definition 1. Our next result shows that our analysis from Section 3 extends to the broader family of scoring rules, and that in a large election, plurality performs as well as any rule in this family.

Theorem 3. *If an environment allows FIE in some rule in Definition 7 then it allows FIE in all rules in Definition 7.*

Theorem 3 extends Proposition 2 in Ahn and Oliveros (2016) which establishes that for three alternatives, an environment allows FIE under Myerson's (A,B)-scoring rule

[14]Our definition for the different scoring rules are standard but normalized to take values between 0 and 1 for simplicity and nothing prevents us from allowing for integer values above 1. For example, storable (or qualitative) voting is a voting system where there is a fixed budget of votes to be allocated across alternatives.

if and only if it allows FIE under (iv). Of course, Theorem 3 is only for large elections: Proposition 1 in Ahn and Oliveros (2016), establishes that, in a finite electorate, for any equilibrium in (vi) there is an equilibrium in (iv) yielding at least as much utility for the voters.[15] From Theorem 3, these differences vanish in large elections.

Unrestricted scoring contains the outcomes in all scoring rules, including some rules not listed in Definition 7. For example, Borda count, with $V^{bc} = \{v \in [0,1]^A : v \text{ is a garbling of } \{\frac{i}{|A|-1} : i = 0, ..., |A|-1\}\} \subset [0,1]^A$ is one such rule. Theorem 3, therefore implies that if an environment allows FIE in Borda count, it allows FIE under any rule in Definition 7.

Our scoring rules are symmetric across alternatives. Asymmetric rules such as non-unanimous supermajority may generally pose additional constraints on FIE. However, for two alternatives this is not the case. For $A = \{a, \hat{a}\}$, let a q-rule be a voting rule where alternative a requires $q \in (0,1)$ share of votes or more to win the election. We establish next that all q-rules are equivalent for FIE.[16]

Proposition 5. *Fix $A = \{a, \hat{a}\}$ and $q \in (0,1)$. An environment allows FIE in a q-rule if and only if the environment allows FIE in plurality.*

As a final consideration, observe that we only consider one-shot simultaneous voting protocols that compute the winner by selecting the alternative with most votes. There are many other mechanisms that have multiple rounds (i.e., runoff elections) or with contingencies that depend on votes and may select more than one winner (i.e., single transferrable vote). Their results are beyond the scope of this paper, but we note that depending on the specificities of these protocols, elections with feedback will tend to expand the set of environments that allow for FIE.

5.2. **Heterogeneous Preferences.**

---

[15] The proof Theorem 3 establishes that this result extends to the other rules in Definition 7.

[16] Our result that all non-unanimous threshold voting rules have the same properties for FIE is mirrored in Feddersen and Pesendorfer (1997), where FIE holds for all such threshold rules. Gerardi and Yariv (2007) also show that if pre-voting deliberation is allowed, then in any common value environment all non-unanimous threshold rules have the same set of equilibrium outcomes.

An environment where payoffs depend on private signals can be captured by a utility function $u : A \times M \times S \to \mathbb{R}$, so the signal not only reveals information about the state but also directly affects a voter's preferences. In such an environment, there are alternative ways of defining majority preference under complete information. We assume that if the state is commonly known each agent votes their top-alternative non-strategically. This can be captured as follows. Set $\tilde{u}(a,\mu,s) = 1$ if $u(a,\mu,s) > u(a',\mu,s)$ for all $a' \neq a$ and $\tilde{u}(a,\mu,s) = 0$ otherwise. Using $\alpha_{\tilde{u}}(\mu) = \arg\max_{a\in A} \sum_{s\in S} \tilde{u}(a,\mu,s)\mu(s)$, we can readily employ the definition of FIE as before, and establish that all our feasibility results apply except for Theorem 2, which is our only result that depends explicitly on voters' incentives and therefore relies on the common-value assumption. Notice that this general setting can encompass many different environments, such as the ones studied in Feddersen and Pesendorfer (1997) and Bhattacharya (2013).

5.3. **Infinite Signals and General State Space.**

Let S be a topological, infinite space of signals, endowed with its Borel sigma-algebra. The primitives are: (i) a state space Θ, which is a measure space endowed with a sigma-algebra; (ii) a transition probability $P : \Theta \to \Delta(S)$, where $\Delta(S)$ is the space of probability measures over S, endowed with the weak* topology and its Borel sigma-algebra;[17] (iii) a prior $\nu \in \Delta(\Theta)$ with supp $\nu = \Theta$, and (iv) a common-value utility function $\tilde{u} : A \times \Theta \to \mathbb{R}$. We assume that P and $\tilde{u}$ are measurable mappings. Also, we assume that $M = P(\Theta)$ is closed and there is a measurable $u : A \times M \to \mathbb{R}$ such that $\alpha_{\tilde{u}} = \alpha_u \circ P$, where $\alpha_{\tilde{u}} : \Theta \to 2^A$ is given by $\alpha_{\tilde{u}}(\theta) = \arg\max_{a\in A} \tilde{u}(a,\theta)$ and $\alpha_u : M \to 2^A$ is given by $\alpha_u(\mu) = \arg\max_{a\in A} u(a,\mu)$, as before. As such, it is without loss to focus on the model with $M \subset \Delta(S)$ as the set of states and common utility $u : A \times M \to \mathbb{R}$.

[17]We abuse notation by using $\Delta(S)$ to denote both the finite-dimensional simplex (when S is finite) as well as the set of probability measures defined on the given sigma-algebra of S (when S is not finite). The (weak*) topology defined in $\Delta(S)$ allows us to derive the Borel sigma-algebra of measurable sets, which is then used to define probability measures over $\Delta(S)$, like the prior λ defined below.

Given ν and P, we define $\lambda \in \Delta(\Delta(S))$ by the formula $\lambda(E) = \nu(P^{-1}(E))$ for each measurable set E in $\Delta(S)$. Observe that $M = \text{supp } \lambda$. A mixed strategy is defined as a Borel measurable function $\sigma : S \to \Delta(A)$, and we use $\sigma^a(s)$ to denote the probability of voting for $a \in A$ at signal s. A linear utility function v is given by $v(a, \mu) = \int b^a(s) d\mu(s)$, where $b^a : S \to \mathbb{R}$ is a bounded measurable function for each $a \in A$. We can now state:

Theorem 4. *An environment $\{u, \lambda\}$ allows FIE if and only if it satisfies the linear refinement condition.*

5.3.1. **Sufficient Conditions with Infinitely many States and Signals**.

