Econometrica, Vol. 59, No. 4 (July, 1991), 997–1021

VIRTUAL IMPLEMENTATION IN NASH EQUILIBRIUM

By Dilip Abreu and Arunava Sen[1]

Reformulate the classical implementation problem à la Maskin (1977) as follows. Think of a social choice correspondence as a mapping from preference profiles to *lotteries* over some finite set of alternatives. Say that a social choice function f is *virtually implementable* in Nash equilibrium if for all $\varepsilon > 0$ there exists a game form G such that for all preference profiles θ, G has a unique equilibrium outcome $x(\theta)$, and $x(\theta)$ is ε-close to $f(\theta)$. (This definition may be directly extended to social choice *correspondences*.) Then (under mild domain restrictions) the following result is true: In societies with at least three individuals *all* social choice correspondences are virtually implementable in Nash equilibrium. This proposition should be contrasted with Maskin's (1977) classic characterization, according to which the nontrivial requirement of *monotonicity* is a necessary condition for exact implementation in Nash equilibrium.

The *two-person* case needs to be considered separately. We provide a complete characterization of virtually implementable two-person social choice functions. While not all two-person social choice functions are virtually implementable, our necessary and sufficient condition is simple. This contrasts with the rather complex necessary and sufficient conditions for exact implementation.

We show how our results can be extended to implementation in *strict* Nash equilibrium and *coalition-proof* Nash equilibrium, to social choice correspondences which map from *cardinal* preference profiles to lotteries, and to environments with a *continuum* of pure alternatives.

Keywords: Virtual implementation, social choice functions, lotteries, Nash equilibrium, coalition-proof, $n \geqslant 2$ players.

1. INTRODUCTION

A social choice correspondence is a mapping which associates with every profile of individual preferences over some set of available alternatives, a subset of socially desirable alternatives. The terminology is derived from the interpretation of the correspondence as representing the preferences of a social planner. More generally we may think of a social choice correspondence as representing the preferences of a "principal" as a function of the preferences of his "agents." The theory of implementation is concerned with the design of *game forms* or *mechanisms* whose equilibrium outcomes are consistent with the social choice correspondence for every profile of individual preferences. It is motivated by the assumption that the planner is uninformed about what individual preferences are. Each individual however knows, or, in a Bayesian

[1] This project was conceived and the basic ideas developed when both authors were visiting the Department of Economics at the Universidad Autónoma de Barcelona in Summer, 1987. We would like to thank the department, and especially Salvador Barbera, for their hospitality. We would also like to thank Faruk Gul, Andreu Mas-Colell, Eric Maskin, David Pearce, Motty Perry, Ariel Rubinstein, Hugo Sonnenschein, Michael Whinston, and a referee for helpful comments. This paper was revised while Abreu was a National Fellow at the Hoover Institution. We thank the latter, the National Science Foundation, and the Sloan Foundation for their support.

After this paper was submitted to *Econometrica* we became aware of prior work by Matsushima (1988) which takes the same basic approach. Though our point of departure is similar we focus on different issues. See the introduction for a comparison. Nevertheless, in terms of the preceding literature, the similarities between our work are more important than the differences, and the present paper should be read as a companion to his.

incomplete information setting, has priors about, the entire preference profile. This abstract formulation embraces a variety of specific applications. These include problems of optimal taxation, choice of public projects, procurement schemes between a buyer and many sellers, optimal auctions, and so on. Another important class of applications concerns the design of constitutions. Indeed a mechanism may be viewed as a generalized voting procedure, implementation theory being concerned with the performance of alternate procedures, in the context of strategic agent behavior. The (extreme) asymmetry assumed between the information of the principal and that of the agents is perhaps most natural in this setting: the mechanism may be in place before individual preferences are actually realized. Recent contributions to the theory of incomplete contracts (see, for instance, Hart and Moore (1988)) provide novel applications in this mode. When some future contingencies are noncontractible (for example, they may not be verifiable) a contract may specify a mechanism involving messages which are verifiable. Such mechanisms may be structured to induce revelation of contingencies which arise ex-post but could not be directly contracted upon ex-ante. This paper belongs to the literature on implementation with *complete information*. For more by way of general background see, for instance, Gibbard (1973), Maskin (1985), and Moulin (1983).

A first issue is what is meant by the requirement that the equilibrium outcomes of a game form be *consistent* with the social choice correspondence in question. Following Maskin (1977) the notion of consistency we use is stringent: for every preference profile the set of equilibrium outcomes must *coincide* with the outcomes chosen by the social choice correspondence. What is critical about this requirement is that for every profile *all* equilibrium outcomes must be elements of the set of socially desirable alternatives.

The problem just described is well posed only after the specification of a solution concept. A large variety have been considered in the literature: dominant strategy equilibrium (Gibbard (1973), Satterthwaite (1975)), Nash equilibrium (Maskin (1977)), sophisticated equilibrium (Farquharson (1957/1969), Moulin (1979)), subgame perfect equilibrium (Moore and Repullo (1988), Abreu and Sen (1990)), and "undominated Nash equilibrium" (Palfrey and Srivastava (1991)). Our results are for Nash equilibrium, though they have implications for implementation using various refinements of this classical concept. We comment later on this point, and the relationship between the present paper and those cited earlier.

The definitive work on Nash implementation is Maskin (1977). He showed that a condition on social choice correspondences called *monotonicity* is necessary for Nash implementation. Furthermore, in societies with at least three individuals any social choice correspondence which satisfies monotonicity and the weak requirement of *no veto power* is Nash implementable.

Unfortunately monotonicity is not a trivial condition. A result due to Muller and Satterthwaite (1977) and Roberts (1979) (see also Dasgupta, Hammond, and Maskin (1979)) makes this point rather forcefully. The result is: "Let f be a social choice function which satisfies citizen sovereignty (that is, every outcome is chosen by the social choice function for some preference profile), whose domain consists of all strict preferences and whose range contains at least three elements. Then f is monotonic if and only if it is dictatorial." See also Saijo (1987) who shows that if the domain of f includes weak orderings and, in particular, contains a profile in which all agents are indifferent amongst all alternatives, then f is monotonic if and only if f is constant. While admitting correspondences and introducing domain restrictions on preferences allows one to escape the full nihilism of the preceding result (for example, the Pareto correspondence, and in economic environments, the correspondence which selects core allocations, is monotonic) a variety of interesting functions and correspondences are not implementable. Examples in "voting" environments include both scoring correspondences such as the plurality rule and the Borda count, and majority rule type correspondences such as the Copeland rule (see Moulin (1983) for details). In "economic" environments the Pareto efficient and egalitarian equivalent correspondence is, for instance, not monotonic. Other interesting economic examples are provided by Moore and Repullo (1988). The recent literature on implementation using *refinements* of Nash equilibrium (Moore and Repullo (1988), Palfrey and Srivastava (1991), Abreu and Sen (1990)) has had as its primary objective enlarging the class of implementable social choice correspondences. This paper achieves the same goal via a quite different route.

The starting point of our work is a restatement of the implementation problem. In our formulation, the set of social alternatives is the set of *lotteries* over some finite set of social states. The social choice correspondence maps to this space of lotteries, which is also the space on which individual preferences are defined. Thus, the planner is able to randomize, and his ability to do so plays an important role in the game forms we present. In this setting there is a natural way to talk about the distance between social alternatives, each social alternative, or lottery, being associated with a point in the simplex of appropriate dimension. Analogously, there is a natural notion of distance between social choice correspondences: we regard two correspondences h and f as being within ε of one another if for all preference profiles θ there exists a bijection between $f(\theta)$ and $h(\theta)$ such that for each element x of $f(\theta)$ the corresponding element x' of $h(\theta)$ is ε-close to x. This definition of closeness is attractive in that for small ε, desired social outcomes are nearly the same under h as under f. We define a social choice correspondence f to be *virtually implementable* in Nash equilibrium if for *all* $\varepsilon > 0$ there exists a social choice correspondence h which is Nash implementable and ε-close to f. Our principal departure from the standard formulation is to require of social choice correspondences only that they be virtually (as opposed to exactly) implementable (in Nash equilibrium).[2] Matsushima (1988) proposes the same basic formulation. Our work was

[2] "Social decision schemes" (Gibbard (1977)), which depend on chance, have a long history and date back to the early work of Zeckhauser (1969) and Gibbard (1977). These papers focus on nonmanipulability, or implementation in dominant strategies, in contrast to the Nash perspective of the present paper. We stress, in addition, our decisive break with the traditional preoccupation with *exact* implementation.

independent. He used the term *ε-implementation*. We prefer virtual implementation in that it is suggestive and avoids confusion with *ε-equilibrium*, a quite distinct idea. In recent papers Matsushima has graciously adopted the "virtual" terminology.

Is virtual implementation good enough? In our view it emphatically is, in that desired social goals can be attained with arbitrarily high probability. For simplicity, think of a social choice *function* which maps to *degenerate* lotteries. If it is virtually implementable, then for any $\varepsilon > 0$ there exists a game form, the Nash equilibria of which yield, for any preference profile, the desired social state with probability at least $(1 - \varepsilon)$. Notice that players' choices are fully optimal. That is, our definition of Nash equilibrium is the standard one—no notions of ε-rationality are involved. We conclude that virtual implementation is, in terms of its substantive consequences, indistinguishable from the traditional requirement of (exact) implementation.

