Discussion Paper No. 1004

CHARACTERIZATIONS OF GAME THEORETIC SOLUTIONS WHICH LEAD TO IMPOSSIBILITY THEOREMS

by

Matthew O. Jackson
and
Sanjay Srivastava

September 1992

Characterizations of Game Theoretic Solutions Which Lead to Impossibility Theorems

Matthew O. Jackson and **Sanjay Srivastava***

First Draft: September 1990

This Draft: September 1992

For some game theoretic solution concepts, such as dominant strategies, Nash equilibrium, and undominated strategies, only dictatorial social choice functions are implementable on a full domain of preferences with at least three alternatives. For other solution concepts, such as the iterative removal of weakly dominated strategies, undominated Nash equilibrium, and maximin, it is possible to implement non-dictatorial social choice functions. Which characteristics of solution concepts account for these differences? We begin by proving a new impossibility theorem. This theorem shows that conditions which are significantly weaker than strategy-proofness, assure that a social choice function is dictatorial on a full domain of preferences. This helps us to identify the essential parts of the impossibility theorems and leads to two characterizations of solution concepts which lead to impossibility results.

Keywords: Implementation, Social Choice.

*Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2001, and Graduate School of Industrial Administration, Carnegie Mellon University, Schenley Park, Pittsburgh, PA 15213-3890, respectively. We gratefully acknowledge financial support from the National Science Foundation. We thank Salvador Barbera, Tilman Börgers, and Larry Jones for helpful conversations.

This version of the paper is substantially changed from the version entitled "Implementing Social Choice Functions: A New Look at Some Impossibility Theorems" which appeared as CMSEMS WP# 965. The impossibility theorem (1) is new, and the direct breaking and positive responsiveness conditions are modified and now both necessary and sufficient.

1. Introduction

The literature on implementation has used a large number of solution concepts to model the behavior of individuals. It is now known that the social choice rules which can be decentralized differ widely across solution concepts. In this paper we ask: which properties of solution concepts are responsible for the differences?[1]

To begin to explore this issue, we consider a very simple setting which has been well–studied: one with a finite set of alternatives and a full domain of preferences. Here the differences across solution concepts are most vivid. The only social choice functions which can be implemented via some solutions are dictatorial, while other solutions can implement non–dictatorial social choice functions .

The first theorem in this area, the Gibbard (1973)–Satterthwaite (1975) Theorem, states that a strategy–proof social choice function on an unrestricted domain of preferences must be dictatorial if it takes on at least three values. Equivalently, the result says that interesting social choice functions cannot be implemented in dominant strategies. This restriction to dictatorial social choice functions is often attributed to the strength of the requirement that there exist a dominant strategy for each agent and every preference profile.

It is interesting, however, that similar results obtain for much weaker solution concepts. Jackson (1992) shows that if a social choice function can be implemented in undominated strategies by a bounded mechanism[2] on a full domain of preferences, then it too must be dictatorial. Undominated strategies is a very weak solution concept, quite the opposite of dominant strategies. Most games have undominated strategies: for instance, all mechanisms with finite action spaces are bounded and have undominated strategies for every preference profile. An impossibility result also holds for Nash equilibrium. The only social choice functions which take on at least three values and are Nash implementable are dictatorial, as shown Dasgupta, Hammond and Maskin (1979). These results indicate that it is not the "strength" of the solution concept which makes it impossible to implement interesting social choice functions .

[1] A systematic exploration of the similarities and differences of various solution concepts should be of interest to game theory in general, beyond questions of implementation.

[2] A mechanism is bounded if, for each weakly dominated action, there exists and undominated action which dominates it. Definitions of various solution concepts and restrictions on mechanisms are provided in Section 2.

In order to understand what makes it impossible to implement interesting social choice functions , it is important to recognize that there are solution concepts which avoid the negative results. Nondictatorial social choice functions can be implemented on a full domain of preferences via undominated Nash equilibria, the iterated removal of weakly dominated strategies, maximin strategies, and other solutions. These observations lead to the following question: Which properties of a solution concept prevent it from implementing interesting social choice functions ?

To answer this question, we begin by proving a new impossibility Theorem. This theorem is stronger than the Gibbard–Satterthwaite Theorem, since the conditions it uses are generally weaker than strategy–proofness. Strategy–proofness is the strong requirement that no individual can ever gain at any preference profile from misrepresenting his or her preferences. The impossibility theorem still holds if strategy–proofness is replaced with two conditions which we call Best and Second Best. Best states that, at a given preference profile, if an individual's most preferred outcome can be obtained by some misrepresention of his or her preferences, then it is also obtained via his or her true preferences. Second Best states that, at a given preference profile, if an individual's most preferred outcome cannot be obtained by any report of his or her preferences, and the individual's second most preferred outcome can be obtained by some misrepresention of his or her preferences, then that second most preferred outcome is also obtained via his or her true preferences. One might think of these two conditions as applying strategy–proofness only in cases where one of an agent's two most preferred outcomes are available (and only at that given profile of other agents' preferences).

This new impossibility theorem provides properties which characterize the solution concepts which lead to impossibility results. A solution implements a dictatorial social choice function via some mechanism if and only if the outcome correspondence associated with the solution concept satisfies the best and second best conditions.

By examining solution concepts which satisfy the best and second best conditions, as well as ones which do not, we are able to develop a second characterization theorem which provides additional insight. This second characterization theorem identifies conditions which are called positive responsiveness and direct breaking, both of which are easilty verified across solutions. Roughly, positive responsiveness states that a solution accounts for improvements available to an agent. The direct breaking condition states that if a set of actions was a solution at a given preference profile, but is no longer a solution when some agent's preferences change, then that agent must have an improving deviation. In addition to being easily verified, the conditions also provide additional intuition into the impossibility results. One may interpret the direct breaking condition as saying that actions to be ruled out are broken directly; or in a sense, "on the proposed equilibrium path". The direct breaking condition is satisfied by undominated strategies (on bounded mechanisms), Nash equilibrium, and the iterative elimination of strictly dominated strategies. However, undominated Nash equilibrium, the iterated removal of weakly dominated strategies, and maximin strategies, which do not lead to impossibility results, do not satisfy the direct breaking condition, and therefore incorporate "off the equilibrium path" information.

The paper is organized as follows. First, we provide definitions and prove the impossibility theorem. Next, we use the intuition from this theorem to address the question of implementation. Examining this theorem with respect to various solutions leads us to define the positive responsiveness and direct breaking conditions and prove a second characterization result. Finally, we discuss potential extensions to allow for indifference. A table at the end of the paper summarizes the results for various solution concepts.

2. Definitions

The finite set of alternatives is denoted A. It is assumed that $\#A \geq 3$.

The society is composed of a finite number, N, of individuals.

Individual preferences are represented by a binary relation which is complete, asymmetric, and transitive. [3] We use the notation P^i to represent such a binary relation for agent i, and for $a \neq b$, let aP^ib mean that i prefers a to b. Let $\mathcal{P}$ denote the set of all such strict preferences over A.

A *social choice function* is a map which associates an alternative to each preference profile. We use F to denote a social choice function, $F : \mathcal{P}^N \to A$.

[3] The combination of completeness and asymmetry rules out indifference. As pointed out by previous authors, considering only strict preferences actually provides for stronger results than allowing for indifference, since it is a more restricted domain. For the implementation problem, the consideration of indifference leads to some difficulties, which we discuss in our concluding remarks.

