VOTING BY COMMITTEES

By Salvador Barberà, Hugo Sonnenschein, and Lin Zhou[1]

Problems of social choice frequently take the following form. There are n voters and a set $K = \{1, 2, \ldots, k\}$ of objects. The voters must choose a subset of K. We define a class of voting schemes called voting by committees. The main result of the paper is a characterization of voting by committees, which is the class of all voting schemes that satisfy voter sovereignty and nonmanipulability on the domain of separable preferences. This result is analogous to the literature on the Groves and Clarke scheme in that it characterizes all of the nonmanipulable voting schemes on an important domain.

Keywords: Social choice, voting scheme, committee, quota, strategy-proofness, voter sovereignty, separable preferences.

1. INTRODUCTION

Problems of social choice frequently take the following form. There are n voters and a set $K = \{1, 2, \ldots, k\}$ of objects. These objects may be bills considered for adoption by a legislature, candidates considered for membership in a club, etc. The voters must choose (adopt) a subset of the set of objects. Assume that each voter has a linear order of the 2^k possible subsets that can be adopted. A voting scheme is a method for passing from n-tuples of these linear orders, called preference profiles, to sets of objects. A scheme is *dictatorial* if there is some voter whose preferred choice is always selected, no matter what the preferences of other voters. A scheme is *manipulable* if, for some profile of individual preferences, a voter can obtain an outcome set that he prefers by misrepresenting his true preferences. Applied to this framework, the Gibbard-Satterthwaite theorem says that if there is more than one issue to consider for adoption, then all nonmanipulable schemes are dictatorial.[2] (An additional hypothesis is required: the range of the scheme must contain at least three sets of objects.)

In this paper we provide a response to the Gibbard-Satterthwaite theorem that is of the same form as the one provided by Groves and Clarke in their classic papers on the free-rider problem. For social choice problems in which the interpretation of an outcome is a public decision and a vector of money transfers, and preferences are of the transferable utility form, the Groves-Clarke mechanisms are nonmanipulable. Furthermore, they are the only nonmanipulable mechanisms on the domain of transferable utility preferences. Here we prove parallel results for the problem of selecting subsets from a set K of objects. We define a class of mechanisms called voting by committees and observe that on an important domain of preferences they are nonmanipulable.

We then establish that voting by committees is the only form of nonmanipulable scheme on this domain.

A familiar special case of voting by committees is voting by quota. In the quota system with quota $Q((1 \leqslant Q \leqslant n)$, the ith voter $(1 \leqslant i \leqslant n)$ votes a subset of the objects $A_i \subseteq K$, and the outcome is the set of objects that belong to at least Q of the sets A_i. This type of voting is used explicitly by some clubs when electing new members: a candidate for membership is elected if his name is checked on some predetermined number (say Q) of ballots. A reduced form of voting by quota operates in many legislative bodies. Elected representatives normally vote on a large number of issues each year. Although the issues are in practice confronted one by one, with, for example, support by two-thirds of the legislators needed for adoption, it is as if all legislators announce at once the issues they support (this might follow some logrolling) and as if issues announced by two-thirds of the legislators are then adopted.

Voting by quota is a particular way of describing which coalitions of voters are winning, in the sense that they can force an object into the final outcome set. A more general class of voting scheme can be described by specifying a family of winning coalitions for each object and then including an object into the outcome set if and only if the set of voters who favor it constitutes a winning coalition. (We assume that if a coalition is winning, then so are its supersets.) This device of specifying voters and winning coalitions is standard in game theory and is known as a (monotonic) simple game, or a committee. Thus, the voting scheme just described, which uses a committee to choose each object, is called voting by committees. The key result of this paper is a characterization of voting by committees.[3]

We now specify the domain of preferences that will be considered. For a preference relation defined on the subsets of K, we will call an object "good" if it is preferred to the empty set. A preference relation is *separable* if for any set $A \subseteq K$, and any object $x \notin A$, $\{x\} \cup A$ is preferred to A if and only if x is good. As a special case, preferences are separable when they have an additive representation.[4] Since separability requires a degree of independence among objects, interaction effects lead to its failure. A conservative legislator may favor the passage of either of two laws by itself ($\{x\} > \varnothing$, and $\{y\} > \varnothing$); the adoption of two together might, in his opinion, represent an intolerable change from the

---

[3] There is some similarity between voting by committees and approval voting, a method that has been studied in detail by Brams and Fishburn (1982). They are similar in that in both methods voters submit a set of objects. However, they are different in virtually every other respect, since the social choice in approval voting is a single object and since approval voting is in general manipulable (after one introduces a tie-breaker rule so that it is well-defined).

[4] A preference relation $>$ on 2^K is additively representable if there exists a function U: $K \cup \varnothing \to R$, with $U(\varnothing) = 0$ such that for all $A, B \subseteq K$,

$$A > B \quad \text{if and only if} \quad \sum_{z \in A} U(z) > \sum_{z \in B} U(z).$$

An important case in which this condition will be satisfied is when (a) each object can be gauged according to a finite number (say n) of characteristics, (b) preferences over bundles of characteristics are represented by a differentiable function U, and (c) objects are infinitesimal changes from a *status quo* $\hat{x}$.

status quo ($\varnothing > \{x, y\}$). Therefore, his preference relation is not separable. If you believe that the appointment of Professor Glutz or Professor Smith would be preferred to nobody being appointed, but you believe that appointing both of them together would be a disaster since it would tear your institute apart, then your preferences are not separable.

