JOURNAL OF ECONOMIC THEORY **61**, 262–289 (1993)

Generalized Median Voter Schemes and Committees*

SALVADOR BARBERÀ

Universitat Autónoma de Barcelona, Barcelona, Spain

AND

FARUK GUL AND ENNIO STACCHETTI[†]

Stanford University, Stanford, California 94305

Received July 19, 1991; revised October 26, 1992

We define a multidimensional analogue of a single-peaked preference and generalize the notion of a median voter scheme. Every onto strategy-proof social choice function on a single-peaked domain is a generalized median voter scheme. Since a single-peaked preference can be identified unequivocally with its bliss point, one can view a social choice function as an Arrowian social welfare function. We show that a social choice function is strategy-proof iff, viewed as a social welfare function, it satisfies a monotonicity property. Finally, we investigate strategic decision making in hierarchical committees. *Journal of Economic Literature* Classification Numbers: C72, D71. © 1993 Academic Press, Inc.

1. INTRODUCTION

In this paper we study multidimensional generalizations of two familiar concepts from social choice theory, median voter schemes and single-peaked preferences, in connection with the problem of designing strategy-proof social choice mechanisms. We prove that when preferences are multidimensional single-peaked, surjective social choice functions are strategy-proof if and only if they are generalized median voter schemes. We also show that these generalized median voter schemes can be nicely

* Barberà's work is partially supported by Research Grant PB 89-0294 from the Dirección General de la Investigación Científica y Técnica, Spanish Ministry of Education. Gul and Stacchetti gratefully acknowledge support from the Alfred P. Sloan Foundation and the National Science Foundation, respectively. We are indebted to Michael de Breton, Matthew Jackson, Jordi Massó, Hervé Moulin, Alejandro Neme, Arunava Sen, Lin Zhou, and especially Hugo Sonnenschein for helpful conversations.

† Present address: The University of Michigan, Ann Arbor, MI 48109.

0022-0531/93 $5.00

decomposed into simple rules, each of which involves the application of median voting to a three-person committee.

When designing criteria for assigning social outcomes to individual preference profiles, it would be interesting to fulfill the following two requirements: the members of society must find it in their interest to truthfully reveal their preferences, and some amount of social compromise should be achieved. The classic result of Gibbard and Satterthwaite (see Gibbard [5], and Satterthwaite [8]) establishes the difficulty of simultaneously meeting both of these objectives. Specifically, the Gibbard–Satterthwaite Theorem states that if the image of a social choice function has at least three elements and each member of the society can have any preference ordering over the alternatives (which are then at least three) then the social choice function is either dictatorial or fails strategy-proofness. That is, for any non-dictatorial social choice function there is some situation in which at least one member of society benefits from misrepresenting his preferences.

Since it is impossible to accommodate truthfulness and compromise in completely general situations, it becomes natural to investigate the possibility of designing strategy-proof social choice functions on more restricted domains. Domains can be restricted by the nature of the alternatives under consideration and by the class of preferences on such alternatives that are considered admissible for agents who participate in the decision-making process. We concentrate on decision problems where the set of alternatives can be represented as the cartesian product of l finite integer intervals, $B_1, B_2, ..., B_l$. An l-tuple $\alpha = (\alpha_1, ..., \alpha_l)$ describes an alternative by specifying the values of all variables of social interest. This description of alternatives as l-tuples of real values is quite standard, and it is appropriate to model many interesting decision problems. For example, each interval may correspond to an issue within a political platform, or to one relevant feature among those describing alternative public projects. Values within an interval would then stand for different positions on the corresponding issue, or for different specifications of the project with respect to the feature in question. A seemingly different kind of problem admitting the same representation for alternatives is that of selecting new members to a club from a collection of l candidates, or which set of new bills should be passed during a legislation. Here the alternatives would be the vertices of the $\{0, 1\}^l$ cube, where a 1 in the jth dimension would represent the decision to admit the jth candidate, or pass the jth bill, and a 0 would represent the decision to reject the jth candidate or bill. Implicitly, we are assuming that any combination of admisible values for each criterion is itself admissible. This means, for example, that any combination of individually feasible specifications for a public project must itself describe a possible project.

The one-dimensional case (i.e., $l = 1$) is the context in which single-peaked preferences are originally defined. Let the set of alternatives be $B = \{a, a+1, ..., b\}$, for some integers a and b. Then a strict preference P over B is said to be single-peaked if there exists a point α, P's most preferred outcome, such that $|\alpha - \gamma| = |\alpha - \beta| + |\beta - \gamma|$ implies β is preferred to γ. Thus, each single-peaked preference has a unique favorite outcome, called its bliss point, and any move away from the favorite outcome leads to less preferred alternatives. Note that this requirement does not restrict the ranking of two alternatives β and γ, when $\beta < \alpha < \gamma$. Alternatively, the array could represent political platforms or choices of location.

In many problems, one dimension will not be rich enough to describe the alternatives. More generally, we could view the set of alternatives as a multidimensional grid, with the alternatives at the nodes and with each edge of length one. The distance between any two alternatives α and β is the length of any shortest path between α and β. This is just the city block metric, induced by the L_1-norm. A preference with bliss point α is *multidimensional single-peaked* if for any two alternatives β and γ with β between α and γ (i.e., with β on a shortest path from α to γ), β is preferred to γ.

Our characterization of strategy-proof social choice functions over multidimensional single-peaked preferences (Theorem 3) is based on a number of different steps, each of which is interesting on its own. We first show that a strategy-proof social choice function can only depend on the bliss points of voters, if the range of the function is itself a cartesian product of intervals. Social choice functions that depend only on the bliss points of agents are called voting schemes.

Next we show that each strategy-proof voting scheme on our domains is separable. That is, it can be decomposed into l strategy-proof voting schemes, one for each dimension of the array of alternatives. Conversely, any choice of l strategy-proof voting schemes will uniquely define a strategy-proof social choice function whose range is a cartesian product.

The last step of our characterization of strategy-proof voting schemes generalizes the notion of median voter scheme. Consider the one-dimensional case and the ordinary median voter scheme. Let the number of agents be $2k + 1$. For any α, a coalition of $k + 1$ agents with bliss points less than or equal to α can guarantee an outcome less than or equal to α. Hence, any group of agents consisting of at least $k + 1$ agents is called a winning coalition at α. In our generalization, we call an arbitrary collection of agents a winning coalition at α if they can guarantee an outcome no greater than α whenever all their bliss points are no greater than α. We show that every strategy-proof voting scheme is uniquely identified by a system of winning coalitions, containing the winning coalitions at each outcome α. For any profile of preferences, the outcome is the smallest alternative β such that the collection of voters with bliss point less than or equal to β is a winning coalition at β.

In Section 3, we exploit the dual interpretation of elements of the l-dimensional box, as outcomes and single-peaked preferences with the corresponding bliss point, to investigate the relationship between social choice functions and social welfare functions. Specifically, we interpret each subbox of the given multidimensional box as a potential choice problem that society might encounter and view a social welfare function as a mapping that specifies for every profile of preferences a rule for making each of those choices. For a fixed subbox we view the implied choices from this subbox, given the profile of individual preferences and the social welfare function, as a social choice function. We prove that the rule for choosing from the entire box is strategy-proof if and only if the rule for choosing from every subbox is strategy-proof. Finally, given the dual interpretation of mappings from n-tuples of bliss points to outcomes (or bliss points) as social functions (scf's) and social welfare functions (swf's), we establish the equivalence of strategy-proofness and the non-negative response property, which is a condition slightly stronger than Arrow's independence of irrelevant alternatives.

In Section 4, we further utilize the duality between scf's and swf's to analyze decentralized decision-making rules and committees. Take m voting schemes that map n profiles of preferences (bliss points) to outcomes (bliss points) and take another voting scheme which maps the resulting profile of m bliss points to bliss points. In this manner we obtain a single voting scheme from profiles of n bliss points to outcomes. We call every such voting scheme a combination of voting schemes. We show that if each of the initial voting schemes from n profiles to bliss points is strategy-proof and the final voting scheme from m profiles to outcomes is strategy-proof, then the combination is strategy-proof. We interpret each initial stage voting scheme as a committee and the outcome of such voting schemes as the representative of the committee. Thus, committees choose representatives, and these representatives vote in other committees to choose representatives, etc. We show that every strategy-proof voting scheme can be decomposed (or decentralized) into a hierarchy of three-person committees such that the voting rule in each committee is the (ordinary) median voter rule.

There are three papers closely related to ours. Moulin [7] provides an extensive treatment of strategy-proof social choice functions over one-dimensional single-peaked preferences. He confines attention to voting schemes and provides two separate characterizations of strategy-proof voting schemes, one for the anonymous case and one for the general case. His characterization also provides an alternative decomposition result for the one-dimensional case, resembling our committee-tree structure. Border and Jordan [3] deal with the case in which the set of alternatives is $\Re^l$ and study a variety of domain restrictions. None of their domain restrictions correspond exactly to our notion of a single-peaked preference. For certain domain restrictions they obtain results similar to our Theorem 1 and Theorem 2 (separability); for other restrictions, they show that dictatorships or constants are the only strategy-proof social choice functions. We provide a more extensive comparison between the results of these two papers and ours in Section 5. Finally, Barberà, Sonnenschein and Zhou [2] deal with the problem we mentioned above. A given society is faced with the problem of deciding on l separate issues. Each decision entails acceptance or rejection. They require that preferences be separable and provide a characterization result for their binary framework. This corresponds to a special case of our representation theorem of social choice functions in terms of generalized median voter schemes. Similar to our first theorem, they also prove that voting schemes are unrestrictive in their setting.

Earlier work by Kramer [6] and Shepsle [9] on majority voting and generalizations of the majority rule also relate to our work. Kramer provides the definition of a voting rule (due to Farquharson [4]). Their voting rules are a subclass of our generalized median voter schemes. An implication of Kramer's results is that voting rules are strategy-proof; we prove the converse.

Shepsle [9] expands the analysis of Kramer [6] by allowing for more institutional detail. He defines the notion of a preference-induced equilibrium which essentially corresponds to the notion of a majority winner. He utilizes the added institutional detail to provide a weaker notion of equilibrium, which he calls a structure-induced equilibrium. In the particular example that he analyzes in detail, he defines an outcome to be a structure-induced equilibrium if, by deviating only on one dimension, no agent is able to obtain a better outcome.[1] Thus, the product of the median of each dimension turns out to be a structure-induced equilibrium. Since only deviations in one dimension are considered, Shepsle does not (and need not) impose the requirement that the most preferred outcome along any dimension is independent of the outcomes of the other dimensions (i.e., what we call multidimensional single-peakedness). Hence, his theorems only require strict quasi-concavity. As a result, the median voter scheme is not strategy-proof on the domain of preferences considered by Shepsle [9] (see Theorem 6).

[1] This follows from what Shepsle calls "the jurisdictional arrangement consisting of the basis vectors" and the "germaneness rule for amendments."

2. Voting Schemes and Single-Peaked Preferences

For a finite set X, $\#X$ denotes its cardinality (i.e., the number of elements contained in X). $\mathbf{Z}$ denotes the set of integer numbers, and $\mathbf{N}$ denotes the set of nonnegative integers. Subsets of $\mathbf{Z}^l$, where $l \in \mathbf{N}$, $l > 0$, are considered as metric subspaces of $\mathfrak{R}^l$ with the L_1-norm:

$$\|\alpha\| := \sum_{j=1}^{l} |\alpha_j| \qquad \text{for any } \alpha \in \mathfrak{R}^l.$$

Definition. A (strict) *preference* on A (the set of social alternatives) is a relation P on A satisfying

(1) completeness: for all α, $\beta \in A$, either $\alpha P \beta$ or $\beta P \alpha$;

(2) antisymmetry: α, $\beta \in A$, $\alpha P \beta$ and $\beta P \alpha$ imply $\alpha = \beta$;

(3) transitivity: for all α, β, $\gamma \in A$, $\alpha P \beta$ and $\beta P \gamma$ imply $\alpha P \gamma$.

