Review of Economic Studies (1999) **66**, 23–38
0034-6527/99/00020023$02.00

Nash Equilibrium and Welfare Optimality*

ERIC MASKIN
Harvard University

First version received June 1998; *final version accepted October* 1998 (*Eds.*)

If A is a set of social alternatives, a social choice rule (SCR) assigns a subset of A to each potential profile of individuals' preferences over A, where the subset is interpreted as the set of "welfare optima". A game form (or "mechanism") implements the social choice rule if, for any potential profile of preferences, (i) any welfare optimum can arise as a Nash equilibrium of the game form (implying, in particular, that a Nash equilibrium exists) and, (ii) *all* Nash equilibria are welfare optimal. The main result of this paper establishes that any SCR that satisfies two properties—monotonicity and no veto power—can be implemented by a game form if there are three or more individuals. The proof is constructive.

I. INTRODUCTION

After society has decided on a social choice rule—a recipe for choosing the optimal social alternative (or alternatives) on the basis of individuals' preferences over the set of all social alternatives—the social planner still faces the problem of how to implement that rule. In particular, the planner may not know individuals' preferences. He might attempt to elicit them, but this may not be an easy task, even abstracting from communication costs. If individuals know the rule by which the planner selects alternatives on the basis of reported preferences, they may have an incentive to report falsely.

One can think of the individuals as playing a game form. They are endowed with strategy spaces coinciding with their sets of possible announcements. The strategies that players choose determine an outcome. Ideally, one might hope to devise game forms which ensure that individuals will always want to announce their true preferences and that the right outcome (*i.e.* the one prescribed by the social choice rule) relative to those preferences is selected. In the case where preferences can be *anything*—that is, when the planner can place no *a priori* restrictions on the nature of individuals' preferences—Gibbard (1973) and Satterthwaite (1975) dash this hope by demonstrating that only dictatorial game forms have the property that players always wish to announce the truth regardless of the strategies of others. In other words, only a game form in which there exists a player who always gets his favourite alternative is strategy-proof.

In view of this negative result, one may be willing to sacrifice the strong incentive-compatibility of strategy-proofness. One may require, for example, only that players be in Nash equilibrium. This weaker stipulation has in fact been pursued by Groves and Ledyard (1977), Hurwicz (1979), and Schmeidler (1980), who construct game forms for the allocation of economic resources—with no restriction on preferences other than the usual convexity, continuity, and monotonicity assumptions—such that Nash equilibria exist and are Pareto optimal. Moreover the game forms constructed are nondictatorial. Indeed, in the Hurwicz and Schmeidler papers the Nash equilibria are not only Pareto optimal, but coincide with the set of Walrasian or Lindahl equilibria.

In this paper I examine the general question of implementation of social choice rules by game forms when Nash equilibrium is the solution concept. The main result asserts that a social choice rule on an arbitrary domain of preferences can be implemented by a game form if it satisfies two arguably reasonable properties: *monotonicity* and *no veto power*.

As I have presented it, implementation theory may appear to be purely a topic in applied welfare economies: Given the desired SCR, how can we go about implementing it? But there is a *positive* aspect to the theory as well. Certain well-known mechanisms—*e.g.* the English auction in the context of selling goods or rank-order voting in the context of electing candidates—are used frequently in practice, and we may wonder what properties the outcomes they give rise to satisfy as individuals' preferences vary. This is a question that the theory can also answer.

I proceed as follows. In the second section I introduce most of the notation and definitions. In the third, I present an "impossibility" result for the case of two players. In the fourth, I discuss the properties of monotonicity and no veto power. I demonstrate that monotonicity is an essential requirement of a social choice rule for implementability. I suggest also that no veto power, though not a necessary condition, is really quite weak and, in fact, is vacuously satisfied in many contexts.

Then in Section V, I present the main result of the paper: a constructive demonstration that monotonicity and no veto power suffice for an SCR of more than two individuals to be implementable in Nash equilibrium. I also show, by example, that the result does not remain true if we drop the no veto power hypothesis.

Finally, in Section VI, I show that we can retain implementability with an even weaker no veto power condition if we impose an individual rationality requirement on the SCR.

II. DEFINITIONS AND NOTATION

Let A be a non-empty, possibly infinite set of social alternatives and let $\mathscr{R}_A$ be the class of all orderings of the elements of A ($\mathscr{R}_A$ is sometimes called the *unrestricted domain* of preferences). If $\mathscr{R}_1, \ldots, \mathscr{R}_n$ are sub-classes of $\mathscr{R}_A$, where n is a positive integer, then f is an n-person *social choice rule* (SCR) on $\boldsymbol{\mathscr{R}} = \mathscr{R}_1 \times \cdots \times \mathscr{R}_n$ if f is a correspondence

$$f: \boldsymbol{\mathscr{R}} \twoheadrightarrow A.$$

One interprets an SCR as selecting a set of "welfare optimal" alternatives $f(\boldsymbol{R})$ for each profile of preferences $\boldsymbol{R} = (R_1, \ldots, R_n) \in \boldsymbol{\mathscr{R}}$, where $R_i (\in \mathscr{R}_i)$ is individual i's preference ordering of A, and $\mathscr{R}_i$ is his domain of possible preference orderings. If $a \in f(\boldsymbol{R})$, we say that a is *f-optimal* for profile $\boldsymbol{R}$.

Prominent examples of SCRs include (i) the (weak) *Pareto correspondence*, which selects all weak Pareto optima corresponding to given profile $\boldsymbol{R}$:

$$f^{\mathrm{PO}}(\boldsymbol{R}) = \{a \mid \text{for all } b \in A \text{ there exists } i \text{ such that } aR_ib\};^{1}$$

(ii) the *Condorcet correspondence*, which, for each profile $\boldsymbol{R}$ of strict preferences,[2] selects each alternative that a (weak) majority prefers to any other alternative:

$$f^{\mathrm{CON}}(\boldsymbol{R}) = \{a \mid \text{for all } b \in A \ \#\{i \mid aR_ib\} \geqq \#\{i \mid bR_ia\}\};^{3}$$

1. The notation "aR_ib" means "a is at least as high as b in the ordering R_i" (*i.e.* a is weakly preferred to b).

2. A preference ordering is *strict* if it ranks no two alternatives as indifferent.

3. The notation $\#\{i \mid aR_ib\}$ denotes the number of individuals who prefer a to b (recall that we are dealing with strict preferences).

and, in a pure exchange economy of l goods, where an alternative a constitutes an *allocation* of goods across individuals (*i.e.* $a = (a_1, \ldots, a_n)$, where $a_i \in \mathbb{R}^l_+$), (iii) the *Walrasian correspondence*, which, given individuals' endowments $(\omega_1, \ldots, \omega_n)$, chooses the set of allocations that can arise in competitive equilibrium:

$$f^{W}(\boldsymbol{R}) = \{a \mid \Sigma a_i = \Sigma \omega_i \text{ and there exists a price vector } \boldsymbol{p} \in \mathbb{R}^l_+ \text{ such that,}$$
$$\text{for all } i,\ a_i \in \mathbb{R}^l_+, \boldsymbol{p} \cdot (a_i - \omega_i) = 0 \text{ and if}$$
$$\text{for some } b_i \in \mathbb{R}^l_+,\ b_i P(R_i) a_i{}^4$$
$$\text{then } \boldsymbol{p} \cdot (b_i - \omega_i) > 0\}.$$

An SCR differs from a *social welfare function* in the sense of Arrow (1951) (a mapping $F\colon \mathscr{R} \to \mathscr{R}_A$) in that it does not rank non-optimal alternatives. Clearly, however, a social welfare function F induces a natural social choice rule: the correspondence which selects the alternatives top-ranked by F for each profile.