It is easy to see that voting by committees is not manipulable on the domain of separable preferences.[5] To see this, fix a profile of preferences and let S denote the set of objects that will be included in the outcome set independently of how the ith voter votes. If voter i has a separable preference relation, then his most preferred subset is his set of good objects, and he would like to add as many of them to S as possible. By truthfully reporting his preferred subset under voting by committees, i adds to S all the objects that he considers good among those he can add. With separable preference this is the best he can do. In addition to nonmanipulability over separable preferences, voting by committees satisfies the classical property of voter sovereignty: no subset of objects is a priori barred from emerging as the outcome.

We can now be more specific about our characterization of voting by committees. It is the only method for passing from n-tuples of individual preferences to sets of objects that satisfies voter sovereignty and nonmanipulability on the domain of separable preferences.

Voting by quota is then easily characterized as the subclass of voting by committee methods that satisfies the additional classical requirements of anonymity and neutrality. Anonymity guarantees the symmetric treatment of individuals, and it formalizes the requirement of "one man one vote." Neutrality means that each object is given a free hand to emerge and that one object is not favored over another simply because of its "name." Under voting by quota, winning coalitions are determined by size only, not by the names of the voters, and the same committee is used for all objects.

The paper proceeds as follows. In Section 2, we introduce notation, definitions, and the main theorems. Then in Section 3 we prove that voting by committees is manipulable on any domain larger than that of separable preferences. Finally, in Section 4 we study the efficiency property of voting by committees.

2. A CHARACTERIZATION OF VOTING BY COMMITTEES

2.1. *Notation and the Statements of the Main Theorems*

The set of voters is $N = \{1, 2, \ldots, n\}$. The set of objects is $K = \{1, 2, \ldots, k\}$. We assume that n and k are at least 2.[6] Subsets of K are denoted by

---

[5] Notice that with separable preferences all triples of sets of objects are free. Thus, there does not exist a social welfare function on this domain that satisfies Arrow's axioms. This should be a warning against misinterpretations of results by Satterthwaite (1975), Kalai and Muller (1979), and others, regarding the close connection between the possibilities of defining Arrowian social welfare functions and strategy-proof decision schemes on a domain.

[6] Our main results can also cover the cases $n = 1$ or $k = 1$. We exclude these trivial cases for ease of exposition since including them would require minor changes in some of our statements.

$A, A', B, B', S, T, \ldots$, with the empty set denoted by $\varnothing$. The ith voter's prefer-
ence relations, denoted by $>_i$, $>_i'$, etc., is an (asymmetric) ordering on 2^K (the
set of subsets of K).[7] Let P denote the set of all orderings on 2^K. Given $>$, a
preference relation in P, and Θ, a class of subsets of K, $\mathrm{argmax}(>, \Theta)$ denotes
the best element of $>$ on Θ; that is, $\mathrm{argmax}(>, \Theta) > B$ for all $A \in \Theta$,
$A \neq \mathrm{argmax}(>, \Theta)$.

We consider voting schemes that have each voter order the 2^k subsets of
objects and for each n-tuple of such orderings produce an outcome of a set of
objects. Formally, we have the following definition.

DEFINITION 1: A *voting scheme f* is a function from P^n to 2^K.

It is important for both theoretical and practical purposes to search for voting
schemes that are simple to execute and also possess certain desirable features.
We believe that the class of schemes of voting by committees is very attractive
on both accounts.

DEFINITION 2: A *committee* (or a *monotonic simple game*) is a pair $C = (N, W)$, where $N = \{1, \ldots, n\}$ is the set of voters, W is a nonempty set of
nonempty coalitions of N, which satisfies $[M \in W$ and $M' \supseteq M] \rightarrow M' \in W$.
Coalitions in W are called *winning*. $M \in W$ is a *minimal winning coalition* if and
only if $M' \subsetneq M \rightarrow M' \notin W$ ($M' \subsetneq M$ means $M' \subseteq M$ and $M' \neq M$).

Let $B(>)$ denote the best element of 2^K according to $>$: $B(>) = \mathrm{argmax}(>, 2^K)$.

DEFINITION 3: A voting scheme $f: P^n \rightarrow 2^K$ is *voting by committees*, if for
each object x, there exists a committee $C_x = (N, W_x)$ such that: for all profiles
$(>_1, >_2, \ldots, >_n)$, $x \in f(>_1, >_2, \ldots, >_n)$ if and only if $\{i \mid x \in B(>_i)\} \in W_x$.

One nice feature of voting by committees is that it depends only on the subset
of K ranked highest by each voter; that is, it is of the form $f: (2^K)^n \rightarrow 2^K$. It is
certainly more economical to execute a scheme of this form than a scheme that
requires information about every detail of voter's preferences. Other attractive
features of voting by committees include its resistance to manipulation and its
respect for voters' preferences.

A voting scheme is *strategy-proof* if it is always in the best interest of voters
to reveal their actual preferences. The formal definition of strategy-proofness
qualifies the above statement by introducing a relevant domain $R = R_1 \times R_2 \times \ldots \times R_n$.

DEFINITION 4: A voting scheme $f: P^n \rightarrow 2^K$ is *manipulable* on R at $(>_1, >_2, \ldots, >_n) \in R$ by $i \in N$ via $>' \in R_i$ if $f(>_1, >_2, \ldots, >', \ldots, >_n) >_i f(>_1,$

---

[7] Requiring strict preference is not too restrictive when K is finite.

$>_2, \ldots, >_n$). As a shorthand, we sometimes simply say that f is manipulable on R. A voting scheme f is *strategy-proof* on R if it is not manipulable on R.

DEFINITION 5: A voting scheme $f: P^n \to 2^K$ satisfies *voter sovereignty* on R if for each $A \subseteq K$ there exists $(>_1, >_2, \ldots, >_n) \in R$ such that $f(>_1, >_2, \ldots, >_n) = A$.

