Manipulation in Elections with Uncertain Preferences

Andrew McLennan
School of Economics
Level 6, Colin Clark Building
University of Queensland
St Lucia, QLD 4072 Australia
mclennan@socsci.umn.edu

November 2007

Abstract

A decision scheme (Gibbard (1977)) is a function mapping profiles of strict preferences over a set of social alternatives to lotteries over the social alternatives. Motivated by conditions typically prevailing in elections with many voters, we say that a decision scheme is *weakly strategy-proof* if it is never possible for a voter to increase expected utility (for some vNM utility function consistent with her true preferences) by misrepresenting her preferences when her belief about the preferences of other voters is generated by a model in which the other voters are i.i.d. draws from a distribution over possible preferences. We show that if there are at least three alternatives, a decision scheme is necessarily a random dictatorship if it is weakly strategy-proof, never assigns positive probability to Pareto dominated alternatives, and is anonymous in the sense of being unaffected by permutations of the components of the profile. This result is established in two settings: a) a model with a fixed set of voters; b) the Poisson voting model of Meyerson (1998a,b, 2000, 2002).

Running Title: **Manipulation in Elections**

Journal of Economic Literature Classification Number D71.

Keywords: Gibbard-Satterthwaite theorem, strategy-proof, manipulation, voting, elections.

1 Introduction

Suppose we are given a finite set V of social alternatives and n voters. A *preference profile* is an n-tuple of strict individual preferences over V. A *social choice function* is a function whose domain is the set of preference profiles and whose range is V. The Gibbard-Satterthwaite theorem (Gibbard (1973), Satterthwaite (1975)) asserts that a social choice function must be dictatorial if there are at least three alternatives, the social alternative selected at a profile of preferences is never Pareto dominated for that profile, and the function is *strategy-proof*, which means that it is never possible for a voter to achieve a preferred outcome by reporting something other than her actual preference ordering.

When the electorate is large, voters typically have quite limited information about the preferences of other voters, so one should consider the possibility that a social choice function might be strategy-proof "in effect" if, in practice, voters are never able to manipulate because they lack sufficiently precise information. This paper develops a weakened notion of strategy-proofness that expresses this perspective. Our main results show that this weaker notion is still strong enough to imply a dictatorial conclusion.

Any mechanism combining the agents' preferences in a nontrivial manner must depart from the framework of the Gibbard-Satterthwaite theorem in some respect, and for this reason the result is fundamental in the theory of mechanism design. The Gibbard-Satterthwaite theorem allows voters to have any strict preferences, but, for example, in the theory of matching (e.g., Roth and Sotomayor (1990)) agents are typically assumed to care only about whether they are matched and, if so, with whom. A *domain restriction* specifies a subset of the set of preference profiles. In the theory of voting the seminal concept of this sort is the notion of single peaked preferences, which leads to the median voter theorem (e.g., Black (1958)). Allowing the voters (but not the mechanism designer) to know each others' preferences, and to behave with greater strategic sophistication, leads to the theory of Nash implementation pioneered by Maskin (1999). In Bayesian implementation the given information, for both the agents and the mechanism designer, includes a prior distribution on the space of n-tuples of agent types, a mechanism determines a Bayesian game, and Bayesian Nash equilibrium, rather than equilibrium in dominant strategies, is the preferred solution concept.

Another relaxation of the Gibbard-Satterthwaite framework, that is more closely related to the work presented here, is to allow a random outcome. Indeed, tied elections are commonly resolved by coin flips, so this extension is very natural, and the Gibbard-Satterthwaite theorem would lose much of its force if there were electoral systems employing randomization that embodied democratic values. Gibbard (1977) defines a *decision scheme* to be a function whose domain is the set of preference profiles and whose range is the set of probability distributions over V. Gibbard's result (a precise explanation is given in Section 2) reinforces the negative conclusion of the Gibbard-Satterthwaite theorem: if a decision scheme is strategy-proof in the sense that manipulation is never beneficial (when evaluated in terms of any von Neumann-Morgenstern utility consistent with the actual preference) and Pareto dominated alternatives never receive any probability, then the scheme must be a random dictatorship.

