9 Notes on Strategy-Proof Social Choice Functions

Salvador Barberà
UNIVERSITAT AUTÒNOMA DE BARCELONA

1 INTRODUCTION

The goodness of a collective decision often depends on the wishes of those affected by it. The candidate with a majority is assumed to be best under democracy. Reasonable proposals on how to finance and carry out a public project should pass the minimal test of efficiency. Satisfactory assignments of students to colleges should take into account the interests of those who are seeking education. All of the above are examples of situations in which knowing the preferences of agents is necessary to determine what is collectively best.

Establishing which collective choices respond best to individual values is a central problem in economics, politics and ethics. Even if agents' preferences were given as data, the principles we have referred to – majority will, efficiency, respect for individual demands – admit different formulations. An essential contribution of Arrow's general possibility theorem (Arrow, 1963) was to stress the pervasiveness of the difficulties in extending the majority principle when more than two alternatives are available. The shortcomings of the Pareto principle of efficiency as a guide to collective action are well known. And Sen's (1970) Paretian liberal paradox has inspired a wide debate on how to model and eventually solve the tensions between the acceptance of individual spheres of action on the one hand, and the unanimity principle on the other. Much of social choice theory has been devoted to proving the robustness and the universality of such difficulties, and to proposing specific rules to (partially) overcome some of them.

In order to take preferences into account, they must be known, directly or indirectly. Some methods of collective decision-making explicitly recognize this and require citizens to vote, as in political elections, 43

or to express their preferences, as when applying to colleges. In other cases, the reference to interested individuals is more indirect. It is not always clear whether and how public decision-makers integrate the preferences of their constituencies into their calculations. But their actions should not be independent of the will of those affected by them, at least from a normative point of view. Hence, we should study the extent to which the behaviour of agents within a given political/economic system reveals its preferences and influences collective choices.

Knowledge about individual preferences is an example of private information, asymmetrically distributed among the members of society. In an extreme version, we may claim that each individual agent knows his or her preferences perfectly, and that all other agents lack any information about them. Less extremely, it is easy to accept that each one knows his or her preferences better than anybody else, and certainly at a lower cost. There is another dimension to the private character of information. In many cases, agents have the right to express any preference or opinion among those considered to be admissible, and to do so even if they do not correspond to their true priorities, regardless of how much other agents know about those.

Revelation of preferences will in general be the result of strategic calculations. It will depend on what individuals think that others are going to do and on the specific rule to be applied when taking such preferences into account. When determining the level of provision of a public good, strategic decisions may give rise to the 'free rider' problem. Similarly, in political elections, strategy may lead to casting 'tactical' votes, rather than supporting the preferred candidate. Hence, there exists an interaction between the rules to be applied once preferences are known and the kind of information about preferences one might expect to obtain. In the absence of strategic behaviour, one could compare alternative methods of collective choice on the basis of the rules leading to choices, given preferences. But then the quality of the information about preferences that a given rule may process is determined by the rule itself.

I will be concerned with the incentives for individuals to reveal their preferences under decision-making mechanisms whose satisfactory performance requires knowledge of them. Specifically, I concentrate on situations and rules under which one should expect agents not to act strategically; or, rather, situations and rules under which it would be rational for individuals to reveal their true preferences, after all relevant strategic calculations.

I will study social choice functions. These are rules that select one alternative, and only one, for each profile of individual preferences. And I will concentrate on situations admitting strategy-proof social choice functions; that is, on environments for which there are rules under which any agent would consider that declaring his or her true preferences is as good as any alternative course of action. In this context, the environment refers to the specific features of the alternatives, and to the set of admissible preferences that agents may reveal. Alternatives may be candidates to a post, distributions of goods among different consumers, levels at which to produce a set of public goods, and so on. They can be modelled in different ways and sets of alternatives can be endowed with some structure. In turn, the interpretation of alternatives and their structure suggest what restrictions on preferences are sensible in each case, hence what is the relevant domain for social choice functions and what strategies are available to each agent.

The question of strategy-proofness for social choice functions may appear to be a rather narrow one within the wide literature on strategy and collective choice. Yet I will argue in what follows that it is not only relevant, but may give rise to a surprisingly rich set of answers, some already known, some certainly worth exploring in the future.

2 THE GIBBARD–SATTERTHWAITE THEOREM

Many social choice functions are manipulable and this has been noted in different contexts by many authors. Yet it was only in the early 1970s that Gibbard (1973) identified a simple framework within which a sharp and general statement about manipulability could be produced. The framework and the result – also proved independently by Satterthwaite (1975) – are as follows.

A is a set of alternatives. $I = \{1, 2, \ldots, n\}$ is a finite set of agents. The preferences of agents are taken to be complete, transitive binary relations on A. U will denote the universal set of preferences on A. R, R', R_i will stand for specific preferences. The associated strict preference relation is defined as usual, and denoted by P, P_i'.