Given strategy spaces $S_1, \ldots, S_n$, an *n-person game form* ("mechanism" and "outcome function" are two synonyms) g *on* A is a mapping

$$g\colon S_1 \times \cdots \times S_n \to A.$$

If players 1 through n choose strategies s_1 through s_n, respectively, then alternative $g(\boldsymbol{s})$, where $\boldsymbol{s} = (s_1, \ldots, s_n)$, is the outcome. Moreover, if players use the vector of mixed strategies $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$[5] we denote the random outcome by $g(\boldsymbol{\mu})$, for which the probability of outcome $g(s_1, \ldots, s_n)$ is $\mu_1(s_1) \cdots \mu_n(s_n)$.

We say that the game form g *implements* the social choice rule f *in Nash equilibrium* if and only if

$$\forall \boldsymbol{R} = (R_1, \ldots, R_n) \in \mathscr{R}\ \forall a \in f(\boldsymbol{R}) \text{ there exists}$$
$$\boldsymbol{s} = (s_1, \ldots, s_n) \in \prod_{j=1}^{n} S_j \text{ such that } g(\boldsymbol{s}) = a \text{ and}$$
$$g(\boldsymbol{s}) R_i g(s_i', \boldsymbol{s}_{-i})^6 \text{ for all } i \in \{1, \ldots, n\} \text{ and all } s_i' \in S_i; \tag{1}$$

and

$$\forall \boldsymbol{R} \in \mathscr{R} \text{ if } \boldsymbol{\mu} \text{ is a mixed-strategy Nash equilibrium}^7 \text{ of } g \text{ with respect to}$$
$$\boldsymbol{R} \text{ then } g(\boldsymbol{s}) \in f(\boldsymbol{R}) \text{ for all realizations } \boldsymbol{s} \text{ in the support of } \boldsymbol{\mu}. \tag{2}$$

Requirement (1) needs little explanation if one's solution concept is Nash equilibrium. It states simply that any welfare-optimal alternative (as defined by f) can arise as a (pure-strategy) Nash equilibrium of the game form.[8] We could alternatively impose the weaker requirement that, for all $\boldsymbol{R} \in \mathscr{R}$, there exists *some* $a \in f(\boldsymbol{R})$ for which there is a Nash equilibrium of g resulting in a. But this would not lead to significantly different results. Indeed, if the game form g implements f using the alternative condition in place of (1),

4. $xP(R_i)y$ means that x is strictly preferred to y under R_i.
5. A mixed strategy μ_i for player i assigns a probability $\mu_i(s_i)$ to each (pure) strategy s_i.
6. The notation "$g(s_i', \boldsymbol{s}_{-i})$" denotes $g(s_1, \ldots, s_{i-i}, s_i', s_{i+1}, \ldots, s_n)$.
7. If $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$ are nondegenerate mixed strategy vectors, then player i's preference between $g(\boldsymbol{\mu})$ and $g(\boldsymbol{\mu}')$ may not be fully specified by his ordinal ranking R_i; we may have to know his risk preferences as well. However, the analysis in this paper holds for any risk preferences consistent with R_i.
8. Requirement (1) is essentially the stipulation that the game form be *unbiased* (see Hurwicz (1979*b*)).

define the subcorrespondence f' as $f'(\boldsymbol{R}) = \{a \in f(\boldsymbol{R}) \mid \text{there exists an equilibrium of } g \text{ resulting in } a\}$. Then g implements f' in the standard sense (*i.e.* using (1)).

Requirement (2), which is essentially the converse of (1), is also quite natural. Given that, in general, there can be multiple Nash equilibria of g and that, in the absence of a theory of how players select among these, one cannot predict which of these will ultimately arise, requirement (2) is necessary to ensure f-optimality of the outcome.

III. THE TWO PLAYER CASE

One might well argue that most social choice rules of interest satisfy the Pareto property.

Pareto property. The SCR $f\colon \mathscr{R} \twoheadrightarrow A$ satisfies the Pareto property if, for all $\boldsymbol{R} \in \mathscr{R}$, $f(\boldsymbol{R}) \subseteq f^{\mathrm{PO}}(\boldsymbol{R})$.

We shall see that the prospects for implementing two-person Pareto optimal SCRs on an unrestricted domain of preferences are quite bleak. We need the following definition:

Dictator. An individual i is a dictator for an SCR

$f\colon \mathscr{R} \twoheadrightarrow A$ if and only if

$$[\forall \boldsymbol{R} \in \mathscr{R}\, \forall a \in A,\ a \in f(\boldsymbol{R}) \text{ if and only if } aR_ib \text{ for all } b \in A].$$

In other words, individual i is a dictator if, for any profile of preferences, the set of welfare-optimal alternatives (with respect to f) consists of the *top-ranked* alternatives for i (the alternatives that i prefers to any other). An SCR that has a dictator shall be called *dictatorial.*

I now show that any Pareto-optimal two-person SCR that is implementable must be dictatorial if it is defined on the unrestricted domain of preferences.

Theorem 1. *Let $f\colon \mathscr{R}_A \times \mathscr{R}_A \twoheadrightarrow A$ be a two-person* SCR *satisfying the Pareto property. Then f can be implemented if and only if it is dictatorial.*[9,10]

Proof. First observe that if f is dictatorial, it is trivially implementable; if i is the dictator, just take the game form in which player i announces an alternative and his announcement is implemented.

To prove the proposition in the other direction, suppose that $g\colon S_1 \times S_2 \to A$ implements f. If A contains only one element, the result is trivial. Therefore assume that A contains at least two elements. For each $s_2^* \in S_2$, let $T_1(s_2^*) = \{a \in A \mid g(s_1, s_2^*) \neq a, \text{ for all } s_1 \in S_1\}$. Define $T_2(s_1^*)$ for $s_1^* \in S_1$ analogously. Notice that $T_i(s_j^*)$ is the set of alternatives that player i cannot induce, given that player j's ($j \neq i$) strategy is s_j^*.

Claim 1. *For any $s_1 \in S_1$ and $s_2 \in S_2$, $T_1(s_2) \cap T_2(s_1) = \phi$. That is, starting from any pair of strategies (s_1, s_2), any alternative a can be reached by a unilateral deviation by some player.*

9. This result has also been obtained, in somewhat different form, by Hurwicz and Schmeidler (1978).

10. Theorem 1 remains true if we replace $\mathscr{R}_A$ with the somewhat smaller domain $\mathscr{R}_A^*$ of *strict* orderings.

Proof of Claim 1. Suppose for $\bar{s}_1 \in S_1$ and $\bar{s}_2 \in S_2$, there exists $a \in T_1(\bar{s}_2) \cap T_2(\bar{s}_1)$. Take $b = g(\bar{s}_1, \bar{s}_2)$. Then, $b \neq a$ by construction. Choose $(\bar{R}_1, \bar{R}_2) \in \mathscr{R}_A \times \mathscr{R}_A$ such that, for all $i \in \{1, 2\}$ and all $c \in A \backslash \{a, b\}$, $aP(\bar{R}_i)bP(\bar{R}_i)c$. Observe that $(\bar{s}_1, \bar{s}_2)$ constitutes a Nash equilibrium for preferences $(\bar{R}_1, \bar{R}_2)$, yet b is not Pareto optimal, a contradiction of f's Pareto optimality. Hence $T_1(\bar{s}_2) \cap T_2(\bar{s}_1) = \phi$. ||

Claim 2. *For all $a \in A$, if $a \notin \bigcup_{s_2 \in S_2} T_1(s_2)$, then there exists $\hat{s}_1 \in S_1$ such that, for all $s_2 \in S_2$, $g(\hat{s}_1, s_2) = a$. Similarly, if $a \notin \bigcup_{s_1 \in S_1} T_2(s_1)$, then there exists $\hat{s}_2 \in S_2$ such that $\forall s_1 \in S_1$, $g(s_1, \hat{s}_2) = a$. That is, if no strategy by player* 2 *prevents player* 1 *from inducing alternative a, then player* 1 *has a strategy that guarantees alternative a (and similarly for player* 2).