The main idea studied here might be thought of as a domain restriction, except that instead of imposing restrictions on the profiles that can occur, we impose restrictions on the voters' *beliefs* about the profile. Specifically, we assume that each voter's belief about the preferences of the other voters can be described by a model in which the other voters' preferences are i.i.d. draws from a common distribution. We are particularly motivated by elections with many voters and more than two candidates such as the primaries used to select the parties' candidates in the U.S. electoral system. Of course voters' beliefs in such a context are never exactly described by an i.i.d. model for various reasons, e.g., the preferences of members of married couples are believed to be correlated. A voter who is contemplating manipulation, but is uncertain about the profile, must average over different ways in which her vote might be pivotal, and our guiding intuition is that the averaging entailed by the i.i.d. assumption is a reasonably accurate approximation of the averaging resulting from the uncertainty voters face in practice. If this is correct, a decision scheme that never rewarded manipulation by voters with i.i.d. beliefs, and was not otherwise flawed, would merit serious consideration, and the existence of such decision schemes would challenge the pertinence of the Gibbard-Satterthwaite theorem.

Conversely, our finding that there are no satisfactory decision schemes of this sort would seem to be a significant strengthening of Gibbard's theorem. Insofar as this is a negative result, it is strengthened by any restriction imposed on the voters beliefs, and in this sense there is no need for us to defend the "realism" of the i.i.d. assumption. In principle attacks on its relevance should take the form of arguments to the effect that certain i.i.d. beliefs do not need to be considered.

Formally, a decision scheme is *weakly strategy-proof* if there is no voter, preference for that voter, von Neumann-Morgenstern utility consistent with that preference, and distribution over orderings of the alternatives, such that the voter can achieve a higher expected utility by manipulating when she regards the other voters' preferences as i.i.d. draws from that distribution. To illustrate this idea concretely, suppose that there are three voters, that from voter 1's point of view the preferences of the other two voters are i.i.d. random variables, and that voter 1 can do better by manipulating when voter 2 has preference P and voter 3 has preference ordering Q. The assumption that the preferences are i.i.d. implies that it is equally likely that voter 2 has preference Q and voter 3 has preference P. In addition, if these two events are much more likely than both voters having preference

P, then in turn both voters having preference Q must be much more likely than either of these events. The fact that there is a profile at which voter 1 can profitably manipulate does not necessarily imply that there is a belief for voter 1 satisfying our assumption at which profitable manipulation is possible.

In fact it is easy to see that there are decision schemes that never assign positive probability to Pareto dominated alternatives and are weakly strategy-proof, but not strategy-proof. For any particular voter with the sorts of beliefs we are allowing, and any particular profile, the voter regards all profiles obtained by permuting the preferences of the other voters as equally likely. For such a voter two decision schemes are effectively equivalent, if, for each profile, the two lotteries obtained by averaging the results of the two decision schemes over all permutations of the other voters' preferences are the same. Since the number $(n-1)!$ of permutations of the other voters grows rapidly with n, one can easily show, simply by counting equations and unknowns, that the set of decision schemes yielding a given system of averages can have high dimension. The *anonymous random dictatorship* is the decision scheme in which the probability of choosing a particular alternative is the fraction of the electorate that have that alternative as their favorite; in effect, a voter is selected according to an equiprobable lottery, and that voter's favorite alternative is the social choice. Starting with this decision scheme, it is not hard to construct examples that are weakly strategy-proof by virtue of giving the same averages, but are not strategy-proof.

These considerations suggest that we should restrict attention to decision schemes that are *anonymous* in the sense that permuting the voter's preferences does not affect the outcome. It is not easy to imagine how an electoral system might be regarded as democratic if it was not anonymous, and anonymity is certainly consistent with the spirit of our assumption concerning agents' beliefs. For any decision scheme there is a derived anonymous decision scheme, which we will call its *anonymization*, in which the lottery assigned to a profile is the average of the lotteries assigned by the given decision scheme to the profiles obtained by permuting the components of the profile. The anonymization of an anonymous decision scheme is the scheme itself, so the process of anonymization partitions the decision schemes into equivalence classes, each of which has an anonymous central element that is the anonymization of every element of the class. A decision scheme never assigns probability to Pareto dominated alternatives if and only if its anonymization also has this property. If a decision scheme is weakly strategy-proof, then so is its anonymization.

We can now state our first main result, which is proved in Section 3: *if a decision scheme is anonymous and weakly strategy-proof, and never assigns positive probability to Pareto dominated alternatives, and there are at least three alternatives, then it is the anonymous random dictatorship.* Since there is the additional hypothesis of anonymity, this is not, strictly speaking, a derivation of Gibbard's conclusion from weaker assumptions, but it seems correct to regard it as such conceptually because the additional possibilities allowed by dropping anonymity are trivial. In particular, there is the following corollary: *if there are at least three alternatives and a decision scheme is weakly strategy-proof and never assigns positive probability to Pareto dominated alternatives, then its anonymization is the anonymous random dictatorship.*

In the discussion to this point we have assumed that the set of voters is fixed. More precisely, we have assumed that each voter has no uncertainty about who the other voters are, even if their preferences are uncertain. In actual elections voters often have quite imprecise knowledge concerning the pool of eligible voters, and in addition there is uncertainty about which ones will actually turn out. Thus it is very natural to consider models in which the size of the electorate is uncertain.