Further properties of preferences may determine whether or not a certain relation is admissible for an agent in a given setting. Such properties, like single-peakedness, continuity, convexity and so on may in turn be based on the particular structure of the set of alternatives. They should be discussed for each particular case, when the general framework presented here is made more specific in order to fit in with particular interpretations. Let T_i stand for the set of preferences which are admissible for agent i in a generic social choice problem.

For R_i in T_i and $B \subseteq A$, the choice of R_i in B is the set $C(R_i, B) = \{x \in B \mid \nexists\, y \in B$ such that $yP_ix\}$. The choice of R_i in A is the ideal, or the top of i on A.

Let $T = T_1 \times T_2 \times \ldots \times T_n$. Elements of T are called admissible preference profiles. Given a preference profile $R = (R_1', \ldots, R_i', \ldots, R_n)$ and a preference R_i', we denote by R/R_i' the preference profile obtained from R by changing agent i's preferences from R_i to R_i' while keeping the preferences of all other agents unchanged.

A function $f: T \to A$ is a social choice function on T. The range of f is denoted by r_f.

A social choice function f on T is manipulable if there exist an admissible preference profile R, an agent i and an admissible preference R_i' for i such that

$$f(R/R_i')\ P_i f(R)$$

where P_i is the strict preference associated with R_i.

A social choice function is dictatorial if there exists an agent i such that, for any R_i, $f(R) = C(R_i, r_f)$.

Theorem Let f be a social choice function on a universal set of preferences. If the range of f contains more than two alternatives, then f is either dictatorial or manipulable.

This result is remarkable for many reasons. A first reason is that it contains a very substantial message, leading to a wide research programme. Given any set of collective decision rules based on the private information of different agents, it will be rational for these agents to condition the information they reveal on the expected actions of other members of society, hence on their knowledge about the interests, information and abilities of others. In short: agents should be expected to play a game, where preferences are strategies and alternatives are the outcomes. The essential message of the Gibbard–Satterthwaite theorem is that, except for trivial social choice functions (with two-element ranges or dictatorial), one cannot expect truth-telling behaviour always to be the best-response function for agents to the strategies of others. Because of that, the theorem is an invitation to a fully fledged analysis of social choice rules from a game-theoretic point of view. Since strategic behaviour is unavoidable under non-trivial rules, the incentives of agents must be taken as important constraints when designing social choice mechanisms. Economists were ready and eager to use game theory as a tool and to insist on incentives as a subject. The vast literature on implementation theory, and the literature I am about to revise as well, show the power of this invitation to pursue a new research agenda.[1]

The Gibbard–Satterthwaite theorem is also remarkable because it is so sharp. Its sharpness derives from the choice of framework. This accounts for much of its success, but also for some of the critical comment that it has been subject to. There are several features which mar the impact of the result. One is that it refers to social choice functions, rather than correspondences. This allows for a clear-cut definition of manipulability, but it forces us into a context where classical rules leading to set-valued choices do not always fit well. At present I do not view this modelling decision as too restrictive, and I will concentrate on social choice functions throughout the paper. Another, very substantial, assumption of the Gibbard–Satterthwaite theorem is that all conceivable preferences on alternatives are admissible. This very strong assumption is justified as a starting point, especially because it leads to a clear-cut impossibility result. But it is also the reason why a further question lies ahead: under what conditions do there exist non-trivial, non-manipulable social choice functions? Are there relevant situations where the structure of the alternatives suggest restrictions on preference domains under which interesting social choice rules may arise?

The sharpness of the Gibbard–Satterthwaite theorem was important for the immediate extension of one of its messages: letting game theory into mechanism design. But it also accounts for some incorrect perceptions. After twenty years of implementation theory,[2] many economists feel that 'almost any social choice rule can be implemented', while 'essentially nothing is strategy-proof'. Both statements are false unless we qualify them carefully. Some social choice correspondences can be implemented by an appropriately chosen mechanism, for some equilibrium concepts and under some further assumptions. Many possibilities have been studied and we know a lot about the trade-offs between the mechanisms, the equilibrium concepts and the characteristics of the environments under which different social choice rules might be implemented. But the same is true for strategy-proofness. Not only dictatorial rules are strategy-proof, in general. Some environments admit non-trivial, strategy-proof social choice functions, and others do not. In this paper I want to describe what we have learned in this direction, and to argue that there are still other interesting questions on strategy-proofness lying ahead.

Before enlarging on the subject, let me digress and mention some of the things we have learned from the proofs of the Gibbard–Satterthwaite theorem. We now have several approaches to the proof of the theorem, and each one of them teaches us something different.