What is the class of virtually implementable social choice correspondences? Presumptions of "continuity" would suggest that this class is identical to the class of social choice correspondences which can be implemented exactly. This intuition is quite misleading. Indeed we show (under a mild domain restriction) that in societies with at least three individuals, *all* social choice correspondences are virtually implementable. This is true for social choice correspondences which map from ordinal preference profiles to lotteries, as well as those which map from cardinal preference profiles to lotteries. Thus, our reformulation dramatically broadens the scope of Nash implementation. Again, it is instructive to think about social choice *functions* which map to degenerate lotteries, and compare our permissive result to the one presented earlier for monotonic social choice functions. There is a "discontinuous" jump from only trivial functions being exactly implementable (given citizen sovereignty, at least three alternatives and universal domain of strict preferences) to all functions being virtually so (when there are at least three players).

What is the source of this discontinuity? Since monotonicity is, by Maskin's (1977) characterization, a necessary condition for exact implementation, it immediately follows that any neighborhood (in terms of the metric defined earlier) of an arbitrary social choice correspondence f must contain a monotonic social choice correspondence h. If preferences over lotteries were completely unrestricted, there is no reason why this must be so. Of course, it is natural to require that these preferences obey certain axioms: we assume that individual preferences over lotteries are *monotone* in the sense that any shift of probability weight from a less preferred to a more preferred pure alternative yields a lottery which is preferred. The axiom that preferences are monotone is, of course, much weaker than the independence axiom, and is implied by the expected utility hypothesis. Given this assumption, any social choice correspondence (defined on strict preference profiles) whose range lies in the *interior* of the simplex, must be monotonic; see the discussion following the proof of Theorem 1. (On the other hand, such a social choice correspondence will *not* satisfy *no veto power*, the additional assumption used by Maskin in his sufficiency proof.) Thus, our results exploit the freedom permitted by virtual, as opposed to exact, implementation, in conjunction with the domain restrictions implied by the weak axiom that preferences over lotteries are monotone.[3]

Nash equilibrium precludes profitable *individual* deviations but is silent on the issue of profitable *coalitional* deviations. It is not obvious how best to address this question. Bernheim, Peleg, and Whinston (1987) argue that the concept of *strong* Nash equilibrium (Aumann (1959)) is in fact too strong and offer instead their definition of *coalition-proof Nash equilibrium*. An attractive feature of our canonical game form is that it virtually implements any social choice function (when no pair of players has the same ordering over all pure alternatives) in coalitional-proof Nash equilibrium also. This is encouraging in that the coalition-proof requirement is frequently hard to satisfy. An analogous result is established for *exact* implementation by Bernheim and Whinston (1987).

The general theory of exact implementation in Nash equilibrium as developed by Maskin, is heavily dependent on *multi-valued* social choice rules or correspondences (recall the result by Muller and Satterthwaite (1977), cited earlier). The associated implementing game forms therefore have *multiple* Nash equilibria for certain preference profiles. This is a source of unease in the context of the view that Nash equilibrium as a solution concept loses much of its plausibility in games with multiple equilibria.[4] This issue has a nice resolution in our work. Under the interpretation that multiple values express a planner's indifference or "neutrality," a social choice correspondence may be formulated naturally in our framework as a social choice *function* which for some profiles maps to *nondegenerate* lotteries. Furthermore, under a weak additional assumption, the canonical implementing game form which we construct has a *unique* Nash equilibrium for every preference profile when the virtually implementable social choice correspondence is actually a *function*.

The results described above apply when there are many (at least three) players. Section 4 addresses the problem of virtual implementation when there are two-players. This special case is actually rather central: a wide range of economic phenomena are essentially bilateral. From the point of view of Nash implementation it has been known since Maskin's work that the two-player case is very different from the many-player case. Maskin-type mechanisms depend critically on being able to view a conflicting announcement by an individual player as a deviation from a majority announcement. This is, of course, not possible when there are only two players. Recently Dutta and Sen (1991) and Moore and Repullo (1990) have provided a complete characterization of exact Nash implementation in the two-person case. We provide a similar result for virtual implementation. Unlike the many-player case, not all two-person social

[3] We are grateful to Ariel Rubinstein who pointed out that our constructions only used the assumption that preferences over lotteries were monotone. We did not use the independence axiom and expected utility. Thanks also to Andreu Mas-Colell and Mike Whinston whose comments clarified the connection between Maskin's results and our own.

[4] We are grateful to Hugo Sonnenschein for emphasizing this issue.

choice functions are virtually implementable. But as in the many-player case the conditions for virtual implementation are much weaker than those for exact implementation, which are complex and cumbersome. We are aware of no results for two person implementation in refinements of Nash equilibrium.

Our permissive results for the three or more player case mirror those obtained for implementation using *refinements* of Nash equilibrium. They go beyond the characterizations for subgame perfect implementation of Moore and Repullo (1988), and following on their work, Abreu and Sen (1990). The (tighter) necessary conditions of the latter, while weak in certain environments, are certainly nontrivial. In terms of the class of "implementable" social choice correspondences our results are closest to those of Palfrey and Srivastava (1991) for implementation in "undominated Nash equilibrium." A drawback of the game forms which Palfrey and Srivastava construct is that for certain strategy combinations of other players, the remaining player may be *unable* to play an undominated best response, because any best response x_n is dominated by another best response x_{n-1} (in his infinite strategy set). We feel that such a method of eliminating equilibria ("tail chasing") is unattractive in that the game form is fundamentally incompatible with the solution concept being advocated. This is akin to insisting that an individual fully optimize and then confronting him with a noncompact and/or discontinuous choice problem. See Jackson (1989) for, among other things, an elaboration of this point of view.[5]

Our analysis may be viewed as being preferable to the contributions discussed above in the following sense also. The perfection approach relies heavily on players behaving with the *exact* degree of "perfection" required under the solution concept in question, so that game forms which fully implement a social choice correspondence for a particular refinement, need not do so for another. An attractive feature of our work, in this respect, is that the mechanism may be adapted so that in all equilibria a player's equilibrium strategy is a *strict* best response. Such Nash equilibria are "perfect" by all the definitions suggested to date, including the various definitions of "stability" proposed by Kohlberg and Mertens (1986). Thus our game forms work for naive Nash players and equally for very "perfect" ones.

We remark that our constructions are rather simple, and in this dimension compare favorably with the quite involved canonical mechanisms of Moore and Repullo (1988), and the even more elaborate general constructions of Palfrey and Srivastava (1991).

The overlap of our work with Matsushima (1988) is primarily in part of Section 3. He does not discuss issues of perfection or coalitional deviations.

[5] Palfrey and Srivastava, however, show that their general construction can be greatly simplified and "tail-chasing" avoided if there is a "holocaust" outcome, that is, an outcome which is "bad" for all players simultaneously. This resolution is problematic in that it violates the *crudest* notions of *renegotiation-proofness* that one might impose. While most of the literature on implementation does not address the renegotiation issue directly, constructions which rely crucially on a uniformly worst outcome would appear to be particularly fragile. See Maskin and Moore (1986) for the first general treatment of implementation with renegotiation.

Neither does he address the two-player case, cardinal social choice correspondences or a continuum of alternatives. On the other hand, his mechanisms are informationally "efficient" in that each player announces only his own preferences and that of his two "neighbors." In this respect his work is similar to that of Saijo (1988) who develops such sparse mechanisms for exact Nash implementation. Saijo credits Hurwicz (1979) and Walker (1981) with first using "cyclic" announcements of the sort that he does.

This paper is organized as follows. After developing some basic notation in Section 2, we present our analysis of the many-player case in Section 3. Subsections 3.2 and 3.3 discuss how our results can be extended to strict Nash equilibrium and coalition-proof Nash equilibrium respectively. Section 4 covers the two-player case. Sections 3 and 4 assume that the social choice correspondence is ordinal. Section 5 provides analogous results for cardinal social choice correspondences. The preceding sections assume that the set of alternatives is finite. Section 6 extends the analysis to the case where this set is an arbitrary subset of a separable metric space. Section 7 concludes.

2. PRELIMINARIES

Let A denote the set of social states which for convenience we take to be a finite, K-element set. Let $\mathscr{L}$ be the set of lotteries over elements of A. We identify $\mathscr{L}$ with the $(K-1)$ dimensional simplex Δ^{K-1}, $x \equiv (x_1, \ldots, x_K) \in \mathscr{L}$ representing the lottery in which the social state a_k occurs with probability x_k. Let $J = \{1, \ldots, N\}$ denote the set of individuals. The set of admissible[6] preference profiles over A and $\mathscr{L}$ are denoted Θ and Γ, respectively. For every $\theta \in \Theta$ and $j \in J$, $R^j(\theta)$ represents individual j's preference ordering over A. The corresponding strict preference and indifference relations are $P^j(\theta)$ and $I^j(\theta)$. For any preference profile γ over elements of $\mathscr{L}$ and $j \in J$, the relations $\tilde{R}^j(\gamma)$, $\tilde{P}^j(\gamma)$, and $\tilde{I}^j(\gamma)$ are similarly defined.

The preference profile $\gamma \in \Gamma$ over elements of $\mathscr{L}$ is said to be consistent with the profile $\theta \in \Theta$ over elements of A if for all $j \in J$ the restriction of $\tilde{R}^j(\gamma)$ to elements of A is $R^j(\theta)$. For given $\theta \in \Theta$ the set of all such consistent profiles is denoted $\Sigma(\theta)$. Conversely, for any $\gamma \in \Gamma$ there exists a unique element of Θ with which Γ is consistent. Denote this element $\sigma(\gamma)$. In the preceding definitions and below, we freely identify elements of A with a degenerate lottery which yields that element with probability 1.