Of course, by the Gibbard-Satterthwaite theorem, voting by committees is manipulable on P^n. We are interested in the restricted domains on which voting by committees becomes strategy-proof. One such restriction is introduced in the following manner. A voter can evaluate each object separately from the other objects. For each $> \in P$, let $G(>) = \{x \in K | \{x\} > \varnothing\}$ be called the set of good objects, and its complement $G^c(>)$ the set of bad objects.

DEFINITION 6: A preference relation $> \in P$ is *separable* if for all $A \subseteq K$ and all $x \notin A$, $A \cup \{x\} > A$ if and only if $x \in G(>)$. The set of all separable preferences is denoted by P^S.

Notice that additive representability implies separability (see footnote 3 for the definition of additive representability). The converse is not true with more than two objects. For example, the preference relation $> : \{x, y, z\} > \{x, y\} > \{x, z\} > \{y, z\} > \{z\} > \{y\} > \{x\} > \varnothing$ is separable but not additively representable. Also notice that whenever $>$ is separable, $B(>) = G(>)$: the best set is the set of all good objects.

It is easy to check that voting by committees is strategy-proof on $(P^S)^n$ and satisfies voter sovereignty. More importantly, our main theorem states that voting by committees is the *only* method to have these two properties.

THEOREM 1: *A voting scheme* $f: (P^S)^n \to 2^K$ *is strategy-proof on* $(P^S)^n$ *and satisfies voter sovereignty if and only if it is voting by committees.*

Theorem 1 is tight. In particular, the requirement of voter sovereignty cannot be dispensed with. For example, consider the voting scheme f under which voter 1 always gets the single object he most prefers. Clearly, it does not satisfy voter sovereignty. Although f is strategy-proof, it is not voting by committees.

Next we specialize our result to anonymous and neutral voting schemes.

DEFINITION 7: A voting scheme f is *anonymous* if $f(>_1, >_2, \ldots, >_n) = f(>_{\sigma(1)}, >_{\sigma(2)}, \ldots, >_{\sigma(n)})$ for all $(>_1, >_2, \ldots, >_n) \in P^n$ and all permutations σ of N.

DEFINITION 8: A voting scheme f is *neutral* if $f(\mu(>_1), \mu(>_2), \ldots, \mu(>_n)) = \mu(f(>_1, >_2, \ldots, >_n))$ for all $(>_1, >_2, \ldots, >_n) \in P^n$ and any permutation μ of K. $\mu(>)$ is the ordering obtained from $>$ by permuting the objects according to μ.

DEFINITION 9: A voting scheme $f: P^n \to 2^K$ is *voting by quota* if there exists Q between 1 and n such that for all $(>_1, >_2, \ldots, >_n)$, we have $x \in f(>_1, >_2, \ldots, >_n)$ if and only if $\#\{i \,|\, x \in B(>_i)\} \geqslant Q$. ($\#L$ denotes the number of the elements in L.)

Voting by quota Q is a special case of voting by committees in which there is one single committee $C = (N, W_Q)$ for all objects and winning coalitions are determined by their sizes. The following corollary of Theorem 1 completely characterizes voting by quota.

COROLLARY 1: *A voting scheme $f: (P^S)^n \to 2^K$ satisfies anonymity, neutrality, voter sovereignty, and is strategy-proof on P^S if and only if it is voting by quota.*

We highlight the implication of our result for the anonymous and neutral case because the resulting methods of voting by quota are simple and intuitive. Of course, many important decision processes can be represented by voting by committees but not voting by quota. Voting mechanisms in the U.N. Security Council and the U.S. Congress illustrate this point.

The last remark we want to make is that Theorem 1 and its corollary still hold when the domain of separable preferences is replaced by that of additively representable preferences.[8]

THEOREM 2: *A voting scheme $f: (P^A)^n \to 2^K$ is strategy-proof on $(P^A)^n$ and satisfies voter sovereignty if and only if it is voting by committees.*

COROLLARY 2: *A voting scheme $f: (P^A)^n \to 2^K$ satisfies anonymity, neutrality, voter sovereignty, and is strategy-proof on $(P^A)^n$ if and only if it is voting by quota.*

2.2. *Proofs of the Theorems*

As mentioned above, voting by committees, although formally defined on preference n-tuples, does not require complete specifications of voters' preference relations. If f is voting by committees, then it depends only on the voters' most preferred sets. We call this property "tops only," and use the following notation:

$$f: (2^K)^n \to 2^K.$$

We first characterize all strategy-proof voting schemes in this restricted subclass. The distinct feature of this characterization is that it does not require any condition other than strategy-proofness. Notice that in general a strategy-proof voting scheme can be such that some object is never elected, or some other always elected. To allow for these possibilities, we extend slightly our

---

[8] Although P^A is a subset of P^S, Theorem 1 and Theorem 2 are logically independent.

previous notion of a committee, admitting an empty set of winning coalitions or the empty coalition of voters as a winning coalition.

DEFINITION 10: A *generalized committee* is a pair $C = (N, W)$, where $N = \{1, \ldots, n\}$ is the set of voters, W is a set of coalitions of N, which satisfies $[M \in W$ and $M' \supseteq M] \rightarrow M' \in W$.

We can now state the following lemma.

LEMMA 1: *A voting scheme* $f: (2^K)^n \rightarrow 2^K$ *is strategy-proof on* $(P^S)^n$ *if and only if it is voting by generalized committees, i.e., for each object x in K, there is a generalized committee* $C_x = (N, W_x)$ *such that* $x \in f(S_1, S_2, \ldots, S_n) \Leftrightarrow \{i | x \in B(>_i)\} \in W_x$.