The betweenness condition is obviously a sufficient condition for FIE with infinitely many signals. As before it is verified if one can solve an appropriate system of equations, with the difference now that it comprises of linear equations in an infinite dimensional setting. Let us extend the arguments leading to Remark 1 to this case. We say that a finite subset $F \subset \Delta(S)$ of probability measures satisfies independence if $\sum_{\mu \in F} \beta_\mu \mu(E) = 0$ for all measurable sets $E \subset S$ implies that $\beta_\mu = 0$ for all $\mu \in F$.

Definition 8. *An environment satisfies* independence* *if every finite subset of M satisfies independence.*

Proposition 6. *If S is compact metric, each $\mu \in M$ has a jointly continuous density with respect to a fixed $\bar{\mu} \in \Delta(S)$, and u is continuous in μ, then the environment allows FIE if independence* is satisfied.*

The following example illustrates.

Example 4. *Suppose $A = \{a, \hat{a}\}$, let $\Theta = [0, 1]^2$ with a uniform prior ν, and assume the common utility function $\tilde{u}$ is such that $\tilde{u}(a, \theta) > \tilde{u}(\hat{a}, \theta)$ (resp. $\tilde{u}(a, \theta) < \tilde{u}(\hat{a}, \theta)$) if $\theta_1 > \theta_2$ (resp. $\theta_1 < \theta_2$), where θ_i is the ith coordinate of θ. Signals are given by $s = \theta + \varepsilon$, where ε is a bivariate normally distributed vector with mean $(0, 0)$, independent of θ. Thus $S = \mathbb{R}^2$ and P sends each θ to the corresponding bivariate normal distribution with mean θ. Define $u : A \times M \to \mathbb{R}$ by $u(a, \mu) = \tilde{u}(a, P^{-1}(\mu))$ and $u(\hat{a}, \mu) = \tilde{u}(\hat{a}, P^{-1}(\mu))$, where $M = P(\Theta)$.*

This environment allows FIE. To verify, let $b : \mathbb{R}^2 \to \mathbb{R}$ *be given by* $b(s) = 1$ *if* $s_1 \geq s_2$ *and* $b(s) = 0$ *otherwise, so* $\int b(s)d\mu(s) = \mu(D)$*, where* $D = \{s \in \mathbb{R}^2 : s_1 \geq s_2\}$*. It follows that* $\int b(s)d\mu(s) > \int b(s)d\hat{\mu}(s)$ *for all* $\mu \in M_a^u$ *and* $\hat{\mu} \in M_{\hat{a}}^u$*, so betweenness is satisfied.*

5.3.2. **Genericity with Infinitely many Signals**.

Finally, we provide an extension of Proposition 3 with infinitely many signals. As in part (c), we will resort to topological statements because the ambient spaces we will consider are infinite dimensional. As a counterpart of part (a), we shall work with $S \subset \mathbb{R}$ and M as a finite subset of the infinite dimensional set of probability measures $\Delta(S)$. In addition, we shall assume that each $\mu \in M$ has an analytic density with respect to a fixed $\bar{\mu} \in \Delta(S)$, so M can be identified with a function $f : S \to \mathbb{R}^{|M|}$ with analytic coordinates (the densities of μ with respect to $\bar{\mu}$). We shall use the $C^\infty(S, \mathbb{R}^{|M|})$ topology when we talk about open sets of subsets M with a fixed cardinality $|M| < \infty$, with M identified with $f \in C^\infty(S, \mathbb{R}^{|M|})$. Let $\mathcal{M}_{\bar{\mu}}^{|M|}(S)$ denote the corresponding space. For the counterpart of Proposition 3(c) we shall again work with the space $\mathcal{M}(S)$ of closed subsets of $\Delta(S)$ endowed with the topology generated by the Hausdorff metric.

Proposition 7.

(a) For any compact $S \subset \mathbb{R}$ *and* $|M| < \infty$*, FIE is allowed in an open and dense subset of* $\mathcal{M}_{\bar{\mu}}^{|M|}(S)$*.*

(b) For any separable metrizable space S*, there exists an open set* $V \subset \mathcal{M}(S)$ *such that FIE fails for each* $M \in V$*.*[18]

6. Conclusion

The existing literature on information aggregation in large elections has largely focused on specific preference and information environments. We instead consider general environments with arbitrary preference and information structures and focus on properties of the environment allowing or precluding information aggregation. We

[18] If S is, in a addition, metrizable with a complete metric, then $\mathcal{M}(S)$ is also a complete metric space and Baire's Theorem implies that FIE cannot be a residual set. As in Section 3.2.3, because FIE is not dense, it cannot be a prevalent set.

characterize environments that allow FIE as those that satisfy a linear refinement condition, that is, existence of a utility function that is linear in distributions of signals and whose top-alternatives agree with the given utility function of the voters. Geometrically, this means that FIE is characterized by special kinds of partitions of the simplex of distributions over signals. However special, such partitions are more general than the partitions associated with previous conditions from the literature. Notwithstanding, FIE is robustly unachievable among all possible information structures for a given fixed set of signals.

We also show that feasibility issues cannot be avoided by changing the scoring rule: FIE is feasible with plurality if and only if it is feasible under the wide class of scoring rules. Furthermore, even though our equilibrium results do not extend to the case of diverse preferences, our feasibility result is not constrained to common preferences.

Our work does not consider contingent voting rules, which have the potential for better information aggregation properties. It can be shown that rules such as sequential or single transferrable voting can circumvent some of our linearity restrictions, opening an exciting avenue of work. Future analysis could consider contingent voting rules, the additional restriction imposed by equilibrium in environments with heterogeneous preferences, correlated signals or environments where information is not invariant to the population size (e.g., if it is costly to acquire information).

7. Appendix

7.1. **Proof of Proposition 1.**

For convenience, let us represent A as $\{1, ..., k\}$, and use indices i, j, l to indicate alternatives so labeled. Given Assumption 1 , let $E \subset M$ with $\lambda(E) = 1$ be such that $|\alpha_u(\mu)| = 1$ for every $\mu \in E$ and $\lambda(E \cap M_i^u) > 0$ for every $i = 1, ..., k$ (since we assume throughout that $\lambda(\{\mu \in M : i \in \alpha_u(\mu)) > 0)$. Suppose that an environment $\{u, \lambda\}$ allows FIE, so we have a refining linear v with the corresponding $b^i \in \mathbb{R}^{|S|}$ for each $i = 1, ..., k$ defining a family $\{h_{ij}\}$ of hyperplanes $h_{ij} = \{x : x \cdot b_{ij} = 0\}$ with $b_{ij} = b^i - b^j$ and hence $h_{ij} = h_{ji}$ and $h_{ij}^+ = h_{ji}^-$, for each pair i, j of distinct alternatives. Let $F \subset E$ with $\lambda(F) = 1$ be such that $\alpha_v(\mu) = \alpha_u(\mu)$ for all $\mu \in F$. To verify (i), if $\mu \in F \cap M_i^u$ and $\mu \notin h_{ij}^+$, then $j \in \alpha_v(\mu)$, contradicting $\alpha_v(\mu) = \alpha_u(\mu) = \{i\}$.