1. Gibbard's original proof, later stylized by Schmeidler and Sonnenschein (1978), exploited the fact that from every strategy-proof social choice function one can define a social welfare function which will be Paretian and independent of irrelevant alternatives over the range of the former. Hence, if this range consists of more than two alternatives, the social welfare function will be dictatorial (by Arrow's theorem), and so will the social choice function from which it derives. This proof brings out a close connection between strategy-proof social choice and Arrovian social welfare functions under the assumption of universal domain. Similar connections can be established in other contexts (see, for example, Barberà, Gul and Stachetti, 1993). However, there are domains admitting strategy-proof non-dictatorial social choice functions for which all Paretian, independent of irrelevant alternatives, social welfare functions would be dictatorial. Conversely, there are domains admitting Arrovian non-dictatorial social welfare functions but for which all strategy-proof social choice functions are dictatorial. Hence, there is no general equivalence between strategy-proofness of social choice functions and Arrow's conditions on social welfare functions.
2. Satterthwaite, and also Schmeidler and Sonnenschein (1978), provided early proofs of the Gibbard–Satterthwaite theorem by induction on the number of agents and alternatives. The major drawback of such proofs is that they do not cover the infinite alternatives case. On the other hand, induction arguments on the number of agents have proven to be very useful in extending other approaches.
3. Another line of proof – see Muller and Satterthwaite (1977) and Moulin (1988) – derives the Gibbard–Satterthwaite result as a consequence of the fact that strategy-proofness requires strong monotonicity, a property that is also necessary for social choice correspondences to be Nash implementable.
4. A fourth approach, introduced in Barberà (1983), and later developed in Barberà and Peleg (1990), concentrates on the role of pivotal voters and the general structure of strategy-proof social choice functions. The set of options for an agent consists of the set of alternatives that might obtain, given the declared preferences of others, depending on what the agent under consideration declares. Some

facts are very general for strategy-proof social choice functions, regardless of domains. For any agent, a strategy-proof social choice function can be described as the result of this agent's best choice over the option set determined by the declared preferences of others. The option set function has additional features: it must contain the best elements for the agents setting their preferences, and depend only on these best elements (this is strictly true in the simplest case when agents are not indifferent between any two alternatives and functions are onto; otherwise, it needs qualification). Additional assumptions, like that of universal domain, generate the classical Gibbard–Satterthwaite result. But the basic structure that is revealed by this form of proof turns out to be very useful for the treatment of restricted domain analysis.

3 STRATEGY-PROOFNESS AND RESTRICTED DOMAINS

A major purpose of social choice theory is to study the trade-offs between different desiderata. Strategy-proofness is certainly a desirable property for social choice functions, especially when compatible with other attractive properties. I want to investigate what other properties it is compatible with, and when.

Let me mention some other requirements of interest. One is coverage: rules are interesting only to the extent that they operate on rich enough preference domains. Another is non-dictatorship; although within the scope of conceivable mechanisms, dictatorial procedures are formally trivial and normatively unattractive. A third requirement is efficiency. We do not need to elaborate on its importance, but it is worth mentioning two related conditions that are useful in understanding why many strategy-proof functions cannot be efficient. The ranges of social choice functions may or may not cover the whole set of alternatives under consideration. Some functions may become strategy-proof at the expense of never choosing some *a priori* feasible alternatives, regardless of individual preferences. The flexibility of a social choice rule, as measured by the extension of its range, will be an important property to check. Feasibility of the social choices is another concern. Some rules may be focused on partial decisions, and one should then worry about the compatibility between such decisions and the overall social resources. For example, the traditional problem of choosing an optimal level for a public decision, along with some transfers of a private good, does not *a priori* require that these transfers should balance.

Hence, feasibility may not be an *a priori* requirement, but yet another desirable condition to be partially or completely fulfilled.

By requiring social choice functions to be defined on a universal set of preferences, the Gibbard–Satterthwaite theorem drastically simplifies the picture. In fact, there is no necessary conflict between efficiency and strategy-proofness in this or any other case, since dictatorial rules are (trivially) efficient! But the remaining strategy-proof rules in this context are those with only two alternatives in the range, which will be typically inefficient when the set of feasible alternatives has more than two elements.

If we relax the coverage requirement, and concentrate on functions defined over restricted sets of preference profiles, then we may find non-dictatorial, strategy-proof social choice functions. In order to understand the trade-offs between strategy-proofness, domain extension, non-dictatorship, efficiency, range dimensionality and feasibility, one must examine different models of interest and obtain specific conclusions for each case.

A strategy of research consists in characterizing the set of all possible strategy-proof social choice functions for given domains. Alternatively, one can require some additional properties, like efficiency, along with strategy-proofness. But the latter strategy is seldom informative, because it leads very quickly to impossibility results. It is best first to understand the structure of all strategy-proof rules, and then to see whether any such rule may be efficient (in full or approximately), non-dictatorial, and so on. In particular, characterization results for strategy-proof rules allow us to understand why these are inefficient in most models (not all!). In some cases, this is clearly due to the fact that the range of social choice functions should be 'small' relative to the set of alternatives. In other cases, rules may be strategy-proof but not always recommend feasible outcomes.