PROOF: (i) First, if $x \in f(S_1, \ldots, S_i, \ldots, S_n)$ and $x \in S_i$, then $x \in T_i$ implies $x \in f(S_1, \ldots, T_i, \ldots, S_n)$. This is because we can always find a separable preference relation $>_i$ such that $G(>_i) = T_i$ and any set of objects including x is preferred to those sets excluding x. (For example, take an additively representable preference relation u_i such that $u_i(x) = 2k$, and $|u_i(y)| < 1$ for any $y \neq x$.) Therefore, f is strategy-proof implies $x \in f(S_1, \ldots, T_i, \ldots, S_n)$ when $x \in T_i$.

(ii) Second, if $x \in f(S_1, \ldots, S_i, \ldots, S_n)$ and $x \notin S_i$, then $x \in f(S_1, \ldots, T_i, \ldots, S_n)$ for every T_i. This is because we can find an additive preference relation $>_i$ such that $G(>_i) = S_i$ and any set of objects excluding x is preferred to $f(S_1, \ldots, S_i, \ldots, S_n)$ which includes x. (For example, let $u_i(x) = -2k$, and $|u_i(y)| < 1$ for any $y \neq x$.) Therefore, f is strategy-proof implies $x \in f(S_1, \ldots, T_i, \ldots, S_n)$ for every T_i.

(iii) For any object x, let $W_x = \{M \subseteq N|$ there exists an $S = (S_1, S_2, \ldots, S_n)$ such that $M = \{i | x \in S_i\}$, and $x \in f(S)\}$. We show that $C_x = (N, W_x)$ is a generalized committee. Suppose $M \in W_x$. There exists an $S = (S_1, S_2, \ldots, S_n)$ such that $M = \{i | x \in S_i\}$, and $x \in f(S)$. Then for any $M' \supseteq M$, construct S' as follows: $S_i' = K$, for $i \in M' \setminus M$; and $S_i' = S_i$, otherwise. When one replaces S_i by S_i' one by one to go from S to S', successive applications of (i) and (ii) lead to $x \in f(S')$. So $M' \in W_x$ since $M' = \{i | x \in S_i'\}$. This shows that $C_x = (N, W_x)$ is a generalized committee.

(iv) Finally, we have to show that for any x and $S = (S_1, S_2, \ldots, S_n)$, $x \in f(S)$ if and only if $M_x = \{i | x \in S_i\} \in W_x$. The "only if" part is just the definition of C_x. On the other hand, if $M_x \in W_x$, then there exists an $S' = (S_1', S_2', \ldots, S_n')$, $x \in S_i' \Leftrightarrow x \in S_i$, and $x \in f(S')$. Again, when one replaces S_i by S_i' one by one to go from S to S', successive applications of (i) and (ii) lead to $x \in f(S)$.   Q.E.D.

The limitation of Lemma 1 is that the voting schemes considered are assumed a priori to have the property of "tops only." The example after Theorem 1 illustrates that "tops only" is not a consequence of strategy-proofness. However, if voter sovereignty is assumed, then we can show that strategy-proofness

implies "tops only." In fact, even a weaker condition than voter sovereignty suffices. To motivate this condition, let us have a closer look at the structure of voting by generalized committees. An object for which the generalized committee has an empty collection of winning coalitions will never appear in any outcome, while an object for which the generalized committee contains even the empty coalition as a winning coalition will appear in all outcomes. Furthermore, if one examines the class of objects that appear in some outcomes but not all, then any subset of this class, union the set of objects that are always chosen, will occur as an outcome. This property is called weak voter sovereignty.

DEFINITION 11: For a voting scheme f, let $U_f = \bigcup_{A \in \text{range}(f)} A$, and $L_f = \bigcap_{A \in \text{range}(f)} A$. We say that f satisfies *weak voter sovereignty* if range(f) $= \{L_f \cup S \mid S \in 2^{U_f}\}$.

We now provide a characterization of voting schemes which are strategy-proof on $(P^S)^n$ and satisfy weak voter sovereignty without the assumption of "tops only."

LEMMA 2: *A voting scheme* $f: (P^S)^n \to 2^K$ *is strategy-proof on* $(P^S)^n$ *and satisfies weak voter sovereignty if and only if it is voting by generalized committees.*

With a voting scheme f that is voting by generalized committees the structures of the generalized committees associated with objects that never come out or always come out are as follows: $x \notin U_f \Leftrightarrow C_x = (N, W_x)$ is empty, and $x \in L_f \Leftrightarrow C_x = (N, W_x)$ contains the empty coalition as a winning coalition. Since voter sovereignty implies weak voter sovereignty, the hypotheses of Theorem 1 imply those of Lemma 2 and together with the fact that voter sovereignty also implies $U_f = K$, and $L_f = \varnothing$ the preceding sentence allows one to deduce Theorem 1 from Lemma 2.[9]

Before we can turn to the proof of Lemma 2, we introduce the concept of option sets as used by Barberà and Peleg (1990).

DEFINITION 12: For a voting scheme f, the *set of options*, given $>_i \in P^S$, is defined by

$$\sigma_{-i}^f(>_i) = \left\{ A \in 2^K \mid \exists >_{-i} \in (P^S)^{n-1} \text{ such that } f(>_i, >_{-i}) = A \right\}.$$

Sometimes it is simply written as $\sigma_{-i}(>_i)$ when it is clear which f is being discussed.

PROOF OF LEMMA 2: It is easy to check that voting by generalized committees is strategy-proof on $(P^S)^n$ and satisfies weak voter sovereignty. To prove the converse, we proceed by induction.

---

[9] Lemma 2 contains little of substance beyond Theorem 1. However, the proof we have adopted proceeds by induction on the number of voters and the weaker hypotheses of Lemma 2 are better adapted to this induction.

Step 1: We begin with the case in which there are two voters. The following facts are direct consequences of the assumption that f is strategy-proof and satisfies voter sovereignty.

Fact 1: For all $>_1, >_2 \in P^S$, $f(>_1, >_2) = \text{argmax}(>_1, \sigma_1(>_2)) = \text{argmax}(>_2, \sigma_2(>_1))$.