From the point of view of the sort of result described above, what qualities are desirable in a random model of the electorate? Since the conclusion is negative—an "impossibility theorem"—the result is more forceful if the conclusion is shown to hold even when voter's beliefs are restricted to a relatively small set, since then the conclusion also holds when the voters' beliefs are less restricted. A bit more subtly, the result is more forceful if the model of beliefs is a natural limit of other models, since the conclusion for the other models can be derived from continuity. Finally, tractable models are preferred, of course. All these considerations strongly recommend the Poisson voting model developed by Meyerson (1998a,b, 2000, 2002) in which the number of voters with each preference is distributed according to a Poisson distribution, and these random variables are statistically independent. Section 4 extends the result described above to that setting.

2 Gibbard's Theorem

Our analysis builds on several lemmas proved in Gibbard (1977), and the reader must refer to that source if she wishes to obtain a complete understanding of the proof. In order to create a package that is as seamless as possible we follow the notation and terminology of that paper quite closely. This sections recapitulates the basic framework, and additional concepts from that paper are introduced in Section 3.

There is a nonempty finite set of *alternatives* V whose elements are denoted by x, y, and z. A (strict) *preference* over V is a complete transitive asymmetric binary relation on V. Such relations are denoted by P, Q, P_k, etc. A *utility scale* is a function $U : V \to \mathbb{R}$. The utility scale U is said to *fit* the preference P if more highly ranked alternatives give greater utility: for all $x, y \in V$, $U(x) > U(y)$ if and only if xPy. For any finite or countable set A, $\Delta(A)$ denotes the space of probability measures on A. A *lottery* is a probability measure on V. A utility is automatically interpreted (in the sense of von Neumann and Morgenstern) as extending linearly to $\Delta(V)$, so that $U(\mu) := \sum_{x \in V} U(x)\mu(x)$ whenever $\mu \in \Delta(V)$.

Society consists of n voters, who are indexed by the integers $1, \dots, n$. A *profile* is an n-tuple $\mathbf{P} = \langle P_1, \dots, P_n \rangle$ assigning a preference to each voter. Let $\mathcal{P}$ be the set of profiles. For each $k = 1, \dots, n$ let $\mathcal{P}_{-k}$ be the set of $(n-1)$-tuples of preferences $\langle P_1, \dots, P_{k-1}, P_{k+1}, \dots, P_n \rangle$, thought of as configurations of preferences of the voters other than k. If $\mathbf{P} \in \mathcal{P}$ is given, $\mathbf{P}_{-k}$ will denote the $(n-1)$-tuple obtained by dropping P_k. If $\mathbf{P}_{-k} \in \mathcal{P}_{-k}$ and P'_k are given,

$$\langle \mathbf{P}_{-k}, P'_k \rangle = \langle P_1, \dots, P_{k-1}, P'_k, P_{k+1}, \dots, P_n \rangle$$

is the profile obtained by combining these objects. If $\mathbf{P} \in \mathcal{P}$ and P'_k are given,

$$\mathbf{P}/_k P'_k := \langle \mathbf{P}_{-k}, P'_k \rangle$$

is the profile obtained from $\mathbf{P}$ by replacing P_k with P'_k.

A *decision scheme* is a function

$$d : \mathcal{P} \to \Delta(V).$$

We denote the probability assigned to alternative x by the decision scheme at profile $\mathbf{P}$ by $d(x, \mathbf{P})$, and for any $X \subset V$ we let $d(X, \mathbf{P}) := \sum_{x \in X} d(x, \mathbf{P})$. We say that d is a *probability mixture* of schemes $d_1, \dots, d_m$ if there are positive numbers $\alpha_1, \dots, \alpha_m$ with $\alpha_1 + \dots + \alpha_m = 1$ such that

$$d(x, \mathbf{P}) = \alpha_1 d_1(x, \mathbf{P}) + \dots + \alpha_m d_m(x, \mathbf{P})$$

for all alternatives x and profiles $\mathbf{P}$.

The decision scheme d is *potentially manipulable* by k at a profile $\mathbf{P}$ if there is a utility scale U that fits P_k and a preference P'_k such that $U(d(\mathbf{P}/_k P'_k)) > U(d(\mathbf{P}))$. We say that d is *manipulable* if it is potentially manipulable by some voter at some profile, and otherwise it is *strategy-proof.* Note that a probability mixture of strategy-proof decision schemes is strategy-proof.