Condition (ii) is immediate, as $x \in h_{ij}^+ \cap h_{jl}^+$ means that $x \cdot b^i > x \cdot b^j$ and $x \cdot b^j > x \cdot b^l$, hence $x \in h_{il}^+$. Similarly, we must also have $h_{ij} \cap h_{jl} \subset h_{il}$. To verify (iii), observe first that because the intersection of the kernels of the linear mappings defined by b_{ij} and b_{jl} is contained in the kernel of the linear mapping defined by b_{il}, it must be that b_{il} is a linear combination of b_{ij} and b_{jl}. That is, there are scalars γ_{ij} and γ_{jl} such that $b_{il} = \gamma_{ij} b_{ij} + \gamma_{jl} b_{jl}$. If $h_{ij} = h_{il}$, then $x \in h_{il}$ implies that $x \in h_{ij}$, so $0 = x \cdot b_{il} = x \cdot (\gamma_{ij} b_{ij} + \gamma_{jl} b_{jl}) = x \cdot \gamma_{jl} b_{jl}$, and hence $x \in h_{jl}$. This means that the hyperplane h_{jl} contains the presumed equal hyperplanes h_{ij} and h_{jl}. As subspaces of co-dimension 1, this can only happen if $h_{jl} = h_{ij} = h_{il} \equiv h$. This implies that one of the three alternatives, say alternative i, can only be optimal for v at states in h. But then $M_i \cap F = \emptyset$, for otherwise $j \in \alpha_v(\mu) = \alpha_u(\mu) = \{i\}$. Hence $\lambda(M_i^u) = 0$ contradicting that $\lambda(E \cap M_i^u) > 0$ for every $i = 1, ..., k$.

For the other direction, suppose that an environment $\{u, \lambda\}$ has a family of hyperplanes $\{h_{ij}\}$ with $h_{ij} = h_{ji}$ and $h_{ij}^+ = h_{ji}^-$ satisfying (i), (ii) and (iii). Let b_{ij} be the normal of h_{ij}, and observe that $b_{ji} = -a b_{ij}$ for some $a > 0$. Let $G \subset E$ with $\lambda(G) = 1$ be such that, for each i, $\lambda(G \cap M_i^u) > 0$ and $G \cap M_i^u \subset h_{ij}^+$ for every $j \neq i$. Observe that $h_{lj}^+ \cap h_{ji}^+ \subset h_{li}^+$ implies $h_{ij}^- \cap h_{jl}^- \subset h_{il}^-$. This means that $h_{ij} \cap h_{jl} \subset h_{il}$. In fact, otherwise there would exist x with $x \cdot b_{ij} = x \cdot b_{jl} = 0$ and $x \cdot b_{il} \neq 0$. Say $x \cdot b_{il} > 0$. Then because $h_{ij} \neq h_{il}$, we would be able to find $\hat{x}$ close to x with $\hat{x} \in h_{ij}^- \cap h_{jl}^-$ and $\hat{x} \in h_{il}^+$, a contradiction. A similar contradiction would be obtained if instead $x \cdot b_{il} < 0$. As above, there are scalars γ_{ij} and γ_{jl} such that $b_{il} = \gamma_{ij} b_{ij} + \gamma_{jl} b_{jl}$. We claim that γ_{il} and γ_{jl} must be positive. If both are negative, then $x \in h_{ij}^+ \cap h_{jl}^+$ implies that $x \cdot b_{il} < 0$, that is, $x \in h_{il}^-$, a contradiction. If they have opposite signs, say, $\gamma_{ij} > 0$ and $\gamma_{jl} < 0$, then take $x \in h_{il} \setminus h_{ij}$, so $\gamma_{ij} x \cdot b_{ij} = -\gamma_{jl} x \cdot b_{jl}$. There are two cases to consider: (a) $x \in h_{ij}^+$ and (b) $x \in h_{ij}^-$. In (a), we must have $x \cdot b_{jl} > 0$, so $x \in h_{ij}^+ \cap h_{jl}^+$ and thus $x \in h_{il}^+$, contradicting $x \in h_{il}$; in (b) we must have $x \cdot b_{jl} < 0$, so $x \in h_{ij}^- \cap h_{jl}^-$ and thus $x \in h_{il}^-$, again contradicting $x \in h_{il}$.

Consider now a family of scalars $\{\gamma_{ij}\}_{i<j}$ and a system of equations $\gamma_{il} b_{il} - \gamma_{ij} b_{ij} - \gamma_{jl} b_{jl} = 0$, for $1 \leq i < j < l \leq k$. There are ${}_kC_3$ equations but only ${}_{(k-1)}C_2$ are linearly independent. In fact, given i, j, l such that $1 < i < j < l$, note that $\gamma_{il} b_{il} - \gamma_{ij} b_{ij} - \gamma_{jl} b_{jl} = (\gamma_{1j} b_{1j} - \gamma_{1i} b_{1i} - \gamma_{ij} b_{ij}) - (\gamma_{1l} b_{1l} - \gamma_{1i} b_{1i} - \gamma_{il} b_{il}) + (\gamma_{1l} b_{1l} - \gamma_{1j} b_{1j} - \gamma_{jl} b_{jl})$. Thus, we can recover any equation that does not include alternative 1 as a combination of the ${}_{(k-1)}C_2$ equations that include alternative 1. Since there are ${}_kC_2$ variables, the system has multiple non trivial solutions. From what we showed above, any such solution will have positive coefficients.

Now choose a strictly positive vector $b^k \in \mathbb{R}^{|S|}$ and let $b^i = \gamma_{ik} b_{ik} + b^k$ for $i = 1, ..., k-1$ so we define a linear utility v with $v(\mu, i) = \mu \cdot b^i$. Observe that $b^i - b^j = \gamma_{ij} b_{ij}$ for all $i < j \leq k$. So for $\mu \in G \cap M_i^u$, we have $\mu \cdot (b^i - b^j) = \gamma_{ij} \mu \cdot b_{ij} > 0$ if $i < j$ and $\mu \cdot (b^i - b^j) = -\mu \cdot (b^j - b^i) = -\gamma_{ji} \mu \cdot b_{ji} = a\gamma_{ji} \mu \cdot b_{ij} > 0$ if $i > j$. It follows that for every $\mu \in G$, $\alpha_v(\mu) = \alpha_u(\mu)$, so v refines u. □

7.2. **Proofs from Section 3.2.**

Proof of Proposition 2.