I will not describe all the work that has already been done regarding the characterization of strategy-proof social choice functions on restricted domains. I will simply make some general remarks on the basis of a particular subset of results. But before doing so in the next section, I note some of the most significant pieces of work within this characterization literature. Table 9.1 brings out the rich variety of situations that have been covered, by classifying the various papers according to the structure of the set of alternatives and the families of admissible preferences they consider. But it is important to note that Table 9.1 is only indicative of the variety of models and results about strategy-proof social choice functions, and it is not exhaustive.[3]

Table 9.1 A gallery of models

Alternatives and leading interpretation	*Admissible preferences*	*Authors*
Common preferences (public)		
Unstructured	Any	Gibbard (1973) Satterthwaite (1975)
$\mathbb{R}$ Level of a public good	Single-peaked	Bossert and Weymark (1993) Kim and Roush (1980) Moulin (1980)
$\mathbb{R}^n$	Continuous	Barberà and Peleg (1990)
Levels of n public goods; characteristics of a candidate	Saturated. Convex upper contour sets	Border and Jordan (1983) Chichilnisky and Heal (1980) Laffond (1980)
	Saturated. Convex. Single-peaked according to L_1, distance	Peters, van der Stel and Storken (1991)
Subsets of integer boxes (vertices of hypercubes). Levels of n public goods characteristics of a candidate	Generalized single-peaked or additively separable	Barberà, Gul and Stachetti (1993) Barberà, Massó and Neme (1994) Barberà, Sonnenschein and Zhou (1991) Le Breton and Sen (1992) Serizawa (1992) Shimomura (1992)
Elements of the simplex lotteries	Von Neumann–Morgenstern	Barberà (1979) Gibbard (1977, 1978) Hylland (1980)
Sets of elements Basis for an unspecified lottery over alternatives	Rankings of sets (possibly incomplete)	Barberà (1977) Duggan and Schwartz (1992) Kelly (1977)
Personalized preferences (private)		
R^{n+1} Level for a public good and n individual transfers of a private good	Quasi-linear	Clarke (1971) Green and Laffont (1979) Groves and Loeb (1975) Roberts (1979) Serizawa (1993)
Simplex Distributions of one private good	Selfish single-peaked	Ching (1992) Sprumont (1991)

Alternatives and leading interpretation	*Admissible preferences*	*Authors*
$\mathbb{R}^{lm}$	Allocations	Bevià and Corchón (1995) Barberà and Jackson (1994) Hurwicz (1972, 1986) Hurwicz and Walker (1990) Jackson (1992) Moreno (1990) Satterthwaite and Sonnenschein (1981) Shenker (1992) Zhou (1991)
Matchings	Selfish	Alcalde and Barberà (1994) Roth (1982)

Finally, note that models are classified in two groups, labelled 'common preferences' and 'personalized preferences', respectively. Models in the first group consider that all agents share the same set of admissible preferences over alternatives, and preference profiles belong to the n-fold Cartesian product of this common set. In the second group of models, each agent has a set of admissible preferences which is personal to him and different from the admissible sets of others. Technically, the treatment of both types of models is different, because unanimous profiles can be used heavily for proof in the first case but not in the second. Conceptually, common preference cases are usually associated with choices involving public goods, while personalized preferences arise as the result of selfish preferences in the presence of private goods. But I prefer the classification used here, to emphasize that public good decisions in the presence of compensatory transfers, for example, belong to the second type.

4 WHAT HAVE WE LEARNED ABOUT STRATEGY-PROOF SOCIAL CHOICE FUNCTIONS?

I will try to summarize what is known about strategy-proof social choice functions in some general statements. These statements are too general and tentative to take the form of theorems, but they are based on facts and regularities which arise in different models. In order to combine my more speculative assertions with others based on solid grounds, I will develop explicitly the analysis of one specific model: the case where alternatives are K-tuples of integer numbers and preferences are generalized single-peaked. Particular facts that hold as theorems for this and other settings will be highlighted, and taken as the basis for general statements or conjectures.

I first introduce the model in its simplest form.

Case 1

(a.1) Alternatives are integer numbers in an interval $[a, b] = [a, a+1, \ldots, b]$.

(a.2) Preferences are single-peaked[4] on $[a, b]$.

To motivate our characterization result, let us first consider the median voter rule when we have an odd number of voters (say five). In that case, the outcome is an alternative x such that (a) the ideal of at least three agents is less than or equal to x, and (b) at most two agents have their ideals below x. The median voter rule assigns coalitions with three or more agents to each possible alternative, and chooses the minimal alternative for which the set of agents desiring this alternative or a smaller one has three agents or more. Generalized median voter schemes extend the same principle, but do not necessarily require that the same coalitions should be attached to different alternatives, nor that these coalitions be identified by their size alone. The notion of a left coalition system will allow for this extension. It can also be viewed as a generalization of monotonic simple games to cases where outcomes take more than two values.

A left-coalition system on $B = [a, b]$ is a correspondence $\mathfrak{L}$ that assigns to every $\alpha \in B$ a collection of non-empty coalitions $\mathfrak{L}(\alpha)$ satisfying

1. If $c \in \mathfrak{L}(\alpha)$ and $c \subset c'$, then $c' \in \mathfrak{L}(\alpha)$.
2. If $\beta > \alpha$ and $c \in \mathfrak{L}(\alpha)$, then $c \in \mathfrak{L}(\beta)$, and
3. $\mathfrak{L}(b) = 2^N - \phi$

For preferences P_i on $[a, b]$, let $\tau(P_i)$ be the maximal element of P_i on $[a, b]$.