A lottery ρ is *Pareto optimal ex post* for profile $\mathbf{P}$ if $\rho(x) = 0$ for any alternative x that is Pareto dominated insofar as there is another alternative y such that yP_ix for all i. The decision scheme d is *Pareto optimific ex post* if, for each profile $\mathbf{P}$, $d(\mathbf{P})$ is Pareto optimal ex post for $\mathbf{P}$. If d is a probability mixture of schemes $d_1, \dots, d_m$, then d is Pareto optimific ex post if and only if each d_j is Pareto optimific ex post.

For a preference P, let $\varphi(P)$ be the top ranked alternative or *favorite.* A decision scheme d is *dictatorial*, or a *dictatorship*, if there is a voter k such that $d(\varphi(P_k), \mathbf{P}) = 1$ for all $\mathbf{P} \in \mathcal{P}$. A *random dictatorship* is a probability mixture of dictatorships.

Gibbard's most general result asserts that a strategy proof decision scheme is a probability mixture of finitely many decision schemes, each of which is nonperverse (this concept is defined in the next section) and either *duple*, meaning that there are two alternatives that are the only alternatives receiving positive probability at any preference profile, or *unilateral*, meaning that it depends only on the preferences of a single voter. Gibbard credits Sonnenschein with the observation that, insofar as a duple decision scheme cannot be Pareto optimific ex post if there are three or more alternatives, and a unilateral decision scheme is Pareto optimific ex post if and only if it is dictatorial, it follows that:

Theorem 1 (Gibbard (1977)). *If there are three or more alternatives and the decision scheme d is strategy-proof and Pareto optimific ex post, then it is a random dictatorship.*

3 The I.I.D. Model

A *model of the electorate* for voter k is a probability measure $\beta \in \Delta(\mathcal{P}_{-k})$. The decision scheme d is *potentially manipulable* by k at a model β if there is a utility scale U that fits P_k and a preference P'_k such that

$$U\Big(\sum_{\mathbf{P}_{-k} \in \mathcal{P}_{-k}} d(\mathbf{P}_{-k}, P'_k)\beta(\mathbf{P}_{-k})\Big) > U\Big(\sum_{\mathbf{P}_{-k} \in \mathcal{P}_{-k}} d(\mathbf{P}_{-k}, P_k)\beta(\mathbf{P}_{-k})\Big). \quad (*)$$

Let $\mathcal{O}$ be the set of all strict orderings of V. The model β is *identically and independently distributed* (i.i.d.) if there is a $\sigma \in \Delta(\mathcal{O})$ such that $\beta(\mathbf{P}_{-k}) = \prod_{i \neq k} \sigma(P_i)$ for all $\mathbf{P}_{-k} \in \mathcal{P}_{-k}$. The decision scheme d is *strongly manipulable* if there is a voter k such that d is potentially manipulable at some i.i.d. model of the electorate for k, and if this is not the case we say that d is *weakly strategy-proof.*

Let $\mathbb{N} := \{0, 1, 2, \dots\}$ be the nonnegative integers, and let $A := \mathbb{N}^{\mathcal{O}}$. An element of A is called an *anonymous profile* because it specifies the number of voters with each preference ordering without attributing those preferences to specific individuals. For $a \in A$ let $|a| = \sum_{P \in \mathcal{O}} a_P$ be the total number of voters, and for $n = 0, 1, 2, \dots$ let $A_n := \{\, a \in A : |a| = n \,\}$. Let $\pi_n : \mathcal{P} \to A_n$ be the function defined by letting the component $\pi_{n,P}(\mathbf{P})$ be the number of k such that $P_k = P$. The decision scheme d is *anonymous* if $d(\mathbf{P})$ depends only on $\pi_n(\mathbf{P})$, so that there is a function $D_n : A_n \to \Delta(V)$ such that $d = D_n \circ \pi_n$. The *anonymous random dictatorship* is the decision scheme d^* given by

$$d^*(x, \mathbf{P}) := \tfrac{1}{n}\#\{\, k : \varphi(P_k) = x \,\}.$$

Theorem 2. *If there are three or more alternatives and the decision scheme d is anonymous, weakly strategy-proof, and Pareto optimific ex post, then it is the anonymous random dictatorship.*

The proof is developed in a sequence of lemmas. Fix a decision scheme d. A set $X \subset V$ *heads* a preference P if xPy for all $x \in X$ and $y \in V \setminus X$. If $d(X, \mathbf{P}) = d(X, \mathbf{P}/_k P'_k)$ for all k, $\mathbf{P}$, and P'_k such that X heads both P_k and P'_k, then d is said to be *localized*. The most innovative step in our argument is:

Lemma 1. *If d is weakly strategy proof and anonymous, then it is localized.*

Proof. Fix a voter k, and let π_{-k} be π_{n-1} reinterpreted as a function with domain $\mathcal{P}_{-k}$: $\pi_{-k,P}(\mathbf{P}_{-k})$ is the number of j such that $P_j = P$. Fixing a set $X \subset V$ and P_k and P'_k such that X heads both P_k and P'_k, let M be the set of $a \in A_{n-1}$ such that

$$d(X, \langle \mathbf{P}_{-k}, P'_k \rangle) > d(X, \langle \mathbf{P}_{-k}, P_k \rangle)$$

for some (hence all, because d is anonymous) $\mathbf{P}_{-k}$ such that $\pi_{-k}(\mathbf{P}_{-k}) = a$. Our goal is to show that $M = \emptyset$.

Supposing otherwise, let $C \subset \mathbb{R}^{\mathcal{O}}$ be the convex hull of M. Then C is the convex hull of its extreme points, each of which is an element of M. Let b be an extreme point. Then b is not an element of the convex hull of $M \setminus \{b\}$, so the separating hyperplane theorem gives a vector $\ell \in \mathbb{R}^{\mathcal{O}}$ such that $\langle \ell, b \rangle < \langle \ell, a \rangle$ for all $a \in M \setminus \{b\}$.

For some $\alpha > 0$ let $\sigma \in \Delta(\mathcal{O})$ be the probability distribution in which the probability of P is proportional to α^{ℓ_P}, so that $\sigma(P) := \alpha^{\ell_P} / \sum_{P' \in \mathcal{O}} \alpha^{\ell_{P'}}$. Let β be the derived i.i.d. model of the electorate for k: $\beta(\mathbf{P}_{-k}) = \prod_{i \neq k} \sigma(P_i)$. Then the probability of $\mathbf{P}_{-k}$ is proportional to

$$\alpha^{\ell_{P_1}} \cdots \alpha^{\ell_{P_{k-1}}} \cdot \alpha^{\ell_{P_{k+1}}} \cdots \alpha^{\ell_{P_n}} = \alpha^{\langle \ell, \pi_{-k}(\mathbf{P}_{-k}) \rangle},$$

so that

$$\beta(\mathbf{P}_{-k}) = \frac{\alpha^{\langle \ell, \pi_{-k}(\mathbf{P}_{-k}) \rangle}}{\sum_{\mathbf{P'}_{-k} \in \mathcal{P}_{-k}} \alpha^{\langle \ell, \pi_{-k}(\mathbf{P'}_{-k}) \rangle}}.$$

Note that if $\pi_{-k}(\mathbf{P}_{-k}) = b$ and $\pi_{-k}(\mathbf{P'}_{-k}) = a \in M \setminus \{b\}$, then

$$\frac{\beta(\mathbf{P}_{-k})}{\beta(\mathbf{P'}_{-k})} = \alpha^{\langle \ell, b-a \rangle} \to \infty \quad \text{as} \quad \alpha \to 0.$$

Therefore

$$\sum_{\mathbf{P}_{-k} \in \mathcal{P}_{-k}} d(X, \langle \mathbf{P}_{-k}, P'_k \rangle) \beta(\mathbf{P}_{-k}) > \sum_{\mathbf{P}_{-k} \in \mathcal{P}_{-k}} d(X, \langle \mathbf{P}_{-k}, P_k \rangle) \beta(\mathbf{P}_{-k})$$

when α is sufficiently small. In this circumstance there is a utility scale U that fits P_k, and which emphasizes the difference between X and $V \setminus X$ while nearly disregarding differences between elements of X and between elements of $V \setminus X$, to such an extent that inequality $(*)$ holds. This contradiction of the assumption that d is weakly strategy-proof completes the proof. □

The remainder of the argument is a matter of marshalling tools developed in Gibbard (1977). We write $xP!y$ to indicate that xPy and that x and y are adjacent in the ranking, so that for all $z \notin \{x, y\}$, zPx if and only if zPy. If this is the case, then P^y denotes the ranking obtained by interchanging x and y without changing the ranking of either in relation to any third alternative z, and we say that P^y is obtained from P by *switching* x and y. Given a profile $\mathbf{P}$ and a voter k with $xP_k!y$, let

$$\mathbf{P}^{ky} := \langle P_1, \dots, P_{k-1}, P_k^y, P_{k+1}, \dots, P_n \rangle.$$

We say that d is *pairwise responsive* if $d(z, \mathbf{P}^{ky}) = d(z, \mathbf{P})$ for all distinct alternatives x, y, and z, all voters k, and all profiles $\mathbf{P}$ such that $xP_k!y$. Of course if this is the case, then $d(\{x, y\}, \mathbf{P}^{ky}) = d(\{x, y\}, \mathbf{P})$ for all x, y, $\mathbf{P}$, and k such that $xP_k!y$.