3. ORDINAL INFORMATION

This section considers social choice correspondences (SCC's) which map from the set of preference profiles over A. An alternative is to regard the domain as being the set of preference profiles over $\mathscr{L}$, the space of lotteries. This is explored in the next section. We start with the ordinal model because this is the

[6] Note that Θ and Γ are not necessarily the set of all possible profiles and in general may embody domain restrictions.

case most frequently considered in the literature, even in settings such as Gibbard (1977) in which the SCC maps to lotteries. Apart from tradition, the ordinal model has special claims to our attention in that the informational requirements of the cardinal setting may be judged to be too demanding. We proceed to details.

First some definitions. An (ordinal) SCC f associates a nonempty set $f(\theta) \subset \mathscr{A}$ with every $\theta \in \Theta$. A game form G is an $(N+1)$-tuple $(S^1, \ldots, S^N; g)$ where S^j is the strategy set (or message space) of individual j and g is the outcome function $g: S^1 \times \ldots \times S^N \rightarrow \mathscr{A}$. A game form G together with a preference profile $\gamma \in \Gamma$ on lotteries defines the game (G, γ).

While the standard definition of a game entails players' *cardinal* preferences being common knowledge, the planner's objectives as formulated in this section (that is, as a mapping from *ordinal* preferences to outcomes) make most sense when players have access only to *ordinal* information about other players. The implementation problem is then to design game forms which players can analyze completely using ordinal information alone. The next definition formalizes this requirement. Let $S = S^1 \times \ldots \times S^N$.

DEFINITION: The game form G is *ordinal* if for all $\theta \in \Theta$, $s \in S$ and γ, $\delta \in \Sigma(\theta)$, s is a Nash equilibrium of (G, γ) if and only if it is a Nash equilibrium of (G, δ).

For the remainder of this section, we confine attention to ordinal game forms G.

We denote by $NE(G, \theta) = \{g(s) | s$ is a Nash equilibrium of (G, γ) for all $\gamma \in \Sigma(\theta)\}$ the set of Nash equilibria of the ordinal game form G under the ordinal preference profile $\theta \in \Theta$. The qualification "ordinal" when applying to game forms will typically be suppressed for the remainder of this section.

To simplify the exposition we assume that individual preferences over alternatives in A are *strict*. This is a domain restriction on Θ. Its simplifying role is explained briefly following the proof of Theorem 1. We will also assume that preferences over lotteries are *monotone* in the sense that shifts in probability mass from less preferred to strictly preferred alternatives in A yield a lottery which is strictly preferred. More precisely, for any $\gamma \in \Gamma$ let $\theta \equiv \sigma(\gamma)$. For any $j \in J$ let p: $M \rightarrow M$ be a permutation of $M \equiv \{1, \ldots, K\}$ such that $a_{p(k)} P^j(\theta) a_{p(k+1)}$, for all $k = 1, \ldots, K-1$. Then for any lotteries $x \equiv (x_1, \ldots, x_K) \neq (y_1, \ldots, y_K) \equiv y$ such that $\Sigma_{k-1}^m x_{p(k)} \geqslant \Sigma_{k-1}^m y_{p(k)}$, $m = 1, \ldots, K$, $x P^j(\gamma) y$. Of course, if individual preferences over lotteries are representable by von Neumann-Morgenstern utility functions they will satisfy the rather weak requirement of monotonicity. To avoid trivial qualifications we will also assume that for all preference profiles there exists *some* pair of individuals who differ in their ranking over *some* pair of pure alternatives. These domain restrictions should be understood below. They are made only for convenience; see Abreu and Sen (1987) for a more general treatment which allows indifference and the very strong form of unanimity excluded here.

We remark that there is no requirement that f map to nondegenerate lotteries. Such truly mixed outcomes may be viewed as having little basis in a setting in which a planner's choices depend only on ordinal information. We note, though, that lotteries may be quite natural in situations of, for instance, perfect symmetry of preferences over some set of alternatives, or, as remarked in the introduction, when the planner is indifferent amongst a number of alternatives, which is presumably the motivation for multi-valued social choice rules.

Definition: The social choice correspondence f is (exactly) implementable in Nash equilibrium if there exists a game form G such that $NE(G, \theta) = f(\theta)$ for all $\theta \in \Theta$.

Let $\rho(x, y)$ denote the Euclidean distance between any pair of lotteries.

Definition: The social choice correspondences f and h are ε-close if for all $\theta \in \Theta$ there exists a bijection τ_θ: $f(\theta) \rightarrow h(\theta)$ such that $\rho(x, \tau_\theta(x)) \leqslant \varepsilon$ for all $x \in f(\theta)$.

Definition: The social choice correspondence f is virtually implementable in Nash equilibrium if for all $\varepsilon > 0$ there exists a social choice correspondence h which is (exactly) implementable in Nash equilibrium and ε-close to f.

3.1. *The Theorem*

The proof of Theorem 1 bears a family resemblance to that of Maskin (1977), elements of which have since become quite standard in the (Nash and its refinements) implementation literature. There are two basic ideas. The first is that players announce preference *profiles*, a profile announced by any $N-1$ of them being taken to be a "reference" profile. The second is to trigger an "unwinnable" competition to be dictator when nonunanimous announcements are made.

Theorem 1: *Let $N \geqslant 3$. Then any social choice correspondence f is virtually implementable in Nash equilibrium.*

Proof: For all $\theta, \varphi \in \Theta$ such that $\theta \neq \varphi$ define $j(\theta, \varphi) \in J$, and $a(\theta, \varphi), b(\theta, \varphi) \in A$ such that $a(\theta, \varphi) P^j(\theta) b(\theta, \varphi)$ and $b(\theta, \varphi) P^j(\varphi) a(\theta, \varphi)$, where $j = j(\theta, \varphi)$. Abusing notation, we will not distinguish between a pure alternative and the lottery which yields that alternative with probability 1. Denote by $\bar{x}$ the completely mixed lottery which gives equal weight to all alternatives in A. We now describe the canonical game form G used to virtually implement f.

Each player simultaneously announces a triplet $(\theta^i, x_i, n_i) \in (\Theta \times \Delta^{k-1} \times Z_+)$ consisting of a preference profile θ^i, a lottery x_i and a nonnegative integer n_i.

Consider arbitrary $\varepsilon > 0$, and let $\eta = \min\{(\varepsilon/2), \frac{1}{2}\}$. The outcome function (corresponding to ε) is defined below.

If $(N-1)$ players announce the same θ and $x \in f(\theta)$ the outcome is

$$L(x,\theta) \equiv (1-2\eta)x + 2\eta\bar{x}$$

unless the remaining agent i announces $\varphi \neq \theta$ and $i = j(\theta,\varphi)$. In this case the outcome is

$$L(x,\theta,\varphi) \equiv (1-2\eta)x + 2\eta\bar{x} + \frac{\eta}{K}[b(\theta,\varphi) - a(\theta,\varphi)].$$

In all other cases the outcome is

$$L^h(\theta^h) \equiv (1-2\eta)a_{p(1)} + \eta\bar{x} + \frac{\eta}{K}(a_{p(1)} - a_{p(K)}) + \eta\sum_k \alpha_k a_{p(k)}$$

where

$$\alpha_k = \frac{K+1-k}{1+2+\ldots+K},$$

p: $M \to M$ is a permutation of $M = \{1,\ldots,K\}$ such that $a_{p(k)}P^h(\theta^h)a_{p(k+1)}$, for all $k = 1,\ldots,K-1$, and $h = \min\{i | n_i \geqslant n_j \text{ for all } j \subset J\}$. That is, the game form chooses the most preferred permutation function (according to his announced preferences and subject to the form of the lottery above) of the player (with the lowest index) who announces the highest integer.

Notice that the game form is well defined in the sense that for all strategy profiles alternatives are assigned nonnegative probabilities which sum to one. Nonnegative probabilities are guaranteed by the term $\eta\bar{x}$, and this is precisely the point at which the notion of *virtual* implementation is being used.

We now argue that the game form G described above implements a social choice correspondence which is ε-close to f. Let the true preference profile be ψ, and consider $x \in f(\psi)$. Then all players announcing $(\psi, x, 0)$ is a Nash equilibrium of G, since a deviating player either does not affect the outcome or obtains the lottery $L(x,\psi,\theta)$ which is, given monotone preferences, dominated for him (given that ψ is true) by the nondeviation outcome $L(x,\psi)$. Hence for any $x \in f(\psi)$ there exists $z = (1-2\eta)x + 2\eta\bar{x} \in NE(G,\psi)$ such that

$$\rho(x,z) \leqslant 2\eta\left(\frac{K-1}{K}\right)^{1/2} \leqslant \varepsilon,$$

where ρ denotes Euclidean distance. To complete the proof we show that in all equilibria players must announce ψ and the same $x \in f(\psi)$. We need to consider three kinds of candidate equilibria.

Case 1: All players announce the same $\theta \neq \psi$ and $x \in f(\theta)$ (and some nonnegative integers). These announcements are not consistent with equilibrium in that player $j(\theta,\psi)$ may profitably deviate by announcing $(\psi, x, 0)$, thereby obtaining the lottery $L(x,\theta,\psi)$ which is preferred to $L(x,\theta)$.