We proceed by induction. Since $|M| = k$, let $M = \{\mu_1, ..., \mu_k\}$. Fix $b \in \mathbb{R}^{|S|}$ and suppose that $b \cdot \mu_1 = b \cdot \mu_2$. There must exist $\hat{s} \in S$ such that $\mu_1(\hat{s}) \neq \mu_2(\hat{s})$. Define $\hat{b} \in \mathbb{R}^{|S|}$ by setting $\hat{b}(s) = b(s)$ for all $s \neq \hat{s}$ and $\hat{b}(\hat{s}) = b(\hat{s}) + \varepsilon$ for some small $\varepsilon > 0$, so that $\hat{b} \cdot \mu_1 \neq \hat{b} \cdot \mu_2$. Now suppose that for some $\hat{k} < k$ it has been established the existence of $b \in \mathbb{R}^{|S|}$ such that $b \cdot \mu_\ell \neq b \cdot \mu_{\ell'}$ for all $\ell, \ell' \leq \hat{k}$ and $\ell \neq \ell'$. Now say that there is $\ell \leq \hat{k}$ such that $b \cdot \mu_\ell = b \cdot \mu_{\hat{k}+1}$. Since $\mu_\ell \neq \mu_{\hat{k}+1}$, there must exist $\hat{s}$ such that $\mu_\ell(\hat{s}) \neq \mu_{\hat{k}+1}(\hat{s})$, and because there are only finitely many μ's, we can find $\varepsilon > 0$ such that $\hat{b} \in \mathbb{R}^{|S|}$ defined by setting $\hat{b}(s) = h(s)$ for all $s \neq \hat{s}$ and $\hat{b}(\hat{s}) = b(\hat{s}) + \varepsilon$ satisfies $\hat{b} \cdot \mu_\ell \neq \hat{b} \cdot \mu_{\ell'}$ for all $\ell, \ell' \leq \hat{k} + 1$ This concludes the induction and establishes the existence of a vector $b \in \mathbb{R}^{|S|}$ such that $b \cdot \mu \neq b \cdot \hat{\mu}$ for every $\mu, \hat{\mu} \in M$. Since there are as many points as alternatives and there exists a family of parallel hyperplanes separating each alternative, the conditions in Proposition 1 are satisfied. □

Proof of Proposition 3.

Part (a) follows from SM. For part (b), consider an environment where M_a^u is equal to the set of vertices of the simplex $\Delta(S)$, and $M_{a'}^u$ is a singleton that lies in a small neighborhood of the barycenter of $\Delta(S)$, for some $a' \neq a$. This is possible because $|M| \geq |S| + |A| - 1$. Obviously M_a^u cannot be separated from $M_{a'}^u$ by the restriction of a hyperplane to the simplex, so FIE fails. Let H denote the corresponding $|M| \times |S|$ matrix, with rows given by the states. It is clear that any $(|M| \times |S|)$-matrix $\hat{H}$

sufficiently close to H, with rows lying in $\Delta(S)$, defines a subset $\hat{M}$ of $\Delta(S)$ with $|M|$ elements. For each such $\hat{M}$, we can then assign an environment $\{\hat{u}, \hat{\lambda}\}$ with $\hat{M}_a^{\hat{u}}$ given by the states closest (or equal) to the vertices of $\Delta(S)$ and $\hat{M}_{a'}^{\hat{u}}$ equal to a singleton and still inside a small neighborhood of the barycenter, for some $a' \neq a$. Each such environment will fail FIE. This means that we can find an open set in $\mathbb{R}^{|M|(|S|-1)}$ where FIE fails. Since it is open, it must have positive Lebesgue measure.

For part (c), let $M = \Delta(S)$ and have M_a^u be the boundary of $\Delta(S)$ (that is, the set of states μ such that $\mu(s) = 0$ for at least one $s \in S$). For $a' \neq a$, let $M_{a'}^u$ be an arbitrary subset of the interior of $\Delta(S)$ (disjoint from M_a^u for $a \neq a'$, obviously). There's no way to separate M_a^u from any such $M_{a'}^u$ with the restriction of a hyperplane to $\Delta(S)$. So FIE fails for this environment. Take any closed subset $\hat{M}$ of $\Delta(S)$ sufficiently close to M in Hausdorff distance, and assign to it an environment $\{\hat{u}, \hat{\lambda}\}$ such that $\hat{M}_a^{\hat{u}} = \arg\max_{\mu \in \hat{M}} p \cdot \mu$ for some $p \in \mathbb{R}^{|S|-1}$ and the remaining $\hat{M}_{a'}^{\hat{u}}$ for $a' \neq a$ as disjoint subsets of $\hat{M} \setminus \hat{M}_a^{\hat{u}}$. It is again not possible to separate $\hat{M}_a^{\hat{u}}$ from $\hat{M}_{a'}^{\hat{u}}$ with the restriction of a hyperplane to the simplex, so FIE fails for all such $\hat{M}$ sufficiently close to M. We can thus construct open set $V \subset \mathcal{M}(S)$ as an open neighborhood of M, where FIE fails. $\square$

7.3. **Proofs from Section 4.**

Proof of Proposition 4.

Observe first that $v(\sigma^{(n)}|u, \lambda)$ can be equivalently expressed as

$$v(\sigma^{(n)}|u, \lambda) = \int_M \sum_{s^{(n)}} \sum_{a^{(n)}} \sum_a u(a, \mu)\varphi_a(a^{(n)}) \prod_{i=1}^{n} \sigma_i(s^i, a^i)\mu^n(s^{(n)})d\lambda(\mu).$$

If σ satisfies FIE then the sequence of symmetric profiles $\{\sigma^n\}_{n>0}$, where each coordinate of σ^n is equal to σ, forms a sequence satisfying FIE*. Indeed, for a.e. $\mu \in M$, we have $\mu \cdot \sigma^a \geq \mu \cdot \sigma^{a'}$ for some $a \in \alpha_u(\mu)$, and all $a' \neq a$, with strict inequality for $a' \notin \alpha_u(\mu)$. Let $Q_{\mu,\sigma} \in \Delta(S \times A)$ be given by $Q_{\mu,\sigma}(s, a) = \mu(s)\sigma^a(s)$, and let $Q_{\mu,\sigma}^n$ and $Q_{\mu,\sigma}^\infty$ denote the n-fold and infinite products of $Q_{\mu,\sigma}$'s, respectively. For a given $\bar{a} \in A$, let $f^{\bar{a}}(s, a) = 1$ if $a = \bar{a}$, and $f^{\bar{a}}(s, a) = 0$ otherwise. Let $f : S \times A \to \{0, 1\}^A$ be given by $f(s, a) = (f^{\bar{a}}(s, a))_{\bar{a} \in A}$. For each $i = 1, 2, ...$, let f_i denote the ith independent draw of f. By the SLLN, $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f_i = (\mu \cdot \sigma^a)_{a \in A}$