Proof of Claim 2. It suffices to prove the statement about player 1's strategy $\hat{s}_1$. Consider $a \in A$ such that $a \notin \bigcup_{s_2 \in S_2} T_1(s_2)$. Choose $\hat{R}_1, \hat{R}_2 \in \mathscr{R}_A$ such that $\forall b \in A \backslash \{a\}$, $aP(\hat{R}_1)b$ and $bP(\hat{R}_2)a$. Let $(\hat{s}_1, \hat{s}_2)$ be a Nash equilibrium for $(\hat{R}_1, \hat{R}_2)$. Because $a \notin T_1(\hat{s}_2)$, there exists $\ _1 \in S_1$ such that $g(\ _1, \hat{s}_2) = a$. For $(\hat{s}_1, \hat{s}_2)$ to be a Nash equilibrium, therefore, we must have $g(\hat{s}_1, \hat{s}_2) = a$. Suppose there exist $b \in A \backslash \{a\}$ and $\ _2 \in S_2$ such that $g(\hat{s}_1, \ _2) =$ b. Then from our choice of $\hat{R}_2$, $(\hat{s}_1, \hat{s}_2)$ cannot be a Nash equilibrium. We infer that, $\forall s_2 \in S_2$, $g(\hat{s}_1, s_2) = a$, as desired. ||

Now, for any $a \in A$, either $a \notin \bigcup_{s_2 \in S_2} T_1(s_2)$ or $a \notin \in_{s_1 \in S_1} T_2(s_1)$, otherwise Claim 1 is violated. Suppose there exist $a, b \in A$, $a \neq b$, such that $a \notin \bigcup_{s_2 \in S_2} T_1(s_2)$ and $b \notin \bigcup_{s_1 \in S_1} T_2(s_1)$. By Claim 2, there exist $\bar{s}_1 \in S_1$ and $\bar{s}_2 \in S_2$ such that $\forall s_2 \in S_2, g(\bar{s}_1, s_2) =$ a and $\forall s_1 \in S_1$, $g(s_1, \bar{s}_2) = b$. But then $g(\bar{s}_1, \bar{s}_2) = a$ and $g(\bar{s}_1, \bar{s}_2) = b$, which is impossible. Therefore, either $\forall a \in A, a \notin \bigcup_{s_2 \in S_2} T_1(s_2)$ or $\forall a \in A, a \notin \bigcup_{s_1 \in S_1} T_2(s_1)$. The first statement implies, by Claim 2, that 1 is a dictator for f, the second that player 2 is a dictator. ||

The negative conclusion of Theorem 1 depends on there being an unrestricted domain of preferences. For restricted domains, results can be quite positive, *e.g.* in the case of "economic preferences," where preferences are required to be increasing, continuous, and convex over allocations of a divisible good (see Dutta and Sen (1991) and Moore and Repullo (1990) for a complete characterization of the implementation possibilities in the $n = 2$ case).

III(i). *An example with more than two players*

When $n > 2$, the conclusion of Theorem 1 no longer holds. Indeed, for this case, it is possible to implement Pareto optimal and nondictatorial SCRs defined on the unrestricted domain of preferences. Consider the following example.

Example 1.[11] For any positive integers m and n, take $A = \{a_1, \dots, a_m\}$, $S_1 = \{2, \dots, n\}$, and $S_2 = \cdots = S_n = A$. Define g: $S_1 \times \cdots \times S_n \to A$ so that $\forall (s_1, \dots, s_n) \in \prod_{j=1}^{n} S_j$, $g(s_1, \dots, s_n) = s_{s_1}$. That is, player 1 chooses a player s_1 (other than himself), and player s_1 chooses the outcome from A. I claim that this game form implements the SCR $f^{\mathrm{KM}}(\boldsymbol{R}) = \{a \mid$ there exists $j \in \{2, \dots, m\}$ such that $aR_j b$ for all $b \in A\}$ when players have preferences in $\mathscr{R}_A$. In other words, f^{KM} chooses each alternative for which there exists some individual other than 1 who top-ranks it.

11. This example is adapted from Hurwicz and Schmeidler (1978), who call the game form we have constructed the "king-maker" mechanism.

To see that g implements f^{KM}, we first note that, for any profile $\boldsymbol{R}=(R_1,\ldots,R_n)$, each $a\in f^{\mathrm{KM}}(\boldsymbol{R})$ corresponds to some Nash equilibrium of g. In particular, if $a\in f^{\mathrm{KM}}(\boldsymbol{R})$ there exists some $j\in\{2,\ldots,m\}$ such that a is top-ranked by R_j. Then the strategy profile $(j,a,\ldots,a)$ is a Nash equilibrium for $\boldsymbol{R}$, and $g(j,a,\ldots,a)=a$, as required. Hence, it remains only to show that all Nash equilibria of g are f-optimal. Suppose $(\mu_1,\ldots,\mu_n)$ is a mixed-strategy equilibrium of g with respect to $\boldsymbol{R}$. Consider $j\in\{2,\ldots,n\}$ to which μ_1 assigns positive probability. Then player j has a positive chance of being able to choose the alternative he wants. Therefore for μ_j to be an equilibrium strategy, it must assign positive probability only to player j's top-ranked alternatives. But this means that only an alternative that is top-ranked for some individual among $2,3,\ldots,n$ can be a realization of $g(\mu_1,\ldots,\mu_n)$, which is what we wanted to show.

Whether or not we take satisfaction from the fact that f^{KM} is implementable, it is only an example. Clearly, what is needed is a set of general criteria for whether *any* given SCR is implementable. It is to this task to which I now turn.

IV. MONOTONICITY AND NO VETO POWER

The condition on SCRs that is central to their implementability is monotonicity.

Monotonicity.[12] The SCR $f:\mathscr{R}\twoheadrightarrow A$ satisfies monotonicity provided that $\forall a\in A$, $\forall \boldsymbol{R},\boldsymbol{R}'\in\mathscr{R}$ if $a\in f(\boldsymbol{R})$ and $[\forall i\in\{1,\ldots,n\}\ \forall b\in A\ aR_ib\Rightarrow aR_i'b]$, then $a\in f(\boldsymbol{R}')$.

In words, monotonicity requires that if alternative a is f-optimal with respect to some profile of preferences and the profile is then altered so that, in each individual's ordering, a does not fall below any alternative that it was not below before, then a remains f-optimal with respect to the new profile.

To see that monotonicity is "reasonable", let us observe that it is satisfied by the prominent SCRs mentioned in Section II. First consider the Pareto correspondence f^{PO}. If a is (weakly) Pareto optimal with respect to $\boldsymbol{R}$ then, for all b, there exists j_* such that $aR_{j_*}b$. But if we replace $\boldsymbol{R}$ by $\boldsymbol{R}'$ such that, for all i, $aR_ib\Rightarrow aR_i'b$, we conclude that $aR'_{j_*}b$. Hence, b is (weakly) Pareto optimal with respect to $\boldsymbol{R}'$, establishing the monotonicity of F^{PO}.

Next, let us examine the Condorcet correspondence f^{CON}. If a is a majority winner for a strict profile (a profile consisting of strict orderings) $\boldsymbol{R}$, then, for any other alternative b, the number of individuals preferring a to b is no less than the number preferring b to a

$$\#\{i\,|\,aR_ib\}\geqq\#\{i\,|\,bR_ia\}. \tag{3}$$

But if $\boldsymbol{R}'$ is a profile such that, for all i, $aR_ib\Rightarrow aR_i'b$, then the left-hand side of (3) cannot fall when we replace $\boldsymbol{R}$ by $\boldsymbol{R}'$. Furthermore, if the right-hand side rises, then we must have aR_ib and $bR_i'a$ for some i, a contradiction of the relation between $\boldsymbol{R}$ and $\boldsymbol{R}'$, given the strictness of preferences. We conclude that (3) continues to hold when $\boldsymbol{R}'$ replaces $\boldsymbol{R}$, and so a is still a majority winner.