A social choice function is *strategy-proof* if for each i, P, and $\widehat{P}^i$ either $F(P) = F(P^{-i}, \widehat{P}^i)$ or $F(P)P^iF(P^{-i}, \widehat{P}^i)$.

A social choice function is *dictatorial* if there exists i such that $F(P)P^ia$ for all $P \in \mathcal{P}$ and $a \neq F(P)$ in the range of F.

A *mechanism* (or game form) is a pair (M, g), where $M = M^1 \times \cdots \times M^N$ and $g : M \to A$. The set of mechanisms to be considered for the implementation problem is denoted $\mathcal{G}$.

A *solution* is a correspondence which indicates the set of actions which might be played for a given game form and profile of preferences. We denote solutions by S where $S : \mathcal{G} \times \mathcal{P}^N \to 2^M$. Thus, $S[(M,g), P]$ is a subset of M and $m \in S[(M,g), P]$ indicates that m is a solution under S to (M, g) at P. The *outcome correspondence* associated with S is $O_S : \mathcal{G} \times \mathcal{P}^N \to 2^A$ is defined by

$$O_S[(M,g), P] = \{a \in A \mid \exists m \in S[(M,g), P] \text{ s.t. } g(m) = a\}.$$

A social choice correspondence F is *implemented* via the solution S and the mechanism (M, g) if $O_S[(M,g), P] = F(P)$ for all $P \in \mathcal{P}^N$.

SOLUTION CONCEPTS.

Although most of the solution concepts we discuss are well–known, we provide definitions for them as used in an implementation context. An (incomplete) list of references for implementation via the solutions we discuss includes: Maskin (1977), (1985) for Nash implementation Palfrey and Srivastava (1991) and Jackson, Palfrey and Srivastava (1990) for undominated Nash implementation, Farquharson (1969), Moulin (1979), (1983), and Abreu and Matsushima (1990), for the iterated removal of weakly dominated strategies, Jackson (1992) and Börgers (1989) for undominated strategies, Thomson (1979) and Moulin (1982) for maximin strategies, Barbera and Dutta (1982) for protective equilibria, Herrero and Srivastava (1992) for implementation via backward induction, and Moore and Repullo (1988) for subgame perfect implementation. Moore (1991) provides an excellent survey of implementation via the solution concepts mentioned above.

DOMINANT STRATEGIES.

An action $m^i \in M^i$ is a *dominant strategy* for agent i at P^i if for each m^{-i} and $\widehat{m}^i$ either $g(m^i, m^{-i}) = g(\widehat{m}^i, m^{-i})$ or $g(m^i, m^{-i}) P^i g(\widehat{m}^i, m^{-i})$.

UNDOMINATED STRATEGIES.

The action $\widehat{m}^i \in M^i$ (weakly) *dominates* $m^i \in M^i$ at P^i if for each m^{-i} either $g(\widehat{m}^i, m^{-i}) = g(m^i, m^{-i})$ or $g(\widehat{m}^i, m^{-i}) P^i g(m^i, m^{-i})$, with the preference being strict for some m^{-i}. The action m^i is *undominated* at P^i if it is not dominated by any other action.

STRICT DOMINATION.

The action $m^i \in M^i$ is *strictly dominated* by $\widehat{m}^i \in M^i$ at P^i if $g(\widehat{m}^i, m^{-i}) P^i g(m^i, m^{-i})$ for each m^{-i}. The action m^i is *strictly undominated* if it is not strictly dominated by any $\widehat{m}^i$.

ITERATIVE REMOVAL OF DOMINATED STRATEGIES.

Given a mechanism (M, g) and sets $X^1 \subseteq M^1, \ldots, X^N \subseteq M^N$, an action $m^i \in X^i$ is dominated by $\widehat{m}^i \in X^i$ at P^i with respect to X if for each $m^{-i} \in X^{-i}$ either $g(\widehat{m}^i, m^{-i}) = g(m^i, m^{-i})$ or $g(\widehat{m}^i, m^{-i}) P^i g(m^i, m^{-i})$, with the preference being strict for some $m^{-i} \in X^{-i}$. Let $D^i(X, P)$ be the set of actions which are not dominated for i at P^i with respect to X, and let $D(X, P) = D^1(X, P) \times \cdots \times D^N(X, P)$. Define a sequence $D_0(M, P), \ldots D_K(M, P) \ldots$ by $D_0(M, P) = D(M, P)$ and $D_K(M, P) = D(D_{K-1}(M, P), P)$. Finally, let $D^*(M, P) = \cap_K D_K(M, P)$. An action $m \in D^*(M, P)$ is said to be *iteratively undominated* at P.

Correspondingly, we can define *iteratively strictly undominated* by using strict domination instead of domination in the above definition.

NASH EQUILIBRIUM.

A (pure strategy) *Nash equilibrium* at P is a profile of actions $m \in M$ such that for each i and $\widehat{m}^i$ either $g(m) = g(\widehat{m}^i, m^{-i})$ or $g(m) P^i g(\widehat{m}^i, m^{-i})$.

UNDOMINATED NASH EQUILIBRIUM.

The actions $m \in M$ form an *undominated Nash equilibrium* at $P \in \mathcal{P}^N$ if m is a Nash equilibrium and each m^i is undominated at P^i.

MAXIMIN.

Let $w(m^i, P^i)$ denote the worst possible outcome for agent i according to P^i if m^i is used. [That is, there exists m^{-i} such that $g(m^i, m^{-i}) = w(m^i, P^i)$ and for every $\widehat{m}^i$ either $g(m^i, \widehat{m}^{-i}) = g(m^i, m^{-i})$ or $g(m^i, \widehat{m}^{-i}) P^i g(m^i, m^{-i})$.] An action $m^i \in M^i$ is a *maximin* action for agent i at P^i, if $w(m^i, P^i) P^i w(\widehat{m}^i, P^i)$ or $w(m^i, P^i) P^i w(\widehat{m}^i, P^i)$, for all $\widehat{m}^i \in M^i$. Under maximin, agents 'rank' their strategies in terms of the worst outcomes they might lead to, and select from among those with the best worst outcome.

BOUNDED MECHANISMS.

A mechanism (M, g) is *bounded at* P if whenever m^i is dominated at P^i, there exists a undominated $\widehat{m}^i$ which dominates it. (M, g) is *bounded* if it is bounded at each $P \in \mathcal{P}$.

3. The Impossibility Theorem

To begin to understand the properties of solution concepts which lead to impossibility results, we first look at the properties of a social choice function which imply that it is dictatorial. Our starting point is the following version of the Gibbard–Satterthwaite Theorem.

THEOREM [GIBBARD (1973), SATTERTHWAITE (1975)]. *If a social choice function has at least three elements in its range, then it is strategy–proof if and only if it is dictatorial.*

It is known that conditions other than strategy–proofness also lead to dictatorial social choice functions . The Muller–Satterthwaite (1977) theorem shows that the following monotonicity condition implies that a social choice function is dictatorial on a full domain of preferences.

A social choice correspondence F is *monotonic* if for each a, P, and $\overline{P}$ such that $a \in F(P)$ and $a \notin F(\overline{P})$, there exists i and b such that $aP^i b$ and $b\overline{P}^i a$. [This condition is also referred to as strong positive association.]