Case 2: $(N-1)$ players announce the same $\theta, x \in f(\theta)$ and the remaining player i announces $(\phi, y) \neq (\theta, x)$. But any $h \neq i$ can now deviate by announcing

ψ, $z \neq x, y$, and $\tilde{n}_h = \Sigma_j n_j + 1$, thereby obtaining $L^h(\psi)$, his strictly most preferred outcome in the range of the outcome function.

Case 3: A candidate equilibrium with outcome of the form $L^h(\psi)$, with some player h winning the integer game. But by our domain restriction there exists at least one player $j \neq h$ who would strictly prefer $L^j(\psi)$ to $L^h(\psi)$ and can obtain the former by announcing $\tilde{n}_j = n_h + 1$.

It is now clear that G is an ordinal game form. Since ε was arbitrary the proof is therefore complete. *Q.E.D.*

The assumption of strict preferences is the only substantial restriction we make. It guarantees that if $\theta \neq \varphi$ there exist a player and a pair of alternatives over which his preferences *strictly* switch. When indifference is permitted it is possible that $\theta \neq \varphi$ and $aR^j(\theta)b$ implies $aR^j(\varphi)b$ for all $a, b \in A$ and $j \in J$. The only change involves strict preference being weakened to indifference. In such a situation our proof does not work. Indeed a necessary condition for virtual implementation is that f satisfy the *reversal property* according to which $f(\varphi) \subseteq f(\theta)$ when θ and φ are as described above. See Abreu and Sen (1987) for further details and also a discussion of the extremely trivial unanimity condition imposed here.

We indicated in the introduction that any social choice correspondence h which maps to the *interior* of the simplex, must be monotonic, *given the domain restrictions implied by our assumption that individuals have monotone preferences over lotteries.*[7] Any neighborhood of an arbitrary social choice correspondence f contains such a social choice correspondence h. This is the essential connection between Maskin's results and our own. Note, however, that Maskin's sufficiency theorem for exact implementation in Nash equilibrium *cannot* be invoked to establish an analogous result for virtual implementation in Nash equilibrium. The reason is that his proof uses the *no veto power* condition, which cannot be satisfied by a social choice correspondence h which maps only to completely mixed lotteries. In the setting of the present paper, no veto power is a much stronger condition than needed. It is replaced by the trivial domain restriction according to which not all individuals have the same ranking over *all* alternatives.

3.2. *Perfect Equilibrium*

The game form may be easily adapted so that all equilibria are *strict*, in the sense that each player's equilibrium strategy is a strict best response to the strategies of other players. Strict Nash equilibria are, of course, robust to all

[7] To see this consider $\theta, \varphi, \theta \neq \varphi$ and $x \in h(\theta)$. Since $\theta \neq \varphi$ there exists an individual j and alternatives a, b such that $aP^j(\theta)b$ and $bP^j(\varphi)a$. Since x is completely mixed $y = x + \delta(b - a)$ is indeed a lottery ($y \geqslant 0$) for small enough δ. By the assumption of monotone preferences over lotteries, $x\bar{P}^j(\theta)y$ and $y\bar{P}^j(\varphi)x$ and h satisfies the monotonicity condition. Recall that a social choice correspondence h is monotonic if for all θ, $\varphi \in \Theta$ and $x \in h(\theta) \setminus h(\varphi)$ there exists an individual j and a lottery (outcome) y such that $x\bar{R}^j(\theta)y$ and $y\bar{P}^j(\varphi)x$.

manner of "trembles," and are therefore "perfect" in a very strong sense. Our basic game form thus implements in Nash equilibrium, as well as in any of a range of refinements of Nash equilibrium.

The proof of Theorem 1 establishes that all equilibria of (G, ψ) involve unanimous announcements of ψ and $x \in f(\psi)$. We may alter the mechanism so that when all players announce $(\psi, x, 0)$, the announcement $(\psi, x, 0)$ is a strict best response. To do so we need only punish a deviating player who announces $(\psi, x, n_i \neq 0)$ by increasing the weight of his least preferred alternative and reducing the weight of his best alternative (according to $P^i(\psi)$) by a corresponding amount. If more than one player announces $n_i \neq 0$, the player to be punished may be picked at random. We leave the details, which are straightforward, to the reader.

3.3. *Coalitional-Proof Nash Equilibrium*

Nash equilibrium guarantees that behavior is self-enforcing in the sense that no individual has an incentive to deviate from proposed behavior given that all the remaining individuals conform with the proposed equilibrium. It leaves open the possibility that coalitions of individuals might find it profitable to deviate holding fixed the behavior of the complementary coalition. It is not obvious what coalitional deviations should be viewed as being admissible. The concept of *strong Nash equilibrium* (Aumann (1959)) admits *any* coalitional deviation which makes all members of the coalition strictly better off than in the proposed equilibrium (holding fixed the behavior of all players outside the coalition). Bernheim, Peleg, and Whinston (1987) argue persuasively that permitting all such coalitional deviations is far too permissive. They propose that allowable coalitional deviations must themselves be self-enforcing in the sense of being immune to self-enforcing deviations by subcoalitions of the original deviating coalition. They formalize these requirements in a concept called *coalition-proof Nash equilibrium*; the reader is referred to their paper for further details and motivation.

A coalition-proof Nash equilibrium (CPNE) is recursively defined. In a two-player game a CPNE is any (weakly) Pareto efficient Nash equilibrium.[8] In a three-player game a Nash equilibrium is coalition-proof if it is a Pareto efficient Nash equilibrium and if no pair of players can deviate profitably to a CPNE of the two-player game obtained by fixing the action of the remaining player at its equilibrium level. In an N-player game a Pareto efficient Nash equilibrium is coalition-proof if no strict subset of players S can profitably deviate to a CPNE of the $|S|$ player game obtained by holding fixed the actions of the remaining players at their equilibrium level.

The definitions of "implementation in CPNE" and "virtual implementation in CPNE" are obvious analogs of the corresponding definitions for Nash equilibrium. They are left to the reader. It turns out that the game form used in

[8] That is a Nash equilibrium which is not Pareto dominated by another Nash equilibrium.

the proof of Theorem 1 virtually implements any social choice *function* in CPNE when we make the mild domain restriction that no two players have identical strong orderings over all pure alternatives. Let $\Theta^* = \{\theta \mid$ for all $i, j \in J$, $i \neq j$, $P^i(\theta) \neq P^j(\theta)\}$.

THEOREM 2: *Let $N \geqslant 3$ and $\Theta \subseteq \Theta^*$. Then any social choice function $f\colon \Theta \to \mathscr{L}$ is virtually implementable in coalition-proof Nash equilibrium.*

PROOF: Consider the game form described in the proof of Theorem 1, and let the true profile be ψ. As argued in the earlier proof all Nash equilibria involve unanimous announcements $(\psi, f(\psi))$ (and possibly varying integers n_i). Thus for all $\psi \in \Theta$ the game form yields a unique Nash equilibrium outcome $(1-2\eta)f(\psi) + 2\eta\bar{x}$. Since any CPNE must be a Nash equilibrium, to complete the proof it suffices to exhibit a CPNE. We show that all players announcing $(\psi, f(\psi), 0)$ is indeed a CPNE. First note that the latter is trivially a Pareto efficient Nash equilibrium. Since it is a Nash equilibrium, single player deviations are not profitable either. Let us now consider deviations by coalitions S where $N-1 \geqslant |S| \geqslant 2$. Such a deviation will yield a lottery of the form $L(f(\theta), \theta)$ (we abstract from the case of a trivial deviation which yields $L(f(\psi), \psi)$) or $L(f(\theta), \theta, \psi)$ or $L(f(\psi), \psi, \theta)$ for some $\theta \neq \psi$, or $L^h(\theta^h)$ for some h and θ^h. In all these cases any member of the coalition $i \in S$ (we also require $i \neq h$ in the last case) can deviate by announcing ψ, some z and $\bar{n}^i$ where $\bar{n}^i$ exceeds all other integer announcements, thereby obtaining his strictly most preferred lottery (in the range of the outcome function) $L^i(\psi)$. Thus there exists no coalitional deviation which is a CPNE of the "remainder game," and the proof is complete.[9] *Q.E.D.*

Bernheim and Whinston (1987) provide an analogous theorem for exact implementation. Their result requires that f be monotonic and Pareto efficient. These assumptions on f may be dispensed with for virtual implementation. In addition they make the domain restriction that no pair of players has the same top-ranked alternative, whereas we only need to assume that players differ in their ranking over some pair of alternatives. Their result is for social choice correspondences, ours for social choice functions. Since we allow arbitrary functions, this does not seem an important difference. The reader may confirm that our result may be extended to correspondences which are "neutral" in the sense that for any $\theta \in \Theta$ no two distinct elements in $f(\theta)$ are Pareto comparable in terms of the preference profile x. Neutrality in this sense is, of course, weaker than Pareto efficiency. Given the variety of principal-agent problems implementation theory embraces, the assumption of Pareto efficiency (relative only to *agent* preferences) is not innocuous.

[9] The reader may have noticed that we could have players simply announce $(\theta^i, n_i) \in \Theta \times Z_+$ instead of $(\theta^i, x_i, n_i) \in \Theta \times \mathscr{L} \times Z_+$. This simplification is possible since f is assumed here to be a function. We thought it best to carry a little extra baggage, rather than to define a new game form.