The generalized median voter scheme on $[a, b]$ defined by a left-coalition system $\mathfrak{L}$, is given by the following rule:

$$F(P_1, P_2, \ldots, P_n) = \beta \quad \text{iff} \quad \{i \mid \tau(P_i) \leq \beta\} \in \mathfrak{L}(\beta)$$
$$\text{and} \quad \{i \mid \tau(P_i) \leq \beta - 1\} \notin \mathfrak{L}(\beta - 1)$$

Result A social choice function on an integer interval $[a, b]$ is strategy-proof (for single-peaked preferences) iff it is a generalized median voter scheme (Moulin, 1980; Barberà and Jackson, 1994; Barberà, Gul and Stachetti, 1993).

Before I derive conclusions, consider a second case.

Case 2

(a.2) Alternatives are K-tuples of integers. Any K-tuple in the integer box obtained as the Cartesian product of K integer intervals will be considered to be admissible.

(b.2) Preferences will be assumed to be generalized single-peaked.

To extend the notion of single-peakedness, endow B with the L_1- norm: for every B, $||\alpha|| = \sum_{k=1}^{k} |\alpha_k|$.

Given α, $\beta \in B$, the minimal box containing α and β is

$$MB(\alpha, \beta) = \{\gamma \in B \mid ||\alpha - \beta|| = ||\alpha - \gamma|| + ||\gamma - \beta||\}$$

A preference P on a box B is multidimensional single-peaked with bliss point α iff (1) $\tau(P) = \alpha$ and (2) $\beta P \gamma$ for all β, $\gamma \in B$ such that $\beta \in MB\ (\tau(P), \gamma)$.

Let $\tau(P)$ denote the best element in B according to P_i, and let $\tau_k(P)$ be the kth component of $\tau(P)$.

The generalized median voter scheme defined by a family of left-coalition systems $\mathbb{L} = \{\mathbb{L}_k\}_{k=1}^{k}$ is given by the following rule:

$$F(P_1, \ldots, P_n) = [F_1(P_1, \ldots, P_n) \ldots F_k(P_1, \ldots, P_n)]$$

where

$$F_k(P_1, \ldots, P_n) = \beta_k \quad \text{iff} \quad \{i \mid \tau_k(P_i) \leq \beta_k\} \in \mathbb{L}_k(\beta_k)$$
$$\text{and} \quad \{i \mid \tau_k(P_i) \leq \beta_k - 1\} \notin \mathbb{L}_k(\beta_k - 1)$$

Result 2 A social choice function onto an integer box is strategy-proof (for generalized single-peaked preferences) iff it is a generalized median voter scheme (Barberà, Gul and Stachetti, 1993; Barberà, Sonnenschein and Zhou, 1991).

I can now comment on some of the implications of the above results.

1. Note that generalized median voter schemes only use information about what alternative is best for each individual. This 'tops only' condition is necessary for strategy-proofness in our contexts, and in many others as well.
2. Another basic feature of generalized median voter schemes is that they are decomposable. The maximal element of each agent is projected into each dimension, projections for each dimension are aggregated through a generalized median voting scheme which is specific to that dimension, and the outcome is obtained as the coordinatewise union of these one-dimensional procedures. The work of Le Breton and Sen (1992) has made it clear that decomposability is a consequence of strategy-proofness on rich domains, and obtains quite independently from other properties. Essentially, rules must be decomposable under any domain that contains additively separable preferences.
3. In addition to sharing the two above properties, strategy-proof social choice functions in our domains are based on simple principles, which generalize the classical median voter rule.
4. Many of the rules we have characterized are efficient in the one-dimensional case. Most of them (except the dictatorial) are inefficient in the multidimensional case.

The above remarks apply to the framework we have considered specifically, but many of them can be extended to many other frameworks.

(a) In general, we find that strategy-proofness requires simple social choice rules; that is, rules which do not process 'too much' information, involving simple functional forms and/or choosing only from limited subsets of alternatives.
(b) Also, non-dictatorial social choice functions cannot in general encompass strategy-proofness with efficiency.

Yet, these remarks, which hold in a strict sense for different frameworks, including the one we start from, may require qualification in other cases.

(i) Note that our version of multidimensional single-peakedness is not the only possible one. In particular, it implies two essential features without which the preceding conclusions would change: (1) the projection of the top alternative into any dimension is the top alternative for the restriction of the preference to that dimension, and (2) any such restriction to one dimension is single-peaked in the classical sense. The class of non-dictatorial strategy-proof social choice functions would be dramatically reduced if preference domains were enlarged so that the preceding properties did not hold.

(ii) If agents were allowed to be indifferent among several 'top' alternatives, strategy-proof rules would still essentially be extensions of the median voter principle for our original framework, but their description would be more complex, and the 'tops only' condition would only hold in a qualified sense (see Barberà and Jackson, 1994). More generally, indifferences among alternatives tend to introduce complications in the characterizations of strategy-proof rules.