THEOREM [MULLER–SATTERTHWAITE (1977)]. *If a social choice function has at least three elements in its range, then it is monotonic if and only if it is dictatorial.*

The impossibility theorem we prove here is based on a different set of conditions which are a priori weaker than strategy–proofness, and distinct from monotonicity.[4] The following conditions also imply that a social choice function is dictatorial.

Let $b_F(P^i)$ denote best element of in the range of F according to P^i. [So $b_F(P^i)P^i a$ for all a in the range of F such that $a \neq b_F(P^i)$.] Let $sb_F(P^i)$ denote the second best element in the range of F according to P^i. [So $b_F(P^i)P^i sb_F(P^i)$ and $sb_F(P^i)P^i a$ for all a in the range of F such that $a \neq b_F(P^i)$ and $a \neq sb_F(P^i)$.]

CONDITION (B).

A social choice function F satisfies the *Best* condition (B) if for any P, i, and $\overline{P}^i$: $F(P) = b_F(\overline{P}^i)$ implies that $F(\overline{P}^i, P^{-i}) = b_F(\overline{P}^i)$.

CONDITION (SB).

A social choice function F satisfies the *Second Best* condition (SB) if for any P, i, and $\overline{P}^i$: $F(\widetilde{P}^i, P^{-i}) \neq b_F(\overline{P}^i)$ for all $\widetilde{P}^i$ and $F(P) = sb_F(\overline{P}^i)$ imply that $F(\overline{P}^i, P^{-i}) = sb_F(\overline{P}^i)$.

(B) requires that at a given preference profile of other agents preferences, if an agent's most preferred alternative is the outcome associated some preferences of the agent, then it is outcome associated with the agent's true preferences. (SB) is similar, except it applies in situations in which the agent's most preferred alternative is not available, but the second most preferred alternative is.

THEOREM 1. *If a social choice function has at least three elements in its range, then it satisfies (B) and (SB) if and only if it is dictatorial.*

Proof of Theorem 1: It is clear that if F is dictatorial then it satisfies (B) and (SB). Here we show the converse for $N = 2$. [5] The extension to $N > 2$ is straightforward and

[4] We use the term "impossibility theorem" loosely. For instance, the above theorem can be restated: A social choice function which has at least three elements in its range cannot be both monotonic and non–dictatorial. We will continue to use the term "impossibility" with this understanding.

[5] The proof presented here follows the same intuition as the simple proof of the Gibbard–Satterthwaite theorem presented by Barbera and Peleg (1990). Some differences are necessary, since (B) and (SB) do not allow us to work with choices from option sets in the same way as Barbera and Peleg.

appears in the appendix. Since (B), (SB), and dictatorial are all defined relative to the range, without loss of generality, we can assume that the range of F is A. We write $b(P^i)$ instead of $b_F(P^i)$.

LEMMA 1. *(Unanimity) If* $b(P^1) = b(P^2) = a$, *then* $F(P) = a$.

PROOF: There exists $\overline{P}$ such that $F(\overline{P}) = a$. By (B) $F(P^1, \overline{P}^2) = a$. By (B) again, $F(P^1, P^2) = a$. ▮

LEMMA 2. *If* $b(P^1) = b(\overline{P}^1)$ *and* $F(P^1, \overline{P}^2) = b(\overline{P}^2)$, *then* $F(\overline{P}^1, \overline{P}^2) = b(\overline{P}^2)$.

PROOF: Let $c = b(\overline{P}^2)$ and $a = b(\overline{P}^1)$. If $a = c$ then this follows from unanimity. So suppose that $a \neq c$ and $F(\overline{P}^1, \overline{P}^2) \neq c$. Then by (B) for player 2, $F(\overline{P}^1, P^2) \neq c$ for all P^2. Let $\tilde{P}^2 = (c, a, \ldots)$. By unanimity $F(\overline{P}^1, P^2) = a$ for some P^2, and so by (SB) $F(\overline{P}^1, \tilde{P}^2) = a$. Since $b(P^1) = b(\overline{P}^1) = a$ it follows from (B) that $F(P^1, \tilde{P}^2) = a$. This is a contradiction since $F(P^1, \overline{P}^2) = c$ and (B) imply that $F(P^1, \tilde{P}^2) = c$. ▮

LEMMA 3. *For any* P^1, *either* $F(P^1, P^2) = b(P^1)\ \forall P^2$, *or* $F(P^1, P^2) = b(P^2)\ \forall P^2$.

PROOF: Suppose the contrary for some P^1. Then $\exists \hat{P}^2$ such that $F(P^1, \hat{P}^2) = c \neq a = b(P^1)$, and there exists $\tilde{P}^2$ such that $F(P^1, \tilde{P}^2) \neq d = b(\tilde{P}^2)$. Without loss of generality by (B), take $b(\hat{P}^2) = c$. Consider $\tilde{P}^1 = (a, d, \ldots)$.[6] From Lemma 2 and the fact that $F(P^1, \hat{P}^2) = c$, it follows that $F(\tilde{P}^1, \hat{P}^2) = c$. From Lemma 2 it also follows that follows that $F(\tilde{P}^1, \tilde{P}^2) \neq d$ (otherwise Lemma 2 would imply that $F(P^1, \tilde{P}^2) = d$). By (B) $F(\tilde{P}^1, P^2) \neq d\ \forall P^2$. Let $\overline{P}^2 = (d, c, \ldots)$ By (SB) it follows that $F(\tilde{P}^1, \overline{P}^2) = c$. Notice that $F(P^1, \overline{P}^2) \neq a\ \forall P^1$ (otherwise (B) would imply that $F(\tilde{P}^1, \overline{P}^2) = a$). Then by unanimity $F(P^1, \overline{P}^2) = d$ for some P^1, and so by (SB) $F(\tilde{P}^1, \overline{P}^2) = d$, contradicting $F(\tilde{P}^1, \overline{P}^2) = c$. ▮

LEMMA 4. *F is dictatorial.*

PROOF: Suppose the contrary. Then by Lemma 3 $\exists P^1$ such that $F(P^1, P^2) = b(P^1)\ \forall P^2$ and $\exists \tilde{P}^1$ such that $F(\tilde{P}^1, P^2) = b(P^2)\ \forall P^2$. By (B) it follows that $b(\tilde{P}^1) = c \neq a = b(P^1)$. Find $d \notin \{a, c\}$ and $\overline{P}^2$ such that $b(\overline{P}^2) = d$. Let $\overline{P}^1 = (c, a, \ldots)$. By Lemma 2 and the fact that $F(\tilde{P}^1, P^2) = b(P^2)\ \forall P^2$. it follows that $F(\overline{P}^1, \overline{P}^2) = d$. This implies by (B) that $F(P^1, \overline{P}^2) \neq c$ for all P^1. However, by (SB) [since $F(P^1, P^2) = b(P^1)\ \forall P^2$] it follows that $F(\overline{P}^1, \overline{P}^2) = a$, a contradiction. ▮

---

[6] Notice that a, c, and d are distinct. By the supposition $a \neq c$ and $c \neq d$. The fact that $a \neq d$ follows from $F(P^1, P^2) \neq d\ \forall P^2$ and unanimity.

Theorem 1 shows that the full force of strategy–proofness is not needed to produce the theorem, but rather only strategy–proofness at particular preference profiles where one of an agent's top alternatives are available. Given Theorem 1, it is clear that these are equivalent on a full domain of preferences, since the social choice function is dictatorial in either case. More generally, however, strategy–proofness will always imply (B) and (SB), but the converse does not hold. This is illustrated in the following example.