Fact 2: For all $> \in P^S$, $f(>, >) = \text{argmax}(>, U_f)$.

(i) We first show: If $x \in B(>_1) \cap U_f$, and $x \notin A \in \sigma_2(>_1)$, then $A \cup \{x\} \in \sigma_2(>_1)$. Assume this is not true. Find a preference relation $>_2$ in P^S such that $A \cup \{x\} >_2 A >_2 D$ for any other D. (For example, choose an additive one with $u_2(y) > 2$ for all $y \in A$, $u_2(x) = 1$, and $u_2(z) < -1$ otherwise.) Fact 1 thus implies $f(>_1, >_2) = A$. Since both $\{x\}$ and A are contained in U_f, so is $A \cup \{x\}$ because f satisfies weak voter sovereignty. Therefore, Fact 2 implies $f(>_2, >_2) = A \cup \{x\}$. But given $A \cup \{x\} >_1 A$, voter 1 then can manipulate at $(>_1, >_2)$ via $>_2$. This is a contradiction.

(ii) Second, if $x \notin B(>_1) \cup L_f$, and $x \in A \in \sigma_2(>_1)$, then $A \setminus \{x\} \in \sigma_2(>_1)$. Assume this is not true. Find a preference relation $>_2$ in P^S such that $A \setminus \{x\} >_2 A >_2 D$ for any other D. (For example, let $u_2(y) > 2$ for all $y \in A$, $u_2(x) = -1$, and $u_2(z) < -1$ otherwise.) Fact 1 thus implies $f(>_1, >_2) = A$. And Fact 2 implies $f(>_2, >_2) = A \setminus \{x\}$. Since $A \setminus \{x\} >_1 A$, voter 1 then can manipulate at $(>_1, >_2)$ via $>_2$. Again a contradiction.

(iii) Third, if $B(>_1) = B(>_1')$, then $\sigma_2(>_1) = \sigma_2(>_1')$. Suppose this is not true. There would exist a set $A \in \sigma_2(>_1)$ such that $A \notin \sigma_2(>_1')$. Decompose A into two disjoint sets:

$$A = E_1 \cup F_1, \quad \text{such that} \quad E_1 \subseteq B(>_1), \quad \text{and} \quad F_1 \subseteq B^C(>_1).$$

Let $>$ in P^S be such that $B(>) = A$. Fact 1 implies $f(>_1, >) = A$. Now consider $f(>_1', >) = D \neq A$, since $A \notin \sigma_2(>_1')$. Also decompose D into two disjoint sets:

$$D = E_2 \cup F_2, \quad \text{such that} \quad E_2 \subseteq B(>_1') = B(>_1), \quad \text{and}$$

$$F_2 \subseteq B^C(>_1') = B^C(>_1).$$

It must be true that $E_1 \subseteq E_2$. If not, let $x \in E_1$ but $x \notin E_2$. Since $x \in E_1 = A \cap B(>_1)$, $x \in B(>_1) = B(>_1')$, and $x \in A = B(>)$. Hence $\{x\} \cup D \in \sigma_2(>_1')$ by (i), and $\{x\} \cup D > D$. Then $f(>_1', >) = D$ contradicts Fact 1. A similar reasoning, using (ii), leads to $F_2 \subseteq F_1$. Therefore, $F \neq E$ implies that at least one of the two inclusions $E_1 \subseteq E_2$ and $F_2 \subseteq F_1$ is strict. But this means $f(>_1', >) = D = E_2 \cup F_2 >_1 E_1 \cup F_1 = A = f(>_1, >)$. Voter 1 then can manipulate at $(>_1, >)$ via $>_1'$.

(iv) Now we show that f depends on the top sets of voter's preferences only. Assume $B(>_1) = B(>_1')$ and $B(>_2) = B(>_2')$, for some $>_1$, $>_1'$, $>_2$, and $>_2' \in P^S$. Since $\sigma_2(>_1) = \sigma_2(>_1')$ by (iii), Fact 1 implies $f(>_1, >_2) = f(>_1', >_2)$. Notice that identical statements of (i), (ii), and (iii) still hold if the roles of the agents are reversed. So we also have $f(>_1', >_2) = f(>_1', >_2')$. Therefore, $f(>_1, >_2) = f(>_1', >_2')$.

(v) Given (iv), we can apply Lemma 1 to conclude that f is voting by generalized committees. And the structures of these generalized committees are obvious.

Step 2: We assume that the theorem is true for all cases with less than n voters. Now we consider a voting scheme f from $(P^S)^n$ to 2^K that satisfies the assumptions of the theorem.

(i) When we fix voter 1's preference relation at $>_1$, we get a voting scheme $f^{>_1}$ from $(P^S)^{n-1}$ to 2^K that is induced by f:

$$f^{>_1}(>_2, \ldots, >_n) = f(>_1, >_2, \ldots, >_n).$$

It is obvious that $f^{>_1}$ is strategy-proof. In addition, there are two sets $D(>_1)$ and $I(>_1)$ such that $\text{range}(f^{>_1}) = \{D(>_1) \cup S \mid S \in 2^{I(>_1)}\}$, and these two sets depend on $B(>_1)$ only. We will have a proof of this statement in (ii). For now let us assume it is true. Therefore, $f^{>_1}$ also satisfies weak voter sovereignty. Then by the induction hypothesis, $f^{>_1}$ is voting by generalized committees.