(iii) Models where the admissible preferences for all agents are drawn from the same set (the public good case) are easier to treat than those where the preferences of each agent come from different sets (the private good case with selfish preferences). One reason for this is that one can expect preference profiles with unanimity in the former, not in the latter, and strategy-proofness requires respect for unanimity. Moreover, selfish agents are systematically indifferent among many alternatives which only differ in aspects about which they are unconcerned, and I have already pointed out that indifferences allow for complicated rules to be strategy-proof; for example, bossy rules for the allocation of private goods, as described in Satterthwaite and Sonnenschein (1981).

In spite of the above qualifications, strategy-proofness is associated with some form of simplicity ('tops only' condition, narrow ranges) and stands in conflict with efficiency when not obtained through dictatorial rules, for most models I know.

Case 3

(a.3) Alternatives will again be K-tuples of integers. However, the set A of alternatives may not be a Cartesian product, but any subset of an integer box.

(b.3) Admissible preferences will be those which are generalized single-peaked on A and have their best alternative in A. Here we define a preference to be single-peaked on A if it is the restriction to A of some single-peaked preference on the minimal box containing A.

Result 3 A social choice function onto a subset A of an integer box is strategy-proof (for generalized single-peaked preferences with top on A) iff it is a generalized median voter scheme (Barberà, Massó and Neme, 1993).

Note that, even when the set of feasible alternatives is not Cartesian, the structural requirements on strategy-proof rules are maintained: they still must be 'tops only', decomposable and take the form of generalized median voter schemes.

Decomposability may now become a source of unfeasibility. Even if all agents vote for a feasible alternative, their votes on each dimension, when combined through arbitrarily chosen one-dimensional generalized median voter schemes, may lead to K-tuples of integers which are not feasible. If we concentrate on feasibility preserving rules, guaranteeing that feasible outcomes will be recommended whenever agents vote for feasible alternatives, then the choice of one-dimensional generalized median voter schemes must guarantee some degree of coordination among the decisions in each dimension. A general condition, which is necessary and sufficient for a generalized median voter scheme to preserve feasibility, given any set A, is described in Barberà, Massó and Neme (1993) under the name of the intersection property. When applied to specific sets A, this general condition yields valuable information about the distributions of decision power among agents that are compatible with strategy-proof rules. As we shall see, the Gibbard–Satterthwaite theorem can be obtained as a corollary of the necessity to respect the intersection property, for an appropriate reinterpretation of the present framework.

5 STRATEGY-PROOFNESS, SINGLE-PEAKEDNESS AND THE DIMENSIONALITY OF THE SET OF ALTERNATIVES

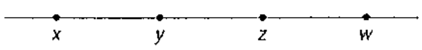
The preceding results were stated for situations where alternatives could be represented as K-tuples of integers and preferences were single-peaked. How restrictive are these two assumptions? I first want to show that, whenever the set of alternatives is finite, they are not restrictive at all. Then it will be seen that the essential assumption behind those results is one of rich domains: the assumption that not only are preferences single-peaked, but also that any single-peaked preference is admissible.

Let A be a finite set of alternatives. Given a set $\mathfrak{P}$ of preference

Admissible preferences | *The space of alternatives*

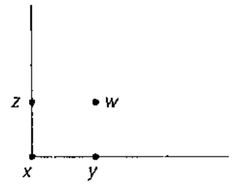
(1)

u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
x	y	y	y	z	z	z	w
y	x	z	z	w	y	y	z
z	z	x	w	y	w	x	y
w	w	w	x	x	x	w	x



(2)

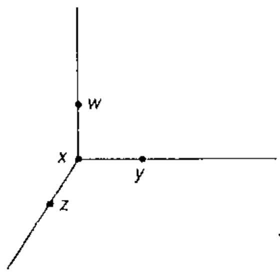
u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
x	x	y	y	z	z	w	w
y	z	x	w	x	w	y	z
z	y	w	x	w	x	z	y
w	w	z	z	y	y	x	x



(3)

u_1	u_2	u_3	u_4	u_5	u_6
x	x	x	x	x	x
y	y	z	z	w	w
z	w	y	w	y	z
w	z	w	y	z	y

u_7	u_8	u_9	u_{10}	u_{11}	u_{12}
y	y	z	z	w	w
x	x	x	x	x	x
z	w	y	w	y	z
w	z	w	y	z	y



(4) *Any preferences* | 4-dimensional unit vectors

Figure 9.1 Four examples with four alternatives (x, y, z, w)

relations on A, can we identify each alternative in A with some K-tuple of integers in such a way that preferences in $\mathfrak{P}$ are single-peaked? Consider the examples in Figure 9.1.

In all of these examples we have found some dimension K such that all the admissible orders of alternatives, when these are identified as points in this space, are indeed single-peaked. Remark that, in fact, given any K alternatives we can identify them with the K unit vectors in the K-dimensional real space. Since none of these unit vectors belongs to the minimal box defined by any other two, the condition of single-peakedness has no bite. Hence, it is always possible to embed any finite set of alternatives in a K-dimensional space, and claim that all the preferences in any set are single-peaked. What is not always possible is to derive our characterization results from such an identification, because our results are based on the assumption that the functions are defined on the set of all possible single-peaked preferences with tops on feasible alternatives. The examples above correspond to cases where choosing a small enough K allows our different sets of preferences to be rich in this sense.