For each preference P on A, $\tau(P)$ denotes the outcome (or social alternative) most preferred by P: $\alpha = \tau(P)$ iff $\alpha P \beta$ for all $\beta \in A$; $\tau(P)$ is called the "top" of P, or the bliss point of P. Clearly, $\tau(P)$ is well defined for every strict preference P.

Definition. For integers $a \leqslant b$, $[a, b]$ will denote the *integer interval* $\{a, a+1, ..., b\}$. An l-dimensional *box* B is a cartesian product of l integer intervals: $B = \times_{j=1}^{l} B_j$, where $B_j = [a_j, b_j]$ and $a_j \leqslant b_j$. A *subbox* of B is any box A contained in B.

Definition. A strict preference[2] on a box B is *multidimensional single-peaked with bliss point* $\alpha \in B$ iff $\tau(P) = \alpha$, and $\beta P \gamma$ for all β, $\gamma \in B$ satisfying $\|\alpha - \gamma\| = \|\alpha - \beta\| + \|\beta - \gamma\|$.

Note that any multidimensional single-peaked preference on the l-dimensional box B is identified by the property that its restriction to any lower dimension subbox is also a single-peaked preference. Furthermore, the bliss point of the restriction corresponds to the projection of the bliss point to the subbox. For example, let P be a single-peaked preference on B and consider the subbox $A := \{x\} \times B_2 \times \cdots \times B_l$, where $x \in B_1$. Suppose α is P's bliss point. Then the restriction of P to A is single-peaked and has bliss point $(x, \alpha_2, ..., \alpha_l)$. Similarly, if $B' \subset B$ is an arbitrary subbox of B, then the restriction of P to B' has bliss point α', where

[2] Our results can be extended to include non-strict preferences, provided that single-peakedness is redefined as follows: P is single-peaked if there exists α such that for all $\beta \neq \gamma$, $\|\alpha - \gamma\| = \|\alpha - \beta\| + \|\beta - \gamma\|$ implies that $\beta P \gamma$ and not $(\gamma P \beta)$. Obviously, such an α is unique. The notation however would become cumbersome.

$\alpha' = \text{argmin}_{\beta \in B'} \|\alpha - \beta\|$. These are easy to prove but important facts, which will be used below. In particular, they are the basis for Corollary 1, and they support the remarks in Section 3 regarding social welfare functions.

Quasi-concavity and separability are other restrictions that have been imposed on preferences. Note that our definition of a single-peaked preference does not require additive separability (or even the weaker notion of separability used in Border and Jordan [3]). For example, let $\alpha = (1, 5)$, $\beta = (1, 0)$, $\gamma = (1, 10)$, $\xi = (2, 0)$, and $\eta = (2, 10)$. There exists a single-peaked preference P with bliss point α such that $\beta P \gamma$ and $\eta P \xi$ (see Lemma 1). However, additive separability (or even separability) requires that $\beta P \gamma$ iff $\xi P \eta$. Also note that multidimensional single-peakedness when applied to $\Re^n$ neither implies quasi-concavity nor is implied by quasi-concavity. However, the set of separable strictly quasi-concave preferences is a strict subset of multidimensional single-peaked preferences.[3]

We identify each outcome α in a box B with the preference class $\mathbf{P}(\alpha)$ of all single-peaked preferences on B with bliss point α. In our analysis, we restrict attention to the collection $\mathbf{P} := \bigcup_{\alpha \in B} \mathbf{P}(\alpha)$ of *all* single-peaked preferences on the box B. Obviously the collection $\mathbf{P}$ depends on the box B, but it will always be clear from the context which box B is referred by $\mathbf{P}$. Throughout this paper, B always denotes an l-dimensional box and $\mathbf{P}$ denotes its corresponding collection of single-peaked preferences. Elements of the box are called the alternatives or outcomes. A *profile* is a preference n-tuple. A social choice function is a rule that assigns an outcome to each preference profile.

DEFINITION. A social choice function (scf) is a map $\varphi: \mathbf{P}^n \to B$.

We now turn to strategy-proofness, a main requirement that one wants to impose on social choice functions. A social choice function is strategy-proof if it is always optimal for all voters to reveal their own preferences, rather than manipulating the social outcome by strategically misrepresenting them at some profiles.

DEFINITION. $\varphi: \mathbf{P}^n \to B$ is a strategy-proof (sp) scf (spscf) if for all P_i', $P_1, ..., P_n \in \mathbf{P}$, $\varphi(P_1, ..., P_i, ..., P_n)\ P_i\, \varphi(P_1, ..., P_i', ..., P_n)$.

Any function $f: B^n \to B$ represents a (unique) social choice function $\varphi: \mathbf{P}^n \to B$ with set of players $N := \{1, ..., n\}$, defined as follows:

$$\varphi(P_1, ..., P_n) := f(\alpha^1, ..., \alpha^n) \qquad \text{for all} \quad P_i \in \mathbf{P}(\alpha^i) \text{ and } i \in N.$$

[3] Provided we extend the notion of a single-peaked preference to $\Re^n$ so that we can speak of quasi-concavity.

Clearly, a general scf $\varphi: \mathbf{P}^n \to B$ does not always admit such a representation because it is possible, for example, that for some $P'_1 \in \mathbf{P}$ and profile $(P_1, ..., P_n) \in \mathbf{P}^n$, with P_1 and P'_1 in the same preference class $\mathbf{P}(\alpha)$,

$$\varphi(P_1, ..., P_n) \neq \varphi(P'_1, ..., P_n).$$

Consider the following example: $B = \{0, 1\}^2$, $n = 3$, and $\varphi: \mathbf{P}^3 \to B$ corresponds to majority rule between $(0, 0)$ and $(1, 1)$. Suppose $\tau(P_2) = (0, 0)$, $\tau(P_3) = (1, 1)$, and $(0, 1)\, P_1 (1, 1)\, P_1 (0, 0)\, P_1 (1, 0)$ while $(0, 1)\, P'_1 (0, 0)\, P'_1 (1, 1)\, P'_1 (1, 0)$. Clearly $\tau(P_1) = \tau(P'_1) = (0, 1)$ and $\varphi(P_1, P_2, P_3) = (1, 1) \neq (0, 0) = \varphi(P'_1, P_2, P_3)$. However, this is not possible if φ is strategy-proof and its range is a subbox of B (see Theorem 1 and its Corollary below).

DEFINITION. A scf $\varphi: \mathbf{P}^n \to B$ is a *voting scheme*[4] if there exists a function $f: B^n \to B$ that represents it. That is,

$$\varphi(P_1, ..., P_n) := f(\tau(P_1), ..., \tau(P_n)) \qquad \text{for all} \quad (P_1, ..., P_n) \in \mathbf{P}^n.$$

The identification of a function $f: B^n \to B$ with its corresponding voting scheme $\varphi: \mathbf{P}^n \to B$ is done routinely. We say, for example, that f is a sp voting scheme if φ is sp.

Hereafter, for any profile $\tilde{\alpha} = (\alpha^1, ..., \alpha^n) \in B^n$, $\tilde{\alpha}_j := (\alpha^1_j, ..., \alpha^n_j) \in B^n_j$ denotes the profile of the jth coordinates. Thus, superindices always refer to players, while subindices refer to coordinates. The following technical lemma will facilitate much of the subsequent analysis. The proof is straightforward and is omitted.

LEMMA 1. (1) *Let P, $P' \in P(\alpha)$ and A be a subbox of B. Then P, P' have the same most preferred element in A. The restriction of P to any subbox A is also single-peaked. Moreover, β is P's most preferred element in A if and only if $\|\alpha - \gamma\| = \|\alpha - \beta\| + \|\beta - \gamma\|$ for all $\gamma \in A$.* (2) *If $\|\alpha - \gamma\| < \|\alpha - \beta\| + \|\beta - \gamma\|$, there exists $P \in P(\alpha)$ such that $\gamma P \beta$.*

Lemma 1 provides a number of important properties of single-peaked preferences which we utilize in subsequent arguments.

When we study surjective scf's, Theorem 1 below shows that it is nonrestrictive to confine attention to voting schemes only. The first main result in Moulin [7] characterizes Pareto *efficient* voting schemes on the domain of single-peaked preferences over a one-dimensional set of alternatives. Clearly, efficiency implies surjectivity. Therefore, Theorem 1 shows that Moulin's characterization result is more general and applies to *all* efficient

[4] Gibbard [5] calls any social choice function a voting scheme. We restrict the use of the latter to refer to a social choice function which only depends on the preferences' bliss points.

spscf's.[5] A similar result is proved by Border and Jordan [3] and Barberà, Sonnenschein, and Zhou [2] for slightly different domain restrictions.

THEOREM 1. *Let B be an l-dimensional box. If the scf $\varphi: \mathbf{P}^n \to B$ is sp and onto, then φ is a voting scheme.*

Proof. See Appendix.

Exploiting the properties of single-peaked preferences further, a stronger version of Theorem 1 can be attained, relaxing the conditions on the range of the scf; this is stated in the following corollary.

COROLLARY 1. *Let B be an l-dimensional box. If the range of the spscf $\varphi: \mathbf{P}^n \to B$ is a subbox of B, then φ is a voting scheme.*

Proof. Let $A = \varphi(\mathbf{P}^n)$. By assumption, A is a subbox. It follows immediately from part 1 of Lemma 1 that the restriction to A of any single-peaked preference on B is single-peaked. Since φ is sp, if the restrictions of $P, P' \in \mathbf{P}$ to A coincide, then $\varphi(P, P_{-i}) = \varphi(P', P_{-i})$ for all $P_{-i} \in \mathbf{P}^{n-1}$. Thus, φ can be viewed as an onto spscf form the set of single-peaked preferences over A into A. Theorem 1 then implies the desired conclusion. Q.E.D.

Throughout this paper we often invoke the close relationship between the L_1-norm, single-peaked preferences, and sp voting schemes. Lemma 2 is straightforward given the definition of joint multidimensional single-peakedness and part 2 of Lemma 1 above.

LEMMA 2. *$f: B^n \to B$ is sp iff for all $\tilde{\alpha} \in B^n$, $\beta \in B$, and $i \in N$, $\|\alpha^i - f(\beta, \alpha^{-i})\| = \|\alpha^i - f(\tilde{\alpha})\| + \|f(\tilde{\alpha}) - f(\beta, \alpha^{-i})\|$.*

Theorem 2 below states that the jth coordinate of the outcome of a sp voting scheme depends only on the jth coordinate of the bliss points of the profile's preferences. Hence, a full characterization of sp voting schemes can be obtained by the characterization of one-dimensional sp voting schemes. This separability result constitutes the second step toward our characterization result.

THEOREM 2 (Separability). *Let $B = \mathsf{X}_{j=1}^{l} B_j$. Then $f: B^n \to B$ is sp iff for each $j \in L := \{1, ..., l\}$ there exist $f_j: B_j^n \to B_j$ such that*

(1) $f(\tilde{\alpha})_j = f_j(\tilde{\alpha}_j)$ *for all* $\tilde{\alpha} = (\alpha^1, ..., \alpha^n) \in B^n$; *and*

(2) f_j *is sp.*

[5] However, in joint multidimensional single-peaked domains, there are no Pareto efficient and strategy-proof mechanisms other than the dictatorial. This observation is also made by Border and Jordan [3] and Barberà, Sonnenschein, and Zhou [2] in related contexts.