Now for any $>_1 \in P^S$, we have a voting scheme $f^{>_1}$ that is voting by generalized committees. Furthermore, $f^{>_1} \equiv f^{>_1'}$ whenever $B(>_1) = B(>_1')$. Suppose this is not true. Then, without loss of generality, there would exist an object $x \in I(>_1)$ and a coalition M such that M is winning in $C_x^{>_1} = (N \setminus \{1\}, W_x)$ but not winning in $C_x^{>_1'} = (N \setminus \{1\}, W_x')$. Consider a preference profile $(>_2, \ldots, >_n)$ in which $B(>_i) = \{x\}$ if $i \in M$, and $B(>_i) = \varnothing$ otherwise. So $f^{>_1}(>_2, \ldots, >_n) = \{x\} \cup D(>_1)$ and $f^{>_1'}(>_2, \ldots, >_n) = D(>_1') = D(>_1)$. But then f is manipulable by voter 1 (either at $(>_1, >_2, \ldots, >_n)$ or at $(>_1', >_2, \ldots, >_n)$), depending on whether x is good or bad for $>_1$, and $>_1'$). This is a contradiction.

To summarize, we have shown that $f^{>_1}$ is voting by generalized committees for any $>_1 \in P^S$, thus depends on the voters' most preferred sets only and that $f^{>_1} \equiv f^{>_1'}$ whenever $B(>_1) = B(>_1')$. Therefore, f must depend on the voters' most preferred sets only. Hence we can now apply Lemma 1 to conclude that f is voting by generalized committees.

(ii) Here we prove the statement made in (i): For any $>_1 \in P^S$, there are two sets $D(>_1)$ and $I(>_1)$ such that $\text{range}(f^{>_1}) = \{D(>_1) \cup S \mid S \in 2^{I(>_1)}\}$, and these two sets depend on $B(>_1)$ only. We will use an important property that any strategy-proof scheme F possesses:

Fact 3: For all $>$, $F(>, >, \ldots, >) = argmax\,(>, range(F))$.

We construct a voting scheme $\hat{f}$ from $(P^S)^2$ to 2^K as follows:

$$\hat{f}(>_1, >_2) = f(>_1, >_2, \ldots, >_2).$$

We first show that $\sigma_2^{\hat{f}}(>_1) = \sigma_{-1}^f(>_1)$ for any $>_1$. It is trivial that $\sigma_2^{\hat{f}}(>_1) \subseteq \sigma_{-1}^f(>_1)$. On the other hand, for any $A \in \sigma_{-1}^f(>_1)$, we find some $>_2 \in P^S$ such that $B(>_2) = A$. Since $f^{>_1}$ is strategy-proof, we can apply Fact 3 to $f^{>_1}$ to get:

$$\hat{f}(>_1, >_2) = f(>_1, >_2, \ldots, >_2) = f^{>_1}(>_2, \ldots, >_2)$$

$$= \text{argmax}\,(>_2, \sigma_{-1}^f(>_1)) = A.$$

This shows $\sigma_{-1}^f(>_1) \subseteq \sigma_2^{\hat{f}}(>_1)$. So, $\sigma_2^{\hat{f}}(>_1) = \sigma_{-1}^f(>_1)$.

We now claim that $\hat{f}$ is also strategy-proof and satisfies weak voter sovereignty. It is obvious that voter 1 in $\hat{f}$ cannot manipulate. And since $\sigma_2^{\hat{f}}(>_1) = \sigma_{-1}^f(>_1)$, Fact 3 leads to $\hat{f}(>_1, >_2) = f(>_1, >_2, \ldots, >_2) = \text{argmax}(>_2, \sigma_{-1}^f(>_1)) = \text{argmax}(>_2, \sigma_2^{\hat{f}}(>_1))$. So voter 2 in $\hat{f}$ cannot manipulate either. Hence $\hat{f}$ is strategy-proof. Because $\text{range}(\hat{f}) = \bigcup \sigma_2^{\hat{f}}(>_1)$ and $\text{range}(f) = \bigcup \sigma_{-1}^f(>_1)$, $\sigma_2^{\hat{f}}(>_1) = \sigma_{-1}^f(>_1)$ for all $>_1$'s leads to $\text{range}(\hat{f}) = \text{range}(f)$. Therefore, $\hat{f}$ satisfies weak voter sovereignty since f does.

By step 1, we know that $f^{>_1}$ is voting by generalized committees. Then it is obvious that for any $>_1 \in P^S$, there are two sets $D(>_1)$ and $I(>_1)$ such that $\sigma_2^{\hat{f}}(>_1) = \{D(>_1) \cup S \mid S \in 2^{I(>_1)}\}$, and these two sets depend on $B(>_1)$ only. So our statement is proved since $\sigma_2^{\hat{f}}(>_1) = \sigma_{-1}^f(>_1) = \text{range}(f^{>_1})$.   Q.E.D.

It is now easy to complete the proofs of the other results. Corollary 1 follows immediately from Theorem 1. If a voting scheme f is strategy-proof on separable preferences and satisfies voter sovereignty, then it is voting by committees by Theorem 1. The additional requirement of neutrality implies that the committee structures for all objects must be the same, and the requirement of anonymity implies that all minimal winning coalitions must have the same size Q. Therefore, f is voting by quota Q. To prove Theorem 2 and Corollary 2, one needs to observe that in our proofs of Lemma 1 and Lemma 2 all preference relations we have used are in fact additively representable. Hence, the exact same proof applies to Theorem 2 and Corollary 2.

3. SEPARABLE PREFERENCES AND THE STRATEGY-PROOFNESS OF VOTING BY COMMITTEES

We already know that voting by committees is the only method that satisfies strategy-proofness and voter sovereignty on the domain of separable preferences. In this section we will show that, in general, voting by committees cannot be strategy-proof on any larger domain. To make this idea precise, we restrict attention to admissible domains that admit a wide enough variety of preferences.

DEFINITION 13: A domain $R = R_1 \times R_2 \times \ldots \times R_n \subseteq P^n$ is *rich* if, for all $i \in N$ and all $A \subseteq K$, there exist $>_i \in R_i$, such that $A = B(>_i)$.