Lemma 2. *d is localized if and only if it is pairwise responsive.*

Proof. This follows from Lemma 1 (p. 672) of Gibbard (1977). □

Given a profile $\mathbf{P}$ and a voter k with $xP_k!y$, the *effect* under d of k's switching y upward is

$$\varepsilon_k^y(d, \mathbf{P}) := d(y, \mathbf{P}^{ky}) - d(y, \mathbf{P}).$$

The decision scheme d is *nonperverse* if $\varepsilon_k^y(d, \mathbf{P}) \geq 0$ for every $\mathbf{P}$, k, and $y \neq \varphi(P_k)$. If P is an ordering and $x, y \in V$, $P{\upharpoonright}\{x, y\}$ is the ordering of $\{x, y\}$ obtained by restricting P to this set. For $\mathbf{P} \in \mathcal{P}$ the derived profile of preferences over $\{x, y\}$ is

$$\mathbf{P}{\upharpoonright}\{x, y\} := \langle P_1{\upharpoonright}\{x, y\}, \dots, P_n{\upharpoonright}\{x, y\} \rangle.$$

We say that d is *pairwise isolated* if

$$\varepsilon_k^y(d, \mathbf{P}) = \varepsilon_k^y(d, \mathbf{P}')$$

for all $\mathbf{P}$, $\mathbf{P}'$, x, and y such that $\mathbf{P}{\upharpoonright}\{x, y\} = \mathbf{P}'{\upharpoonright}\{x, y\}$ and all k such that $P_k = P'_k$ and $xP_k!y$. The decision scheme d is *decomposable* if, for any fixed k, x, and y with $x \neq y$, there are functions γ and δ such that for all $\mathbf{P}$ with $xP_k!y$,

$$\varepsilon_k^y(d, \mathbf{P}) = \gamma(\mathbf{P}{\upharpoonright}\{x, y\}) + \delta(P_k).$$

Lemma 3. *If d is localized, then it is pairwise isolated and decomposable.*

Proof. This is Lemma 3 (p. 673) of Gibbard (1977). □

For a profile $\mathbf{P}$, $\varphi(\mathbf{P}) := \langle \varphi(P_1), \dots, \varphi(P_n) \rangle$. We say that d *depends only on favorites* if $d(\mathbf{P}') = d(\mathbf{P})$ for all profiles $\mathbf{P}$ and $\mathbf{P}'$ such that $\varphi(\mathbf{P}') = \varphi(\mathbf{P})$.

Lemma 4. *If d is localized and Pareto optimific ex post, and there are three or more alternatives, then d depends only on favorites. If, in addition, d is anonymous, then it is the anonymous random dictatorship.*

Proof. First consider particular k, x, and y with $x \neq y$. Since d is localized, it is decomposable; let γ and δ be the functions given by the definition of decomposability. Since there are three distinct alternatives, for some $z \notin \{x, y\}$ there are profiles $\mathbf{P}$ with $\varphi(P_i) = z$ for all i. Since d is Pareto optimific ex post, $d(z, \mathbf{P}) = 1$ for any such $\mathbf{P}$. By allowing $\mathbf{P}$ to vary in the set of such profiles we can deduce that γ is identically zero, and that $\delta(P_k) = 0$ whenever $xP_k!y$ and $x \neq \varphi(P_k)$. Since d is pairwise responsive, and it is possible to move between any two $\mathbf{P}$ and $\mathbf{P}'$ with $\varphi(\mathbf{P}) = \varphi(\mathbf{P}')$ through a sequence of switches that do not affect the vector of favorites, d depends only on favorites.

It now follows that there are numbers $\epsilon_k(x, y) \in [0, 1]$ such that $\varepsilon_k^y(d, \mathbf{P}) = \epsilon_k(x, y)$ whenever $x = \varphi(P_k)$ and $xP_k!y$. Consider a profile $\mathbf{P}$ in which all voters rank x first and y second. Since d is Pareto optimific ex post, by switching x with y one voter at a time we obtain

$$\epsilon_1(x, y) + \cdots + \epsilon_n(x, y) = 1.$$

If d is anonymous, then $\epsilon_k(x, y)$ does not depend on k, so $\epsilon_k(x, y) = 1/n$ for all k, x, and y. □

Theorem 2 follows from Lemmas 1 and 4.