Proof. We only prove necessity; sufficiency is easy to verify. First note that (1) is equivalent to

$$f(\tilde{\alpha})_j = f(\tilde{\beta})_j \qquad \text{whenever} \quad \tilde{\alpha}_j = \tilde{\beta}_j. \tag{1'}$$

Clearly (1′) implies (1″) below

$$f(\alpha, \gamma^{-i})_j = f(\beta, \gamma^{-i})_j \qquad \text{for all } i, \text{ whenever } \alpha_j = \beta_j. \tag{1''}$$

To see that (1″) implies (1′) and hence is equivalent to (1′), for any $\tilde{\alpha}$ and $\tilde{\beta}$ satisfying $\tilde{\alpha}_j = \tilde{\beta}_j$, construct $\tilde{\gamma}(0), \tilde{\gamma}(1), ..., \tilde{\gamma}(n)$ as follows. Set $\tilde{\gamma}(0) = \tilde{\alpha}$ and sequentially replace the preferences in $\tilde{\alpha}$ with those of $\tilde{\beta}$, one-by-one, so that $\gamma(n) = \tilde{\beta}$. Applying (1″) at each step yields $f(\tilde{\gamma}(k))_j = f(\tilde{\gamma}(k+1))_j$ for all $k = 0, 1, ..., n-1$, and hence $f(\tilde{\alpha})_j = f(\tilde{\beta})_j$.

To prove that strategy-proofness implies (1″), let $\alpha' = f(\alpha, \gamma^{-i})$ and $\beta' = f(\beta, \gamma^{-i})$. Assume by contradiction that $\alpha_j = \beta_j$ and $\alpha'_j \neq \beta'_j$. First, consider the case in which $\|\beta - \beta'\| + \|\beta' - \alpha'\| > \|\beta - \alpha'\|$. Then, by part 2 of Lemma 1, there exists $P \in \mathbf{P}(\beta)$ such that $\alpha' P \beta'$, which contradicts f being sp. If $\|\beta - \beta'\| + \|\beta' - \alpha'\| = \|\beta - \alpha'\|$, then $|\beta_j - \beta'_j| + |\beta'_j - \alpha'_j| = |\beta_j - \alpha'_j|$. Since $\alpha_j = \beta_j$, $|\alpha_j - \beta'_j| + |\beta'_j - \alpha'_j| = |\alpha_j - \alpha'_j|$. This implies that $|\alpha_j - \alpha'_j| + |\alpha'_j - \beta'_j| = |\alpha_j - \beta'_j| + |\beta'_j - \alpha'_j| + |\alpha'_j - \beta'_j| > |\alpha_j - \beta'_j|$ since $\alpha'_j \neq \beta'_j$. But then $\|\alpha - \alpha'\| + \|\alpha' - \beta'\| > \|\alpha - \beta'\|$, and a symmetric argument to the one made for the case $\|\beta - \beta'\| + \|\beta' - \alpha'\| > \|\beta - \alpha'\|$ above yields a contradiction to the strategy-proofness.

To show, for example, that f_1 is sp, let $\tilde{\gamma} \in B^n$ and $\alpha, \beta \in B$ be such that for all $k \neq i$ and $j \neq 1$, $\gamma^k_j = \alpha_j = \beta_j$. As before, let $\alpha' = f(\alpha, \gamma^{-i})$ and $\beta' = f(\beta, \gamma^{-i})$. Since f is sp, $\|\alpha - \alpha'\| + \|\alpha' - \beta'\| = \|\alpha - \beta'\|$, and thus $|\alpha_1 - \alpha'_1| + |\alpha'_1 - \beta'_1| = |\alpha_1 - \beta'_1|$, which establishes that f_1 is sp by Lemma 2.

Q.E.D.

Since spscf's must be voting schemes, and sp voting schemes can be decomposed into l one-dimensional sp voting schemes, it will suffice to characterize the latter. Our representation is an extension of median voter schemes; we now reexamine this familiar class of scf's on the domain of single-peaked preferences[6] over a one-dimensional box $B = [a, b]$. Let $n = 2k + 1$ be any positive odd number. It is easy to see that associated with every $\tilde{\alpha} \in B^n$, there is a unique median value $\mu(\tilde{\alpha}) \in B$. This median value can be defined in three equivalent ways:

(1) $\mu(\tilde{\alpha}) := \text{argmin}_x \sum_{i=1}^{n} |\alpha^i - x|$.

(2) $\mu(\tilde{\alpha}) := \beta$ such that $\#\{i \mid \alpha^i < \beta\} \leqslant k$ and $\#\{i \mid \alpha^i > \beta\} \leqslant k$.

(3) $\mu(\tilde{\alpha}) := \min\{x \mid \#\{i \mid \alpha^i \leqslant x\} \geqslant k+1\}$.

[6] An early study of single-peaked preferences and group decision making can be found in Black [3a].

The equivalence of the three definitions and the fact that the median is well defined for any $\tilde{\alpha} \in B^n$, whenever n is odd, is well known and simple to verify. It is easy to see that the median $\mu: B^n \to B$, viewed as a voting scheme, is sp. For our purposes, the most useful definition of the median is (3) above. Note that according with this definition, the median is the smallest value x for which the number of bliss points at or to the left of this value is no less than $k+1$. Loosely speaking, the median voter scheme highlights the following feature of any sp voting scheme: every agent i whose bliss point α^i is to the left of x "pulls" the outcome away from $x+1$, and toward his own bliss point. Starting with $x = b$, note that the size of the coalition exerting force to pull x to the left diminishes as x decreases. As we move x to the left, eventually we find a value of x such that the number of bliss points $\alpha^i \leqslant x$ is at least $k+1$, but the number of bliss points $\alpha^i \leqslant x-1$ is less than $k+1$. This x is the median.[7]

Lemma 3 states that this characterization of the median is the appropriate one to describe general sp voting schemes. The lemma generalizes the median voter scheme described above in two ways: first, it allows for the possibility that the scf is not anynomous, in the sense that different agents need not have the same influence over the outcomes (hence, in terms of the above description, different agents may "exert" different levels of force in pulling the outcome to the left); second, different alternatives may receive a differential treatment—neutrality may fail. That is, the amount of force needed to pull the outcome below x may be different than the amount required to pull it below y. Lemma 3 establishes that modulo these two modifications, a voting scheme is sp iff it is a (generalized) median voter scheme. A few definitions are needed to make these changes precise.

DEFINITION. Let $B = [a, b]$ be a one-dimensional box and $N = \{1, ..., n\}$. A *left-coalition system* on B is a correspondence $\mathfrak{C}: B \to 2^N$ (i.e., $\mathfrak{C}(\xi)$ is a collection of coalitions for each $\xi \in B$) satisfying the following conditions:

(1) if $C \in \mathfrak{C}(\xi)$ and $C \subset D$, then $D \in \mathfrak{C}(\xi)$;

(2) if $\eta > \xi$ and $C \in \mathfrak{C}(\xi)$, then $C \in \mathfrak{C}(\eta)$; and

(3) $\mathfrak{C}(b) = 2^N$.

C is a (left) *winning coalition at* ξ if $C \in \mathfrak{C}(\xi)$. There is a unique *right-coalition system* $\bar{\mathfrak{C}}: B \to 2^N$ associated with each left-coalition $\mathfrak{C}: B \to 2^N$: $C \in \bar{\mathfrak{C}}(\xi)$ iff $C^c = N \backslash C \in \mathfrak{C}(\xi)$.

[7] This particular description of the median could just as well have been stated in terms of the bliss points being to the right of x. Indeed, one could present an alternative definition (3) based on the cardinality of the set $\{i \mid \alpha^i \geqslant x\}$.

Left-coalition systems can be used to induce voting schemes in a natural way. For each $\tilde{\alpha} \in B^n$ and $\xi \in B$, let $C(\tilde{\alpha}; \xi) := \{i \in N \mid \alpha^i \leqslant \xi\}$ be the coalition to the left of ξ.

DEFINITION. Let $B = [a, b]$ be an integer interval and $\mathfrak{C}$ be a left-coalition system on B. The voting scheme $f\colon B^n \to B$, defined as follows:

$$f(\tilde{\alpha}) = \min\{\xi \mid C(\tilde{\alpha}; \xi) \in \mathfrak{C}(\xi)\}, \qquad (*)$$

is called the *generalized median voter scheme* (GMVS) induced by $\mathfrak{C}$. When $B = \mathsf{X}_{j=1}^{l} B_j$ is an l-dimensional box, the voting scheme $f\colon B^n \to B$ is a GMVS if $f = (f_1, ..., f_l)$ and each f_j is the GMVS induced by some left-coalition system $\mathfrak{C}_j$ in B_j.

Observe that by (*) two different left coalition systems always define different voting schemes. Hence, any GMVS f induces a unique family of coalition systems $\{\mathfrak{C}_j\}_{j=1}^{l}$ (i.e., f_j is the GMVS defined by $\mathfrak{C}_j$)

Comparing (*) to the third definition of the median, we conclude that $f(\tilde{\alpha}) = \mu(\tilde{\alpha})$ if and only if f is induced by $\mathfrak{C}$, where $\mathfrak{C}(\alpha) = \{C \subset N \mid \# C \geqslant k+1\}$ for all α. This is our motivation for calling a scf defined by (*) a GMVS: a winning coalition of players (i.e., $C \in \mathfrak{C}(\alpha)$) can guarantee an outcome no greater than α whenever their bliss points are no greater than α, irrespective of the preferences of the remaining agents. Moreover, the outcome is the lowest integer ξ such that there exists some winning coalition of voters who prefer an outcome ξ or less.

Lemma 3 below establishes that every sp voting scheme is a GMVS.

LEMMA 3. *Let $B = [a, b]$ be an integer interval. A voting scheme $f\colon B^n \to B$ is strategy-proof iff it is a GMVS.*

Proof. Let $\mathfrak{C}$ be a left-coalition system in B and $f\colon B^n \to B$ be defined by (*). Pick any $i \in N$, $\alpha^{-i} \in B^{n-1}$, and $\beta, \gamma \in B$. Let $\xi := f(\beta, \alpha^{-i})$ and $\eta := f(\gamma, \alpha^{-i})$. Suppose $\gamma < \beta$. Then $C((\gamma, \alpha^{-i}); \delta) \supset C((\beta, \alpha^{-i}); \delta)$ for all $\delta \in B$, and $C((\gamma, \alpha^{-i}); \delta) = C((\beta, \alpha^{-i}); \delta)$ for all $\delta \in B \setminus [\gamma, \beta)$. Therefore, if $\xi \geqslant \beta$, $\eta = \xi$; and if $\xi < \beta$, $\eta \leqslant \xi$. In either case, $|\beta - \xi| + |\xi - \eta| = |\beta - \eta|$. Now assume $\gamma \geqslant \beta$. Then $C((\gamma, \alpha^{-i}); \delta) \subset C((\beta, \alpha^{-i}); \delta)$ for all $\delta \in B$, and $C((\gamma, \alpha^{-i}); \delta) = C((\beta, \alpha^{-i}); \delta)$ for all $\delta \in B \setminus [\beta, \gamma)$. Therefore, if $\xi < \beta$, $\eta = \xi$; and if $\xi \geqslant \beta$, $\eta \geqslant \xi$. Again, $|\beta - \xi| + |\xi - \eta| = |\beta - \eta|$. By Lemma 2, this shows that f is sp.