EXAMPLE 1.

There are two individuals $\{1,2\}$ and four alternatives $\{a,b,c,d\}$. The possible preference profiles are (P^1, P^2), $(P^1, \overline{P}^2)$, $(\widetilde{P}^1, \widetilde{P}^2)$,. $(\widehat{P}^1, \widehat{P}^2)$ These are represented below where the vertical order indicates an agent's preference.

$P^1 = P^2$	$\overline{P}^2$	$\widetilde{P}^1 = \widetilde{P}^2$	$\widehat{P}^1 = \widehat{P}^2$
a	b	c	c
b	a	a	b
c	d	b	a
d	c	d	d

The social choice function F is defined by $F(P^1, P^2) = d$, $F(P^1, \overline{P}^2) = c$, $F(\widetilde{P}^1, \widetilde{P}^2) = a$, $F(\widehat{P}^1, \widehat{P}^2) = b$. F satsifies (B) and (SB) since agents' best and second best alternatives are not available at any given preference profile because of the restricted domain. However it is not strategy–proof, since at (P^1, P^2) individual 2 prefers the outcome associated with $(P^1, \overline{P}^2)$.

We should also compare (B) and (SB) to the monotonicity condition identified by the Muller–Satterthwaite theorem, which has a different inuition and on restricted domains has no logical relation to (B) and (SB). Example 1 demonstrates that monotonicity is not implied by (B) and (SB). The social choice function in Example 1 satsifies (B) and (SB), but does not satisfy monotonicity when we consider $F(P^1, P^2)$ and $F(P^1, \overline{P}^2)$. The following example shows the converse.

EXAMPLE 2.

There are two individuals $\{1,2\}$ and four alternatives $\{a,b,c,d\}$. The possible preferences are as follows: $aP^1bP^1cP^1d$, $aP^2bP^2cP^2d$, and $b\overline{P}^2a\overline{P}^2d\overline{P}^2c$, which are represented below.

$P^1 = P^2$	$\overline{P}^2$
a	b
b	a
c	d
d	c

The social choice function F is defined by $F(P^1,P^2) = a$ and $F(P^1,\overline{P}^2) = d$. F is monotonic, yet it does not satsify (B) since $a = b_F(\overline{P}^2)$ is available to agent 2 and yet $a \neq F(P^1,\overline{P}^2)$.

4. Implementation and A Comparison of Solution Concepts

Let us now turn to the implementation problem. Theorem 1 provides a characterizations of solution concepts which lead to impossibility results. (B) and (SB) have the following counterparts which apply to solution concepts.

A solution S satisfies the (B^I) relative to (M,g) if for any P, i, and $\overline{P}^i$: $b(\overline{P}^i) \in O_S[(M,g),P]$ implies $b(\overline{P}^i) \in O_S[(M,g),P^{-i},\overline{P}^i]$.

A solution S satisfies the (SB^I) relative to (M,g) if for any P, i, and $\overline{P}^i$: that $sb(\overline{P}^i) \notin O_S[(M,g),P^{-i},\widetilde{P}^i]$ for all $\widetilde{P}^i$ and $sb_F(\overline{P}^i) \in O_S[(M,g),P]$ implies that $sb(\overline{P}^i) \in O_S[(M,g),P^{-i},\overline{P}^i]$.

Theorem 2 then follows directly from Theorem 1.[7]

THEOREM 2. *The solution S satisfies (B^I) and (SB^I) relative to a mechanism via which it implements a social choice function (with at least three outcomes in its range), if, and only if, the social choice function is dictatorial.*

We illustrate the above theorem in the following examples.

[7] (B^I) and (SB^I) are now defined relative to the range of a mechanism, although we no longer include subscripts.

EXAMPLE 3. *Undominated Strategies*

If we restrict attention to the class of bounded mechanisms, then the solution of undominated strategies satisfies (B^I) and (SB^I). Consider a bounded mechanism (M, g) and any i, $P \in \mathcal{P}^N$, $\overline{P}^i \in \mathcal{P}$, where $m \in S[(M,g), P]$ and $g(m) = b(\overline{P}^i)$. If m^i is undominated at $\overline{P}^i$, then $m \in S[(M,g), P^{-i}, \overline{P}^i]$, and ($B^I$) is satisfied. If m^i is dominated at P, then it is dominated by an undominated $\overline{m}^i$ and so $g(\overline{m}^i, m^{-i}) = g(m)$ and (B^I) is satisfied. To verify (SB^I), consider a bounded mechanism (M, g) and any i, $P \in \mathcal{P}^N$, $\overline{P}^i \in \mathcal{P}$, where $m \in S[(M,g), P]$. $g(m) = sb(\overline{P}^i)$ and $b(\overline{P}^i) \notin O_S[(M,g), P^{-i}, \tilde{P}^i]$ for all $\tilde{P}^i$. If m^i is undominated at $\overline{P}^i$, then $m \in S[(M,g), P^{-i}, \overline{P}^i]$, and ($SB^I$) is satisfied. If m^i is dominated at P, then it is dominated by an undominated $\overline{m}^i$. Since $b(\overline{P}^i) \notin O_S[(M,g), P]$ it follows that $g(\overline{m}^i, m^{-i}) \neq b(\overline{P}^i)$. Thus, $g(\overline{m}^i, m^{-i}) = g(m)$ and (SB^I) is satisfied.

Implementation in undominated strategies shows that the possibility of non–trivial implementation depends critically on the domain of possible mechanisms $\mathcal{G}$. If $\mathcal{G}$ includes all mechanisms, then any social choice function is implementable in undominated strategies [Theorem 1 in Jackson (1992)]. The above argument breaks down in trying to find the appropriate $\overline{m}^i$ because for an unbounded mechanism, there exist infinite strings of strategies, with each strategy dominating the previous one, but none of which are undominated. For such mechanisms, an agent might find that a dominated strategy provides a better outcome than all of the undominated strategies, against a particular set of strategies of other agents [See example 1 in Jackson (1992)]. For such a mechanism, however, it seems unreasonable to argue that agents will only play undominated strategies. Thus the restriction to bounded mechanisms is appropriate.

The same arguments as in Example 3 show that the solution of dominant strategies satisfies (B^I) and (SB^I).