4. TWO PLAYERS

This section provides a complete characterization of *two*-person social choice functions which are virtually implementable in Nash equilibrium. A separate treatment is necessary because the constructions of the previous sections depend in an essential way on the assumption that the number of players is at least three. In such an environment if $(N-1)$ players announce the same profile θ, then θ can be taken to be a "reference" profile, and the mechanism can treat a conflicting announcement by the remaining player as a deviation from the reference profile θ. In a two-person setting, of course, either player is the remaining player, and conflicting announcements are symmetric. The treatment of the two-person problem is therefore rather different and, in fact, somewhat more complicated.

The two-person implementation problem is important. A variety of economic phenomena are basically bilateral, and this is even more true of economic models. Moreover, as Moore and Repullo (1990) have emphasized the two-person case is central in a contractual setting; in the context of contractual incompleteness, optimal contracts may well embody mechanisms which depend on messages sent by symmetrically informed parties.

We shall restrict attention to ordinal social choice correspondences and retain the notation and assumptions (individuals have strict preferences over pure alternatives, monotone preferences over lotteries, and for no profile do all individuals have the same ranking over all alternatives) of the previous section. We will continue to require that the implementing game form be ordinal. Some additional notation is useful. For any $j \in J$, $x \in \mathscr{L}$, and $\theta \in \Theta$. let $W(j, x, \theta) = \{y \in \mathscr{L} \mid x\hat{R}^j(\gamma)y$ for all $\gamma \in \Sigma(\theta)\}$ be the set of lotteries which, given monotone preferences, are weakly dominated by x, for individual j, under preferences θ.

The following nontrivial condition is central to the theory of two-person implementation. It is one of *many* necessary conditions for exact implementation (see Dutta and Sen (1991) and Moore and Repullo (1990)); it is both necessary and sufficient for virtual implementation.

Definition: The (two-person) social choice correspondence f satisfies the (lower contour set) *intersection property* if for all $\theta, \varphi \in \Theta$, and $x, y \in \mathscr{L}$ such that $x \in f(\theta)$ and $y \in f(\varphi)$, $W(1, y, \varphi) \cap W(2, x, \theta) \neq \varnothing$.

For exact two person implementation the necessity of this condition is obvious. Fix the implementing game form G and consider the matrix representation below where player 1 is row and player 2 column:

$$\begin{array}{lcccccccc}
 & s_2(x,\theta) & s_2(y,\varphi) & \cdot & \cdot & \cdot & \cdot & \\
s_1(x,\theta) & x & z & \cdot & \cdot & \cdot & \cdot & \\
s_1(y,\varphi) & z' & y & \cdot & \cdot & \cdot & \cdot & G \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot &
\end{array}$$

Suppose the strategy combination $s(x,\theta)$ is a Nash equilibrium of (G,θ) and $s(y,\varphi)$ an equilibrium of (G,φ). Let g be the outcome function. Then $z = g(s_1(x,\theta), s_2(y,\varphi)) \in W(2,x,\theta)$ since column may deviate to $s_2(y,\varphi)$ when row plays $s_1(x,\theta)$. Conversely, since $s(y,\varphi)$ is an equilibrium, and row may deviate to $s_1(x,\theta)$, it follows that $z \in W(1,y,\varphi)$. By considering a sequence of implementing game forms we show that this property is a necessary condition for virtual implementation also.

Theorem 3: *If the two-person SCC f is virtually implementable in Nash equilibrium, then it satisfies the intersection property.*

Proof: Fix $\theta, \varphi \in \Theta$ and let $x, y \in \mathscr{L}$ be such that $x \in f(\theta)$ and $y \in f(\varphi)$. Since f is virtually implementable there exists a game form G^n and outcomes $x^n, y^n \in \mathscr{L}$ with $\rho(x^n, x) < (1/n)$, $\rho(y^n, y) < (1/n)$, $x^n \in NE(G^n, \theta)$ and $y^n \in NE(G^n, \varphi)$ for all positive integers n. Let $s^n(x^n, \theta)$ and $s^n(y^n, \varphi)$ denote, respectively, the Nash equilibrium profiles under θ and φ such that $g^n(s^n(x^n,\theta)) = x^n$ and $g^n(s^n(y^n,\varphi)) = y^n$. Let $z^n \in \mathscr{L}$ denote the outcome $g^n(s_1^n(x^n,\theta), s_2^n(y^n,\varphi))$. Observe that player 1 can obtain z^n by deviating unilaterally from $s^n(y^n,\varphi)$, and that player 2 can obtain z^n by deviating unilaterally from $s^n(x^n,\theta)$. It follows from the definition of $s^n(x^n,\theta)$ and $s^n(y^n,\varphi)$ and the assumption that G^n is ordinal, that $z^n \in W(1, y^n, \varphi) \cap W(2, x^n, \theta)$. We now prove the result by taking appropriate limits.

Let p, q be permutations of $\{1, \ldots, K\}$ such that $a_{p(k)} P^1(\varphi) a_{p(k+1)}$ and $a_{q(k)} P^2(\theta) a_{q(k+1)}$ for all $k = 1, \ldots, K-1$. Since $z^n \in W(1, y^n, \varphi) \cap W(2, x^n, \theta)$ it follows that $\sum_{k=1}^m y^n_{p(k)} \geqslant \sum_{k=1}^m z^n_{p(k)}$ and $\sum_{k=1}^m x^n_{q(k)} \geqslant \sum_{k=1}^m z^n_{q(k)}$ for all $m = 1, \ldots, K$. Assume w.l.o.g. that z^n converges to some $z \in \mathscr{L}$. Of course, $x^n \to x$ and $y^n \to y$. But then $\sum_{k=1}^m y_{p(k)} \geqslant \sum_{k=1}^m z_{p(k)}$ and $\sum_{k=1}^m x_{q(k)} \geqslant \sum_{k=1}^m z_{q(k)}$. That is, $z \in W(1, y, \varphi) \cap W(2, x, \theta)$, so that $W(1, y, \varphi) \cap W(2, x, \theta) \neq \varnothing$, as required.

Q.E.D.

For any $\theta \in \Theta$, $a \in A$, and $j \in J$, let $\overline{W}(j, a, \theta) = \{b \in A \mid \sim bP^j(\theta)a\}$ be the set of (pure) alternatives which are not strictly better than a, according to $P^j(\theta)$. If x and y are degenerate lotteries which yield the pure alternatives a and b respectively, then it may be directly checked that $W(1, y, \varphi) \cap W(2, x, \theta) \neq \varnothing$ implies that there exists an alternative $c \in A$ such that $bP^1(\varphi)c$ and $aP^2(\theta)c$. That is, $\overline{W}(1, b, \varphi) \cap \overline{W}(2, a, \theta) \neq \varnothing$. Thus we have the following corollary to Theorem 2.

Corollary: *Let f be a two-person SCC which maps only to subsets of degenerate lotteries over A. If f is virtually implementable, then for all $\theta, \varphi \in \Theta$, $a \in f(\theta)$, and $b \in f(\varphi)$, $\overline{W}(1, b, \varphi) \cap \overline{W}(2, a, \theta) \neq \varnothing$.*

Remark: Let f be a selection from a social choice correspondence h. That is $f(\theta) \subseteq h(\theta)$ and $f(\theta) \neq \varnothing$ for all $\theta \in \Theta$. It is obvious that if h satisfies the intersection property, then so does f. Theorem 4 establishes that if a social choice *function* f satisfies the intersection property, then it is virtually implementable. We do not have a sufficiency result for social choice *correspondences.* This is not a drawback given the standard interpretation of a social choice correspondence as a mapping to outcomes which are equally desirable for the mechanism designer, and the fact that the necessary condition for virtual implementation is inherited by subcorrespondences. It is instructive to compare this feature of the intersection property with monotonicity. Selections from a monotonic correspondence are not necessarily monotonic; this in part explains why correspondences have appeared so prominently in the theory of Nash implementation.

THEOREM 4: *Let f be a two-person social choice function which satisfies the intersection property. Then f is virtually implementable in Nash equilibrium.*

PROOF: For any ordered pair of profiles (θ^1, θ^2) let $Q(\theta^1,\theta^2) = \{(a,b) \in A^2 | aP^i(\theta^j)b,\ i \neq j\}$. Notice that $Q(\theta^1,\theta^2) = \varnothing$ implies that the preferences of player 1 under θ^2 are exactly opposed to those of player 2 under θ^1. It follows that $\theta^1 \neq \theta^2$ and $Q(\theta^1,\theta^1) = Q(\theta^2,\theta^2) = \varnothing$, together imply $Q(\theta^1,\theta^2) \neq \varnothing$. $(Q(\theta^1,\theta^1) = Q(\theta^2,\theta^2) = Q(\theta^1,\theta^2) = \varnothing$ together imply $P^1(\theta^1) = P^1(\theta^2)$ and $P^2(\theta^1) = P^2(\theta^2)$. That is, $\theta^1 = \theta^2$.) If $Q(\theta^1,\theta^2) \neq \varnothing$, let $(a(\theta^1,\theta^2), b(\theta^1,\theta^2))$ denote some element of $Q(\theta^1,\theta^2)$.

The proof is constructive. In the canonical game form, each player simultaneously announces a quintuplet $(\theta^i, \alpha_i, a(i), b(i), n_i) \in (\Theta \times \{0,1\} \times A^2 \times Z_+)$ consisting of a preference profile θ^i, nonnegative integers α_i, n_i and alternatives $a(i), b(i)$ drawn from the indicated sets. For arbitrary $\varepsilon > 0$ let

$$\eta = \min\left\{\frac{\varepsilon}{2}, \frac{1}{2}\right\}.$$

The outcome function is defined below.