4 The Poisson Model

In earlier sections a decision scheme was a function whose argument was an assignment of preferences to a fixed set of "names." When the set of voters is variable, any attempt to keep track of names would be cumbersome at best, and irrelevant to our aims, so we adopt a definition that embeds the assumption of anonymity. An *extended anonymous decision scheme* (EADS) is a function

$$D : A \to \Delta(V)$$

from anonymous profiles to lotteries.

Fix such a D. Let $D(x, a)$ be the probability that x is chosen when the anonymous profile is a. The *extended anonymous random dictatorship* is the EADS D^* given by

$$D^*(x, a) := \frac{1}{|a|} \sum_{\varphi(P)=x} a_P.$$

We say that D is *Pareto optimific ex post* if $D(x, a) = 0$ whenever there is y such that yPx for all P with $a_P > 0$. As in the last section, we wish to develop a mild notion of strategy proofness based on restrictions on the beliefs a voter may hold about the preferences of other voters.

For a random variable taking values in $\mathbb{N}$ that is distributed according to the Poisson distribution with mean μ, the probability that the variable takes on value m is

$$f_\mu(m) := \mu^m e^{-\mu}/m!.$$

For $\lambda \in (0, \infty)^{\mathcal{O}}$ and $a \in A$ let

$$F_\lambda(a) := \prod_{P \in \mathcal{O}} f_{\lambda_P}(a_P).$$

Then F_λ specifies a model of the electorate in which the numbers a_P of voters with each preference are independent random variables and each a_P has a Poisson distribution with mean λ_P.

Poisson models of elections and more general games have been studied extensively by Meyerson (1998a,b, 2000, 2002). The Poisson distribution with mean μ is the limit as $N \to \infty$ of the distribution of the number of heads among N independent coin flips, each of which comes up heads with probability μ/N. Insofar as Poisson models are limits of related or more general models, results such as Theorem 3 below imply, by continuity, corresponding results for "nearby" models. In part because they are limits, Poisson models tend to be especially tractable.

A particularly pleasant property, which we take as the basis of our analysis, is called *environmental equivalence*. Suppose that a voter believes that the probability of being one of $m+1$ voters with preference P is proportional to $m + 1$ times the probability that there are $m + 1$ voters with preference P. Then the probability, conditional on being a voter with preference P, that there are m other voters with preference P, should be

$$\frac{(m+1)f_{\lambda_P}(m+1)}{\sum_{j=1}^{\infty} j f_{\lambda_P}(j)} = \frac{\lambda_P^{m+1}/m!}{\sum_{j=1}^{\infty} \lambda_P^j/(j-1)!} = \frac{\lambda_P^m/m!}{e^{\lambda_P}} = f_{\lambda_P}(m).$$

Thus the voter's belief concerning the number of *other* voters with preference P coincides with the given distribution of the *total* number of voters with this preference. Since the numbers a_P are statistically independent, being a voter with preference P conveys no information about the number of voters with any other preference. Therefore the voter's belief about the rest of the electorate should coincide with the model's description of the entire electorate. These calculations are heuristic, appealing to intuitions concerning the perspective of one member of an infinite pool of potential voters, each of whom is chosen with infinitesimal probability, but they can be made precise by taking the limit of a sequence of models in which, for each P, N_P potential voters with preference P each have probability λ_P/N_P of being included in the electorate.

An *extended model of the electorate* is a probability measure $B \in \Delta(A)$. For each P let $\mathbf{e}_P$ be the element of A whose P-coordinate is 1 and whose other coordinates are 0. We say that D is *potentially manipulable* by preference ordering P at B if there is a utility scale U that fits P and a preference P' such that

$$U\Big(\sum_{a\in A} D(a+\mathbf{e}_{P'})B(a)\Big) > U\Big(\sum_{a\in A} D(a+\mathbf{e}_{P})B(a)\Big) \qquad (**).$$

The EADS D is *strongly manipulable* if there is a preference P such that D is potentially manipulable by P at some Poisson model F_λ, and if this is not the case we say that D is *weakly strategy-proof.*

Theorem 3. *If there are three or more alternatives and the EADS D is weakly strategy-proof and Pareto optimific ex post, then it is the extended anonymous random dictatorship.*

For any $X \subset V$ let $D(X,a) := \sum_{x\in X} D(x,a)$. We say that the EADS D is *localized* if $D(X, a+\mathbf{e}_P) = D(X, a+\mathbf{e}_{P'})$ for all $a \in A$ and $P, P' \in \mathcal{O}$ such that X heads both P and P'. For each $n = 1, 2, \ldots$ there is a unique anonymous decision scheme d_n such that $d_n = D \circ \pi_n$. Clearly D is localized if and only if each d_n is localized, D is Pareto optimific ex post if and only if each d_n is Pareto optimific ex post, and D is the extended random anonymous dictatorship if and only if each d_n is the anonymous random dictatorship for that n. Therefore Theorem 3 follows from Lemma 4 and the following analogue of Lemma 1.