EXAMPLE 4. *Nash Implementation.*

We now verify that the Nash solution satisfies (B^I) and (SB^I) relative to any mechanism on which it is single valued. Consider a mechanism (M, g) and any i, $P \in \mathcal{P}^N$, $\overline{P}^i \in \mathcal{P}$, where m is a Nash equilibrium at P and $g(m) = b(\overline{P}^i)$. It follows that m is a Nash equilibrium at $P^{-i}, \overline{P}^i]$, and (B^I) is satisfied. To verify (SB^I), consider a mechanism (M, g) (on which Nash outcomes are single valued) and any i, $P \in \mathcal{P}^N$, $\overline{P}^i \in \mathcal{P}$, where m is a

Nash equilibrium at P, $g(m) = a = sb(\overrightarrow{P}^i)$ and $b(\overrightarrow{P}^i) \notin O_S[(M,g), P^{-i}, \widetilde{P}^i]$ for all $\widetilde{P}^i$. Let $\overline{m}$ be a Nash equilibrium at $P^{-i}, \overrightarrow{P}^i$ and suppose that $g(\overline{m}) = c \neq a = g(m)$. Since $b(\overrightarrow{P}^i) \notin O_S[(M,g), P^{-i}, \overrightarrow{P}^i]$, it must be that $c \neq b(\overrightarrow{P}^i)$ and so $a\overrightarrow{P}^i c$. Consider $\widetilde{P}^i$ such that $b(\widetilde{P}^i) = a$ and $sb(\widetilde{P}^i) = c$. Since $\overline{m}$ is a Nash equilibrium at $P^{-i}, \overrightarrow{P}^i$ it follows that $\overline{m}$ is a Nash equilibrium at $P^{-i}, \widetilde{P}^i$. However it also follows that m is a Nash equilibrium at $P^{-i}, \widetilde{P}^i$. This contradicts the fact that the outcome correspondence is single valued. Thus our supposition was wrong and so $g(\overline{m}) = g(m)$ and so (SB^I) is satisfied.

As we see above, for some solutions it is (SB^I) is not trivial to check, and may depend on whether the outcome correspondence is single valued or not. Thus (SB^I) may not be easy to verify generally. With this in mind, we will develop an alternative characterization. The intuition for this second characterization comes from identifying what it is about solutions which makes sure that they do or do not satisify (B^I) and (SB^I).

EXAMPLE 5. *Undominated Nash implementation.*

The following mechanism allows both agents some say in the selection of an outcome, and yet it has a unique undominated Nash equilibrium for any preference profile. Thus this illustrates that undominated Nash equilibrium is a solution which does not satsify (B^I) and (SB^I) relative to some mechanisms.

	m^2	$\overline{m}^2$
m^1	a	a
$\overline{m}^1$	b	c

The mechanism represented above always has a unique undominated Nash equilibrium. The column player always has a unique undominated action, depending on the preference between b and c. The row player has a unique best response to this action, which completes the equilibrium. Notice that an iterated elimination of (weakly) dominated strategies will lead to the same predictions as the undominated Nash equilibria for this mechanism.[8]

[8] More discussion of interesting social choice functions which can be implemented by an iterated elimination of weakly dominated strategies on a full domain of preferences is given in Moulin (1982), (1983), and Herrero and Srivastava (1992).

If we examine other solution concepts applied to the above mechanism, such as undominated strategies, Nash equilibrium, or dominant strategies, they do not lead to a unique prediction for the above mechanism at some preference profiles. At some profiles there are more than one predicted outcomes for the undominated strategy or Nash solution concepts, while agent 1 has no dominant strategy.

Let us see where (B^I) fails. If $P^1 = (c, a, b)$ and $P^2 = (a, b, c)$, then the solution is m, with outcome a. If agent 2's preferences change to $\overline{P}^2 = (a, c, b)$, then the solution is bottom right, with outcome c. Thus (B^I) fails, since agent 2 would rather have preferences P^2, when he or she has preferences $\overline{P}^2$. Indeed, the social choice function implemented by the mechanism of Example 1 is not dictatorial. Similar examples can be constructed to illustrate that (SB^I) is not satsified.

Remark that undominated Nash equilibrium is stronger than either Nash equilibrium or undominated strategies, and weaker than dominant strategies. This indicates that whether or not a solution produces an impossibility result is not related to the strength of the solution concept in a simple way. In the above example, dominant strategies is too strong so that it produces no outcomes at P^1, P^2. In contrast, Nash equilibria or undominated strategies are too weak and produce multiple outcomes at P^1, P^2. Undominated Nash is weak enough to have the outcome m at P^1, P^2, and yet strong enough to rule out other outcomes, thus implementing a non-dictatorial social choice function .

Intuition is drawn from this example by examining why agent 2 chooses $\overline{m}^2$ instead of m^2 at $P^1, \overline{P}^2$. The solution m is not broken "directly", but rather indirectly. It is what happens in the bottom row which causes agent 2 to prefer $\overline{m}^2$ over m^2. However, this new choice for agent 2 now influences the choice of agent 1 who would rather choose the bottom row. Thus m is broken by "off the equilibrium path" considerations. Neither agent has an improving deviation away from m, but other considerations make agent 2 choose $\overline{m}^2$ and then it is best for agent 1 to change to $\overline{m}^1$.

Another solution concept which permits a positive implementation result is maximin strategies.[9]

[9] For more discussion of implementation via maximin see Thomson (1979). Example 6 also works for implementation via protective equilibria [Barbera and Dutta (1982)], which is a refinement of the set of maximin strategies [see Barbera and Jackson (1988)].

EXAMPLE 6. *Maximin Strategies.*

The following mechanism shows that the maximin solution is single valued on a mechanism which is not dictatorial. In fact, the maximin outcome function for the mechanism below is anonymous. One way to think of this mechanism is that each agent can veto a single outcome. The unique maximin solution is to veto your worst outcome.

	m^2	$\overline{m}^2$	$\tilde{m}^2$
m^1	a	a	b
$\overline{m}^1$	a	c	c
$\tilde{m}^1$	b	c	c

Let us see where (B^I) fails. If $P^1 = (c, a, b)$ and $P^2 = (a, b, c)$, then the solution is $\overline{m}^1, m^2$, with outcome a. If agent 2's preferences change to $\overline{P}^2 = (a, c, b)$, then the solution is $\overline{m}^1, \overline{m}^2$, with outcome c. Thus (B^I) fails, since agent 2 would rather have preferences P^2, when he or she has preferences $\overline{P}^2$.

Again, the failure to satisfy the (B^I) come from "off the equilibrium path" considerations. The change in behavior for agent 2 does not result from a better opportunity against $\overline{m}^1$, but rather from considerations of how agent 2's actions do against the rows more generally.

Before moving on with this intuition, we want to remark that this failure to satisfy the best condtion is not necessarily a shortcoming. Although the best condition seems like a compelling condition for a solution to satisfy, we should be careful to consider its interpretation under different information structures. For solutions which operate in incomplete information settings, such as undominated strategies or dominant strategies, it seems natural since agents do not know the preferences of others and thus choose actions independently of the actions or preferences of others. The only change in actions from a change from P^i to $\overline{P}^i$ is due to a change by agent i. The agent should not choose actions which do uniformly worse against the actions of the other agents. However, when we move to a complete information setting, the preceeding argument can no longer be made. A solution such as undominated Nash equilibrium, looks for a stable point given that all agents know each others' preferences. In Example 1, m is ruled out at $P^1, \overline{P}^2$ since agent 1 knows that it is a dominant strategy for agent 2 to play $\overline{m}^2$. Given this, agent 1 should play $\overline{m}^1$, even though the agent would prefer that both agents play m.

The intuition obtained from the preceeding examples is captured in the following condition which we call the Direct Breaking condition. Loosely, it states that a solution which leads to dictatorial outcomes on a full domain pays attention directly to the equilibrium. A change implies that something better is available via some deviation. Thus equilibria which are broken are broken directly.

DIRECT BREAKING.

A solution S satisfies *direct breaking* with respect to the mechanism (M, g) if for each P, and $\overline{P}$ such that $O_S[(M,g),P] \cap O_S[(M,g),\overline{P}] = \emptyset$, either

(i) there exists i with $b(P^i) \in O_S[(M,g),P]$ and $b(\overline{P}^i) \in O_S[(M,g)\overline{P}]$, or

(ii) there exists i $m \in S[(M,g),P]$, and $\overline{m}^i$ such that $g(m^{-i}, \overline{m}^i)\overline{P}^i g(m)$.