If $\alpha_1 = \alpha_2 = 0$ and $\theta^1 = \theta^2 = \theta$:

$$L(\theta^1,\theta^2) \equiv \left[\begin{array}{l} (1-2\eta)f(\theta) + 2\eta\bar{x} + \dfrac{\eta}{2K}(a(\theta,\theta) - b(\theta,\theta)) \\ \qquad \text{if } Q(\theta,\theta) \neq \varnothing, \\ (1-2\eta)f(\theta) + 2\eta\bar{x} \qquad \text{otherwise;} \end{array}\right.$$

if $\alpha_1 = \alpha_2 = 0$ and $\theta^1 \neq \theta^2$:

$$L(\theta^1,\theta^2) \equiv \left[\begin{array}{l} (1-2\eta)w + 2\eta\bar{x} + \dfrac{\eta}{2K}(b(\theta^1,\theta^2) - a(\theta^1,\theta^2)) \\ \qquad \text{if } Q(\theta^1,\theta^1) = Q(\theta^2,\theta^2) = \varnothing, \\ (1-2\eta)w + 2\eta\bar{x} \qquad \text{otherwise;} \end{array}\right.$$

where $w \in W(1, f(\theta^2), \theta^2) \cap W(2, f(\theta^1), \theta^1)$ which is nonempty by assumption, and we use the facts noted above, that $Q(\theta^1,\theta^2) \neq \varnothing$ if $\theta^1 \neq \theta^2$ and $Q(\theta^1,\theta^1) = Q(\theta^2,\theta^2) = \varnothing$.

Thus $L(\theta^1, \theta^2)$ as specified above is the outcome when $\alpha_1 = \alpha_2 = 0$ in the two cases $\theta^1 = \theta^2$ and $\theta^1 \neq \theta^2$.

If $\alpha_1 = 1$ and $\alpha_2 = 0$, the outcome is $L(\theta^1, \theta^2)$ unless $a(1), b(1)$ satisfy

(i) $a(1) P^1(\theta^2) b(1)$ or

(ii) $a(1) = b(\theta^2, \theta^2)$, $b(1) = a(\theta^2, \theta^2)$, when $\theta^1 \neq \theta^2$
and $Q(\theta^2, \theta^2) \neq \varnothing$, or

(iii) $a(1) = b(\theta^1, \theta^2)$, $b(1) = a(\theta^1, \theta^2)$, when $\theta^1 \neq \theta^2$
and $Q(\theta^1, \theta^1) = Q(\theta^2, \theta^2) = \varnothing$,

in which case the outcome is

$$L(\theta^1, \theta^2) + \frac{\eta}{2K}(b(1) - a(1)).$$

The case $\alpha_1 = 0$, $\alpha_2 = 1$ is defined symmetrically.

If $\alpha_1 = \alpha_2 = 1$ we are in the "integer game," the outcome being $L^h(\theta^h)$ where $h = 1$, if $n_1 \geqslant n_2$ and $h = 2$ otherwise, and $L^h(\theta^h)$ is as defined in the proof of Theorem 1. The rest of the argument follows the same pattern as the latter proof, and we leave details to the reader.

Let the true preference profile be ψ. Then inspection of the game form clarifies that any strategy profile in which players' announcements satisfy $\theta^1 = \theta^2 = \psi$ and $\alpha^1 = \alpha^2 = 0$, is an equilibrium profile with associated outcome $(1 - 2\eta) f(\psi) + 2\eta \bar{x}$, which is, of course, ε-close to $f(\psi)$.

It is also easy to see that there is no equilibrium with $\alpha_1 = \alpha_2 = 0$ and $\theta^1 = \theta^2 \neq \psi$. In this case either $P^1(\psi) \neq P^1(\theta^2)$ or $P^2(\psi) \neq P^2(\theta^1)$. In the former case player 1 can profitably deviate by announcing $(\theta^1, 1, a'(1), b'(1), 0)$ where $a'(1) P^1(\theta^2) b'(1)$ and $b'(1) P^1(\psi) a'(1)$. This deviation adds the term $(\eta/2K)(b'(1) - a'(1))$ to the lottery he would otherwise obtain (see (i) in the description of the outcome function above) and is therefore profitable. A similar deviation is available to player 2 if $P^2(\psi) \neq P^2(\theta^1)$.

As before there is no equilibrium with $\alpha_1 = \alpha_2 = 1$, since either player can win the integer game, given the other player's n_i, and has a strict incentive to do so. If $\alpha_1 = 1$ and $\alpha_2 = 0$, player 2 will be strictly better off by playing $(\psi, 1, a_1, a_1, n_1 + 1)$, and symmetrically for the case $\alpha_1 = 0$ and $\alpha_2 = 1$.

The only remaining possibility is $\alpha_1 = \alpha_2 = 0$ and $\theta^1 \neq \theta^2$. If $P^1(\psi) \neq P^1(\theta^2)$, then player 1 can deviate profitably as in the earlier discussion of the case $\alpha_1 = \alpha_2 = 0$ and $\theta^1 = \theta^2 \neq \psi$. Similarly for player 2 if $P^2(\psi) \neq P^2(\theta^1)$. Now suppose $P^1(\psi) = P^1(\theta^2)$ and $P^2(\psi) = P^2(\theta^1)$. If $Q(\theta^2, \theta^2) \neq \varnothing$, then player 1 can profitably deviate by announcing $(\theta^1, 1, b(\theta^2, \theta^2), a(\theta^2, \theta^2), 0)$, thereby adding the term $(\eta/2K)(a(\theta^2, \theta^2) - b(\theta^2, \theta^2))$ to the lottery he would otherwise obtain (see (ii) in the description of the outcome function above). A symmetric argument applies to player 2 when $Q(\theta^1, \theta^1) \neq \varnothing$. Finally suppose $Q(\theta^1, \theta^1) = Q(\theta^2, \theta^2) = \varnothing$. Then since $\theta^1 \neq \theta^2$, it follows that $Q(\theta^1, \theta^2) \neq \varnothing$. Now player 1 can profitably deviate by announcing $(\theta^1, 1, b(\theta^1, \theta^2), a(\theta^1, \theta^2), 0)$ which adds the term $(\eta/2K)(a(\theta^1,\theta^2)-b(\theta^1,\theta^2))$ to the lottery he would otherwise obtain (see (iii) in the description of the game-form above).

Hence in all equilibria $\theta^1=\theta^2=\psi$ and $\alpha_1=\alpha_2=0$. *Q.E.D.*

The intersection property is much weaker than the sufficient conditions for exact implementation in two-person environments. Necessary and sufficient conditions for exact implementation in the two-person case were discovered independently by Dutta and Sen (1991) and Moore and Repullo (1990). We state this condition below (called Condition β in Dutta-Sen and Condition $\mu 2$ in Moore-Repullo) and compare it with the intersection property. These papers work within the traditional deterministic framework and the definition below applies to social choice correspondences which map to sets of pure alternatives.

Definition: The social choice correspondence f satisfies *Condition* β (or *Condition* $\mu 2$) if for all profiles $\theta\in\Theta$, $a\in f(\theta)$, and $j\in\{1,2\}$, there exists a set $C(j,a,\theta)\subseteq\overline{W}(j,a,\theta)$ with the following properties:

i. For all $\varphi\in\Theta$ and $b\in f(\varphi)$: (a) $C(j,a,\theta)\cap C(i,b,\varphi)\neq\varnothing$ $(i\neq j)$; (b) there exists $e\in C(j,a,\theta)\cap C(i,b,\varphi)$ $(i\neq j)$ such that for all $\psi\in\Theta$, $[C(j,a,\theta)\subseteq\overline{W}(j,e,\psi)$ and $C(i,b,\phi)\subseteq\overline{W}(i,e,\psi)]$ implies that $e\in f(\psi)$.

ii. For all $\varphi\in\Theta$, $a\notin f(\varphi)$ implies that there exists an individual j and an alternative $b\in A$ such that $b\in C(j,a,\theta)$ and $b\notin\overline{W}(j,a,\varphi)$.

iii. For all $\varphi\in\Theta$, if $e\in C(j,a,\theta)$ is such that $C(j,a,\theta)\subseteq\overline{W}(j,e,\varphi)$ and $A\subseteq\overline{W}(i,e,\varphi)$ $(i\neq j)$, then $e\in f(\varphi)$.

iv. For all $\varphi\in\Theta$, if e is such that $A\subseteq\overline{W}(j,e,\varphi)$, $j=1,2$, then $e\in f(\varphi)$.

The intersection property is implied by (i)(a) of Condition β. Virtual implementation therefore allows us to do without parts (i)(b), (ii), (iii), and (iv). Of these, part (ii) is Maskin monotonicity and it is perhaps not surprising that it is now no longer necessary, in view of our results in the many-person ($N\geqslant 3$) case. An attractive aspect of virtual implementation is that it drops monotonicity (of f) and also dispenses with the cumbersome and less easily interpretable parts (i)(b) and (iii).

The intersection property has another significant advantage over Condition β. A major practical difficulty in verifying whether a social choice correspondence satisfies Condition β is that the $C(\cdot)$ sets may be strict subsets of the appropriate $\overline{W}(\cdot)$ sets. For any given (j,a,θ) the set $\overline{W}(j,a,\theta)$ is unique; however, there may be several candidate $C(j,a,\theta)$ sets. Since the $C(\cdot)$ sets have to be constructed for all (j,a,θ) checking for Condition β is often a tedious and complicated exercise. On the other hand, the nonemptiness condition, by focusing only on the $\overline{W}(\cdot)$ sets, avoids this difficulty completely.