$Q^{\infty}_{\mu,\sigma}$-a.s. By Sanov's theorem,

$$\lim_{n\to\infty} Q^n_{\mu,\sigma}(E^n) = 1,$$

where $E^n = \left\{(s^{(n)}, a^{(n)}) \in S^n \times A^n : \sum_{a\in\alpha_u(\mu)} \varphi_a(a^{(n)}) = 1\right\}$. It follows that

$$1 \geq \lim_{n\to\infty} \sum_{s^{(n)}} \sum_{a^{(n)}} \sum_{a\in\alpha_u(\mu)} \varphi_a(a^{(n)}) Q^n_{\mu,\sigma}(s^{(n)}, a^{(n)}) \geq \lim_{n\to\infty} Q^n_{\mu,\sigma}(E^n) = 1,$$

which implies that

$$\lim_{n\to\infty} \sum_{s^{(n)}} \sum_{a^{(n)}} \sum_{a\notin\alpha_u(\mu)} \varphi_a(a^{(n)}) Q^n_{\mu,\sigma}(s^{(n)}, a^{(n)}) = 0,$$

and then

$$\lim_{n\to\infty} \sum_{s^{(n)}} \sum_{a^{(n)}} \sum_{a} u(a,\mu)\varphi_a(a^{(n)}) Q^n_{\mu,\sigma}(s^{(n)}, a^{(n)}) = \max_a u(a,\mu).$$

As

$$v(\sigma^n|u,\lambda) = \int_M \sum_{s^{(n)}} \sum_{a^{(n)}} \sum_{a} u(a,\mu)\varphi_a(a^{(n)}) Q^n_{\mu,\sigma}(s^{(n)}, a^{(n)}) d\lambda(\mu),$$

Lebesgue Dominated Convergence ensures that

$$\lim_{n\to\infty} v(\sigma^n|u,\lambda) = \int_M \max_a u(a,\mu) d\lambda(\mu) \equiv \bar{u},$$

so the sequence $\{\sigma^n\}_{n>0}$ satisfies FIE*.

Conversely, let $\{\sigma^{(n)}\}_{n>0}$ be a sequence of possibly asymmetric strategy profiles satisfying FIE*. For λ-a.e. μ, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{s} \mu(s)\sigma_i^{(n)}(s, a(\mu)) \geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{s} \mu(s)\sigma_i^{(n)}(s, a')$$

for some $a(\mu) \in \alpha_u(\mu)$, and all $a' \neq a$, with strict inequality for $a' \notin \alpha_u(\mu)$, by Kolmogorov's SLLN. Take a subsequence such that $\frac{1}{n}\sum_{i=1}^n \sigma_i^{(n)}$ converges and, still using n to index the subsequence, define $\tilde{\sigma} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \sigma_i^{(n)}$. It follows that, for λ-a.e μ, $\mu\cdot\tilde{\sigma}^{a(\mu)} \geq \mu\cdot\tilde{\sigma}^{a'}$, for some $a(\mu)\in\alpha_u(\mu)$ and all $a'\neq a$, with strict inequality for $a'\notin\alpha_u(\mu)$. Therefore, $\max_a \mu\cdot\tilde{\sigma}^a \subset \alpha_u(\mu)$ a.s., for otherwise there would exist a set E with $\lambda(E) > 0$ and $a' \notin \alpha_u(\mu)$ with $\mu\cdot\tilde{\sigma}^{a'} \geq \mu\cdot\tilde{\sigma}^{a(\mu)}$ for all $\mu\in E$, contradicting what we just established. That is, $\tilde{\sigma}$ satisfies FIE. $\square$

Proof of Theorem 2.

Let $\{\sigma^{(n)}\}_{n>0}$ be a sequence of profiles of strategies such that $\sigma^{(n)} \in \arg\max v(\sigma^{(n)}|u,\lambda)$. Part (i) follows from McLennan (1998)'s Theorem 1. Let σ satisfy FIE and let σ^n denote the symmetric profile with all entries equal to σ. We then have

$$\bar{u} \geq \limsup_{n\to\infty} v(\sigma^{(n)}|u,\lambda) \geq \liminf_{n\to\infty} v(\sigma^{(n)}|u,\lambda) \geq \liminf_{n\to\infty} v(\sigma^{n}|u,\lambda) = \bar{u},$$

where the last equality follows from Proposition 4. This establishes part (ii). □

7.4. **Proofs from Section 5.**

Proof of Theorem 3.

In terms of restrictions imposed on strategy sets, and using "$(vi) \Rightarrow (v)$" to denote that if an environment allows FIE under (vi) then it also allows FIE under (v), the implications $(vi) \Rightarrow (v) \Rightarrow (iv) \Rightarrow (ii) \Rightarrow (i)$ and $(v) \Rightarrow (iii) \Rightarrow (ii)$ are immediate.[19] It remains to show that $(i) \Rightarrow (vi)$. Suppose σ allows FIE under (i). Define $R(s) = 1 - \frac{1}{|A|}\sum_{a\in S}\bar{\sigma}_a(s)$, and let $\sigma_a^{\mathrm{PV}}(s) \equiv \frac{1}{|A|}[\bar{\sigma}_a(s) + R(s)]$. For each s, we have $\sigma_a^{\mathrm{PV}}(s) \in [0,1]$ and $\sum_{a\in A}\sigma_i^{\mathrm{PV}}(s) = \frac{1}{|A|}\sum_{a\in A}\left(\bar{\sigma}_a(s) + 1 - \frac{1}{|A|}\sum_{a\in A}\bar{\sigma}_a(s)\right) = 1$, so $\sigma^{\mathrm{PV}}(s) \in \Delta(A)$ for every s. Take $\sigma_a^{\mathrm{PV}}(s)$ to be the probability that a voter with signal s gives a score of 1 to alternative a under voting rule (vi). As $\sum_s(\sigma_a^{\mathrm{PV}}(s) - \sigma_{\hat{a}}^{\mathrm{PV}}(s))\mu(s) = \frac{1}{|A|}\sum_s(\bar{\sigma}_a(s) - \bar{\sigma}_{\hat{a}})\mu(s)$, if a wins under σ under rule (i) then a wins under σ^{PV} under rule (vi) in a large election, by the SLLN. □

Proof of Proposition 5.