To establish the converse, assume f is sp and for each $\xi \in B$ define

$$\mathfrak{C}(\xi) := \{C \subset N \mid f(\tilde{\alpha}) \leqslant \xi \text{ for some } \tilde{\alpha} \in B^n \text{ with } \alpha^i \leqslant \xi \text{ for } i \in C \text{ and } \alpha^i > \xi \text{ for } i \notin C\}.$$

We prove that this family is a left-coalition system. Let $\xi \in B$, $C \in \mathfrak{C}(\xi)$, and $\tilde{\beta} \in B^n$ be such that $\beta^i \leqslant \xi$ for all $i \in C$. We first show that $f(\tilde{\beta}) \leqslant \xi$. By contradiction, suppose $f(\tilde{\beta}) > \xi$. Let $\tilde{\alpha} \in B^n$ be such that $\alpha^i \leqslant \xi$ for $i \in N$, $\alpha^i > \xi$ for $i \notin N$, and $f(\tilde{\alpha}) \leqslant \xi$. To simplify the notation, w.o.l.g. assume that $C = \{1, ..., k\}$. Starting with profile $\tilde{\beta}$, sequentially change β^i to α^i for each $i \leqslant k$. If at some point the outcome switches to an alternative less than or equal to ξ, we would have a contradiction. Indeed, if for some $m < k$,

$$\gamma := f(\alpha^1, ..., \alpha^m, \beta^{m+1}, ..., \beta^n) > \xi$$

and

$$\delta := f(\alpha^1, ..., \alpha^{m+1}, \beta^{m+2}, ..., \beta^n) \leqslant \xi,$$

then there exists $P \in \mathbf{P}(\beta^{m+1})$ with $\delta P \gamma$ (recall that $\beta^{m+1} \leqslant \xi$). But a player $m+1$ with such a preference would prefer to vote like a player with a preference in $\mathbf{P}(\alpha^{m+1})$ when his opponent's profile is $(\alpha^1, ..., \alpha^m, \beta^{m+2}, ..., \beta^n)$, and f would not be sp. Therefore,

$$f(\alpha^1, ..., \alpha^k, \beta^{k+1}, ..., \beta^n) > \xi.$$

Starting with profile $(\alpha^1, ..., \alpha^k, \beta^{k+1}, ..., \beta^n)$, sequentially switch β^i to α^i for each $i > k$. Since $f(\tilde{\alpha}) \leqslant \xi$, eventually the outcome will switch to a point less than or equal to ξ. Hence, for some $m \geqslant k$,

$$\gamma := f(\alpha^1, ..., \alpha^m, \beta^{m+1}, ..., \beta^n) > \xi$$

and

$$\delta := f(\alpha^1, ..., \alpha^{m+1}, \beta^{m+2}, ..., \beta^n) \leqslant \xi.$$

Since $\alpha^{m+1} > \xi$, there exists $P \in \mathbf{P}(\alpha^{m+1})$ such that $\gamma P \delta$. A player $m+1$ with such a preference would prefer to vote like a player with a preference in $\mathbf{P}(\beta^{m+1})$ when his opponent's profile is $(\alpha^1, ..., \alpha^m, \beta^{m+2}, ..., \beta^n)$. This contradicts the fact that f is sp.

Suppose $\xi \in B$, $\tilde{\beta} \in B^n$ and $C(\tilde{\beta}; \xi) \notin \mathfrak{C}(\xi)$. We now show that $f(\tilde{\beta}) > \xi$. If $\mathfrak{C}(\xi) = \varnothing$, then $f(\tilde{\beta}) > \xi$. Otherwise, let $C \in \mathfrak{C}(\xi)$ be a minimal coalition containing $C(\tilde{\beta}; \xi)$, and let $\tilde{\alpha} \in B^n$ with $f(\tilde{\alpha}) \leqslant \xi$, $\alpha^i \leqslant \xi$ for all $i \in C$ and $\alpha^i > \xi$ for all $i \notin C$. Again, assume w.l.o.g. that $C = \{1, ..., k\}$. Starting with profile $\tilde{\beta}$, sequentially switch β^i to α^i for each $i \leqslant k$. Since $f(\tilde{\alpha}) \leqslant \xi$, there exists $m < k$ such that

$$f(\alpha^1, ..., \alpha^m, \beta^{m+1}, ..., \beta^n) > \xi \quad \text{and} \quad f(\alpha^1, ..., \alpha^{m+1}, \beta^{m+2}, ..., \beta^n) \leqslant \xi.$$

This contradicts the fact that f is sp. Therefore $f(\tilde{\beta}) \leqslant \xi$ iff $C(\tilde{\beta}; \xi) \in \mathfrak{C}(\xi)$. Q.E.D.

It is easy to verify that the image of a one-dimensional GMVS is a one-dimensional subbox. Clearly, products of subboxes are subboxes. Hence, it follows from Corollary 1, Theorem 2, and Lemma 3 that a spscf is a separable voting scheme if and only if its image is a subbox.

Remark. For future reference (see the proof of Lemma 6) we record the following observation. Pick any $i \in N$, $\alpha^{-i} \in B^{n-1}$, and $\beta, \gamma \in B$, and let $\xi := f(\beta, \alpha^{-i})$ and $\eta := f(\gamma, \alpha^{-i})$. In the proof of part (1) we have shown that if $\beta < \xi$, then

$$\eta = \xi \text{ for all } \gamma \leqslant \beta \qquad \text{and} \qquad \eta \geqslant \xi \text{ for all } \gamma > \beta,$$

while if $\beta > \xi$, then

$$\eta = \xi \text{ for all } \gamma \geqslant \beta \qquad \text{and} \qquad \eta \leqslant \xi \text{ for all } \gamma < \beta.$$

We summarize the results of this section in the following theorem.

THEOREM 3. *Suppose the range of the* scf $\varphi: \mathbf{P}^n \to B$ *is a subbox of* B. *Then,* φ *is sp iff it is a multidimensional GMVS.*

Theorem 3 provides a convenient tool to ascertain properties of spscf's. For example, this particular characterization provides relatively simple proofs of Theorems 5 and 7 below.,

3. Social Welfare Functions

In this section we study the problem of preference aggregation in multidimensional single-peaked domains. So far we have concentrated exclusively on assigning a (social) outcome to each profile of preferences. In many problems, one would like a rule for choosing a social outcome among the feasible alternatives before learning the constraints on the feasible set. Thus the object of investigation is a set B, a collection $\mathbf{P}$ of preferences on B, and a mapping $F: \mathbf{P}^n \to \mathbf{P}$. Such a function F will be called a social welfare function (swf). For each preference profile, F picks a social ordering, which after learning the constraints on the feasible set is used to select the social outcome. We address two issues. Given that the agents know the procedure to select a final outcome, do they have correct incentives to truthfully reveal their preferences for *every* possible feasible set? Theorem 4 below shows that in multidimensional single-peaked domains, if the collection of feasible sets is the collection of all subboxes of an l-dimensional box, this stronger strategy-proofness requirement entails no new restriction. That is, if the scf defined by the most preferred outcome of $F(\tilde{P})$ is sp, then, for any subbox A, the scf defined by the most preferred outcome of $F(\tilde{P})$ in A is sp. The second issue involves the relationship between the normative considerations that have motivated Arrow's General Possibility Theorem [1] and the normative and strategic considerations underlying the Gibbard–Satterthwaite Theorem [5, 8]. This issue is addressed by Theorem 5. A precise statement and analysis of Theorem 5 require a few definitions.

For the definitions below, let $\mathfrak{B}$ denote any collection of subsets of some arbitrary set B, and $\mathbf{P}$ denote a collection of preferences on B.

DEFINITION. $F: \mathbf{P}^n \to \mathbf{P}$ satisfies unanimity if whenever there is x such that $xP_i y$ for all $\{x, y\} \in \mathfrak{B}$ and $i \in N$, then $xF(\tilde{P})y$ for all $\{x, y\} \in \mathfrak{B}$.

DEFINITION. F is dictatorial if there exists i such that for all $\{x, y\} \in \mathfrak{B}$, not $yP_i x$ implies not $yF(\tilde{P})x$.

DEFINITION. F satisfies independence of irrelevant alternatives (IIA) if for $\{x, y\} \in \mathfrak{B}$ and any two profiles $\tilde{P}$ and $\tilde{P}'$, $(xP_i y \Leftrightarrow xP'_i y$ for all $i \in N)$ implies $(xF(\tilde{P})y \Rightarrow xF(\tilde{P}')y)$. A slightly stronger condition, obtained from IIA by replacing "$\Leftrightarrow$" with "$\Rightarrow$" will be called non-negative response (NNR).

Observe that if $\mathfrak{B}$ includes all two-element subsets of B, then Arrow's celebrated theorem establishes that if F satisfies unanimity and IIA then it is dictorial. In the current framework we have modified Arrow's restrictions to allow for the possibility that the set of binary comparisons for which the conditions apply may be a strict subset of the set of all possible pairs. This is motivated by our desire to view the swf as a rule for making decisions in every relevant contingency and to restrict the feasible sets to the collection of all subboxes. Note that for any single-peaked preference P, it is enough to know how P compares contiguous pairs (i.e., pairs such that $\{x, y\}$ is a subbox) to know how P would choose from every subbox A. Hence if the problem is one of choosing from every subbox, then preferences over all contiguous pairs are the relevant binary comparisons.

For the remainder of this section we concentrate exclusively on the case in which B is an l-dimensional box, $\mathbf{P}$ is the set of all multidimensional single-peaked preferences, and $\mathfrak{B}$ is the set of all subboxes of B. Recall that $A = \times_{j=1}^{l} A_j$ is a subbox of B, if for each j, A_j is an integer subinterval of B_j. By Lemma 1, two single-peaked preferences with the same bliss point have the same most preferred element in every subbox. Since the bliss point of a single-peaked preference uniquely determines the most preferred outcome in each subbox, we can write $F: \mathbf{P}^n \to B$ instead of $F: \mathbf{P}^n \to \mathbf{P}$. Lemma 4 below states that if F satisfies IIA then it must be sensitive only to the bliss points of the preferences in the profile $\tilde{P}$. Therefore, we can view a swf F satisfying IIA as a map from B^n into B. Such a swf will be called a representative voter scheme. The relationship between a representative voter scheme F and the corresponding swf is obviously analogous to the relationship between a voting scheme and the corresponding scf. Note that for every $A \in \mathfrak{B}$ the representative voter scheme F defines a voting scheme f_A, where $f_A(\tilde{\alpha})$ is defined to be the most preferred outcome of $F(\tilde{\alpha})$ in A. Theorem 4 below shows that F viewed as a voting scheme (i.e., f_B) is sp iff f_A is sp for all $A \in \mathfrak{B}$. That is, it is a dominant strategy for all agents to report truthfully across all possible realizations of other agents' preferences and across all possible realizations of society's costraints.

LEMMA 4. *Let F: $\mathbf{P}^n \to B$ be a swf satisfying IIA. Then, $F(\tilde{P}) = F(\tilde{P}')$ for any two profiles $\tilde{P}$ and $\tilde{P}'$ such that $\tau(P_i) = \tau(P_i')$ for all i.*

Since for a swf F satisfying IIA, $F(P)$ depends only on $(\tau(P_1), ..., \tau(P_n))$, it will be convenient to adopt the following notation. For any x, y, $\alpha \in B$, we write $x\alpha y$ if xPy for all $P \in \mathbf{P}(\alpha)$.