We noted in the Introduction a result due to Muller and Satterthwaite (1977) and Roberts (1979) according to which a social choice *function* defined over a universal domain of strict preferences was monotonic if and only if it was dictatorial. This result is independent of the number of players and has damaging implications for two-person settings since monotonicity is a necessary condition for exact implementation in Nash equilibrium also. In fact, in this setting, a negative result is available for two-person social choice *correspondences* also. This result was proved by Hurwicz and Schmeidler (1978) and Maskin (1977) (see also Moore and Repullo (1990)).

DEFINITION: A social choice correspondence $f\colon \Theta \to A$ is *Pareto efficient* if for all $\theta \in \Theta$ there does not exist $b \in A$ such that $bP^j(\theta)f(\theta)$ for all $j \in I$.

THEOREM (Hurwicz-Schmeidler (1978), Maskin (1977)): *A two-person Pareto efficient social choice correspondence f defined on an unrestricted domain of strong orderings is Nash implementable if and only if it is dictatorial.*

This result strongly suggests that exact implementation is very demanding, at least in the absence of domain restrictions. Such a result is no longer true for virtual implementation: the intersection property is much weaker than the necessary and sufficient conditions for exact implementation.

While the intersection property is simple and easily stated it is by no means trivial and it is not hard to specify social choice correspondences which violate it.

EXAMPLE: There are three alternatives $\{a, b, c\}$ and two profiles θ, φ:

θ		φ	
1	2	1	2
a	c	b	a
b	b	c	c
c	a	a	b

and

$$f(\theta) = c \qquad f(\varphi) = b.$$

Then $\overline{W}(1, c, \theta) \cap \overline{W}(2, b, \varphi) = \varnothing$ in violation of the intersection property. The SCF of this example satisfies monotonicity, so that this example also demonstrates that monotonicity does not imply the intersection property.

This sort of example can be generalized. Let f^T be the social choice correspondence which selects each player's top-ranked alternative. That is, $f^T(\theta) = \{a \mid aP^i(\theta)b$ for all $b \in A \setminus \{a\}, i = 1, 2\}$. Then we have the following result.

Proposition: *Let f be an SCC defined on an unrestricted domain of strict preferences and suppose that $f(\theta) \subset f^1(\theta)$ for all θ. Assume that $|A| \geqslant 3$. Then f is virtually implementable if and only if it is dictatorial.*

Proof: See Appendix.

The next example demonstrates that the intersection property can easily be satisfied when monotonicity fails.

Example: There are two profiles θ, φ and four alternatives:

θ		φ	
1	2	1	2
a	b	a	b
b	a	b	a
c	c	z	z
z	z	c	c
$f(\theta) = a$		$f(\varphi) = b$	

Clearly f is not monotonic and the reader may easily check that the intersection property is satisfied.

More generally, when might we expect the intersection property to hold? One simple case is when there is some bad outcome z which is worst for all players and across all preference profiles. This assumption of a holocaust outcome is also helpful in the exact implementation context and has been invoked there by Moore-Repullo (1990). This sort of domain restriction may, of course, be refined: what we need are outcomes which are "bad" (not necessarily worst) for both players relative to the range of f. In a variety of natural examples the latter may well be a small subset of A. This kind of domain restriction does not seem particularly strong given that "money" can be transferred by both players to an uninformed (in the informed case we would be in the completely permissive world of many-player ($N \geqslant 3$) virtual implementation) third party.

An alternative route to guaranteeing the intersection property is to make restrictions on f. One reasonable condition which works is that players be allowed to "veto" a large enough number of their least preferred alternatives. Specifically let k_1, k_2 be nonnegative integers such that $k_1 + k_2 = |A| - 1 = K - 1$. Assume that f maps to degenerate lotteries. Then it follows almost immediately that if for both players, and all $\theta \in \Theta$, $f(\theta)$ does not contain any of player i's k_i least preferred alternatives according to $P^i(\theta)$, then f satisfies the intersection property.

Let $B_i(\theta, k)$ be defined by $|B_i(\theta, k)| = K - k$ and $B_i(\theta, k) = \{a \mid aP^i(\theta)b$ for all $b \notin B_i(\theta, k)\}$; $B_i(\theta, k)$ is the set of player i's $(K - k)$ top-ranked alternatives according to $P^i(\theta)$. Then we have the following result.

Proposition: *Let f be a two-person SCF such that for all $\theta \in \Theta$ $f(\theta) \in B_1(\theta, k_1) \cap B_2(\theta, k_2)$ for some nonnegative integers k_1, k_2 which satisfy $k_1 + k_2 = K - 1$. Then f is virtually implementable in Nash equilibrium.*

5. CARDINAL INFORMATION

We now briefly consider social choice correspondences which map from *cardinal* preference profiles $\gamma \in \Gamma$ to lotteries over social states: a cardinal social choice correspondence f associates a nonempty set $f(\gamma) \subset \mathscr{L}$ with every $\gamma \in \Gamma$. Players are assumed here to know one another's *cardinal* preferences, and we drop the restriction to ordinal game forms. We now need to assume that players' preferences satisfy the von Neumann-Morgenstern axioms. This is because the implementing game form we use involves adding and subtracting *nondegenerate* lotteries, thereby yielding compound lotteries which are not comparable in terms of the partial ordering yielded by the axiom that preferences over lotteries are monotone. Note that all this is in keeping with the traditional game-theoretic informational assumption: players' *von Neumann-Morgenstern* utilities are *common knowledge*.

In this richer informational setting much more can be implemented, social choice being potentially sensitive to changes in cardinal preferences even when ordinal preferences are unaltered. However, the formal structure of the analysis is almost identical, $\mathscr{L}$ here playing the role of the finite set of social states A. We confine attention to the many-player ($N \geqslant 3$) case. As the analogies between Section 3 and the present one are very strong, we proceed rapidly, omitting details. Reformulations of definitions will largely be left to the reader, and typically require that θ, Θ be replaced by γ, Γ, respectively.

For simplicity we will continue to assume that individual preferences over *pure* alternatives are strict, and that for all preference profiles, some pair of individuals differ in their ranking over some pair of *lotteries*.

Theorem 5: *Let $N \geqslant 3$. Then any cardinal social choice correspondence f is virtually implementable in Nash equilibrium.*

Proof: The proof is almost identical to that of Theorem 1, and is only sketched. For all $\gamma, \xi \in \Gamma$ such that $\gamma \neq \xi$ there exists $j(\gamma, \xi) \in J$ and $x(\gamma, \xi), y(\gamma, \xi) \in \mathscr{L}$ such that $x(\gamma, \xi) \tilde{P}^j(\gamma) y(\gamma, \xi)$ and $y(\gamma, \xi) \tilde{P}^j(\xi) x(\gamma, \xi)$. The proof proceeds as before except that the term $\eta \Sigma \alpha_k a_{p(k)}$ is replaced by $(\eta(1-\beta)/\beta)\Sigma_{n=1}^{\infty} \beta^n x_{p(n)}$, where $\beta \in (0,1)$, $\{x_n\}_{n=1}^{\infty}$ is a countable dense subset of $\mathscr{L} = \Delta^{K-1}$ and p: $Z_+ \to Z_+$ is a bijection such that $((1-\beta)/\beta)\Sigma \beta^n x_{p(n)} \bar{R}^h(\gamma^h)((1-\beta)/\beta)\Sigma \beta^n x_{q(n)}$ for any other bijection q. As before, h is the lowest indexed winner of the integer game. With these modifications the rest of the earlier proof may be easily mimicked to complete the argument. *Q.E.D.*

6. A CONTINUUM OF ALTERNATIVES

The preceding sections assumed that A was a finite set. This assumption simplifies the analysis, but is not essential. Here we show how the results may be extended to the case where A is an arbitrary (nonfinite) subset of a separable metric space. This extension is useful in that it is often natural and convenient to pose economic questions in nondiscrete models.

Let $A^* = \{a_1, a_2, \ldots\}$ be a countable dense subset of A and let Θ be the set of admissible preference profiles on A. We assume that for any pair of distinct profiles $\theta, \varphi \in \Theta$, there exists $j \in J$ and $a, b \in A^*$ such that $aP^j(\theta)b$ and $bP^j(\varphi)a$. We denote these elements $j(\theta, \varphi)$, $a(\theta, \varphi)$ and $b(\theta, \varphi)$ respectively. We also assume that for all $\theta, \varphi \in \Theta$ there exists a pair $i, j \in J$ and alternatives $a, b \in A^*$ such that $aP^i(\theta)b$ and $bP^j(\varphi)a$. The first assumption plays the same role as the assumption of strict preferences over pure alternatives in earlier sections, and the latter corresponds to the assumption that all players do not have the same preference ordering.

To avoid inessential technical qualifications we take $\mathscr{L}$ to be the set of discrete lotteries (that is, lotteries with countable support) on A.