Suppose an environment allows FIE for some $q \in (0,1)$ and let σ be the strategy that satisfies FIE, with the interpretation that $\sigma(s)$ is the probability of voting for a given signal $s \in S$. Now consider any other $q' \in (0,1)$. Replacing σ by σ' given by $\sigma'(s) = q' + \varepsilon(\sigma(s) - q)$, where $\varepsilon > 0$ is chosen so that σ' is a strategy, we ensure it satisfies FIE given voting rule q'. □

Proof of Theorem 4.

[19] In addition, for a finite electorate and any two rules connected by "⇒", Theorem 2 in McLennan (1998) establishes that if there is an equilibrium in the first then there is an equilibrium in the second yielding at least as much utility – this extends Proposition 1 in Ahn and Oliveros (2016) to all rules in Definition 7.

For the "only if" direction, let σ satisfy FIE and define $b^a : S \to \mathbb{R}$ as $b^a(s) = \sigma^a(s)$ for every $s \in S$. As b^a is bounded and measurable, $v(a, \mu) = \int b^a(s)d\mu(s)$ defines a linear utility function refining u.

For the "if" direction, suppose that there is a linear utility function $v(a, \mu) = \int b^a(s)d\mu(s)$ that refines u. Consider a linear $\hat{v} \in \mathcal{U}$ given by $\hat{v}(a, \mu) = \int \hat{b}^a(s)d\mu(s)$, where $\hat{b}^a(s) = \gamma b^a(s) + \frac{1}{|A|}(1 - \sum_a \gamma b^a(s))$ for every $s \in S$, with $\gamma > 0$ chosen to ensure that $\hat{b}^a(s) > 0$ for all $a \in A$ and $s \in S$. Indeed, since there exists $c > 0$ such that $b^a(s) \geq -c$ for every $s \in S$, any γ satisfying $0 < \gamma < \frac{1}{2c|A|}$ will do. Since $\hat{v}$ is affine to v, $\alpha_{\hat{v}}(\mu) = \alpha_v(\mu)$ for all μ, and hence $\alpha_{\hat{v}}(\mu) \subset \alpha_u(\mu)$ a.s. Define a mixed strategy by setting $\sigma^a(s) = \hat{b}^a(s)$ for all $a \in A$ and $s \in S$. Observe that indeed $\sigma(s) \in \Delta(A)$ and that σ is Borel measurable. Since $\arg\max_a E_\mu(\sigma^a) = \alpha_{\hat{v}}(\mu)$, σ satisfies FIE.

Proof of Proposition 6.

Consider a sequence of finite subsets M^m of M such that (i) $M^m \subset M^{m+1}$, (ii) $M^m \to M$ in Hausdorff sense, and (iii) the densities f_μ of state μ with respect to $\bar{\mu}$ are independent, for all $\mu \in \mathcal{M}^m$. We can do this because the environment satisfies independence. Again use $A = \{1, ..., k\}$.

Fix a list $c_1 > \cdots > c_k > 0$. By independence, for each m there is $b^m \in L_\infty(\bar{\mu})$ (in fact, we can choose b^m to have range in $[-1, 1]$) such that $\int b^m(s)f_\mu(s)d\bar{\mu}(s) = c_i$, for all $\mu \in M_i^{u,m}$, for all i, where $M_a^{u,m} = M^m \cap M_a^u$.

Now, because the Borel sigma-algebra in X is countably generated, $L_1(\bar{\mu})$ is separable and the weak* topology in $L_\infty(\bar{\mu})$ is metrizable. Also, because $\bar{\mu}$ is a probability measure, the norm dual of $L_1(\bar{\mu})$ is $L_\infty(\bar{\mu})$ (Aliprantis and Border (2006), Theorem 13.28), and hence by Alaoglu's theorem (Aliprantis and Border (2006), Theorem 6.21), the so constructed sequence b^m has a weak*-convergent subsequence. Let b be its limit. As $M_a^{u,m} \subset M_a^{u,m'}$ for $m' > m$, for each $\mu \in M_a^{u,m}$ for each i, we have $\int b(s)f_\mu(s)d\bar{\mu}(s) = c_a$.

Fix $\mu \in M_a^u$. Since u is continuous, we will have a sequence $(\mu^m)_m$ with $\mu^m \in M_a^{u,m}$ and $\mu^m \to \mu$. By continuity of f, for each s we have $f_{\mu^m}(s) \to f_\mu(s)$. So, by Lebesgue Dominated Convergence, we have $\int b(s)f_{\mu^m}(s)d\bar{\mu}(s) \to \int b(s)f_\mu(s)d\bar{\mu}(s)$ so it must be that $\int b(s)f_\mu(s)d\bar{\mu}(s) = c_a$, and we are done. $\square$

Proof of Proposition 7.

For (a), we will use the fact that a finite set of analytic functions $\{f_1, ..., f_{|M|}\}$, with $f_j : S \to \mathbb{R}$ for $j = 1, ..., |M|$, is linearly independent if and only if the Wronskian $W(f)$ evaluated at some s is not zero. Fix $s \in S$. Consider the set in $D = \{f \in \mathcal{M}_{\bar{\mu}}^{|M|}(S) : W(f)(s) \neq 0\}$. As the mapping $f \mapsto W(f)(s)$ is continuous, D is an open set. Now pick $g \in \mathcal{M}_{\bar{\mu}}^{|M|}(S)$ with $W(g)(s) \neq 0$. For any $f \in \mathcal{M}_{\bar{\mu}}^{|M|}(S)$ and $0 < \varepsilon \leq 1$, consider $(1-\varepsilon)f + \varepsilon g$, also an element of $\mathcal{M}_{\bar{\mu}}^{r}(S)$. The function $\varepsilon \mapsto W((1-\varepsilon)f+\varepsilon g)(s)$ is a polynomial function which is not identically equal to zero because it's non-zero when $\varepsilon = 1$. It has finitely many zeros, so for any ε as close as one pleases to zero, $W((1-\varepsilon)f+\varepsilon g)(s) \neq 0$. As $(1-\varepsilon)f+\varepsilon g$ converges to f in the $C^\infty(S, \mathbb{R}^{|M|})$ topology as $\varepsilon \to 0$, this establishes that D is dense. Hence the set of sets of states in $\mathcal{M}_{\bar{\mu}}^{|M|}(S)$ satisfying independence is open and dense. Now apply Proposition 6.