THEOREM 4. *Let F: $B^n \to B$ be a representative voter scheme. Then, F viewed as a voting scheme is sp iff the voting scheme f_A is sp for each subbox $A \in \mathfrak{B}$.*

Proof. Suppose f_A is sp for each $A \in \mathfrak{B}$. Then f_B is sp, and since $F(\tilde{\alpha}) = f_B(\tilde{\alpha})$ for each $\tilde{\alpha} \in B^n$, F is a sp voting scheme. Conversely, assume F is a sp voting scheme. Let $\tilde{\alpha} \in B^n$ and i be arbitrary. W.l.o.g. we can assume that $F(\tilde{\alpha})$ is greater than or equal to α^i in every dimension. Then $\|F(\alpha^{-i}, \beta^i) - \alpha^i\| = \|F(\alpha^{-i}, \beta^i) - F(\tilde{\alpha})\| + \|F(\tilde{\alpha}) - \alpha^i\|$ for each $\beta^i \in B$. Thus, $F(\alpha^{-i}, \beta^i)$ is greater than or equal to $F(\tilde{\alpha})$ in every dimension. Hence, $f_A(\alpha^{-i}, \beta^i)$, the unique minimizer of $\|F(\alpha^{-i}, \beta^i) - x\|$ over $x \in A$, is greater than or equal to $f_A(\tilde{\alpha})$, the unique minimizer of $\|F(\tilde{\alpha}) - x\|$ over $x \in A$, in every dimension. Thus, $\|f_A(\alpha^{-i}, \beta^i) - \alpha^i\| = \|f_A(\alpha^{-i}, \beta^i) - f_A(\tilde{\alpha})\| + \|f_A(\tilde{\alpha}) - \alpha^i\|$, which by Lemma 2 establishes that $f_A(\tilde{\alpha})\, \alpha^i f_A(\alpha^{-i}, \beta^i)$, as desired. Q.E.D.

We now turn to the issue of relating Arrow's normative conditions on swf's to strategy-proofness. Arrow's General Possibility Theorem establishes that on unrestricted domains no swf can respect IIA, choose the unambiguously most preferred social outcome when one exists, and still achieve some level of compromise across the members of society. The intuitive appeal of IIA is clear. Nevertheless, Arrow's Theorem shows that in unrestricted domains one can not insist on IIA without confronting obviously less attractive consequences such as the failure of unanimity or dictatorship. A separate argument in favor of IIA is the following. Many proofs of the Gibbard–Satterthwaite Theorem utilize Arrow's Theorem. Thus, one might suspect that IIA is related to sp. The theorem below establishes this relationship. It shows that on single-peaked domains, NNR, a condition slightly stronger than IIA is equivalent to sp, provided one restricts the binary comparisons to the set of contiguous outcomes. Thus, single-peaked preferences provide an interesting example of a domain that affords non-trivial spscf's and non-trivial swf's that satisfy a version of IIA. The theorem below shows that these are in fact the same functions.

THEOREM 5. *The swf $F: B^n \to B$ satisfies NNR iff viewed as a scf it is sp.*

Proof. Pick j in $\{1, ..., l\}$ and $\xi \in B$ such that $\xi_j < b_j$. Let $\eta_k = \xi_k$ for all $k \neq j$, and $\eta_j = \xi_j + 1$. Define $\mathfrak{C}_j(\xi)$ as follows: $D \in \mathfrak{C}_j(\xi)$ iff $\xi F(\tilde{\alpha}) \eta$ when $\alpha^i = \xi$ for all $i \in D$ and $\alpha^i = \eta$ for all $i \notin D$. By NNR, (1) if $C \in \mathfrak{C}_j(\xi)$ and $C \subset D$, then $D \in \mathfrak{C}_j(\xi)$. By IIA (which is implied by NNR), for any $\tilde{\alpha}$, $\{i \mid \xi \alpha^i \eta\} \in \mathfrak{C}_j(\xi)$ iff $\xi F(\tilde{\alpha}) \eta$. By single-peakedness, this implies that

$$\{i \mid \xi \alpha^i \eta\} \in \mathfrak{C}_j(\xi) \quad \text{iff} \quad F(\tilde{\alpha})_j \leqslant \xi_j. \qquad (*)$$

It follows from (*) that (2) if $\eta_j < \xi_j$ and $C \in \mathfrak{C}_j(\eta)$, then $C \in \mathfrak{C}_j(\xi)$. Hence, the collection $\mathfrak{C}_j(\xi)$ depends only on ξ_j and not on other components of ξ. Therefore, we denote it by $\mathfrak{C}_j(\xi_j)$ hereafter. Define (3) $\mathfrak{C}_j(b_j) = 2^N$. By (1), (2), and (3), $\mathfrak{C}_j$ is a left coalition system $j = 1, ..., l$. By (*), F is induced by $\{\mathfrak{C}_j\}_{j=1}^l$. Hence F is a GMVS and by Theorem 4, F interpreted as a scf is sp.

To prove the converse, note that if F is sp, then $\{\xi, \eta\} \in B$ (i.e., ξ and η are contiguous) implies $f_{\{\xi, \eta\}}$ is sp (Theorem 5). Therefore, $\{i \mid \xi \alpha^i \eta\} \subset \{i \mid \xi \beta^i \eta\}$ and $f_{\{\xi, \eta\}}(\tilde{\alpha}) = \xi$ implies that $f_{\{\xi, \eta\}}(\tilde{\beta}) = \xi$. Moreover, $f_{\{\xi, \eta\}}(\tilde{\alpha}) = \xi$ iff $\xi F(\tilde{\alpha}) \eta$. Therefore F satisfies NNR. Q.E.D.

In the next section we exploit this dual interpretation of functions from B to B as swf's and scf's to analyze committees and the decentralization of the decision-making process.

4. COMMITTEES AND SIMPLE CHOICE RULES

In this section we explore the consequences of interpreting voting schemes as swf's for disaggregating the social decision process. In particular, we have in mind the following decision process. A given subset of the agents will form a committee and determine a representative as the outcome of their committee. Another (possibly overlapping) committee will form its own representative. The two representatives will perhaps join a third committee to determine a social ordering and/or social outcome. We are interested in the following questions. If the rules to construct representatives and to aggregate their preferences are sp, is the overall decision process sp? To what extent can we characterize the overall decision rules by studying the decision rules within the committees? We have already noted the possibility of interpreting an element $\alpha \in B$ in two different ways: as a set of preferences and as an outcome. Clearly, for elements in the domain (of both scf's and swf's), the first interpretation is appropriate; in terms of the range, the first is appropriate for swf's and the second for scf's. We no longer remark on the various interpretations and simply call functions from B^n to B voting schemes. We leave it to the reader to note which interpretation is adequate in the given context.

To formalize our notion of voting in committees and committee representative, we need the following definitions.

DEFINITION. The function $M: \bigcup_{k=1}^{\infty} (\mathbf{Z}^l)^{2k+1} \to \mathbf{Z}^l$, defined by

$$M(\alpha^1, ..., \alpha^{2k+1}) := (\mu(\alpha_1^1, ..., \alpha_1^{2k+1}), ..., \mu(\alpha_l^1, ..., \alpha_l^{2k+1})),$$

is called the *multidimensional median value function*. That is, $M(\alpha^1, ..., \alpha^{2k+1})$ is the coordinate-wise median value of the vectors $\alpha^1, ..., \alpha^{2k+1} \in \mathbf{Z}^l$. If n is odd, M_n denotes the restriction of M to $(\mathbf{Z}^l)^n$ (M_n is the n-players median value function).

It follows from Theorem 2 and the fact that one-dimensional medians are sp that M_n, interpreted as a voting scheme over the set of all single-peaked preferences on B, is sp.[8] Our notion of single-peaked preferences has a claim to being the appropriate generalization of Moulin's to multidimensional arrays of outcomes, provided one wants to preserve the close relationship between spscf's and voting schemes. Our set of single-peaked preferences is the largest collection of strict preferences for which the voting scheme $M_n(\tau(P_1), ..., \tau(P_n))$ is sp for any $n > 1$ (provided for all $\alpha \in B$ there exists some P in the collection such that $\tau(P) = \alpha$). This is established in Theorem 6 below.

THEOREM 6. *Let* $\mathbf{P}$ *be any collection of preferences such that* $M_n: \mathbf{P}^n \to B$ *is strategy-proof for* $n = 2k+1$, $k > 0$. *Suppose* $\tau(\mathbf{P}) = B$ (*i.e., for every* $\alpha \in B$ *there exists* $P \in \mathbf{P}$ *such that* $\tau(P) = \alpha$). *Then, every* $P \in \mathbf{P}$ *must be a single-peaked preference.*

Proof. Assume to the contrary that $P \in \mathbf{P}$ is not single-peaked, and let $\tau(P) = \alpha$. Then, there exists $\beta, \gamma \in B$ such that $\|\alpha - \beta\| + \|\beta - \alpha\| = \|\alpha - \gamma\|$ and not $\beta P \gamma$. Consider any profile in which there are k individuals who prefer β the most, k individuals who prefer γ the most, and one person with preference P. The social outcome for this profile would be β. If P were to misrepresent his preference and claim that his most preferred outcome is γ, then the social outcome would have been γ, which contradicts the strategy-proofness of M_n. Q.E.D.

[8] Essentially the same observation is made by Border and Jordan [3].

Theorem 7 below establishes that every sp voting scheme f can be represented by a ternary tree of the following kind: there is an initial node which leads to three subsequent nodes. Each subsequent node either is a terminal node or leads to three subsequent nodes, and so on. Every terminal node is labeled by either a constant outcome or the bliss point α^i of an agent's preference. To construct the social outcome, we "solve the tree backward." At any node that leads to three labeled nodes, we compute the median of the labels and label the node with this outcome. Repeat the process until the initial node is labeled. The label assigned to the initial node corresponds to the social outcome $f(\tilde{\alpha})$. Contrast this result for voting schemes over domains of single-peaked preferences with that for scf's over unrestricted domains. For the latter, only constant and dictatorial rules are sp (by the Gibbard–Satterthwaite Theorem). For the former, one needs to include the median voter scheme, and hence all combinations that can be obtained with it and the constant and dictatorial rules. The result is also interesting in that it shows how any spscf can be reconstructed with an appropriate set up of committees and the use of the median voter scheme.

In the following scheme sequence of definitions and lemmas, let B be an l-dimensional box.

DEFINITION. The sets of all n-players voting schemes and sp voting schemes on B are $\mathfrak{F}_n^* := \{f \mid f: B^n \to B\}$ and $\mathfrak{F}_n^{**} := \{f \in \mathfrak{F}_n^* \mid f \text{ is sp}\}$, respectively. The sets of all voting schemes and all sp voting schemes on B are $\mathfrak{F}^* := \bigcup_{n \geqslant 3} \mathfrak{F}_n^*$ and $\mathfrak{F}^{**} := \bigcup_{n \geqslant 3} \mathfrak{F}_n^{**}$, respectively.

In analyzing the committee decision process, we consider three different simple choice functions: dictatorships, constants, and median voter schemes. We have already defined median voter schemes and noted they are sp; that dictatorital and constant social function are sp is obvious.

DEFINITION. For each $n \geqslant 3$ we have the following special scf on B^n:

(1) for each $i \in N$, $I_n^i(\alpha) := \alpha^i$ for all $\alpha \in B^n$ is player i's dictatorial rule;

(2) for each $\xi \in B$, $\xi_n^*(\alpha) := \xi$ for all $\alpha \in B^n$ is the ξ-constant rule.

DEFINITION. Let $f_0 \in \mathfrak{F}_m^*$ and $f_k \in \mathfrak{F}_n^*$, $k = 1, ..., m$. The *combination* $f_0 * (f_1, ..., f_m)$ is the scf on B^n defined by:

$$f_0 * (f_1, ..., f_m)(\tilde{\alpha}) := f_0(f_1(\tilde{\alpha}), ..., f_m(\tilde{\alpha})) \qquad \text{for each} \quad \tilde{\alpha} = (\alpha^1, ..., \alpha^n) \in B^n.$$

DEFINITION. Let $\mathfrak{F} \subset \mathfrak{F}^*$ and for each $n \geqslant 3$ define $\mathfrak{F}_n := \mathfrak{F} \cap \mathfrak{F}_n^*$. The *span* $\langle \mathfrak{F} \rangle$ of $\mathfrak{F}$ is the smallest collection of scf closed under combinations that contains $\mathfrak{F}$. Recursively define

$$K^1(\mathfrak{F}) := \{f_0 * (f_1, ..., f_m) \mid \text{for some } n \text{ and } m, f_0 \in \mathfrak{F}_m \text{ and } f_k \in \mathfrak{F}_n, k = 1, ..., m\},$$

$$\text{and for } k \geqslant 1,\ K^{k+1}(\mathfrak{F}) := K^1\left[\bigcup_{j=0}^{k} K^j(\mathfrak{F})\right],$$

where $K^0(\mathfrak{F}) := \mathfrak{F}$. Finally, let $K(\mathfrak{F}) := \bigcup_{j=0}^{\infty} K^j(\mathfrak{F})$.