The direct breaking condition may be interpreted as follows. Suppose that a change in preferences leads to a complete change in outcomes. Either some agent is getting his or her best outcome in each case, or else the actions leading to some original outcome are not stable at the agents' new preferences. That is, some agent could benefit from deviating from at least one of the solutions associated with the original preferences.

We make two remarks about the above condition. First, it is slightly more complicated than the simple intuition we obtained from Examples 5 and 6. In particular, part (i) is extra. This is due to the fact that we desire a full characterization result and there are some somewhat pathological solution concepts to be covered. The role of part (i) is illustrated in Example 7 below. Second, the above condition is not sufficient for a solution to only implement dictatorial social choice functions . The reason is that the solution may satisfy the direct breaking condition, and yet not result in a dictatorial outcome, becasue it is picking worst outcomes instead of best ones. That is, there is nothing in the direct breaking condition which assures that agents in any way attempt to maximize their preferences, instead of minimizing them. Thus to obtain a characterization, we need to add a condition which says that agents are acting in accordance with their preferences.

POSITIVE RESPONSIVENESS.

A solution S satisfies *positive responsiveness* with respect to the mechanism (M, g) if for every i, $m \in S[(M,g),P]$, and $\widetilde{m}^i \in M^i$, such that $g(m^{-i}, \widetilde{m}^i)P^i g(m)$ either

(i) there exists $\overline{m} \in S[(M,g),P]$ such that $g(\overline{m})P^i g(m^{-i}, \widetilde{m}^i)$ or $g(\overline{m}) = g(m^{-i}, \widetilde{m}^i)$, or

(ii) there exists $j \neq i$ such that $O_S[(M,g),\widetilde{P}] = b(\widetilde{P}^j)$ for all $\widetilde{P}^j$.

The positive responsiveness condition states that either the solution accounts for improvements available to an agent, or else the agent is essentially a dummy agent who does not have any affect over the outcome. This condition turns out to be fairly weak and is satisfied by almost all solution concepts. If we consider any solution which is stable in a Nash equilibrium sense, then this condition is satisfied almost vacuously: there can exist no such improvement $\widetilde{m}^i$ for i. For solutions which work by means of domination arguments, the condition is also satisfied, but only when we restrict attention to bounded mechanisms. For example, if we consider undominated strategies, then such an action $\widetilde{m}^i$ is either undominated itself, or dominated by an undominated action which then must lead to at least as good an outcome for agent i as $\widetilde{m}^i$. If the mechanism is not bounded, then this is no longer true.[10] Two solutions which do not satisfy positive responsiveness are maximin and the protective criterion. Both solutions rely on information about the worst outcomes which an action may lead to, and do not account for the outcome of an action against particular actions of the other agents.

Both the direct breaking condition and the positive responsiveness condition consisted of two parts. The following example illustrates why these conditions need two parts. For instance, the positive responsivenss condition states that either an individual responds positvely to improvements which are available to him or her, or else the agent is effectively a dummy agent. In the following example some agents act as dummies – always acting in accordance with another agent's wishes, and the solution produces a dictatorial outcome.

[10] Example 1 in Jackson (1992) provides an illustration. In that example there are deviations $\widetilde{m}^i$ which are strict improvements, but they form a string each one dominating the previous one.

EXAMPLE 7. *A Dictatorial Solution.*

Consider S defined by

$$S[(M,g),P] = \{m \mid g(m)P^1 g(\overline{m})\forall \overline{m} \in M\}.$$

S is the somewhat pathological solution concept which assumes that all agents choose actions which are best for the first agent. Clearly, the social choice functions which are implementable via S are dictatorial. Yet, S does not satisfy part (i) of positive responsiveness and part (ii) of direct breaking with respect to the following mechanism.

	m^2	$\overline{m}^2$
m^1	a	b
$\overline{m}^1$	c	d

Let $P^1 = (a,b,c,d)$, $\overline{P}^1 = (d,a,b,c)$, and $P^2(c,d,a,b)$. The solution at P is m and the solution at $\overline{P}^1, P^2$ is $\overline{m}$. Part (i) of positive responsiveness is not satisfied since $g(\overline{m}^1, m^2)P^2 g(\overline{m})$. (This is part of what makes the solution so unappealing.) Part (ii) of direct breaking is not satisfied since neither agent has an improving deviation away from m.

In this example, we see also why the direct breaking condition needs two parts. Things are not broken directly here, since agent 2 acts in agent 1's interest, rather than in his or her own interest. Thus part (i) of Positive Responsivenss and Part (ii) of direct breaking are needed to provide a full characterization result, since pathological solution concepts, such as the one above, are covered. However, as we shall see in the examples below, these parts of these conditions can effectively be ignored when we deal with any of the standard solution concepts in which agents act in their own best interest.

THEOREM 3. *A solution satisfies positive responsiveness and direct breaking for a mechanism via which it implements a social choice function (with at least three elements in its range), if and only if the social choice function is dictatorial.*

PROOF: Let F be a social choice function which has at least three elements in its range and is implemented via the solution S by the mechanism (M,g). It is easy to check that if F is dictatorial then S satisfies both positive responsiveness and direct breaking relative to (M,g). We show the converse. So suppose that S satisfies positive responsiveness and direct breaking relative to (M,g). We show that S satisfies (B^I) and (SB^I) with respect to (M,g). Thus by Theorem 2, F is dictatorial.

First, we verify (B^I). Consider $m \in S(P)$ such that $g(m) = b(\overline{P}^i)$. Suppose that $b(\overline{P}^i) \notin O_S[(M,g), P^{-i}, \overline{P}^i]$. By direct breaking it follows that there exists some k, $m \in S(P)$, and $\overline{m}^k$ such that $g(m^{-k}, \overline{m}^k)\overline{P}^k g(m)$ (where if $k \neq i$ then $\overline{P}^k = P^k$). Since $g(m) = b(\overline{P}^i)$, it must be that $k \neq i$. Then $\overline{P}^k = P^k$ and so $g(m^{-k}, \overline{m}^k)P^k g(m)$. By positive responsiveness either the outcome function is dictatorial and so (B^I) is statisfied, or else there exists $\tilde{m} \in S[(M,g), P]$ such that $g(\tilde{m})P^k g(m)$, contradicting the fact that the outcomes are single valued. Thus $b(\overline{P}^i) \in O_S[(M,g), P^{-i}, \overline{P}^i]$, satisfying (B^I).