It should be clear, then, that the models above are not completely general, but that they are more general than apparent at first sight. The general features of strategy-proof rules that I have discussed do cover a substantial ground. Of course, different economic and political contexts will lead us to the choice of different models. But I hope this paper serves to show that it is often worth examining how far one can go in designing strategy-proof mechanisms before jumping to other much more complex and hard-to-interpret forms of implementation.

Notes

1. The work of Hurwicz (for example, 1972, 1986) was equally influential in the same direction.
2. The paper by B. Dutta (1996) in this volume provides a lucid account of implementation theory.
3. For surveys of part of the literature, see Muller and Satterthwaite (1985) and Sprumont (1995).
4. In this section I assume strict preferences: no alternative is indifferent to any other. A preference P on $[a, b]$ is single-peaked if (1) P has a unique bliss point $\tau(P)$, and (2) whenever $\beta \neq \gamma$ and $\beta \in [\gamma, \tau(P)]$ then $\beta P \gamma$. Single-peaked preferences in one dimension were introduced by Black (1948), who had already discussed their role in connection with strategic voting. Different extensions to multidimensional spaces have been considered in the literature.

References

Alcalde, J. and Barberà, S. (1994) 'Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems', *Economic Theory*, vol. 4, pp. 417–35.

Arrow, K. J. (1963) *Social Choice and Individual Values*, 2nd edn (New York: Wiley).

Barberà, S. (1977) 'The Manipulation of Social Choice Mechanisms That Do Not Leave "Too Much" to Chance', *Econometrica*, vol. 45, pp. 1573–88.

Barberà, S. (1979) 'Majority and Positional Voting in a Probabilistic Framework', *Review of Economic Studies*, vol. 46, pp. 389–97.

Barberà, S. (1983) 'Strategy-Proofness and Pivotal Voters: A Direct Proof of the Gibbard–Satterthwaite Theorem', *International Economic Review*, vol. 24, pp. 413–7.

Barberà, S. and Jackson, M. (1994) 'A Characterization of Strategy-Proof Social Choice Functions for Economies with Pure Public Goods', *Social Choice and Welfare*, vol. 11, pp. 241–52.

Barberà, S. and Peleg, B. (1990) 'Strategy-Proof Voting Schemes with Continuous Preferences', *Social Choice and Welfare*, vol. 7, pp. 31–8.

Barberà, S., Gul, F. and Stachetti, E. (1993) 'Generalized Median Voters Schemes and Committees', *Journal of Economic Theory*, vol. 61, pp. 262–89.

Barberà, S., Massó, J. and Neme, A. (1993) 'Voting under Constraints', working paper no. 200–92, Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona.

Barberà, S., Sonnenschein, H. and Zhou, L. (1991) 'Voting by Committees', *Econometrica*, vol. 59, pp. 595–609.

Bevià, C. and Corchón, L. (1995) 'On the Generic Impossibility of Truthful Behavior', *Economic Theory*, vol. 6, pp. 365–91.

Black, D. (1948) 'On the Rationale of Group Decision Making', *Journal of Political Economy*, vol. 56, pp. 23–34.

Border, K. and Jordan, J. (1983) 'Straightforward Elections, Unanimity and Phantom Agents', *Review of Economic Studies*, vol. 50, pp. 153–70.

Bossert, W. and Weymark, J. (1993) 'Generalized Median Social Welfare Functions', *Social Choice and Welfare*, vol. 10, pp. 17–34.

Chichilnisky, G. and Heal, G. (1980) 'Incentive Compatibility and Local Simplicity', mimeo, University of Essex, UK.

Ching, S. (1992) 'A Simple Characterization of the Uniform Rule', *Economics Letters*, vol. 40, pp. 57–60.

Clarke, E. H. (1971) 'Multipart Pricing of Public Goods', *Public Choice*, vol. II, pp. 17–33.

Duggan, J. and Schwartz, T. (1992) 'Strategic Manipulability is Inescapable: Gibbard–Satterthwaite without Resoluteness', mimeo, Department of Political Science, University of California, Los Angeles, USA.

Dutta, B. (1996) 'Reasonable Mechanisms and Nash Implementation', this volume, pp. 3–23.

Gibbard, A. (1973) 'Manipulation of Voting Schemes: A General Result', *Econometrica*, vol. 41, pp. 587–601.

Gibbard, A. (1977) 'Manipulation of Schemes that Mix Voting with Chance', *Econometrica*, vol. 45, pp. 665–81.

Gibbard, A. (1978) 'Straightforwardness of Game Forms with Lotteries as Outcomes', *Econometrica*, vol. 46, pp. 595–614.

Green, J. and Laffont, J.-J. (1979) '*Incentives in Public Decision-Making* (Amsterdam: North-Holland).

Groves, T. and Loeb, M. (1975) 'Incentives and Public Inputs', *Journal of Public Economics*, vol. 4, pp. 211–26.