12. Monotonicity is called "strong positive association" by Muller and Satterthwaite (1977), who show that when f is *single-valued* and the domain consists of all strict preferences $\mathscr{R}_A^*$ then monotonicity is necessary and sufficient for implementation in dominant strategies. However, more generally, when f is either nonsingle-valued or the domain of preferences admits indifference or is more highly restricted than $\mathscr{R}_A^*$, this characterization result no longer obtains (see Dasgupta, Hammond, and Maskin (1979)).

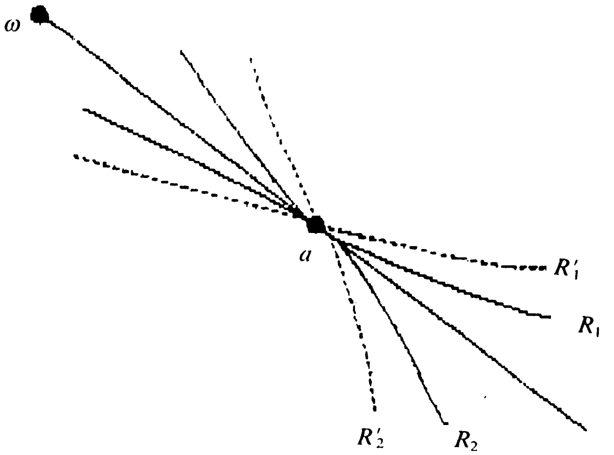

FIGURE 1
The Walrasian correspondence

As for the Walrasian correspondence, refer to the Edgeworth Box in Figure 1. In this two-consumer, two-good economy, allocation a is a competitive equilibrium allocation with respect to the endowments ω and the preference profile $\boldsymbol{R} = (R_1, R_2)$. If we now alter $\boldsymbol{R}$ so that any allocation that was worse than a for consumer i remains worse than a, we obtain profile $\boldsymbol{R}' = (R_1', R_2')$, with respect to which a remains a competitive equilibrium. Hence, f^{W} is monotonic.[13]

We should also point out that monotonicity is *automatically* satisfied by any SCR whose domain of preferences is among certain classes of preferences often studied in the literature. For example, this is true of classes of preferences satisfying the "single-crossing" property (*i.e.* the "Spence/Mirrlees" condition). A set of preferences $\mathscr{R}_i$ satisfies this property if no two indifference curves in the class intersect more than once. Notice that this means (refer to Figure 2) that if $R_i, R_i' \in \mathscr{R}_i$ and $a, b \in A$ are such that aR_ib and $aR_i'b$, then there exists another alternative $b' \in A$ such that aR_ib' but $b'P(R_i')a$. Hence, the hypothesis of the monotonicity condition cannot be satisfied, and so the condition holds vacuously.

A particularly interesting case in which single-crossing holds is that in which alternatives are nondegenerate lotteries over a set of possible outcomes. If individuals' preferences over lotteries satisfy the von Neumann–Morgenstern axioms, then indifference curves in probability space are straight, parallel lines, and so clearly indifference curves corresponding to distinct preferences can intersect only once. This insight figures prominently in the literature on "virtual implementation" (see Abreu and Sen (1991) and Abreu and Matsushita (1992)).

For an example of a well-known SCR that fails to satisfy monotonicity, consider the Borda Court (*i.e.* rank-order voting) SCR f^{BC}. For each individual, according to f^{BC}, points are assigned to each of the m alternatives available: m points are assigned to his favourite alternative, $m-1$ to his next favourite, and so on. The alternative (or alternatives) chosen by f^{BC} is the one for which the sum of points over individuals is highest. Suppose that $A = \{a, b, c, d\}$ (*i.e.* $m = 4$) and $n = 2$. Consider the profile $\boldsymbol{R} = (R_1, R_2)$:

13. This argument relies on competitive allocations like a being *interior* allocations. For what can go wrong if a competitive allocation occurs on the boundary, see Hurwicz, Maskin, and Postlewaite (1995).

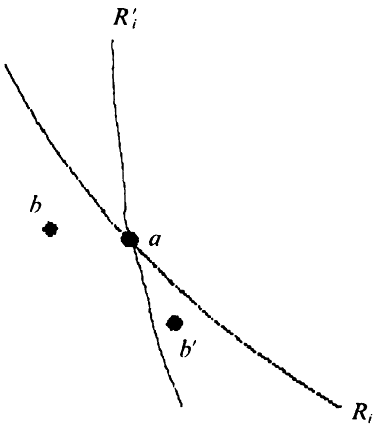

FIGURE 2
Single-crossing preferences

$\boldsymbol{R}_1$	$\boldsymbol{R}_2$
a	c
d	b
b	a
c	d

Note that in this profile, alternative a garners the most points (6), and so is chosen by f^{BC}. Next consider the profile $\boldsymbol{R}' = (R_1', R_2')$:

$\boldsymbol{R}_1'$	$\boldsymbol{R}_2'$
a	b
b	c
d	a
c	d

Notice that in going from R_i to R_i', a does not fall *vis-à-vis* any other alternative. Thus, monotonicity would require that it still be chosen for profile $\boldsymbol{R}'$. However, a no longer attracts the most points; alternative b does (7). Hence monotonicity is violated.

Whether one accepts monotonicity as natural or has qualms about its restrictiveness, it is an inescapable requirement for implementability in Nash equilibrium, as the following result shows.

Theorem 2. *If $f: \mathscr{R} \twoheadrightarrow A$ is an* SCR *that is implementable in Nash equilibrium, then it is monotonic.*

Proof. Suppose that f is implementable in Nash equilibrium by the game form $g: S_1 \times \cdots \times S_n \to A$. For some profile $\boldsymbol{R} \in \mathscr{R}$ consider $a \in f(\boldsymbol{R})$. Then there exists a Nash equilibrium $\boldsymbol{s}$ of g with respect to $\boldsymbol{R}$ such that $g(\boldsymbol{s}) = a$. Consider profile $\boldsymbol{R}' \in \mathscr{R}$ such that

$$\text{for all } i \text{ and all } b \in A \qquad aR_i b \Rightarrow aR'_i b. \tag{4}$$

If there exist i and s'_i such that $g(s'_i, s_{-i})P(R'_i)g(s) = a$, then from (4), $g(s'_i, s_{-i})P(R_i)a$, a contradiction of the assumption that $\boldsymbol{s}$ is a Nash equilibrium with respect to $\boldsymbol{R}$. Hence $\boldsymbol{s}$ is also a Nash equilibrium with respect to $\boldsymbol{R}'$. From requirement (2) of the definition of implementability, therefore, we conclude that $a \in f(\boldsymbol{R}')$. Thus f is monotonic. ||

I will show below (Theorem 3) that not only is monotonicity a necessary condition for implementability, as just demonstrated, but almost a sufficient condition as well. Nevertheless monotonicity by itself does not suffice to ensure implementability (see Example 2 below). One weak condition that we can add to monotonicity to do the trick is no veto power:

No Veto Power (NVP). An SCR $f: \mathscr{R} \twoheadrightarrow A$ satisfies NVP if, for all $\boldsymbol{R} \in \mathscr{R}$ and all $a \in A$, whenever there exists $i \in \{1, \dots, n\}$ such that, for all $j \neq i$ and all $b \in A$, $aR_j b$, then $a \in f(\boldsymbol{R})$.

NVP says that if an alternative is at the top of $n-1$ individuals' preference orderings, then the last individual cannot prevent the alternative from being f-optimal (*i.e.* he cannot "veto" it).