If voting by committees is strategy-proof on a rich domain R, then what does this tell about the structure of R? Not much can be deduced if the power structures of the committees are extreme. Consider the example of voting by a single committee C. When C is dictatorial, voting by committee C is strategy-proof on the universal domain, and there is no restriction at all on the admissible preferences. Dictatorship is an example of one of the cases we consider extreme: some voter can single-handedly prevent other voters from forming a winning coalition. The other extreme is when some voter never contributes to any winning coalition: for any winning coalition he belongs to, some voters within it—excluding him—can form a winning coalition on their own. We now present a condition that is used to eliminate such situations.

DEFINITION 14: Assume that we have a committee $C = (N, W)$ with minimal winning coalitions N_α's. A voter i is a *vetoer* if $i \in \bigcap_\alpha N_\alpha$. A voter i is a *dummy* if $i \notin \bigcup_\alpha N_\alpha$. A committee $C = (N, W)$ is nonextreme if it contains neither veto voters nor dummy voters.

THEOREM 3: *Assume that voting scheme f is voting by nonextreme committees. If f is strategy-proof on a rich domain* $R = R_1 \times R_2 \times \ldots \times R_n$, *then* $R_i \subseteq P^S$ *for all* i.[10]

PROOF: We start with an equivalent specification of P^S: A preference relation $>$ is separable on K if and only if,

(**)   for all x and all $A \subset K (x \notin A)$, $A \cup \{x\} > A$ if and only if $x \in B(>)$.

The necessity of (**) is obvious. On the other hand, suppose (**) holds for some $>$. If we let $A = \emptyset$ in (**), then (**) says $\{x\} > \emptyset$ if and only $x \in B(>)$. Hence $B(>) = G(>)$. But (**) is then simply the definition of separability when $B(>)$ is replaced by $G(>)$. So $>$ is separable on K.

Now we show that $R_i \subseteq P^S$ for all i. Given a preference $>_i \in R_i$, take any x, and A such that $x \notin A$.

(i) Suppose $x \in B(>_i)$. Since C_x is nonextreme, there is a minimal winning coalition N_1 in C_x such that $i \in N_1$. Again since C_x is nonextreme, $N \setminus \{i\}$ is winning in C_x. Because R is rich, we can choose $>_j$ for $j \in N \setminus \{i\}$ such that

$$B(>_j) = A \cup \{x\}, \qquad j \in N_1 \setminus \{i\},$$

$$B(>_j) = A, \qquad \text{otherwise.}$$

With preferences defined this way,

$$f[B(>_1), \ldots, B(>_i), \ldots, B(>_n)] = A \cup \{x\}.$$

But if voter i announces $E = B(>_i) \setminus \{x\}$,

$$f[B(>_1), \ldots, E, \ldots, B(>_n)] = A.$$

f is strategy-proof implies $A \cup \{x\} >_i A$.

(ii) Suppose $x \notin B(>_i)$. By the same method, we can show that f is strategy-proof implies $A >_i A \cup \{x\}$.

Hence by the specification of P^S provided above, $>_i \in P^S$.   Q.E.D.

4. EFFICIENCY AND STRATEGY-PROOFNESS ON RESTRICTED DOMAINS[11]

We now apply Theorem 1 to demonstrate that, in general, there exists no efficient, strategy-proof and nondictatorial voting scheme on $(P^S)^n$. The point here is that the constraints associated with incentive compatibility prevent one from achieving the Pareto frontier even on this restricted domain.

---

[10] A more general result accommodating extreme committees is available. But such a treatment will blur the main picture without providing more insights.

[11] This section was motivated by discussions with L. Corchón and A. Postlewaite, who pointed out to us that voting by quota is not Pareto optimal.

DEFINITION 15: A voting scheme f is *efficient* if for all $(>_1, >_2, \ldots, >_n)$ in its domain, the outcome is Pareto optimal, i.e., for all $A \subseteq K$, $A \neq f(>_1, >_2, \ldots, >_n)$, $f(>_1, >_2, \ldots, >_n) >_i A$ for at least one i.

DEFINITION 16: A voting scheme f is *dictatorial* if there exists a voter i such that $f(>_1, >_2, \ldots, >_n) = \text{argmax}(>_i, 2^K)$ for all $(>_1, >_2, \ldots, >_n)$ in its domain.

THEOREM 4: *For $k \geqslant 3$, there exists no efficient, strategy-proof, and nondictatorial voting scheme on $(P^S)^n$.*

Efficiency obviously implies voter sovereignty; therefore, by Theorem 1, any voting scheme that is efficient and strategy-proof must be voting by committees. Furthermore, we can show that voting by committees is generally not efficient unless it is dictatorial. As an introductory example, we consider a situation where $K = \{x, y\}$, $N = \{1, 2\}$, and the voting scheme is voting by quota $Q = 1$. Voter 1 prefers x to the empty set, but dislikes y so much that even *x and y together* are worse than the empty set; and voter 2 prefers y to the empty set, but dislikes x so much that *x and y together* are also worse than the empty set. Voting by quota $Q = 1$ produces the outcome $\{xy\}$, which is Pareto dominated by $\varnothing$. The second example is: $K = \{x, y\}$, $N = \{1, 2, 3\}$, and the voting scheme is voting by quota $Q = 2$. In this case the voting scheme turns out to be efficient. However, the second example represents a rare case. It occurs only when $k = 2$ and the committees in the voting schemes have some special structures.[12] The negative result prevails when $k \geqslant 3$.

PROOF OF THEOREM 4: We will provide a proof for the case in which K contains three distinct objects only: x, y, and z. The proof for the general case is similar. We show that the existence of such a voting scheme leads to a contradiction.