Lemma 5. *If D is weakly strategy proof, then it is localized.*

Proof. Fixing a nonempty $X \subset V$ and P and P' such that X heads both P and P', let M be the set of $a \in A$ such that

$$D(X, a+\mathbf{e}_{P'}) > D(X, a+\mathbf{e}_P).$$

Our goal is to show that $M = \emptyset$. Supposing otherwise, let n be the minimum value of $|a|$ for $a \in M$, let L be the set of $a \in M$ such that $|a| = n$, and let $N := M \cup \{\, a \in A : |a| > n \,\}$.

Let B be the convex hull of L, and let C be the convex hull of N. Then B is the convex hull of its extreme points, each of which is an element of L. Let b be one of these extreme points. Then b is also an extreme point of C because any representation of b as convex combination of elements of C must assign positive weight only to points whose components sum to n. The separating hyperplane theorem gives a vector $\ell \in \mathbb{R}^{\mathcal{O}}$ such that $\langle \ell, b\rangle < \langle \ell, a\rangle$ for all $a \in N \setminus \{b\}$. Since $a \in N$ whenever $a_Q \geq b_Q$ for all Q, all components of ℓ are positive.

For some $\alpha > 0$ let λ be the vector with components $\lambda_Q := \alpha^{\ell_Q}$. For $a \in A$ we have $\prod_Q \lambda_Q^{a_Q} = \alpha^{\sum_Q \ell_Q a_Q} = \alpha^{\langle \ell, a \rangle}$ and

$$F_\lambda(a) = \prod_{Q \in \mathcal{O}} \lambda_Q^{a_Q} e^{-\lambda_Q}/a_Q! = \alpha^{\langle \ell, a \rangle} e^{-\sum_{Q \in \mathcal{O}} \lambda_Q} \Big/ \prod_{Q \in \mathcal{O}} a_Q!.$$

In particular

$$\frac{F_\lambda(a)}{F_\lambda(b)} = \Big(\prod_{Q \in \mathcal{O}} b_Q!/a_Q! \Big) \frac{\alpha^{\langle \ell, a \rangle}}{\alpha^{\langle \ell, b \rangle}}.$$

For each $a \in N \setminus \{b\}$ this quantity goes to zero as $\alpha \to 0$, but in fact a stronger statement is true and relevant. Since the components of ℓ are positive, $\langle \ell, a \rangle$ is bounded below by a positive multiple of $|a|$. The number of $a \in A$ with $|a| = n'$ is bounded above by a polynomial function of n', and $\prod_{Q \in \mathcal{O}} b_Q!/a_Q!$ is bounded above by $\prod_{Q \in \mathcal{O}} b_Q!$. Therefore

$$\frac{\sum_{a \in M \setminus \{b\}} F_\lambda(a)}{F_\lambda(b)} \leq \frac{\sum_{a \in N \setminus \{b\}} F_\lambda(a)}{F_\lambda(b)} \to 0 \quad \text{as} \quad \alpha \to 0,$$

so

$$\sum_{a \in A} D(X, a + \mathbf{e}_{P'}) F_\lambda(a) > \sum_{a \in A} D(X, a + \mathbf{e}_P) F_\lambda(a)$$

when α is sufficiently small, in which case there is a utility scale U that fits P such that $(**)$ holds when $B = F_\lambda$. This contradiction of the assumption that D is weakly strategy-proof completes the proof. □

References

D. Black. *The Theory of Committees and Elections*. Cambridge University Press, Cambridge, MA, 1958.

A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41:587–602, 1973.

A. Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45:665–681, 1977.

Eric Maskin. Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66:23–38, 1999.

R.B. Meyerson. Extended Poisson games and the Condorcet jury theorem. *Games and Economic Behavior*, 25:111–131, 1998a.

R.B. Meyerson. Population uncertainty and Poisson games. *International Journal of Game Theory*, 27:375–392, 1998b.

R.B. Meyerson. Large Poisson games. *Journal of Economic Theory*, 94: 7–45, 2000.

R.B. Meyerson. Comparison of scoring rules in Poisson voting games. *Journal of Economic Theory*, 103:219–251, 2002.

A. E. Roth and M. Sotomayor. *Two Sided Matching: A Study in Game-Theoretic Modelling and Analysis*. Cambridge University Press, Campbridge, MA, 1990.

M. A. Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.