NVP is satisfied by virtually all "standard" SCRs (including the Pareto and Condorcet correspondences). It is also often automatically satisfied by any SCR when preferences are restricted. Consider, for example, a pure exchange economy with at least three consumers, in which an alternative corresponds to an allocation of goods. If there exists at least one good that is transferable and desirable, that gives rise to no externalities, and that is available in a positive quantity, then it is impossible to find an alternative that all but one consumer rank at the top of their preference orderings. This is because, for an individual to prefer a given allocation to all others, the allocation must assign him all of the good in question; if any other individuals got some of this good, he would be better off receiving their portions. Clearly, no other individual could also rank this allocation first, since it cannot be the case the two consumers each receive all of the good in question. Therefore the NVP property is satisfied vacuously.

V. SUFFICIENT CONDITIONS FOR IMPLEMENTATION

I now present the main result of the paper.

Theorem 3. *If* $n \geqq 3$ *and* $f: \mathscr{R} \rightarrow A$ *is a n-person* SCR *satisfying monotonicity and* NVP, *then it is implementable in Nash equilibrium.*

Proof. The proof is by construction. I first show that we can construct a game form all of whose *pure-strategy* equilibria satisfy (1) and (2).[14] In the appendix I show that the construction can be extended to handle mixed strategies. For each player i, define the strategy space

$$S_i = \mathscr{R} \times A \times \mathscr{N},$$

14. This elegant proof is due essentially to Repullo (1987).

where $\mathcal{N}$ consists of the nonnegative integers. In other words, player i chooses as a strategy a triple consisting of a preference profile $\boldsymbol{R}^i$ in $\mathcal{R}$ (not necessarily the true one), an alternative a^i from A, and a number m^i in $\mathcal{N}$ (the numbers serve only to break ties).

For all i, all $R_i \in \mathcal{R}_i$ and all $a \in A$ define

$$L(a, R_i) = \{b \mid aR_ib\}.$$

$L(a, R_i)$ is the lower contour set of R_i at alternative a: the set of alternatives that are no better than a according to R_i.

I will construct an implementing game form g that implements f in three steps:

(i) If, for some $\boldsymbol{R}$, a, and m,

$$s_1 = \cdots = s_n = (\boldsymbol{R}, a, m) \text{ and } a \in f(\boldsymbol{R}), \qquad \text{take } g(s_1, \dots, s_n) = a. \tag{5}$$

In words, if players are unanimous in their strategy, and their proposed alternative a is f-optimal given their proposed profile $\boldsymbol{R}$, the outcome is a.

(ii) If, for all $j \neq i$, $s_j = (\boldsymbol{R}, a, m)$, $s_i = (\boldsymbol{R}^i, a^i, m^i) \neq (\boldsymbol{R}, a, m)$, and $a \in f(\boldsymbol{R})$ take

$$g(s_1, \dots, s_n) = \begin{cases} a^i, & \text{if } a^i \in L(a, R_i) \\ a, & \text{if } a^i \notin L(a, R_i). \end{cases}$$

That is, suppose that all players but one play the same strategy and, given their proposed profile, their proposed alternative a is f-optimal. Then, the odd-man-out, gets his proposed alternative, provided that it is in the lower contour set at a of the ordering that the *other* players propose for him; otherwise, the outcome is a.

(iii) If neither (i) nor (ii) applies, then

$$g(s_1, \dots, s_n) = a^{i*}, \tag{6}$$

where $i^* = \max\{i \mid m^i = \max_j m^j\}$. In other words, when neither (i) nor (ii) applies, the outcome is the alternative proposed by the player with the highest index among those whose proposed number is maximal.

It remains to show that this game form implements f. I first claim that, for all $\boldsymbol{R} \in \mathcal{R}$ and all $a \in A$, if $a \in f(\boldsymbol{R})$, then, for any $m \in \mathcal{N}$, the strategy profile $(s_1, \dots, s_n)$ satisfying (5) constitutes a Nash equilibrium with respect to $\boldsymbol{R}$. To see this, note from (i) that the outcome from this strategy profile is a. Moreover, from (ii), any player i who deviates unilaterally from $(s_1, \dots, s_n)$ gets an alternative in $L(a, R_i)$, which, by definition of the lower contour set, is no better for him than a. Thus I have established requirement (1)—that there is a Nash equilibrium of g corresponding to each f-optimal alternative—in the definition of implementability.

To establish (2)—that every Nash equilibrium of g is f-optimal—consider first a Nash equilibrium $(s_1, \dots, s_n)$ in which (5) holds and $a \in f(\boldsymbol{R})$, but where the true preference profile is $\boldsymbol{R}'$. From (i), the equilibrium outcome is a. Moreover, because $(s_1, \dots, s_n)$ is an equilibrium with respect to $\boldsymbol{R}'$, (ii) implies that

$$\text{for all } i \text{ and all } b \in L(a, R_i), \qquad aR'_ib. \tag{7}$$

(To understand why (7) holds, note that if instead we had $bP(R'_i)a$ for some i and $b \in L(a, R_i)$, it would pay player i to deviate from s_i and induce b, which (ii) implies he could do. But this would contradict the assumption that $(s_1, \dots, s_n)$ is an equilibrium.) But (7) can be rewritten as

$$\text{for all } i \text{ and all } b \in A,\ aR_ib \Rightarrow aR'_ib. \tag{8}$$

Hence because f satisfies monotonicity, (8) and the fact that $a \in f(\boldsymbol{R})$ imply that $a \in f(\boldsymbol{R}')$, *i.e.* the equilibrium outcome is f-optimal.

Next let us consider a Nash equilibrium $(s_1, \dots, s_n)$ for $\boldsymbol{R}'$ in which, for all $j \neq i$,

$$s_j = (\boldsymbol{R}, a, m),$$

where $a \in f(\boldsymbol{R})$, but $s_i \neq (\boldsymbol{R}, a, m)$, *i.e.* the strategy profile is such that (ii) applies. Let the outcome from this profile be a'. From (iii), each player $j \neq i$ could deviate from s_j and induce any alternative $a \in A$ he wishes by choosing m^j high enough (*i.e.* higher than $\max_{k \neq j} m^k$). Hence, the fact that $(s_1, \dots, s_n)$ is a Nash equilibrium for $\boldsymbol{R}'$ implies that, for all $j \neq i$,

$$a' R'_j b \quad \text{for all } b \in A. \tag{9}$$

We conclude that NVP together with (9) ensures that $a' \in f(\boldsymbol{R}')$, *i.e.* the equilibrium outcome is again f-optimal.

The same argument as in the preceding paragraph applies if $(s_1, \dots, s_n)$ is a Nash equilibrium for which (iii) applies. ||

Remark 1. The game form constructed in the proof of Theorem 3 may be considered rather complicated. However, much of the complexity derives from its generality—the fact that it is supposed to work for a vast array of possible SCRs. For a *specific* SCR, by contrast, it is often possible to find an implementation that is quite simple (*e.g.* the mechanism in Example 1).

Remark 2. Even if the set of alternatives A is finite, the game form constructed in the proof of Theorem 3 has an unbounded strategy space, since $\mathcal{N}$ is unbounded. Jackson (1992) points out, however, that, for some solution concepts, the set of SCRs implementable by unbounded game forms is strictly larger than the limit of those implementable by bounded game forms as the bound goes to infinity. It remains an open question whether this is so for Nash equilibrium.

We have argued that no veto power is a weak condition. It is nevertheless restrictive, and so it is of some interest understanding its role in Theorem 3.[15] As we will see below (Theorem 4) NVP is not necessary for implementability. However, as the following example establishes, monotonicity by itself does not suffice.