Hurwicz, L. (1972) 'On Informationally Decentralized Systems', in McGuire, C. B. and Radner, R. (eds), *Decision and Organization* (Amsterdam: North-Holland).

Hurwicz, L. (1986) 'On Informational Decentralization and Efficiency in Resource Allocation Mechanisms', in Reiter, S. (ed.), *Studies in Mathematical Economics*, Mathematics Association of America.

Hurwicz, L. and Walker, M. (1990) 'On the Generic Nonoptimality of Dominant-Strategy Allocation Mechanisms: A General Theorem that Includes Pure Exchange Economies', *Econometrica*, vol. 58, pp. 683–704.

Hylland, A. (1980) 'Strategy-Proofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies', mimeo, University of Oslo, Norway.

Jackson, M. (1992) 'Competitive Allocations and Incentive Compatibility', *Economics Letters*, vol. 40, pp. 299–302.

Kelly, J. (1977) 'Strategy-Proofness and Social Choice Functions without Singlevaluedness', *Econometrica*, vol. 45, pp. 439–46.

Kim, K. W. and Roush, F. (1980) 'Special Domains and Nonmanipulability', *Mathematical Social Sciences*, vol. 1, no. 1, pp. 85–92.

Laffond, G. (1980) 'Révélation des preferences et utilités unimodales', Centre National des Arts et Métiers, Paris, France.

Le Breton, M. and Sen, A. (1992) 'Strategyproof Social Choice Functions over Product Domains with Unconditional Preferences', mimeo, Groupe de Recherche en Economie Mathématique et Quantitative, Marseille, France.

Moreno, D. (1990) 'Nonmanipulable Decision Mechanisms for Economic Environments', discussion paper no. 91–6, College of Business and Public Administration, University of Arizona, USA.

Moulin, H. (1980) 'On Strategy-Proofness and Single Peakedness', *Public Choice*, vol. 35, pp. 437–56.

Moulin, H. (1988) *Axioms of Cooperative Decision Making* (Cambridge: Cambridge University Press).

Muller, E. and Satterthwaite, M. (1977) 'The Equivalence of Strong Positive Association and Strategy-Proofness', *Journal of Economic Theory*, vol. 14, pp. 412–8.

Muller, E. and Satterthwaite, M. (1985) 'Strategyproofness: The Existence of Dominant Strategy Mechanisms', in Hurwicz, L., Schmeidler, D. and Sonnenschein, H. (eds), *Social Goals and Social Organization*, (Cambridge: Cambridge University Press).

Peters, H., van der Stel, H. and Storken, T. (1991) 'On Uncompromisingness and Strategy-Proofness', Report M 91–15, University of Limburg, Holland.

Roberts, K. (1979) 'The Characterization of Implementable Choice Rules', in Laffont, J.-J. (ed.), *Aggregation and Revelation of Preferences, Studies in Public Economics* (Amsterdam: North-Holland).

Roth, A. E. (1982) 'The Economics of Matching: Stability and Incentives', *Mathematical Operations Research*, vol. 7, pp. 617–28.

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*, vol. 10, pp. 187–217.

Satterthwaite, M. A. and Sonnenschein, H. (1981) 'Strategy-Proof Allocation Mechanisms at Differentiable Points', *Review of Economic Studies*, vol. 48, pp. 587–97.

Schmeidler, D. and Sonnenschein, H. F. (1978) 'Two Proofs of the Gibbard–Satterthwaite Theorem on the Possibility of a Strategy-Proof Social Choice Function', in Essler, W. and Gottinger, H. (eds), *Proceedings of a Conference on Decision Theory and Social Ethics* (Dordrecht: Reidel).

Sen, A. K. (1970) 'The Impossibility of a Paretian Liberal', *Journal of Political Economy*, vol. 78, no. 1, pp. 152–7.
Serizawa, S. (1992) 'Cross-Shaped Preferences and Voting by Committees', mimeo, University of Rochester, New York, USA.
Serizawa, S. (1993) 'Strategy-Proof and Individually Rational Social Choice Functions for Public Good Economies', mimeo, University of Rochester, New York, USA, and Osaka University, Japan and forthcoming in *Economic Theory*.
Shenker, S. (1992) 'On the Strategy-Proof and Smooth Allocation of Private Goods in Production Economies', mimeo, Palo Alto Research Center, Xerox Corporation, California, USA.
Shimomura, K.-I. (1992) 'Partially Efficient Voting by Committees', mimeo, University of Rochester, New York, USA.
Sprumont, Y. (1991) 'The Division Problem with Single Peaked Preferences: A Characterization of the Uniform Allocation Rules', *Econometrica*, vol. 59, pp. 509–19.
Sprumont, Y. (1995) 'Strategyproof Collective Choice in Economic and Political Environments', *Canadian Journal of Economics*, vol. 28, pp. 68–107.
Zhou, L. (1991) 'Inefficiency of Strategy-Proof Allocation Mechanisms in Pure Exchange Economies', *Social Choice and Welfare*, vol. 8, pp. 247–54.