THEOREM 6: *Let $N \geqslant 3$. Then any SCC $f\colon \Theta \to A$ is virtually implementable in Nash equilibrium.*

PROOF: The argument is very similar to that of Theorem 1, the main difference being that the lotteries $L(x, \theta)$ etc., are slightly modified. Specifically,

$$L(x,\theta) \equiv (1-2\eta)x + 2\eta\frac{1-\beta}{\beta}\sum_{n=1}^{\infty}\beta^n a_n$$

where $\beta \in (0,1)$ and $A^* \equiv \{a_1, a_2, \ldots\}$ is the countable dense subset of A defined earlier.

$$L(x,\theta,\varphi) = (1-2\eta)x + 2\eta\frac{1-\beta}{\beta}\sum_{n=1}\beta^n a_n$$

$$+\eta\frac{1-\beta}{\beta}\beta^k[b(\theta,\varphi) - a(\theta,\varphi)]$$

where k is defined by $a_k = a(\theta, \varphi)$. Finally,

$$L^h(\theta^h) = (1-2\eta)a_{p(1)} + \eta\frac{1-\beta}{\beta}\sum_{n=1}\beta^n a_n$$

$$+\eta\frac{1-\beta}{\beta}\beta^k(a_{p(1)} - a_k) + \eta\frac{1-\beta}{\beta}\sum_{n=1}\beta^n a_{p(n)}$$

where player h, the player with the lowest index who announces the highest integer may "pick"[10] any $k \in Z_+$ and bijection $p\colon Z \to Z$. The latter satisfies $a_{p(k)} R^h(\theta^h) a_{p(k-1)}$ for $k = 1, 2, \ldots$. With these changes the rest of the proof may be completed by the reader. *Q.E.D.*

7. CONCLUSION

This paper reformulates the implementation problem, allowing planners to *randomize* and requiring that social goals be only *virtually*, as opposed to *exactly*, attained by implementing game forms. This perspective dramatically expands the class of implementable social choice correspondences.

For environments with at least three players we obtain the following permissive result: (under mild domain restrictions) *all* social choice correspondences are virtually implementable in Nash equilibrium. This theorem should be contrasted with the classic characterization of Maskin (1977), according to which *monotonicity* is a necessary condition for exact implementation in Nash equilibrium. The restrictive nature of this requirement has motivated a number of recent papers on implementation using *refinements* of Nash equilibrium. These papers (see the Introduction for references) extend, with increasing success, the class of social choice correspondences which can be implemented within the general framework of (Nash) equilibrium theory. The present paper also succeeds in evading the monotonicity requirement, but in a manner which is technically simpler, and more transparent.

A natural question is why our results parallel rather closely those obtained in the refinements literature. The fundamental point is that *any* preference reversal may *potentially* be used to eliminate an unwanted equilibrium. In the virtual approach, equilibrium destroying deviations based on preference reversals operate in a very direct way. With subgame perfection deviations are initially triggered in subgames possibly *off* the equilibrium path and these work their way backwards through the tree to yield a deviation *on* the equilibrium path. In undominated Nash equilibrium a previously undominated best response becomes dominated when a preference reversal occurs. The virtual perspective provides a clean and simple expression of the equilibrium destroying potential of the most minimal differences in preferences.

The case of three or more players is very different from the two-player case. We provide a simple necessary and sufficient condition (the "intersection property") for two-person virtual implementation. Again the virtual perspective leads to simpler and more permissive results. Finally we discuss how the analysis may be extended to virtual implementation in *strict* Nash equilibrium and *coalition-proof* Nash equilibrium, to social choice correspondences which map from *cardinal* preference profiles to lotteries, and to environments with a *continuum* of pure alternatives.

[10] Strictly speaking, we now need players to announce a quadruplet $(\theta_i, x_i, n_i, k_i)$ where $k_i \in Z_+$ and corresponds to the k "picked" by the winner of the integer game.

Our results exploit the extra freedom permitted by virtual as opposed to exact implementation, together with the domain restrictions that preferences over lotteries satisfy. This basic approach should prove fruitful in other applications in the general area of social choice and implementation theory.

Department of Economics, Princeton University, Princeton, NJ 08544-1021, U.S.A.

and

Indian Statistical Institute, 7, SJS Sansanwal Marg, New Delhi, India

Manuscript received December, 1987; final revision received October, 1990.

APPENDIX A

The following proof was omitted from the text.

PROPOSITION: *Let f be an SCC defined on an unrestricted domain of strict preferences and suppose that $f(\theta) \subset f^T(\theta)$ for all θ. Assume that $|A| \geq 3$. Then f is virtually implementable if and only if it is dictatorial.*

PROOF: Let f be a virtually implementable SCC which satisfies the hypotheses of the proposition. We show that f must be dictatorial. Since the intersection property is inherited by subcorrespondences it suffices to prove the result for the case in which f is a *function*.

Let $\theta \in \Theta$ be such that $a_1 P^1(\theta) \ldots P^1(\theta) a_K$ and $a_K P^2(\theta) \ldots P^2(\theta) a_1$. Then $f^T(\theta) = \{a_1, a_K\}$. Assume, without loss of generality, that $f(\theta) = a_1$. We now claim that individual 1 must be a dictator.

Since $|A| \geq 3$, we assume that the alternatives a_1, a_2, and a_K are distinct. Let $a_r \neq a_2$ and let φ denote the profile for which $a_r P^1(\varphi) \ldots P^1(\varphi) a_2$ and $a_2 P^2(\varphi) \ldots P^2(\varphi) a_r$. Either $a_2 = f(\varphi)$ or $a_r = f(\varphi)$. However, $\overline{W}(1, a_2, \varphi) \cap \overline{W}(2, a_1, \theta) = \varnothing$. Therefore, $a_r = f(\varphi)$.

Let ξ be any profile in which individual 1 ranks a_r first. Observe that $\overline{W}(1, a_s, \xi) \cap \overline{W}(2, a_r, \varphi) = \varnothing$ for all $a_s \neq a_r$. Therefore $a_r = f(\xi)$.

Let ψ be the profile such that $a_2 P^1(\psi) \ldots P^1(\psi) a_K$ and $a_K P^2(\psi) \ldots P^2(\psi) a_2$. Either $a_K = f(\psi)$ or $a_2 = f(\psi)$. However, $\overline{W}(1, a_K, \psi) \cap \overline{W}(2, a_1, \theta) = \varnothing$. Therefore, $a_2 = f(\psi)$.

Finally, for any profile λ in which individual 1 ranks a_2 first, $f(\lambda) = a_s \neq a_2$ implies that $\overline{W}(1, a_s, \lambda) \cap \overline{W}(2, a_2, \psi) = \varnothing$. Therefore, $a_2 = f(\lambda)$. That is, individual 1 is a dictator. *Q.E.D.*

REFERENCES

ABREU, D., AND A. SEN (1987): "Virtual Implementation in Nash Equilibrium," mimeo, Harvard University.

——— (1990): "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition," *Journal of Economic Theory*, 50, 285–299.

AUMANN, R. (1959): "Acceptable Points in General Cooperative n-person Games," in *Contributions to the Theory of Games IV*. Princeton, NJ: Princeton University Press.

BERNHEIM, D., B. PELEG, AND M. WHINSTON (1987): "Coalition-Proof Nash Equilibria. I: Concepts," *Journal of Economic Theory*, 42, 1–12.

BERNHEIM, D., AND M. WHINSTON (1987): "Coalition-Proof Nash Equilibria, II: Applications," *Journal of Economic Theory*, 42, 13–29.

DASGUPTA, P., P. HAMMOND, AND E. MASKIN (1979): "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Review of Economic Studies*, 46, 185–216.

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

FARQUHARSON, R. (1957/1969): *Theory of Voting*. New Haven: Yale University Press.

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

——— (1977): "Manipulation of Schemes that Mix Voting with Chance," *Econometrica*, 45, 665–681.

Hart, O., and J. Moore (1988): "Incomplete Contracts and Renegotiation," *Econometrica*, 56, 755–758.

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

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

Jackson, M. (1989): "Implementation in Undominated Strategies: A Look at Bounded Mechanisms," mimeo, Northwestern University.

Kohlberg, E., and J-F Mertens (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1037.

Maskin, E. (1977): "Nash Equilibrium and Welfare Optimality," mimeo, M.I.T.

——— (1985): "The Theory of Implementation in Nash Equilibrium: A Survey," in *Social Goals and Social Organization*, ed. by L. Hurwicz, D. Schmeidler, H. Sonnenschein. Cambridge: Cambridge University Press.

Maskin, E., and J. Moore (1986): "Implementation and Renegotiation," mimeo, M.I.T.

Matsushima, H. (1988): "A New Approach to the Implementation Problem," *Journal of Economic Theory*, 45, 128–144.

Moore, J., and R. Repullo (1988): "Subgame Perfect Implementation," *Econometrica*, 56, 1191–1220.

——— (1990): "Nash Implementation: A Full Characterization," *Econometrica*, 58, 1083–1100.

Moulin, H. (1979): "Dominance Solvable Voting Schemes," *Econometrica*, 47, 1337–1351.

——— (1983): *The Strategy of Social Choice*. Amsterdam: North-Holland Publishing Company.

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

Palfrey, T., and S. Srivastava (1991): "Nash Implementation Using Undominated Strategies," *Econometrica*, 59, 479–501.

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

Saijo, T. (1987): "On Constant Maskin Monotonic Social Choice Functions, *Journal of Economic Theory*, 42, 382–386.

——— (1988): "Strategy Space Reductions in Maskin's Theorem: Sufficient Conditions for Nash Implementation," *Econometrica*, 56, 693–700.

Satterthwaite, M. A. (1975): "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187–217.

Walker, M. (1981): "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica*, 49, 65–73.

Zeckhauser, R. (1969): "Majority Rule with Lotteries on Alternatives," *Quarterly Journal of Economics*, 83, 696–703.