The following technical result states that the span of any collection of voting schemes can be obtained recursively.

LEMMA 5. *$K(\mathfrak{F}) = \langle \mathfrak{F} \rangle$ for all $\mathfrak{F} \subset \mathfrak{F}^*$.*

Proof. Let $\mathfrak{F} \subset \mathfrak{F}^*$. Clearly $K(\mathfrak{F}) \subset \langle \mathfrak{F} \rangle$. Thus we need only show that $K(\mathfrak{F}) \supset \langle \mathfrak{F} \rangle$, or equivalently, that $K(\mathfrak{F})$ is closed under combinations. But if for some n and m, $f_0 \in \langle \mathfrak{F} \rangle_m$ and $f_k \in \langle \mathfrak{F} \rangle_n$, $k = 1, ..., m$, then there exists t such that $f_k \in \bigcup_{j=0}^{t} K^j(\mathfrak{F})$ for all $k = 0, 1, ..., m$. Therefore

$$f_0 * (f_1, ..., f_m) \in K^1\left[\bigcup_{j=0}^{t} K^j(\mathfrak{F})\right] = K^{t+1}(\mathfrak{F}) \subset K(\mathfrak{F}). \qquad \text{Q.E.D.}$$

The following lemma demonstrates the importance of the combination operation. It shows that combinations of sp voting schemes are sp.

LEMMA 6. *$K^1(\mathfrak{F}) \subset \mathfrak{F}^{**}$ and $K(\mathfrak{F}) \subset \mathfrak{F}^{**}$ for all $\mathfrak{F} \subset \mathfrak{F}^{**}$.*

Proof. We first show the following result: let $f \in \mathfrak{F}_n^{**}$, $\tilde{\alpha} \in B^n$, β, $\gamma \in B$, and $i \in N$, be such that $\|\alpha^i - \gamma\| = \|\alpha^i - \beta\| + \|\beta - \gamma\|$. Then, if $\xi := f(\beta, \alpha^{-i})$ and $\eta := f(\gamma, \alpha^{-i})$, we have $\|\alpha^i - \eta\| = \|\alpha^i - \xi\| + \|\xi - \eta\|$. Recall the remark following Theorem 3; for each $j = 1, ..., l$, we obtain the following cases:

if $\alpha_j^i \leqslant \beta_j \leqslant \gamma_j$ then [if $\xi_j < \beta_j$, then $\eta_j = \xi_j$, and if $\xi_j \geqslant \beta_j$, then $\eta_j \geqslant \xi_j$];

if $\alpha_j^i \geqslant \beta_j \geqslant \gamma_j$ then [if $\xi_j > \beta_j$, then $\eta_j = \xi_j$, and if $\xi_j \leqslant \beta_j$, then $\eta_j \leqslant \xi_j$].

In all cases, $|\alpha_j^i - \eta_j| = |\alpha_j^i - \xi_j| + |\xi_j - \eta_j|$; this proves the equality above.

Let $\mathfrak{F} \subset \mathfrak{F}^{**}$, and for some n and m, let $f_0 \in \mathfrak{F}_m$ and $f_k \in \mathfrak{F}_n$, $k = 1, ..., m$. Pick any $i \in N := \{1, ..., n\}$, $\alpha^{-i} \in B^{n-1}$, and β, $\gamma \in B$. Let

$$\xi^k := f_k(\beta, \alpha^{-i}), \qquad \eta^k := f_k(\gamma, \alpha^{-i}), \qquad k = 1, ..., m,$$

$$\xi^0 := f_0(\xi^1, ..., \xi^m), \qquad \text{and} \qquad \eta^0 := f_0(\eta^1, ..., \eta^m).$$

To conclude, we need to show that $\|\beta - \eta^0\| = \|\beta - \xi^0\| + \|\xi^0 - \eta^0\|$. Let $\delta^0 := \eta^0$ and $\delta^k := f_0(\xi^1, ..., \xi^k, \eta^{k+1}, ..., \eta^m)$, $k = 1, ..., m$ (note that $\delta^m = \xi^0$). Since each f_k is sp, $\|\beta - \eta^k\| = \|\beta - \xi^k\| + \|\xi^k - \eta^k\|$, and from the result above, we get that $\|\beta-\delta^k\| = \|\beta-\delta^{k+1}\| + \|\delta^{k+1}-\delta^k\|$ for each $k=0, 1, ..., m-1$. Adding these equalities we obtain:

$$\|\beta-\delta^0\| = \|\beta-\delta^n\| + \sum_{k=0}^{m-1} \|\delta^{k+1}-\delta^k\|. \qquad (*)$$

Recall that $\delta^0 := \eta^0$ and $\delta^n = \xi^0$, and note that

$$\|\xi^0-\eta^0\| \leqslant \sum_{k=0}^{m-1} \|\delta^{k+1}-\delta^k\| \quad \text{and} \quad \|\beta-\eta^0\| \leqslant \|\beta-\xi^0\| + \|\xi^0-\eta^0\|;$$

these two inequalities, together with $(*)$, imply that

$$\|\xi^0-\eta^0\| = \sum_{k=0}^{m-1} \|\delta^{k+1}-\delta^k\| \quad \text{and} \quad \|\beta-\eta^0\| = \|\beta-\xi^0\| + \|\xi^0-\eta^0\|;$$

Q.E.D.

Note that for any n and $k<n$, stacking the k dictatorial rules $I_n^1, ..., I_n^k$ and constructing the combination $f_0 * (I_n^1, ..., I_n^k)$ results in the aggregate rule f_0 for the committee $\{1, ..., k\}$. Thus, the concept of a dictatorial rule is an integral part of our notion of a committee. A voter within the committee can be an agent (represented by the rule I_n^k), a constant, or the outcome of some other committee. Theorem 7 establishes that every sp voting scheme can be represented by a committee structure in which every committee has exactly three members and aggregates the preferences of its members according to the median voter scheme. It is worth noting that our notion of a committee does not allow for separate committees for each dimension. Hence, Theorem 7 cannot be proved by considering the one-dimensional case and appealing to Theorem 2 as we did in proving Theorem 3.

DEFINITION. Let $\hat{\mathfrak{F}} := [\bigcup_{n\geqslant 3} \hat{\mathfrak{F}}_n]$, where

$$\hat{\mathfrak{F}}_3 := \{I_3^i \mid i=1, 2, 3\} \cup \{\xi_3^* \mid \xi \in B\} \cup \{M_3\},$$
$$\hat{\mathfrak{F}}_n := \{I_n^i \mid i=1, ..., n\} \cup \{\xi_n^* \mid \xi \in B\}, n \geqslant 4.$$

THEOREM 7. $\langle \hat{\mathfrak{F}} \rangle = \mathfrak{F}^{**}$.

Proof. See Appendix.

5. CONCLUSION

Having established all our results and stated the related definitions, we now are in a position to compare them to those provided by Moulin [7] and Border and Jordan [3]. As noted earlier, Moulin [7] provides an extensive analysis of single-peaked preferences in the one-dimensional case (i.e., when $B=\mathfrak{R}$). His characterization of anonymous strategy-proof voting schemes can be expressed as follows: if $f\colon \mathfrak{R}^n \to \mathfrak{R}$ is an anonymous sp voting scheme, then there exists $\tilde{\beta} \in \mathfrak{R}^{n+1}$ such that

$$f(\tilde{\alpha}) = \mu(\tilde{\alpha}, \tilde{\beta}) \qquad \text{for all} \quad \tilde{\alpha} \in \mathfrak{R}^n.$$

His general characterization of sp voting schemes states that if $f\colon \mathfrak{R}^n \to \mathfrak{R}$ is sp, then for each coalition $C \subset N$, there exists $\alpha_C \in \mathfrak{R} \cup \{-\infty, +\infty\}$ such that

$$f(\tilde{\alpha}) = \inf\{\sup\{x \mid x = \alpha_C \text{ or } x = \alpha^i \text{ for some } i \in C\} \mid C \subset N\}.$$

Moulin also makes the following observations:

(1) infs and sups can be expressed as combinations of constants and medians, and medians can be expressed as combinations of infs, sups, and dictatorial voting schemes.

(2) Combinations of sp voting schemes are sp.

(3) Combinations can be interpreted as committee representation and viewed as a decentralization of the decision-making process.

Thus, his characterization of the non-anonymous voting schemes combined with (3) results in a conclusion similar to our Theorem 7. Our work is complementary to his in the following ways:

(a) We show that his restriction to voting schemes is without loss of generality, provided the range of the scf is a subbox (i.e., an interval in the one-dimensional case).

(b) We extend his results to the multidimensional case.

(c) We provide an alternative characterization of sp voting schemes (i.e., GMVS) which includes the work of Barberà, Sonnenschein, and Zhou [2] and the Farquharson [4] and Kramer [6] characterizations of a voting rule as special cases.

(d) Note that our separability property (Theorem 2) does not depend on the set of alternatives being discrete and that combinations of medians are GMVSs. Hence, Moulin's observation (1) implies that our characterization of sp voting schemes in terms of GMVSs can be extended to include the case in which $B=\mathfrak{R}^l$, or B is any box in $\mathfrak{R}^l$.

(e) Minor modifications of the proof of Theorem 7 (i.e., committee representation) and the observation that sp voting schemes are GMVSs, will also lead to an extension of Theorem 7 that includes the case in which $B=\mathfrak{R}^l$ or B is a subbox of $\mathfrak{R}^l$.

Border and Jordan [3] also extend Moulin's work to $\mathfrak{R}^l$. Their main theorem for the multidimensional case states the following result. If $f: \Sigma^n \to \mathfrak{R}^l$ is a surjective scf, where Σ denotes the set of all separable preferences on $\mathfrak{R}^l$, then f is sp iff it is separable in each coordinate. A lemma used to prove this result provides a result similar to our Theorem 1. In addition, they show that for the multidimensional case, efficiency and strategy-proofness imply the scf must be dictatorial, and that even small departures from the separable case may lead to dictatorial scf's. Our work complements theirs in at least two ways:

(a) The set of all separable preferences is strictly smaller than the set of (multidimensional) single-peaked preferences. In fact, we show that our notion of single-peaked preferences is the appropriate generalization of Moulin's for multidimensional sets of alternatives, in that such preferences constitute the largest class which extends Moulin's results (i.e., the largest class for which all GMVSs are sp).

(b) We are able to weaken the surjectivity requirement of Theorem 1.

Finally, we consider swf's on single-peaked domains and relate these swf's to the spscf's considered earlier. In particular, we show that swf's viewed as contingent choices are strategy-proof if and only if the unrestricted choice (i.e., the bliss point) implied by the swf is sp. Moreover, we show that the non-negative response property of the swf is equivalent to strategy-proofness and establish the relationship between Arrow's normative conditions and strategy-proofness in a framework that does not rule out non-trivial scf's and swf's.

Appendix

Proof of Theorem 1

Definition. For any subset $A \subset B$, let $\hat{B}(A)$ be the minimal subbox of B that contains it.

Lemma 1.1. *For every profile* $\tilde{P} = (P_1, ..., P_n)$, $\varphi(\tilde{P}) \in \hat{B}(\{\tau(P_i) \mid i = 1, ..., n\})$.