Move to (b). Let d denote the metric on S and consider two open balls $B(s_1, r)$ and $B(s_2, r)$ of radius $r > 0$, where $d(s_1, s_2) > 4r$. Set F_j to be the closure of $B(s_j, p)$ in S, for $j = 1, 2$. Consider their corresponding spaces of Borel probability measures $\Delta(F_1)$ and $\Delta(F_2)$. Let $M = \Delta(S)$ and consider an environment $\{u, \lambda\}$ such that $u(a, \mu)$ is the indicator function of the set $\Delta(F_1) \cup \Delta(F_2)$, $u(a', \mu) > 0$ for all $\mu \in \Delta(S) \setminus (\Delta(F_1) \cup \Delta(F_2))$, for all $a' \neq a$, and $\lambda(\Delta(F_j)) > 0$ for $j = 1, 2$. Thus $M_a^u = \Delta(F_1) \cup \Delta(F_2)$ and $M_{a'}^u \subset M \setminus M_a^u$ for all $a' \neq a$. Observe that for given $\mu_1 \in \Delta(F_1)$ and $\mu_2 \in \Delta(F_2)$, $\mu = \beta\mu_1 + (1-\beta)\mu_2 \notin M_a^u$ for any $\beta \in (0, 1)$. If FIE was possible, for λ-a.e. μ_1 and μ_2 in $\Delta(F_1)$ and $\Delta(F_2)$, respectively, and λ-a.e. $\mu = \beta\mu_1 + (1-\beta)\mu_2$ and $\beta \in (0, 1)$, there would exist a function $f : S \to \mathbb{R}$ such that $\int f d\mu_1 > \int f d\mu$ and $\int f d\mu_2 > \int f d\mu$. But $\int f d\mu = \beta \int f d\mu_1 + (1-\beta) \int f d\mu_2$, so we would have $\int f d\mu_1 > \int f d\mu_2 > \int f d\mu_1$, an absurd. So M fails FIE.

Given $0 < \varepsilon < r/2$, consider an ε-open neighborhood of M in $\mathcal{M}(S)$, that is, the set $V = \{\hat{M} \in \mathcal{M}(S) : d_H(\hat{M}, M) < \varepsilon\}$, where d_H denotes the Hausdorff distance. For each such $\hat{M}$, we must have $\hat{M} \cap \Delta(F_j) \neq \emptyset$, $j = 1, 2$, and $\hat{M} \cap M_{a'}^u \neq \emptyset$ for some $a' \neq a$. Indeed, because $\Delta(S)$ is endowed with the weak* topology, $\hat{M}$ contains a set $\{\delta_{\hat{s}_1}, \delta_{\hat{s}_2}, \delta_{\hat{z}}\}$, where $\max\{d(s_1, \hat{s}_1), d(s_2, \hat{s}_2), d(z, \hat{z})\} < \varepsilon$ for some z with $\min\{d(z, s_1), d(z, s_2)\} > 2r$. For each $\hat{M}$, pick $\hat{\lambda}$ with supp $\hat{\lambda} = \hat{M}$, $\hat{\lambda}(\Delta(F_j)) > 0$,

$j = 1, 2$, and and let $\hat{u}$ be the restriction of u to $\hat{M}$. The computations above show that FIE fails for any such $\hat{M} \in V$. □

7.5. **Green and Osband (1991).**

In this section, we connect our results to Green and Osband's. Let $\beta : M \to A$ and $w : A \times S \to \mathbb{R}$, with $w^a(s) \equiv w(a, s)$. Green and Osband (GO henceforth) employ the following notion:

Definition 9. *A function w rationalizes β if $\forall \mu \in M$, $E_\mu(w^{\beta(\mu)}) = \max_a E_\mu(w^a)$ and $\forall a \in A, \exists \mu \in M$ such that $\beta(\mu) = a$ and $\forall \hat{a} \neq a$, $E_\mu(w^{\hat{a}}) < E_\mu(w^a)$.*

GO provide conditions under which a given function β can be rationalized by some w, by focusing on the geometry of the partition defined by β, in which each element is given by $\pi_a = \{\mu \in M : \mu \in \beta^{-1}(a)\}$. We use Π to denote the partition $\{\pi_a\}_{a \in A}$, defined by β. We start by summarizing their condition.

For convenience, let us write $A = \{1, ..., k\}$ and use i and j to denote two generic elements of A. Let $\pi_{ij} \subset M$ be the relative interior of the intersection of the closures of π_i and π_j, if the latter has dimension $|S| - 2$ (it is equal to the empty set otherwise). Say that π_i and π_j are adjacent when $\pi_{ij} \neq \emptyset$. For any $m > 1$, a *circuit* is a function $\gamma_m : \{0, ..., m\} \to \Pi$ such that $\gamma_m(i) = \gamma_m(j)$ if and only if $i = 0$ and $j = m$, and $\pi_{\gamma_m(i)\gamma_m(i+1)} \neq \emptyset$ for all $i < m$. Set $L_{ij} = \{y \in \mathbb{R}^{|S|} : y = \alpha\omega + \beta(\mu - \hat{\mu})\}$ for α and β in $\mathbb{R}$, μ and $\hat{\mu}$ in π_{ij}, and ω equal to the unit vector in $\mathbb{R}^{|S|}$. A *flow* is a function $d : \Pi \times \Pi \to \mathbb{R}^{|S|}$ such that whenever $\pi_{ij} \neq \emptyset$, (i) $d(\pi_i, \pi_j)$ is orthogonal to L_{ij}, (ii) $(\mu - \hat{\mu}) \cdot d(\pi_i, \pi_j) > 0$ for μ, $\hat{\mu}$ in π_{ij}, and (iii) $d(\pi_i, \pi_j) = -d(\pi_j, \pi_i)$. Π satisfies the *integrability condition* if there exists a flow d such that for every $m > 1$ and for every elementary circuit γ_m, we have $\sum_{i=0}^{m} d(\gamma_m(i), \gamma_m(i+1)) = 0$.

A polyhedral convex subset of an affine space is defined to be the intersection of finitely many closed affine half-spaces. Let $H = \{x \in \mathbb{R}^{|S|} : \sum_{i=1}^{|S|} x_i = 1\}$. We say that the partition Π satisfies the *convex separating partition condition* if, for each $a \in A$, the relative interior of π_a is non-empty and there is a polyhedral convex set $C \subset H$ such that $\text{int}(C) \cap M \subset \pi_a \subset C$. That is, boundary of each partition element π_a is composed of finitely many pieces, each contained on a hyperplane in H.