Next, we verify (SB^I). Consider $m \in S(P)$ such that $g(m) = sb(\overline{P}^i)$ and $b(\overline{P}^i) \notin O_S[(M,g), P^{-i}, \tilde{P}^i]$ for all $\tilde{P}^i$. Without loss of generality, given that we have shown (B^I), we can assume that $g(m) = b(P^i)$, and that $O_S[(M,g), P^{-i}, \overline{P}^i] = sb(P^i)$. Suppose that $sb(\overline{P}^i) \notin O_S[(M,g), P^{-i}, \overline{P}^i]$. By direct breaking it follows that there exists some k, $\overline{m} \in S(P^{-i}, \overline{P}^i)$, and $\tilde{m}^k$ such that $g(\overline{m}^{-k}, \tilde{m}^k)P^k g(\overline{m})$. If $k \neq i$, then by positive responsiveness either the outcome function is dictatorial and so (SB^I) is statisfied, or else there exists $\tilde{m} \in S[(M,g), P^{-i}, \overline{P}^i]$ such that $g(\tilde{m})P^k g(\overline{m})$, contradicting the fact that the outcomes are single valued. Thus it must be that $k = i$ and $g(\overline{m}^{-i}, \tilde{m}^i)P^i g(\overline{m})$. Since $O_S[(M,g), P^{-i}, \overline{P}^i] = sb(P^i)$, it follows that $g(\overline{m}^{-i}, \tilde{m}^i) = b(P^i)$. This implies that $g(\overline{m}^{-i}, \tilde{m}^i) = sb(\overline{P}^i)$. Since $g(\overline{m}) \neq b(\overline{P}^i)$, it follows that $g(\overline{m}^{-i}, \tilde{m}^i)\overline{P}^i g(\overline{m})$. By positive responsiveness either the outcome function is dictatorial and so (SB^I) is statisfied, or else there exists $\tilde{m} \in S[(M,g), P^{-i}, \overline{P}^i]$ such that $g(\tilde{m})\overline{P}^i g(\overline{m})$, which contradicts our supposition. Thus $sb(\overline{P}^i) \in O_S[(M,g), P^{-i}, \overline{P}^i]$, satisfying (SB^I).

The direct breaking condition may seem somewhat similar to requiring that the outcome correspondence associated with a solution be monotonic. There are important differences, however, and the direct breaking condition is weaker. The direct breaking condition is binding only when all outcomes change due to a switch in preferences by some agent. In contrast, monotonicity is binding when any outcome changes due to a change in the preferences of an agent. Further, monotonicity then requires a preference switch between the outcome and some alternative for the agent whose preferences have changed. Direct breaking only requires that some agent have an improving deviation from some original solution. These important differences are evident in the following example.

EXAMPLE 8. *The Iterated Removal of Strictly Dominated Strategies.*

Only dictatorial social choice functions are implementable via the iterated removal of strictly dominated strategies.[11] Positive responsiveness and direct breaking are easily verified as follows:

First we check positive responsiveness. Consider i, $m \in S[(M,g),P]$ and $\tilde{m}^i$ such that $g(m^{-i},\tilde{m}^i)P^i g(m)$. Since $\#A$ is finite, there exists $\overline{m}^i$ such that for each $\hat{m}^i$ either $g(m^{-i},\overline{m}^i)P^i g(m^{-i},\hat{m}^i)$ or $g(m^{-i},\overline{m}^i) = g(m^{-i},\hat{m}^i)$. It follows that $m^{-i},\overline{m}^i$ is iteratively undominated at P, and hence satisfies the requirement of positive responsiveness.

Checking the direct breaking condition is as straightforward. Consider m which is left after the iterated elimination of strictly dominated strategies at P. If m is not a solution at $\overline{P}$, then there is a first stage such that m^j is strictly dominated by $\overline{m}^j$ for some j. This implies that $g(m^{-j},\overline{m}^j)\overline{P}^j g(m)$.

Now we show that although the iterative elimination of strictly dominated strategies satisfies positive responsiveness and direct breaking, it does not always have a monotonic outcome correspondence. Consider the following mechanism.

	m^2	$\overline{m}^2$
m^1	a	a
$\overline{m}^1$	b	c

Let $P^1 = (b,a,c)$, $\overline{P}^1 = (b,c,a)$, and $P^2 = (c,b,a)$. At P, neither agent can remove a strategy and so the set of outcomes is $\{a,b,c\}$. At $\overline{P}^1, P^2$, agent 1 can remove m^1 since it is strictly dominated by $\overline{m}^1$. This then allows agent 2 to remove m^2. The solution is $\overline{m}$ with outcome c. This is inconsistent with monotonicity: the relative ordering of b remains unchanged, and yet it is dropped as an outcome. It is consistent, however, with direct breaking since $g(\overline{m})P^2 g(\overline{m}^1,m^2)$.

EXAMPLE 9. *Nash Implementation Revisited.*

Although it was somewhat difficult to check that the Nash equilibrium solution (SB^I) for mechanisms on which it is single-valued, we can easily check that it satisfies both positive responsiveness and direct breaking for any mechanism. Positive responsiveness is satisfied since by the definition of Nash equilibrium there can never exist $\hat{m}^i$ such that $g(m^{-i},\hat{m}^i)P^i g(m)$, when m is a Nash equilibrium at P. Direct breaking is satisfied since if m is a Nash equilibrium at P but not at $P^{-i},\tilde{P}^i$, then m^i must no longer be a best response for player i.

EXAMPLE 10. *Undominated Strategies Revisited.*

[11] A similar result for a slightly differnet solution was obtained by Börgers (1989).

We close this section by verifying that the solution of undominated strategies satisfies both positive responsiveness and direct breaking for any bounded mechanism. Consider P, m which is undominated at P, and i and $\widehat{m}^i$ such that $g(m^{-i}, \widehat{m}^i)P^i g(m)$. Either $\widehat{m}^i$ is undominated, or it is dominated by an undominated action $\overline{m}^i$. Positive responsiveness is thus satisfied by either $(m^{-i}, \widehat{m}^i)$ or $(m^{-i}, \overline{m}^i)$, respectively. To verify direct breaking, consider P and $\overline{P}^i$ such that $O_S[(M,g), P] \cap O_S[(M,g), P^{-i}, \overline{P}^i] = \emptyset$. Let m be undominated at P. It follows that m^i is dominated at $\overline{P}^i$ by an action $\overline{m}^i$ which is undominated at $\overline{P}^i$. Thus, $g(m) \neq g(m^{-i}, \overline{m}^i)$, and so $g(m^{-i}, \overline{m}^i)\overline{P}^i g(m)$. Therefore, direct breaking is satisfied.

We remark here on a sublety concerning the direct breaking condition and our comments concerning "paying attention to off the equilibrium path outcomes." Theorem 3 has shown that direct breaking (together with positive responsiveness) is necessary and sufficient for a solution to implement only dictatorial social choice functions . This means that a solution which incorporates off the equilibrium path information might break equilibria indirectly and thus avoid the direct breaking condition and the impossibility result. This, however, does not mean that a solution which incorporates off the equilibriu path information *necessarily* avoids the direct breaking condition. For instance, the solution of undominated strategies pays attention to all outcomes which can be obtained by any combination of other agents actions. Thus the set of actions which form a solution can change due only to a change in ordering of alternatives which are not available to any agent given the equilibrium actions of the other agents. Yet, undominated strategies still satisfies direct breaking relative to bounded mechanisms, since a change in the outcome does require that some agent have an alternative action which is better given the actions of the other agents.[12]

5. Concluding Remarks

In this paper we have examined properties of solution concepts which limit their ability to implement non-dictatorial social choice functions on a full domain of preferences. The solutions with limited implementation results had the common trait of "breaking" certain equilibria directly by requiring that some agent have an improving deviation against the actions of the other agents. In contrast, solutions which permit implementation of interesting social choice functions on a full domain of preferences incorporate information which permits them to break equilibria without requiring that any agent an improving deviation directly against the actions of the other agents. [Table 1 summarizes the results for various solutions.]