Proof. By contradiction, assume not. Let $\tilde{P}$ be a profile for which $\varphi(\tilde{P}) =: \alpha \notin X := \hat{B}(|\tau(P_i) \mid i = 1, ..., \})$. Let α^* be the point in X closest to α. Note that $\alpha^* P_i \alpha$ for each $i = 1, ..., n$. Pick any $\bar{P}$ with bliss point α^* and such that $\beta \bar{P} \gamma$ for all $\beta \in Y$ and $\gamma \notin Y$, where $Y := \hat{B}(\{\alpha, \alpha^*\})$. Sequentially replace each P_i by $\bar{P}$. Since φ is sp, $\varphi(\bar{P}, P_{-1}) \in Y$, otherwise player 1 with preference $\bar{P}$ would report P_1. In fact, $\varphi(\bar{P}, P_{-1}) = \alpha$, otherwise player 1 with preference P_1 would prefer to report $\bar{P}$ when his opponent's profile is P_{-1}. Repeating this argument for each player, we obtain $\varphi(\bar{P}, ..., \bar{P}) = \alpha$. But this is a contradiction: φ would violate unanimity. Q.E.D.

To prove the theorem, by contradiction suppose $\tilde{P}^*$ is a profile for which there exists j and P_j^+ such that $\tau(P_j^+) = \tau(P_j^*)$, $\alpha^* := \varphi(\tilde{P}^*) \neq \varphi(P^*_{-j}, P_j^+) =: \alpha^+$. For the rest of this proof, j always refers to this voter. In addition, suppose $\tilde{P}^*$ was chosen so that the box $X^* := \hat{B}(\{\tau(P_i^*) \mid i = 1, ..., n\})$, is minimal, in the sense that for no profile $\tilde{P}$, with $\hat{B}(\{\tau(P_i) \mid i = 1, ..., n\})$ strictly contained in X^*, one can switch the preference of a player k to a preference P'_k in such a way that $\tau(P'_k) = \tau(P_k)$ and $\varphi(\tilde{P}) \neq \varphi(P_{-k}, P'_k)$. By the previous lemma, we know that $Y := \hat{B}(\{\alpha^*, \alpha^+\}) \subset X^*$. As an intermediate step, we first show that $Y = X^*$, i.e., that the social outcomes α^* and α^+ must be kitty-corners (i.e., diagonal opposites) of the box X^*.

LEMMA 1.2. $Y = X^*$.

Proof. For every preference P with $\tau(P) \notin Y$, let α be the point in Y closest to $\tau(P)$, and denote by $\pi(P)$ the collection of preferences with bliss point α, which rank each outcome in Y above every outcome not in Y and keep the same relative ranking as P over outcomes in Y. That is, $\pi(P) := \{\bar{P} \in \mathfrak{P}(\alpha) \mid \text{for every } \beta, \gamma \in Y \text{ and } \delta \notin Y, \beta \bar{P} \delta, \text{ and } \beta \bar{P} \gamma \text{ iff } \beta P \gamma\}$. Suppose by contradiction that there exists i such that $\tau(P_i^*) \notin Y$. If $\tau(P_j^*) \in Y$, let $\bar{P}_j = P_j^*$ and $\bar{P}_j^+ = P_j^+$; otherwise, pick any $\bar{P}_j \in \pi(P_j^*)$ and $\bar{P}_j^+ \in \pi(P_j^+)$. Since φ is sp, $\varphi(P^*_{-j}, \bar{P}_j) = \alpha^*$ and $\varphi(P^*_{-j}, \bar{P}_j^+) = \alpha^+$ in either case. Let $i \neq j$ be such that $\tau(P_i^*) \notin Y$, and pick any $\bar{P}_i$ in $\pi(P_i^*)$. Since φ is sp, $\varphi((P^*_{-j}, \bar{P}_j)_{-i}, \bar{P}_i) = \alpha^*$ and $\varphi((P^*_{-j}, \bar{P}_j^+)_{-i}, \bar{P}_i) = \alpha^+$. Sequentially replace every preference P_i with $\tau(P_i^*) \notin Y$ by a preference $\bar{P}_i$ in $\pi(P_i^*)$, and let $\bar{P}_i = P_i^*$ whenever $\tau(P_i^*) \in Y$. Then $\varphi(\bar{P}_1, ..., \bar{P}_n) = \alpha^*$ and $\varphi(\bar{P}_{-j}, \bar{P}_j^+) = \alpha^+$. But this contradicts the minimality of the box X^*. Hence $\tau(P_i^*) \in Y$ and $Y = X^*$. Q.E.D.

Consider any two preferences Q and Q' such that $\tau(Q) = \alpha^*$ and $\tau(Q') = \alpha^+$. Starting with the profile $\tilde{P}^*$, for each $i \neq j$, sequentially switch P_i^* to Q if $\alpha^* P_i^* \alpha^+$, and to Q' otherwise; denote the final profile by $\tilde{P}$ (so $P_j = P_j^*$, and for $i \neq j$, either $P_i = Q$ or $P_i = Q'$). The social outcomes along the sequence thus generated, and along the sequence when player j's preference is replaced by P_j^+, remain constant. For example, if $j \neq 1$ and $\alpha^* P_1^* \alpha^+$, since φ is sp and $\varphi(\tilde{P}^*) = \alpha^*$, we must have $\varphi(P^*_{-1}, Q) = \alpha^*$. Also, we must have $\varphi((P^*_{-j}, P_j^+)_{-1}, Q) \neq \alpha^*$, for otherwise a player 1 with preference P_1^* would rather report Q when his opponents' profile is $(P^*_{-j}, P_j^+)_{-1}$, contradicting strategy-proofness. Since $\varphi((P^*_{-j}, P_j^+)_{-1}, Q) \neq \alpha^*$, the kitty-corner condition just established and the minimality of the box X^* imply $\varphi((P^*_{-j}, P_j^+)_{-1}, Q)=\alpha^+$. Thus, $\varphi(\tilde{P})=\alpha^*$ and $\varphi(P_{-i}, P_i^+)=\alpha^+$.

To conclude, starting with profile $\tilde{P}$, sequentially switch every preference P_i which is equal to Q to the preference P_j^*; denote the resulting final profile $\bar{P}$. Suppose, for example, that $j\neq 1$ and $P_1=Q$. Then $\varphi(P_{-1}, P_j^*)\neq\alpha^+$, for otherwise a player 1 with preference P_j^* would prefer to report Q when his opponents' profile is P_{-1}. Hence, along the sequence from $\tilde{P}$ to $\bar{P}$ the social outcome is never equal to α^+, and thus $\varphi(\bar{P})\neq\alpha^+$. Since $\beta Q\alpha^+$ for all $\beta\in Y$, $\varphi((P_{-1}, P_j^*)_{-j}, P_j^+)=\alpha^+$. Otherwise a player 1 with preference Q would prefer to report P_j^* when his opponent's profile is $(P_{-j}, P_j^+)_{-1}$. Therefore, along the sequence from $\tilde{P}$ to $\bar{P}$, when player j's preference is replaced by P_j^+, the social outcome remains constant at α^+, and $\varphi(\bar{P}_{-j}, P_j^+)=\alpha^+$. By the kitty-corner condition, we must have $\varphi(\bar{P})=\alpha^*$. By Lemma 1.2, $\bar{\alpha}:=\tau(P_j^*)=\tau(P_j^+)\in B(\{\alpha^*, \alpha^+\})=X^*$. Since $\alpha^*\neq\bar{\alpha}$ (otherwise P_j^+ would misrepresent), $\hat{B}(\{\tau(\bar{P}_i)\mid i=1, ..., n\})=B(\{\bar{\alpha}, \alpha^+\})$ is strictly contained in X^*. This contradicts the minimality of X^*. Q.E.D.

Proof of Theorem 7

LEMMA 7. *$M_n\in\langle\hat{\mathfrak{F}}\rangle$ for all $n\geqslant 3$ odd.*

Proof. The proof is by induction. However, to avoid tedious notation, we only show how to construct M_5 from M_3, and M_7 from M_5, and ask the reader to infer the general step. For any four different players i, j, k, $m\in\{1, 2, 3, 4, 5\}$, let

$$f_{ijk}:=M_3*(I_5^i, I_5^j, I_5^k) \qquad\text{and}\qquad g_{ijkm}:=M_3*(f_{ijm}, f_{jkm}, f_{ikm}).$$

Clearly, $f_{ijk}\in\langle\hat{\mathfrak{F}}\rangle$, and therefore $g_{ijkm}\in\langle\hat{\mathfrak{F}}\rangle$.

We now show that $M_5=M_3*(g_{1234}, g_{1235}, f_{123})$, and therefore $M_5\in\langle\hat{\mathfrak{F}}\rangle$. Fix a coordinate $j\in\{1, 2, ..., l\}$. We show that for all $\tilde{\alpha}\in B^5$, the jth coordinate of $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})$ equals the jth coordinate of $M_5(\tilde{\alpha})$. Let $x:=M(\tilde{\alpha})_j$, $C^-:=\{i\mid\alpha_j^i<x\}$ and $C^+:=\{i\mid\alpha_j^i>x\}$. Then, $\#C^-\leqslant 2$ and $\#C^+\leqslant 2$. Suppose $g_{1234}(\tilde{\alpha})_j<x$; since $g_{1234}(\tilde{\alpha})_j$ is the median of $f_{124}(\tilde{\alpha})_j$, $f_{234}(\tilde{\alpha})_j$ and $f_{134}(\tilde{\alpha})_j$, we must have $4\in C^-$ and at least one more element from $\{1, 2, 3\}$ must be in C^-. But $\#C^-\leqslant 2$, so exactly one element of $\{1, 2, 3\}$ must be in C^-, and $5\notin C^-$. Therefore, $f_{123}(\tilde{\alpha})_j\geqslant x$ and $g_{1235}(\tilde{\alpha})_j\geqslant x$, so $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j\geqslant x$. A symmetric argument establishes that $g_{1235}(\tilde{\alpha})_j<x$ implies $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j\geqslant x$. But if both $g_{1234}(\tilde{\alpha})_j\geqslant x$ and $g_{1235}(\tilde{\alpha})_j\geqslant x$, then obviously $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j\geqslant x$. So, $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j\geqslant x$. Again a symmetric argument establishes that $g_{1234}(\tilde{\alpha})_j>x$ or $g_{1235}(\tilde{\alpha})_j>x$ implies $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j\leqslant x$. So $M_3*(g_{1234}, g_{1235}, f_{123})(\tilde{\alpha})_j=x$, as desired.

Similarly, for any six different players $i, j, k, m, n, q \in \{1,, ..., 7\}$, let

$$f_{ijkmn} := M_5 * (I_7^i, ..., I_7^n)$$

and

$$g_{ijkmnq} := M_5 * (f_{ijkmq}, f_{ijknq}, f_{ijmnq}, f_{ikmnq}, f_{jkmnq}).$$

Again, $f_{ijkmn} \in \langle \hat{\mathfrak{F}} \rangle$ and $g_{ijkmnq} \in \langle \hat{\mathfrak{F}} \rangle$. Thus $M_3 * (g_{123456}, g_{123457}, f_{12345}) \in \langle \hat{\mathfrak{F}} \rangle$. An argument similar to that above establishes that $M_3 * (g_{123456}, g_{123457}, f_{12345}) = M_7$. In particular, for any j, there are at most three agents such that their bliss points on the jth coordinate are below $M_7(\tilde{\alpha})_j$. Hence, if $g_{123456}(\tilde{\alpha})_j < M_7(\tilde{\alpha})_j$, $g_{123457}(\tilde{\alpha})_j \geqslant M_7(\tilde{\alpha})_j$ and $f_{12345}(\tilde{\alpha})_j \geqslant M_7(\tilde{\alpha})_j$. Therefore, $M_3 * (g_{123456}, g_{123457}, f_{12345})(\tilde{\alpha})_j \geqslant M_7(\tilde{\alpha})_j$. As before, symmetric arguments establish that $M_7 = M_3 * (g_{123456}, g_{123457}, f_{12345}) \in \langle \hat{\mathfrak{F}} \rangle$. Q.E.D.