Suppose f is a voting scheme that is efficient, strategy-proof, and nondictatorial on $(P^S)^n$. By Theorem 1, we know that f must be voting by committees. Let C_x, C_y, and C_z be the corresponding committees for x, y, and z. For any nonempty partition of voters, $N = N_1 \cup N_2$ ($N_1 \cap N_2 = \varnothing$):

(i) If N_1 is winning for one object, say x, then N_2 cannot be winning for another object, say y. Otherwise, let $>_1$ and $>_2$ in P^S be the following:

$$\{x\} >_1 \varnothing >_1 \{xy\} >_1 \{y\} >_1 \ldots, \quad \text{and}$$

$$\{y\} >_2 \varnothing >_2 \{xy\} >_2 \{x\} >_2 \ldots.$$

(For example, take two additively representable preference relations u_1 and u_2 such that $u_1(x) = 1$, $u_1(y) = -2$, $u_1(z) = -3$, and $u_1(\varnothing) = 0$; and $u_2(x) = 2$, $u_2(y) = -1$, $u_2(z) = -3$, and $u_2(\varnothing) = 0$.) For the preference profile in which every voter in N_1 has $>_1$ and everyone in N_2 has $>_2$, f produces the outcome $\{xy\}$, which is Pareto dominated by $\varnothing$.

---

[12] An earlier version of our paper (Barberà, Sonnenschein, and Zhou (1988)) contained a more detailed discussion on this point.

(ii) If N_1 is not winning for one object, say x, then N_2 must be winning for any other object, say y. Otherwise, let $>_1$ and $>_2$ in P^S be the following:

$$\{x\} >_1 \{xy\} >_1 \varnothing >_1 \{y\} >_1 \dots, \quad \text{and}$$

$$\{y\} >_2 \{xy\} >_2 \varnothing >_2 \{x\} >_2 \dots.$$

For the preference profile in which every voter in N_1 has $>_1$ and everyone in N_2 has $>_2$, f produces the outcome $\varnothing$, which is Pareto dominated by $\{xy\}$.

Similarly, if $N = N_1 \cup N_2 \cup N_3$ (N_i's are exclusive), then:

(iii) It cannot be true that N_1 is not winning for x, N_2 not winning for y, and N_3 not winning for z.

We now define $r_x = \min\{\#M | M$ is winning in $C_x\}$. r_y and r_z are defined similarly. Finally, let $r = \min\{r_x, r_y, r_z\}$.

(1) Suppose $r = 1$. Without loss of generality, we assume that voter 1 is winning for x. According to (i), $N\setminus\{1\}$ is not winning for y or z. Then, according to (ii), voter 1 is winning for both z and y. Again, according to (i), $N\setminus\{1\}$ is not winning for x. But now voter 1 is a dictator. This is a contradiction.

(2) Suppose $2 \leqslant r < n$. Assume M is winning for x and $\#M = r$. According to (i), $N\setminus M$ cannot be winning for z. Since $\#M = r \geqslant 2$, we can divide M into two disjoint nonempty coalitions: $M = M_1 \cup M_2$. By the definition of r, M_1 is not winning for x and M_2 not winning for y. Together with the fact that $N\setminus M$ is not winning for z, we have a contradiction to (iii).

(3) Suppose $r = n$. Simply divide N into two disjoint nonempty coalitions: $N = N_1 \cup N_2$. By the definition of r, N_1 is not winning for x and N_2 not winning for y. This contradicts (ii).   Q.E.D.

Departament d'Economia i d'Historia Econòmica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain;

Dean's Office, 116 College Hall, University of Pennsylvania, Philadelphia, PA 19104-6377, U.S.A.;

Cowles Foundation for Research in Economics, Department of Economics, P.O. Box 2125, Yale Station, New Haven, CT 06520-2125, U.S.A.

REFERENCES

ARROW, K. J. (1963): *Social Choice and Individual Values*. New Haven: Yale University Press.
BARBERÀ, S., AND B. PELEG (1990): "Strategy-Proof Voting Schemes with Continuous Preferences," *Social Choice and Welfare*, 7, 31–38.
BARBERÀ, S., H. SONNENSCHEIN, AND L. ZHOU (1989): "Voting by Quota and Committee," *Southern European Economics Discussion Series*, D.P. 64.
BRAMS, S., AND P. FISHBURN (1982): *Approval Voting*. Boston: Birkhause.
GIBBARD, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587–602.
GREEN, J., AND J. -J. LAFFONT (1979): *Incentives in Public Decision-Making*. Amsterdam: North-Holland.
GROVES, T., AND M. LOEB (1975): "Incentives and Public Inputs," *Journal of Public Economics*, 43, 211–226.

HOLMSTRÖM, B. (1979): "Groves Scheme on Restricted Domains," *Econometrica*, 47, 1137–1144.
KALAI, E., AND E. MULLER (1979): "Characterization of Domains Admitting Non-dictatorial Social Welfare Functions and Nonmanipulable Voting Procedures," *Journal of Economic Theory*, 16, 457–469.
PELEG, B. (1984): *Game Theoretic Analysis of Voting in Committees*. Cambridge: Cambridge University Press, 1984.
SATTERTHWAITE, M. A. (1975): "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187–217.
SCHMEIDLER, D., AND H. SONNENSCHEIN (1978): "Two Proofs of the Gibbard-Satterthwaite Theorem on the Possibility of a Strategy-Proof Social Choice Function," in *Decision Theory and Social Ethics: Issues in Social Choice*, ed. by H. W. Gottinger and W. Leinfellner. Dordrecht: D. Reidel.
SHAPLEY, L. S. (1962): "Simple Games: An Outline of the Descriptive Theory," *Behavioral Science*, 7, 59–66.