Example 2. Suppose that $n = 3$ and $A = \{a, b, c\}$. For all i, let $\mathscr{R}_i = \mathscr{R}^*_A$ (*i.e.* the domain consists of all strict orderings). Define f^* such that, for all $\boldsymbol{R} \in \mathscr{R}$,

for each $x \in \{a, b\}$, $x \in f^*(\boldsymbol{R})$ if and only if x is Pareto-optimal and top-ranked for individual 1

$c \in f^*(\boldsymbol{R})$ if and only if c is Pareto optimal and not bottom-ranked for individual 1.

It is easy to verify that f^* is monotonic. However, it does not satisfy NVP because if individual 1 bottom-ranks alternative c, it fails to be f^*-optimal even if individuals 2 and 3 top-rank c.

15. For conditions that are necessary and sufficient for implementability, see Moore and Repullo (1990).

Consider the following three profiles $\boldsymbol{R}^*, \boldsymbol{R}^{**}, \boldsymbol{R}^{***}$:

$$\begin{aligned} \boldsymbol{R}^* &= ([b, c, a], [c, a, b], [c, a, b]) \\ \boldsymbol{R}^{**} &= ([a, b, c], [c, b, a], [c, a, b]) \\ \boldsymbol{R}^{***} &= ([b, a, c], [a, b, c], [a, b, c]), \end{aligned}$$

where "$[x, y, z]$" denotes the ordering in which x is preferred to y, and y is preferred to z. Then

$$f^*(\boldsymbol{R}^*) = \{b, c\}, f^*(\boldsymbol{R}^{**}) = \{a\}, f^*(\boldsymbol{R}^{***}) = \{b\}.$$

If f^* were implementable, there would exist a game form g and a Nash equilibrium $\boldsymbol{s}^* = (s_1^*, s_2^*, s_3^*)$ with respect to $\boldsymbol{R}^*$ such that $g(\boldsymbol{s}^*) = c$. Because $bP(R_1^*)c$ there does not exist $s_1'' \in S_1$ such that $g(s_1'', s_2^*, s_3^*) = b$.

If there existed $s_1' \in S_1$ such that $g(s_1', s_2^*, s_3^*) = a$, then (s_1', s_2^*, s_3^*) would be a Nash equilibrium for $\boldsymbol{R}^{***}$, a contradiction since $a \notin f(\boldsymbol{R}^{***})$. Hence, s_1' cannot exist. We conclude that (s_1^*, s_2^*, s_3^*) is a Nash equilibrium for $\boldsymbol{R}^{**}$, which contradicts the fact that $c \notin f(\boldsymbol{R}^{**})$. Hence, f^* is not implementable.

VI. INDIVIDUAL RATIONALITY

We will say that an SCR $f: \mathscr{R} \twoheadrightarrow A$ is *individually rational* (IR) with respect to some alternative $a^0 \in A$ if for all $\boldsymbol{R} \in \mathscr{R}$, all $a \in f(\boldsymbol{R})$, and all i, $aR_i a^0$. That is, if a is f-optimal, it must be weakly preferred by all individuals to a^0. In general, an SCR satisfying IR does not satisfy NVP because if, for some profile preference, everyone but individual i top-ranks alternative a, then NVP would require that a be f-optimal with respect to that profile. But f would then violate IR if i strictly preferred a^0 to a.

Nevertheless, many SCRs satisfying IR are implementable. One example is the "Individual Rationality" correspondence: for all $\boldsymbol{R} \in \mathscr{R}$

$$f^{\mathrm{IR}}(\boldsymbol{R}) = \{a \in A \mid aR_i a^0 \text{ for all } i\}.$$

This SCR is implemented by the game form g^{IR} such that $S_i = A$ for all i and

$$g^{\mathrm{IR}}(s_1, \ldots, s_n) = \begin{cases} a, & \text{if } s_1 = \cdots = s_n = a \text{ for some } a \in A \\ a^0, & \text{otherwise.} \end{cases}$$

The example of f^{IR} suggests that if we relax NVP so that it applies only to individually rational alternatives, we might obtain a general result.

Weak no veto power (WNVP). An SCR $f: \mathscr{R} \twoheadrightarrow A$ satisfies WNVP with respect to a^0 if, for all $\boldsymbol{R} \in \mathscr{R}$ and all $a \in A$, whenever there exists i such that for all $j \neq i$ $aR_j b$ for all b and $aR_i a^0$, then $a \in f(\boldsymbol{R})$.

Theorem 4. *If $n \geqq 3$ and $f: \mathscr{R} \twoheadrightarrow A$ is an* SCR *satisfying monotonicity,* WNVP, *and* IR *with respect to a^0, then it is implementable in Nash equilibrium.*

Proof. The proof of Theorem 4 uses exactly the same construction as that of Theorem 3. The only thing to show is that Nash equilibria satisfying configurations (ii) or (iii) in the proof of Theorem 3 are individually rational. This enables us to apply WNVP and complete the proof.

Thus, consider a configuration (ii) equilibrium $(s_1, \ldots, s_n)$ with respect to profile $\boldsymbol{R}'$. That is, there exist i, a, m, and $\boldsymbol{R}$ such that $a \in f(\boldsymbol{R})$ and, for all $j \neq i$, $s_j = (\boldsymbol{R}, a, m)$. We must show that $g(s_1, \ldots, s_n) R'_k a^0$ for all k. This is immediate for all $j \neq i$, since each of those players can deviate from s_j and induce his top-ranked alternative. Hence, the fact that $(s_1, \ldots, s_n)$ constitutes an equilibrium means that $g(s_1, \ldots, s_n)$ itself must be top-ranked and so $g(s_1, \ldots, s_n) R'_j a^0$. As for player i, note that because f satisfies IR with respect to a^0, $a R_i a^0$, *i.e.* $a^0 \in L(R_i, a)$. Hence, by construction of g, player i can induce a^0 when the other players all use strategy $(\boldsymbol{R}, a, m)$. Thus because $(s_1, \ldots, s_n)$ is an equilibrium implies that $g(s_1, \ldots, s_n) R'_i a^0$. The argument is virtually identical for configuration (iii). ||

APPENDIX

Proof of Theorem 3 *for mixed strategies*

The argument that all Nash equilibria of the mechanism in Theorem 3 are f-optimal does not carry over to mixed strategies.

To see the problem consider a mixed-strategy equilibrium $(\mu_1, \ldots, \mu_n)$ (for profile $\boldsymbol{R}'$) for which one possible realization is $\boldsymbol{s} = (s_1, \ldots, s_n)$ where, for some j, $\boldsymbol{R} \in \mathscr{R}$, and $a \in f(\boldsymbol{R})$,

$$s_i = (\boldsymbol{R}, a, 0) \quad \text{for all } i \neq j,$$

but $s_j \neq (\boldsymbol{R}, a, 0)$. In the proof of Theorem 3, I showed that the outcome corresponding to $\boldsymbol{s}$ must be f-optimal since, by deviating from s_i, each player $i \neq j$ could induce his favourite alternative a^i (NVP then would imply f-optimality of the outcome). But if there are other possible realizations of μ_{-i}, then player i might suffer by trying to induce a^i. Suppose, for example, that $\boldsymbol{s}'_{-i}$ is a realization in which, for some $\boldsymbol{R}' \in \mathscr{R}$ and $a' \in f(\boldsymbol{R}')$

$$s'_k = (\boldsymbol{R}', a', 0) \quad \text{for all } k \neq i.$$

Assume, furthermore, that

$$a^i P(R'_i) a'. \tag{*}$$

Then, although individual i can induce a^i against $\boldsymbol{s}_{-i}$, formula (*) and the construction of Theorem 3 imply that he cannot induce a^i against $\boldsymbol{s}'_{-i}$. Indeed, if he tries to do so, the outcome will be a', which may be a very bad alternative for him.