For any subsets C and D, let $C \Delta D := [C \backslash D] \cup [D \backslash C\}$, and if C and D are finite, let

$$\Delta(C, D) := \#(C \Delta D) \qquad \text{(Hausdorff "metric")}.$$

For any sp voting scheme $f: B^n \to B$, let $\mathfrak{C}^f = (\mathfrak{C}_1^f, ..., \mathfrak{C}_l^f)$ denote the coalition system induced by f. If $f, g: B^n \to B$, let

$$\delta(f, g) := \sum_{j=1}^{l} \sum_{\xi = a_j}^{b_j} \Delta(\mathfrak{C}_j^f(\xi), \mathfrak{C}_j^f(\xi)).$$

Clearly, if f and g are sp, $f = g$ iff $\delta(f, g) = 0$.

Let $f \in \mathfrak{F}_n^{**}$ and $g \in \langle \hat{\mathfrak{F}} \rangle_n$ be such that $\delta(f, g) > 0$. We construct a scf $h \in \langle \hat{\mathfrak{F}} \rangle_n$ such that $\delta(f, h) < \delta(f, g)$. Therefore, if $\hat{f}$ is the "best approximation" to f in $\langle \hat{\mathfrak{F}} \rangle$ according to the metric δ, we must have $\delta(f, \hat{f}) = 0$, and $f = \hat{f} \in \langle \hat{\mathfrak{F}} \rangle$. For notational convenience we only present the proof for the case in which B is a two-dimensional box (i.e., $l = 2$). W.l.o.g. suppose that $\mathfrak{C}_1^f \neq \mathfrak{C}_1^g$. Let $\xi \in B_1$ be the smallest point for which $\mathfrak{C}_1^f(\xi) \backslash \mathfrak{C}_1^g(\xi) \neq \emptyset$. If $\mathfrak{C}_1^f(\eta) \subset \mathfrak{C}_1^g(\eta)$ for all $\eta \in B_1$, we must have $\mathfrak{C}_1^g(\eta) \backslash \mathfrak{C}_1^f(\eta) \neq \emptyset$ for some $\eta \in B_1$; we consider this case later. Let $C \subset \mathfrak{C}_1^f(\xi) \backslash \mathfrak{C}_1^g(\xi)$. The scf h is constructed so as to satisfy the following conditions:

(i) $\mathfrak{C}_2^h = \mathfrak{C}_2^g$ (that is, $h_2 = g_2$);

(ii) $\mathfrak{C}_1^h(\eta) = \mathfrak{C}_1^g(\eta)$ for all $\eta \in B_1$ with $\eta < \xi$;

(iii) $\mathfrak{C}_1^h(\xi) = \mathfrak{C}_1^g(\xi) \cup \{D \in 2^N \mid C \subset D\}$;

(iv) $\eta > \xi$, $D \notin \mathfrak{C}_1^g(\eta)$ and $C \not\subset D$ imply $D \notin \mathfrak{C}_1^h(\eta)$.

Condition (i) says that the coalition system induced by h in B_2 is the same as that induced by g, while conditions (ii)–(iv) guarantee that the coalition system induced by h on B_1 is the minimal coalition system that coincides with that of g up to $\xi - 1$, and includes $\mathfrak{C}_1^g(\xi) \cup \{C\}$ at ξ. Let $m = \# C$ and define:

$$h := M_{4m+3} * (g, ..., g, (\xi, b_2)_n^*, (a_1, a_2)_n^*, I_n^C, I_n^C),$$

where I_n^C denotes the n-dimensional vector of dictatorial rules whose components are I_n^i with $i \in C$. The following observation is used repeatedly below: if $\mathfrak{C}_j^M$ denotes the coalition system induced by M_{4m+3} on B_j, $j = 1, 2$, then $D \in \mathfrak{C}_j^M(\xi)$ iff $\# D \geqslant 2m + 2$.

CLAIM 1. $h_2 = g_2$.

Pick $\tilde{\alpha} \in B_2^n$ and let $\eta = g_2(\tilde{\alpha})$ and $\tilde{\beta} = (\eta, ..., \eta, \xi, a_2, \alpha^C, \alpha^C)$, where α^C is the profile with components α^i, $i \in C$ (so $h_2(\tilde{\alpha}) = M_{4m+3}(\hat{\beta})$). Clearly $C(\hat{\beta}; \eta) = 2m + 2$. Therefore $C(\tilde{\beta}; \eta) \in \mathfrak{C}_2^M(\eta)$ and $h_2(\tilde{\alpha}) \leqslant \eta$. If $\eta = a_2$, then $h_2(\tilde{\alpha}) = \eta = g_2(\tilde{\alpha})$. Otherwise, note that $C(\hat{\beta}; \eta - 1) \leqslant 2m + 1$ because the first $2m + 2$ entries of $\tilde{\beta}$ are greater than η; hence $h_2(\tilde{\alpha}) > \eta - 1$, and, together with the earlier inequality, this implies $h_2(\tilde{\alpha}) = \eta = g_2(\tilde{\alpha})$.

CLAIM 2. *$\mathfrak{C}_1^g(\eta) \subset \mathfrak{C}_1^h(\eta)$ for all $\eta \in B_1$.*

Let $\eta \in B_1$ and $D \in \mathfrak{C}_1^g(\eta)$. Pick any $\tilde{\alpha} \in B_1^n$ such that $C(\tilde{\alpha}; \eta) = D$. Then $g_1(\tilde{\alpha}) \leqslant \eta$ and $(a_1)_n^* (\tilde{\alpha}) = a_1 \leqslant \eta$. Hence $h_1(\tilde{\alpha}) \leqslant \eta$ and $D \in \mathfrak{C}_1^h(\eta)$.

CLAIM 3. *$\mathfrak{C}_1^g(\eta) \supset \mathfrak{C}_1^h(\eta)$ for all $\eta \in B_1$ with $\eta < \xi$.*

Let $\eta < \xi$ and $D \notin \mathfrak{C}_1^g(\eta)$. Pick any $\tilde{\alpha} \in B_1^n$ such that $C(\tilde{\alpha}; \eta) = D$. Then $g_1(\tilde{\alpha}) > \eta$ and $\xi_n^*(\tilde{\alpha}) = \xi > \eta$. Hence $h_1(\tilde{\alpha}) > \eta$ and $D \notin \mathfrak{C}_1^h(\eta)$.

CLAIM 4. $C \in \mathfrak{C}_1^h(\xi)$.

Pick any $\tilde{\alpha} \in B_1^n$ with $C(\tilde{\alpha}; \xi) = C$. Then $I_n^i(\tilde{\alpha}) = \alpha^i \leqslant \xi$ for all $i \in C$, $(a_1)_n^* (\tilde{\alpha}) = a_1 \leqslant \xi$, and $\xi_n^*(\tilde{\alpha}) = \xi$. Hence $h_1(\tilde{\alpha}) \leqslant \xi$ and $C \in \mathfrak{C}_1^h(\xi)$.

CLAIM 5. *$D \notin \mathfrak{C}_1^g(\xi)$ and $C \not\subset D$ imply $D \in \mathfrak{C}_1^h(\xi)$.*

Let $D \notin \mathfrak{C}_1^g(\xi)$ be such that $C \not\subset D$. Pick any $\tilde{\alpha} \in B_1^n$ with $C(\tilde{\alpha}; \xi) = D$. Then $g_1(\tilde{\alpha}) > \xi$ and there exists $i \in C \setminus D$. Since for each $i \notin D$, $I_n^i(\tilde{\alpha}) = \alpha^i > \xi$, $h_1(\tilde{\alpha}) > \xi$ and $D \notin \mathfrak{C}_1^h(\xi)$.

Claims 1–5 imply that $\Delta(\mathfrak{C}_2^f(\eta), \mathfrak{C}_2^h(\eta)) = \Delta(\mathfrak{C}_2^f(\eta), \mathfrak{C}_2^g(\eta))$ for all $\eta \in B_2$, $\Delta(\mathfrak{C}_1^f(\eta), \mathfrak{C}_1^h(\eta)) = \Delta(\mathfrak{C}_1^f(\eta), \mathfrak{C}_1^g(\eta))$ for all $\eta \in B_1$ with $\eta < \xi$, and $\Delta(\mathfrak{C}_1^f(\eta), \mathfrak{C}_1^h(\eta)) \leqslant \Delta(\mathfrak{C}_1^f(\eta), \mathfrak{C}_1^g(\eta)) - 1$ for all $\eta \in B_1$ with $\eta \geqslant \xi$. Therefore $\delta(f, h) \leqslant \delta(g, h) - 1$.

If $\mathfrak{C}_1^f(\eta) \subset \mathfrak{C}_1^g(\eta)$ for all $\eta \in B_1$, let ξ be the largest point for which $\mathfrak{C}_1^g(\xi) \setminus \mathfrak{C}_1^f(\xi) \neq \emptyset$; let $C \in \mathfrak{C}_1^g(\xi) \setminus \mathfrak{C}_1^f(\xi)$. In this case we construct the scf h so as to satisfy the following properties:

(i) $\mathfrak{C}_2^h = \mathfrak{C}_2^g$;

(ii) $\mathfrak{C}_1^h(\eta) = \mathfrak{C}_1^g(\eta)$ for all $\eta > \xi$;

(iii) $\mathfrak{C}_1^h(\xi) = \mathfrak{C}_1^g(\xi) \setminus \{D \in 2^N \mid C \subset D\}$;

(iv) $\eta < \xi$, $D \notin \mathfrak{C}_1^g(\eta)$ and $C \subset D$ imply $D \notin \mathfrak{C}_1^h(\eta)$.

Let $D = N \setminus C$ and $m = \# D$. Define:

$$h := M_{4m+3} * (g, ..., g, (\xi, a_2)_n^*, (b_1, b_2)_n^*, I_n^D, I_n^D),$$

and note that if $\bar{\mathfrak{C}}_1^M$ denotes the right-coalition system induced by M_{4m+3} on B_1, then $E \in \bar{\mathfrak{C}}_1^M(\eta)$ iff $\# E \leqslant 2m+1$. A similar argument to that above demonstrates that $\delta(f, h) \leqslant \delta(g, h) - 1$ for this case as well.

References

1. K. J. Arrow, "Social Choice and Individual Values," Cowles Foundation Monographs, Yale Univ. Press, New Haven, 1963.
2. S. Barberà, H. Sonnenschein, and L. Zhou, Voting by committees, *Econometrica* **59**, (1991), 595–609.
3. K. Border and J. Jordan, Straightforward elections, unanimity and phantom vectors, *Rev. Econ. Stud.* **50** (1983), 153–170.

3a. D. Black, "The Theory of Committees and Elections," Cambridge Univ. Press, 1958.

4. R. Farquharson, "Theory of Voting," Yale Univ. Press, New Haven, 1983.
5. A. Gibbard, Manipulation of voting schemes: A general result, *Econometrica* **41** (1973), 587–601.
6. G. H. Kramer, Sophisticated voting over multidimensional choice spaces, *J. Math. Sociol.* **2** (1972), 165–180.
7. H. Moulin, On strategy-proofness and single-peakedness, *Public Choice* **35** (1980), 437–455.
8. M. A. Satterthwaite, Strategy-proofness and Arrow's condition: Existence and correspondence theorems for voting procedures and social welfare functions, *J. Econ. Theory* **10** (1975), 187–217.
9. K. A. Shepsle, Institutional arrangements and equilibrium in multidimensional voting models, *Amer. J. Polit. Sci.* **23** (1979), 27–60.