Proposition 8. *An environment $\{u, \lambda\}$ with $M = supp\ \lambda$ convex and $|S| - 1$ dimensional, and λ absolutely continuous w.r.t. the Lebesgue measure allows FIE if and only if $\exists \beta : M \to A$, with $\beta(\mu) = \alpha_u(\mu)$ a.s., such that the convex separating partition and integrability conditions hold w.r.t. the partition generated by β.*

The proof follows from the following corollary and GO's Proposition 4:

Corollary 1. *An environment $\{u, \lambda\}$ with $M = supp\ \lambda$ convex and $|S| - 1$ dimensional, and λ absolutely continuous w.r.t. the Lebesgue measure allows FIE if and only if there exists w that rationalizes some β such that $\beta(\mu) = \alpha_u(\mu)$ a.s.*

Proof of Corollary 1. Let $v(a, \mu) = b^a \cdot \mu$ refine u. Consider any β such that $\beta(\mu) = \alpha_u(\mu)$ a.s. Let $F \subset M$ be the set of states where $\beta(\mu) = \alpha_u(\mu)$, $\alpha_v(\mu) \subset \alpha_u(\mu)$, and $|\alpha_u(\mu)| = 1$ for every $\mu \in F$. Now define β' by

$$\beta'(\mu) = \begin{cases} \beta(\mu) & \text{if } \mu \in F \\ \alpha_v(\mu) & \text{if } \mu \in M \setminus F. \end{cases}$$

Then, $\lambda(F) = 1$ and it follows that w given by $w(a, s) = b^a(s)$ rationalizes β' on M. Conversely, suppose $\beta(\mu) = \alpha_u(\mu)$ a.s. and let w rationalize β. Set $v(a, \mu) = w^a \cdot \mu$ so that $\beta(\mu) = \alpha_v(\mu)$ for all $\mu \in M$. Now let $F \subset M$ be the set of states where $\beta(\mu) = \alpha_u(\mu)$. Since $\lambda(F) = 1$, it follows that v refines u. □

References

Acharya, A. (2016): "Information Aggregation Failure in a Model of Social Mobility," *Games and Economic Behavior*, 100, 257 – 272.

Ahn, D. S. and S. Oliveros (2016): "Approval Voting and Scoring Rules with Common Values," *Journal of Economic Theory*, 166, 304 – 310.

Ali, S. N., M. Mihm, and L. Siga (2019): "Adverse Selection in Distributive Politics," Working paper.

Aliprantis, C. D. and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer-Verlag Berlin Heidelberg, 3 ed.

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

Bhattacharya, S. (2013): "Preference Monotonicity and Information Aggregation in Elections," *Econometrica*, 81, 1229–1247.

——— (2018): "Condorcet Jury Theorem in a Spatial Model of Elections," Working paper.

Bouton, L. and M. Castanheira (2012): "One Person, Many Votes: Divided Majority and Information Aggregation," *Econometrica*, 80, 43–87.

Callander, S. (2008): "Majority rule when voters like to win," *Games and Economic Behavior*, 64, 393 – 420, special Issue in Honor of Michael B. Maschler.

Chakraborty, A. and P. Ghosh (2003): "Efficient Equilbria and Information Aggregation in Common Interest Voting Games," Working paper.

Condorcet, M. d. (1785): "Essai sur l'application de l'analyse a la probabilité des decisions rendues a la pluralité des voix." Trans. McLean Iain and Hewitt Fiona, 1994. Paris.

Coughlan, P. J. (2000): "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting," *The American Political Science Review*, 94, 375–393.

Downs, A. (1957): *A n Economic Theory of Democracy*, New York: Harper & Bros.

Duggan, J. and C. Martinelli (2001): "A Bayesian Model of Voting in Juries," *Games and Economic Behavior*, 37, 259 – 294.

Ekmekci, M. and S. Lauermann (2020): "Manipulated Electorates and Information Aggregation," *The Review of Economic Studies*, 87, 997–1033.

——— (2022): "Information Aggregation in Poisson-Elections," *Theoretical Economics*, 17, 1–23.

Feddersen, T. and W. Pesendorfer (1997): "Voting Behavior and Information Aggregation in Elections With Private Information," *Econometrica*, 65, 1029–1058.

——— (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *The American Political Science Review*, 92, 23–35.

Gerardi, D. and L. Yariv (2007): "Deliberative Voting," *Journal of Economic Theory*, 134, 317 – 338.

Goertz, J. M. and F. Maniquet (2011): "On the Informational Efficiency of Simple Scoring Rules," *Journal of Economic Theory*, 146, 1464 – 1480.

Green, E. and K. Osband (1991): "A Revealed Preference Theory for Expected Utility," *The Review of Economic Studies*, 58, 677–696.

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

Krishna, V. and J. Morgan (2012): "Voluntary Voting: Costs and Benefits," *Journal of Economic Theory*, 147, 2083 – 2123.

Mandler, M. (2012): "The Fragility of Information Aggregation in Large Elections," *Games and Economic Behavior*, 74, 257 – 268.

Martinelli, C. (2002): "Convergence Results for Unanimous Voting," *Journal of Economic Theory*, 105, 278 – 297.

——— (2006): "Would Rational Voters Acquire Costly Information?" *Journal of Economic Theory*, 129, 225 – 251.

McLennan, A. (1998): "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents," *The American Political Science Review*, 92, 413–418.

McMurray, J. (2013): "Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses," *The Review of Economic Studies*, 80, 277–312.

——— (2017): "Ideology as Opinion: A Spatial Model of Common-Value Elections," *American Economic Journal: Microeconomics*, 9, 108–140.

Mihm, M. and L. Siga (2021): "A Behavioral Characterization of the Likelihood Ratio Order," *American Economic Review: Insights*, 3, 353–66.

Myerson, R. B. (1998): "Extended Poisson Games and the Condorcet Jury Theorem," *Games and Economic Behavior*, 25, 111 – 131.

——— (2002): "Comparison of Scoring Rules in Poisson Voting Games," *Journal of Economic Theory*, 103, 219 – 251.

Ott, W. and J. A. Yorke (2005): "Prevalence," *Bulletin of the American Mathematical Society*, 42, 263–290.

Persico, N. (2004): "Committee Design with Endogenous Information," *The Review of Economic Studies*, 71, 165–191.

Pesendorfer, W. and J. M. Swinkels (1997): "The Loser's Curse and Information Aggregation in Common Value Auctions," *Econometrica*, 65, 1247–1281.

Razin, R. (2003): "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates," *Econometrica*, 71, 1083–1119.

Siga, L. and M. Mihm (2021): "Information Aggregation in Competitive Markets," *Theoretical Economics*, 16, 161–196.

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

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

Indian Institute of Management Calcutta, Email:sourav@iimcal.ac.in.

Department of Economics, University of Essex, Email: lucas.siga@essex.ac.uk.