I now show, however, that the game form from Theorem 3 can be modified to circumvent this difficulty. For each player i, define the strategy space

$$S_i = \mathscr{R} \times A \times \{\alpha \mid \alpha : \mathscr{R}^n \times A^n \to A\} \times \mathscr{N}.$$

In other words, player i chooses as a strategy a quadruple consisting of a preference profile $\boldsymbol{R}^i \in \mathscr{R}$, an alternative $a^i \in A$, a function $\alpha^i(\cdot)$ mapping each possible vector of announced profiles and alternatives $(\boldsymbol{R}^1, \ldots, \boldsymbol{R}^n, a^1, \ldots, a^n)$ to an alternative $\alpha^i(\boldsymbol{R}^1, \ldots, \boldsymbol{R}^n, a^1, \ldots, a^n) \in A$, and an integer $m^i \in \mathscr{N}$.

As in the proof of Theorem 3, I will construct the implementing game form g in three steps:

(i) If $s_1 = \cdots = s_n = (\boldsymbol{R}, a, \alpha(\cdot), m)$ and $\alpha(\boldsymbol{R}, \ldots, \boldsymbol{R}, a, \ldots, a) = a \in f(\boldsymbol{R})$, take $g(s_1, \ldots, s_n) = a$.

In other words, if players are unanimous in their strategy, and their proposed alternative a is prescribed by their proposed function $\alpha(\cdot)$ and f-optimal given their proposed profile $\boldsymbol{R}$, the outcome is a.

(ii) If, for all $j \neq i$, $s_j = (\boldsymbol{R}, a, \alpha(\cdot), m)$ with $\alpha(\boldsymbol{R}, \ldots, \boldsymbol{R}, a, \ldots, a) = a \in f(\boldsymbol{R})$ but $s_i = (\boldsymbol{R}^i, a^i, \alpha^i(\cdot), m^i) \neq (\boldsymbol{R}, a, \alpha(\cdot), m)$, take

$$g(s_1, \ldots, s_n) = \begin{cases} \alpha^i(\boldsymbol{R}, \ldots, \boldsymbol{R}^i, \ldots, \boldsymbol{R}, a, \ldots, a^i, \ldots a), & \text{if } \alpha^i(\boldsymbol{R}, \ldots, \boldsymbol{R}^i, \ldots, \boldsymbol{R}, a, \ldots, a^i, \ldots, a) \in L(a, R_i), \\ a, & \text{otherwise.} \end{cases}$$

That is, suppose that all players but player i play the same strategy and their proposed alternative a is prescribed by their proposed function $\alpha(\cdot)$ and f-optimal, given their proposed profile $\boldsymbol{R}$. Then, player i gets the alternative prescribed by his proposed function $\alpha^i(\cdot)$ (given the vector of proposed profiles and alternatives

$(\boldsymbol{R},\ldots,\boldsymbol{R}^i,\ldots,\boldsymbol{R},a,\ldots,a^i,\ldots,a))$ provided that it is in the lower contour set at a of the ordering that the other players propose for him; otherwise, the outcome is a.

(iii) If neither (i) nor (ii) applies, then

$$g(s_1,\ldots,s_n)=\alpha^{i^*}(\boldsymbol{R}^1,\ldots,\boldsymbol{R}^n,a^1,\ldots,a^n),\tag{A1}$$

where $i^*=\max\{i \mid m^i=\max_j m^j\}$. That is, the outcome is the alternative prescribed by the proposed function of the player whose index is highest among those proposing the maximal number.

I must show that this game form implements f. Note first that, for all $\boldsymbol{R}\in\mathscr{R}$ and all $a\in A$, if $a\in f(\boldsymbol{R})$, then, for any $m\in\mathscr{N}$, the strategy profile $(s_1,\ldots,s_n)$ where

$$s_1=\cdots=s_n=(\boldsymbol{R},a,\alpha(\cdot),m) \quad \text{and} \quad \alpha(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a)=a,\tag{A2}$$

constitutes a Nash equilibrium with respect to $\boldsymbol{R}$. To see this, note from (i) that the outcome from this strategy profile is a. Moreover, from (ii), any player i who deviates unilaterally from $(s_1,\ldots,s_n)$ induces an alternative in $L(a,R_i)$, which by definition of $L(a,R_i)$ is no better for player i (with preference ordering R_i). Thus the profile (A2) indeed constitutes a Nash equilibrium with respect to $\boldsymbol{R}$, and so I have established that, for every f-optimal alternative, there is a Nash equilibrium of g giving rise to that alternative.

It remains to show that if $(\mu_1,\ldots,\mu_n)$ is a Nash equilibrium for g with respect to profile $\boldsymbol{R}'$, then the outcome corresponding to each realization $(s_1^*,\ldots,s_n^*)$ in the support of $(\mu_1,\ldots,\mu_n)$ is f-optimal. Suppose first that $(s_1^*,\ldots,s_n^*)$ is a realization for which (A2) holds and $a\in f(\boldsymbol{R})$, but that the profile $\boldsymbol{R}$ differs from the true profile $\boldsymbol{R}'$. From (i), the equilibrium outcome is a. For any player i, consider $b\in A$ such that aR_ib. Now imagine that player i plays $s_i=(\boldsymbol{R}^i,a^i,\alpha^i,m^i)$ such that

$$(\boldsymbol{R}^i,a^i,m^i)=(\boldsymbol{R},a,m),\tag{A3}$$

and

$$\alpha^i(\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n)=\begin{cases} b, & \text{if } (\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n)=(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a),\\ \alpha(\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n), & \text{otherwise.}\end{cases}\tag{A4}$$

That is, s_i is the same as $s_i^*=(\boldsymbol{R},a,\alpha(\cdot),m)$ except for the function $\alpha^i(\cdot)$, which, in turn, is the same as $\alpha(\cdot)$ except at the point $(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a)$. Now because $b\in L(a,R_i)$, (ii), (A3), and (A4) together imply that

$$g(s_i,\boldsymbol{s}_{-i}^*)=b.\tag{A5}$$

Moreover, because $\alpha^i(\cdot)$ is the same as $\alpha(\cdot)$ except at $(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a)$,

$$g(s_i,\hat{\boldsymbol{s}}_{-i})=g(s_i^*,\hat{\boldsymbol{s}}_{-i})\tag{A6}$$

for any realization $\hat{\boldsymbol{s}}_{-i}=(\hat{\boldsymbol{R}}^{-i},\hat{\boldsymbol{a}}^{-i},\hat{\boldsymbol{a}}^{-i}(\cdot),\hat{\boldsymbol{m}}^{-i})$ of μ_{-i} such that $(\hat{\boldsymbol{R}}^{-i},\hat{\boldsymbol{a}}^{-i})\neq(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a)$. Hence, if $bP(R_i')a$, (A5) and (A6) imply that player i is better off using s_i rather than s_i^* against μ_{-i}. We conclude that

$$\text{for all } i \text{ and all } b \qquad aR_ib\Rightarrow aR_i'b.\tag{A7}$$

Hence, because f is monotonic (A7) and the fact that $a\in f(\boldsymbol{R})$ imply that $a\in f(\boldsymbol{R}')$. That is, the outcome $g(s_1^*,\ldots,s_n^*)$ is f-optimal, as required.