[12] This is where the restriction to bounded mechanisms is important. Without this restriction, an action could be ruled out by an infinite string of actions each dominating the previous one, but none of which are undominated. Non-dictatorial social choice functions are implementable through unobunded mechanisms.

We have focussed attention on full preference domains and on the implementation of social choice functions . Comparisons across solution concepts might also prove useful in understanding implementation in more structured environments, where there are additional restrictions on the set of preferences considered, and where it is possible to implement correspondences instead of just functions.

Another extension would allow for the possibility of indifference in preferences. Considering a full domain of preferences with the possibility of indifference, produces difficulties for the implementation of social choice functions .[13] For almost any solution we consider, there are no non-constant social choice functions which are implementable on a full domain of preferences where indifference is allowed. [Thus not even a dictatorial social choice function is implementable on such a domain.] This is easily seen by noting that when all agents are completely indifferent, then all actions will be possible under almost any solution concept. An implemented social choice function must then take on all values at such a preference profile. Even if complete indifference is ruled out, allowing for some indifference will produce multiple outcomes for some preference profiles. Thus to extend the discussion of implementation to the domain of indifference, one has to consider social choice correspondences.

[13] Theorem 1, however, could be extended to allow for indifference, given that preferences which have best and second best alternatives are admissible. An extension beyond that, to say a metric space of alternatives and a restriction to continuous preferences [as in Barbera and Peleg (1990)], would require a change of the (SB) condition.

TABLE 1

D – Only dictatorial functions are implementable.

N – Non–dictatorial functions are implementable.

Solution	$N = 2$ $\#A = 2$	$N \geq 3$ $\#A = 2$	$N \geq 2$ $\#A \geq 3$
Dominant Strategies	N	N	D
Undominated Strategies (unbounded mechanisms)	N	N	N
Undominated Strategies (bounded mechanisms)	N	N	D
Iterated Elimination: Weakly Dominated	N	N	N
Iterated Elimination: Strictly Dominated	D	D	D
Nash Equilibria	D	N	D
Undominated Nash	N	N	N
Perfect Equilibria	N	N	N
Maximin	D	N	N

References

ABREU, D. AND H. MATSUSHIMA (1990), "Exact Implementation," mimeo: Princeton University.

BARBERÀ, S. AND B. DUTTA (1982), "Implementation via Protective Equilibria," *Journal of Mathematical Economics,* **4**, pp. 49–65.

BARBERÀ, S. AND M. JACKSON (1988), "Maximin, Leximin, and the Protective Criterion: Characterizations and Comparisons," *Journal of Economic Theory,* **46**, pp. 34–44.

BARBERÀ, S. AND B. PELEG (1990), "Strategy-Proof Voting Schemes with Continuous Preferences," *Social Choice and Welfare,* **7**, pp. 31–38.

BÖRGERS, T. (1989), "Undominated Strategies and Coordination in Normalform games," mimeo.

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**, pp. 185–216.

DUTTA, B. AND A. SEN (1991), "Implementation Under Strong Equilibria: A Complete Characterization," *Journal of Mathematical Economics,* **20**, pp. 49–68.

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

HERRERO, M. AND S. SRIVASTAVA (1992), "Implementation via Backward Induction," *Journal of Economic Theory,* **56**, pp. 70–88.

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

JACKSON, M., T. PALFREY, AND S. SRIVASTAVA (1990), "Undominated Nash Implementation in Bounded Mechanisms," forthcoming: *Games and Economic Behavior.*

JACKSON, M. AND S. SRIVASTAVA (1992), "On Two Person Nash Implementable Choice Functions," *Social Choice and Welfare,* **9**, pp. 263–264.

MASKIN, E. (1977), "Nash Equilibrium and Welfare Optimality," Department of Economics, MIT.

MASKIN, E. (1985), "The Theory of Implementation in Nash Equilibrium." In L. Hurwicz, D. Schmeidler, and H. Sonnenschein, *Social Goals and Social Organization: Essays in Honor of Elisha A. Pazner*, Cambridge: Cambridge University Press.

MOORE, J. (1991), "Implementation in Environments with Complete Information," forthcoming in *Advances in Economic Theory: The Proceedings of the Sixth World Congress of the Econometric Society* ed. J.-J. Laffont, Cambridge University Press.

MOORE, J. AND R. REPULLO (1988), "Subgame Perfect Implementation," *Econometrica,* **56**, pp. 1191–1220.

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

MOULIN, H. (1982), "Non-Cooperative Implementation: A Survey of Recent Results," *Mathematical Social Sciences*, **3**, pp. 243–257.

MOULIN, H. (1983), *The Strategy of Social Choice*, North Holland: Amsterdam.

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

PALFREY, T., AND S. SRIVASTAVA (1991), "Nash Implementation Using Undominated Strategies," *Econometrica*, **59**, pp. 479–502.

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

SCHMEIDLER, D. AND H. SONNENSCHEIN (1978), " Two Proofs of the Gibbard–Satterth waite Theorem on the possibility of a Strategy-proof Social Choice Function." In Gottinger H. and W. Leinfellner, *Decision Theory and Social Ethics*, Reidel, Dordrecht.

THOMSON, W. (1979), "Maximin Strategies and the Elicitation of Preferences." In J.J. Laffont, *Aggreagation and the Revelation of Preferences*, North Holland, Amsterdam.

APPENDIX

Proof of Theorem 1 for $N > 2$. [14]

A subset of agents S has *veto power* if aP^ib for all $i \in S$ implies $F(P) \neq b$.

(i) Partition agents into S and S^c. Either S or S^c has veto power.

Consider preference profiles for which all agents in S have identical preferences, and all agents in S^c have identical preferences. From the proof of the Theorem for $N = 2$ and the fact that (B) and (SB) work in coalitional versions (when all members of a coalition have identical preferences), F gives the most preferred outcome of either S or S^c on this restricted domain. Say it is S. Suppose that for some P, $F(P) = b$ while aP^ib for all $i \in S$. Consider $\overline{P}$ such that every agent in S has identical preferences with a most preferred and b second, and all agents in S^c have identical preferences with b most preferred. The outcome is the most preferred outcome of S, so $F(\overline{P}) = a$. Since $F(P) = b$ by (B) it follows that $F(P^S, \overline{P}^{S^c}) = b$. Then by (SB) $F(\overline{P}^S, \overline{P}^{S^c}) = b$. This is a contradiction.

(ii) If S has veto power and $j \in S$, then either $S - j$ or j has veto power.

Suppose not. Then by (i) it follows that $S^c \cup j$ and $S - j \cup S^c$ have veto power. Let $S - j$ have preferences $(a, b, c, \ldots)$, j have preferences $(b, c, a, \ldots)$, and S^c have preferences $(c, a, b, \ldots)$. $F(P) \neq b$, since aP^ib for all $i \neq j$. $F(P) \neq c$, since bP^ic for all $i \in S$. $F(P) \neq a$, since cP^ia for all $i \in S^c \cup j$. This is a contradiction since any coalition will veto all outcomes not in $\{a, b, c\}$.

(iii) F is dictatorial.

Begin by (i) and then apply (ii). If j has veto power, then j is a dictator. If not, then pick $k \in S - j$ and apply (ii) again. Repeat until k is a dictator.

[14] This portion of the proof is partly based on a proof in Schmeidler and Sonnenschein (1978).