Next let us consider a realization $(s_1^*,\ldots,s_n^*)$ in the support of $(\mu_1,\ldots,\mu_n)$ in which, for all $j\neq i$,

$$s_j^*=(\boldsymbol{R},a,\alpha(\cdot),m),$$

where $\alpha(\boldsymbol{R},\ldots,\boldsymbol{R},a,\ldots,a)=a\in f(\boldsymbol{R})$ but $s_i^*=(\boldsymbol{R}^i,a^i,\alpha^i(\cdot),m^i)\neq(\boldsymbol{R},a,\alpha(\cdot),m)$. That is, the strategy profile is such that (ii) applies. Let the outcome from this profile be a'. For any $j\neq i$, choose $b^j\in A$ such that

$$b^jR_j'b \quad \text{for all } b\in A.\tag{A8}$$

Then, consider $s_j=(\boldsymbol{R}^j,a^j,\alpha^j(\cdot),m^j)$ such that

$$(\boldsymbol{R}^j,a^j)=(\boldsymbol{R},a),\tag{A9}$$

$$m^j>\max\{m^i,m\},\tag{A10}$$

and

$$\alpha^j(\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n)$$
$$=\begin{cases} b^j, & \text{if } (\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n)=(\boldsymbol{R},\ldots,\boldsymbol{R}^i,\ldots,\boldsymbol{R},a,\ldots,a^i,\ldots,a),\\ \alpha(\hat{\boldsymbol{R}}^1,\ldots,\hat{\boldsymbol{R}}^n,\hat{a}^1,\ldots,\hat{a}^n), & \text{otherwise.}\end{cases}\tag{A11}$$

That is, s_j is the same as $s_j^* = (\boldsymbol{R}, a, \alpha(\cdot), m)$ except for the integer m_j (which we take to be bigger than the other numbers in s_{-j}^*) and the function $\alpha^j(\cdot)$, which, in turn, is the same as $\alpha(\cdot)$ except at the point $(\boldsymbol{R}, \ldots, \boldsymbol{R}^i, \ldots, \boldsymbol{R}, a, \ldots, a^i, \ldots, a)$. From (A9)–(A11),

$$g(s_j, s_{-j}^*) = b^j. \tag{A12}$$

Moreover, for each $\hat{s}_{-j} \neq s_{-j}^*$ (A9)–(A11) ensure that either

$$g(s_j, \hat{s}_{-j}) = b^j, \tag{A13}$$

or

$$g(s_j, \hat{s}_{-j}) = g(s_j^*, \hat{s}_{-j}). \tag{A14}$$

Hence, from (A8) and (A12)–(A14), we conclude that player j does strictly better with s_j than with s_j^* against μ_{-j}, a contradiction, unless player j does not strictly prefer b^j to a', *i.e.* unless

$$a' R_j' b \quad \text{for all } b \in A. \tag{A15}$$

Thus (A15) must hold for all $j \neq i$, and so, from NVP, $a' \in f(\boldsymbol{R}')$, as required.

The same argument as in the preceding paragraph applies if $(s_1, \ldots, s_n)$ is a realization in the support of $(\mu_1, \ldots, \mu_n)$ to which (iii) applies. ||

Remark. The reason for having players report functions $\alpha^i(\cdot)$ rather than merely fixed alternatives is to be able to accommodate mixed strategies. Which alternative is best for a player to propose will, in general, depend on the profiles and alternatives that the other players propose. But if the others are playing mixed strategies, then a player may not be able to forecast (except probabilitically) what these proposals will be. Allowing him to propose a function enables him, in effect, to propose an alternative on a contingent basis.

Acknowledgements. This work was supported by a grant from the National Science Foundation and a research fellowship from Jesus College, Cambridge. The paper was first presented at the Summer Workshop of the Econometric Society in Paris, June, 1977.

For useful conversations about the original 1977 version I thank Frank Hahn, Peter Hammond, Jerry Hausman, Leo Hurwicz, Roy Radner, and William Thomson. I am especially indebted to Karl Vind, whose suggestions led to a substantial strengthening and simplification of my results. I thank a referee for comments on the 1998 version.

I am most grateful to the Editor, Patrick Bolton, for giving me the opportunity to publish this elderly paper. The literature on implementation is now very large. But because there are a number of excellent recent surveys (see Moore (1993), Palfrey (1993) and (1998), Chapter 10 of Osborne and Rubinstein (1994), and Corchon (1996); see also my old survey Maskin (1985)), I have not made a systematic attempt to bring the references up to date. Indeed, I have made as few changes as possible to the 1977 text.

REFERENCES

ABREU, D. and SEN, A. (1991), "Virtual Implementation in Nash Equilibrium", *Econometrica*, **59**, 997–1021.

ABREU, D. and MATSUSHIMA, H. (1992), "Virtual Implementation in Iteratively Undominated Strategies: Complete Information", *Econometrica*, **60**, 993–1008.

ARROW, K. (1951) *Social Choice and Individual Values* (New York: John Wiley).

CORCHON, L. (1996) *The Theory of Implementation of Socially Optimal Decisions in Economics* (New York: St. Martin's Press).

DASGUPTA, P., HAMMOND, P. and MASKIN, E. (1979), "The Implementation of Social Choice Rules", *Review of Economic Studies*, **46**, 185–216.

DUTTA, B. and SEN, A. (1991), "A Necessary and Sufficient Condition for Two-Person Nash Implementation", *Review of Economic Studies*, **58**, 121–128.

GIBBARD, A. (1973), "Manipulation of Voting Schemes: A General Result", *Econometrica*, **41**, 587–602.

GROVES, T. and LEDYARD, J. (1977), "Optimal Allocation of Public Goods: A Solution to the Free Rider Problem", *Econometrica*, **41**, 617–631.

HURWICZ, L. (1979*a*), "Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium", *Review of Economic Studies*, **46**, 217–225.

HURWICZ, L. (1979*b*), "On Allocations Attainable Through Nash Equilibria", *Journal of Economic Theory*, **21**, 140–165.

HURWICZ, L., MASKIN, E. and POSTLEWAITE, A. (1995), "Feasible Nash Implementation of Social Choice Correspondences when the Designer Does Not Know Endowments or Production Sets", in J. Ledyard (ed.), *The Economics of Informational Decentralization: Complexity, Efficiency and Stability* (Dordrecht: Kluwer Academic Publishers).

HURWICZ, L. and SCHMEIDLER, D. (1978), "Outcome Functions which Guarantee the Existence and Pareto Optimality of Nash Equilibria", *Econometrica*, **46**, 144–174.

JACKSON, M. (1992), "Implementation in Undominated Strategies: A Look at Bounded Mechanisms", *Review of Economic Studies*, **59**, 757–775.

MASKIN, E. (1985), "The Theory of Nash Equilibrium: A Survey", in L. Hurwicz, D. Schmeidler, and H. Sonnenschein, *Social Goals and Social Organization* (Cambridge: Cambridge University Press), 173–204.

MOORE, J. (1993), "Implementation, Contracts, and Renegotiation in Environments with Complete Information", in J.-J. Laffont (ed.), *Advances in Economic Theory* (Cambridge: Cambridge University Press), 182–282.

MOORE, J. and REPULLO, R. (1990), "Nash Implementation: A Full Characterization", *Econometrica*, **58**, 1083–1099.

MULLER, E. and SATTERTHWAITE, M. (1977), "The Equivalence of Strong Positive Association and Strategy Proofness", *Journal of Economic Theory*, **14**, 412–418.

OSBORNE, M. and RUBINSTEIN, A. (1994) *A Course in Game Theory* (Cambridge: MIT Press).

PALFREY, T. (1993), "Implementation in Bayesian Equilibrium: The Multiple Equilibrium Problem in Mechanism Design", in J.-J. Laffont (ed.), *Advances in Economic Theory* (Cambridge: Cambridge University Press).

PALFREY, T. (1998), "Implementation Theory", forthcoming in R. Aumann and S. Hart (eds.), *Handbook of Game Theory, Vol. 3* (Amsterdam: North-Holland).

REPULLO, R. (1987), "A Simple Proof of Maskin's Theorem on Nash Implementation", *Social Choice and Welfare*, **4**, 39–41.

SATTERTHWAITE, M. (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. (1980), "Walrasian Analysis via Strategic Outcome Functions", *Econometrica*, **48**, 1585–1595.