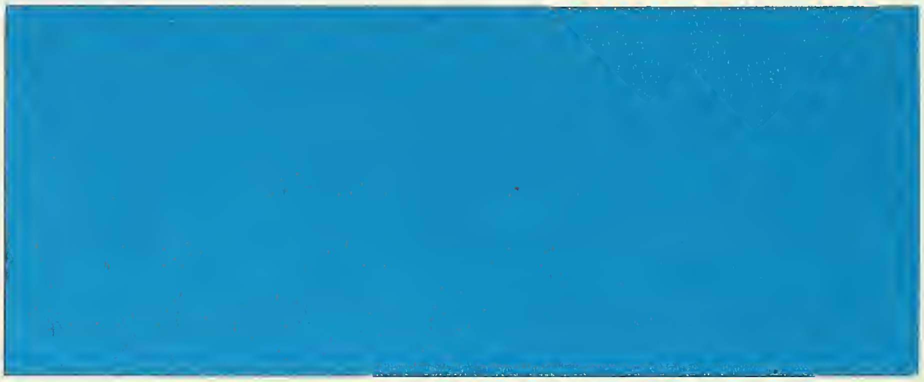
working paper
department
of economics

IMPLEMENTATION AND STRONG NASH EQUILIBRIUM

Eric Maskin

Number 216 January 1978

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139



IMPLEMENTATION AND STRONG NASH EQUILIBRIUM

Eric Maskin

Number 216 January 1978

I am grateful for the financial support of the National Science Foundation.

A social choice correspondence (SCC) is a mapping which associates each possible profile of individuals' preferences with a set of feasible alternatives (the set of f-optima). To implement an SCC, f, is to construct a game form g such that, for all preference profiles the equilibrium set of g (with respect to some solution concept) coincides with the f-optimal set.

In a recent study [1], I examined the general question of implementing social choice correspondences when Nash equilibrium is the solution concept. Nash equilibrium, of course, is a strictly noncooperative notion, and so it is natural to consider the extent to which the results carry over when coalitions can form. The cooperative counterpart of Nash is the strong equilibrium due to Aumann. Whereas Nash defines equilibrium in terms of deviations only by single individuals, Aumann's equilibrium incorporates deviations by every conceivable coalition. This paper considers implementation for strong equilibrium.

The results of my previous paper were positive. If an SCC satisfies a monotonicity property and a much weaker requirement called no veto power, it can be implemented by Nash equilibrium. The results for strong equilibrium, on the other hand, are on the whole negative. I show (theorem 2) that SCC's satisfying no veto power cannot in general be implemented for strong equilibrium when the number of alternatives is at least three.

There are, of course, some circumstances in which one is willing to forego no veto power, weak though it is. If one can identify some alternative as the status quo, for example, then the property individual rationality may have some appeal. Individual rationality implies nonetheless that non-status-quo alternatives can be vetoed. I show (theorem 3) that one in fact can implement the SCC f_0 which selects all Pareto optima that no one finds less desirable than the status quo. Unfortunately, as theorem 3 demonstrates, f_0 is the only individually rational SCC which is implementable. I begin with a section on terminology.

1. Notation and Definitions

Let A be a set of social alternatives containing, to avoid trivialities, at least two elements. For convenience, I shall assume A to be finite throughout, but all results may be extended to the infinite case. Let $\mathcal{R}_A$ be the class of all orderings of the elements of A. An n-person social choice correspondence (SCC) on $(\mathcal{R}_1,\ldots,\mathcal{R}_n)$, where $\mathcal{R}_1,\ldots,\mathcal{R}_n \subseteq \mathcal{R}_A$, is a correspondence

$$f : \mathcal{R}_1 \times \ldots \times \mathcal{R}_n \twoheadrightarrow A$$

where $\forall\ (R_1,\ldots,R_n) \in \Pi\mathcal{R}_j$, $f(R_1,\ldots,R_n)$ is nonempty. In this study, I shall be concerned primarily with the case where $\mathcal{R}_i = \mathcal{R}_A$ for all i; i.e., with the case where no a priori restriction can be placed on preferences. For any profile $(R_1,\ldots R_n) \in \prod_{j=1}^{n} \mathcal{R}_j$, the set $f(R_1,\ldots,R_n)$ is called the set of f-optima. I shall assume throughout that

$$\bigcup_{(R_1,\ldots,R_n) \in \prod_{j=1}^{n}\mathcal{R}_j} f(R_1,\ldots R_n) = A,$$

Otherwise we may delete those elements from A which can never be f-optima.

Three properties which shall concern me and which may be desirable in SCC's are monotonicity, no veto power, and individual rationality.

Monotonicity: $f : \Pi \mathcal{R}_j \rightrightarrows A$ is monotonic iff $\forall (R_1,\ldots,R_n), (R'_1,\ldots,R'_n) \in \Pi \mathcal{R}_j$ $a \in f(R_1,\ldots,R_n)$ and $[\forall i\ \forall b\ aR_ib \Rightarrow aR'_ib]$ imply $a \in f(R'_1,\ldots,R'_n)$.

No Veto Power: f satisfies no veto power iff $\forall (R_1,\ldots,R_n) \in \Pi \mathcal{R}_j$ $\forall\ a \in A$ if $\exists\ j \in \{1,\ldots,n\}$ such that $\forall\ i \neq j$ aR_ib for all b, then $a \in f(R_1,\ldots,R_n)$.

Individual Rationality: Let some alternative a_o be identified as the status quo. The SCC $f : \Pi_j \rightrightarrows A$ satisfies individual rationality iff

$$\forall (R_1,\ldots,R_n) \in \Pi \mathcal{R}_j \quad \forall\ a \in f(R_1,\ldots,R_n) \quad aR_ia_o \text{ for all } i.$$

The alternative $a \in A$ is (weakly) Pareto optimal in A with respect to $(R_1,\ldots,R_n)$ iff there does not exist $b \in A$ such that bP_ia[1/] for all i.

An n-person game form for the set A is a mapping

$$g : S_1 \times \ldots \times S_n \rightarrow A$$

where S_i is agent i's strategy space.

<u>Nash Equilibrium</u>: $\bar{s} \in \Pi S_j$ is a Nash equilibrium for the game form $g : \Pi S_j \to A$ with respect to the preference profile $(R_1,\ldots,R_n)$ iff $\forall\, i \;\; \forall\, s_i \in S_i \;\; g(\bar{s})R_i g(s_i, \bar{s}_{-i})$.[2/]

<u>Strong Equilibrium</u>: $\bar{s} \in \Pi S_j$ is a strong equilibrium for the game form $g : \Pi S_j \to A$ with respect to the profile $(R_1,\ldots,R_n)$ iff $\forall\, C \subseteq \{1,\ldots,n\} \;\forall\, s_C \in \prod_{j \in C} S_j$, $\exists\, i \in C$ such that $g(\bar{s})R_i g(s_C, \bar{s}_{-C})$.[3/]

The game form $g : \Pi S_j \to A$ is said to <u>implement</u> the SCC $f : \Pi R_j \to A$ for <u>Nash equilibrium</u> iff $\forall (R_1,\ldots,R_n) \in \Pi R_j$ $\; NE_g(R_1,\ldots,R_n) = f(R_1,\ldots,R_n)$, where $NE_g(R_1,\ldots,R_n)$ is the set of Nash equilibria for the game form g with respect to the preferences $(R_1,\ldots,R_n)$. Analogously, g implements f for strong equilibrium iff $\forall (R_1,\ldots,R_n) \in \Pi S_j$ $\; SE_g(R_1,\ldots,R_n) = f(R_1,\ldots,R_n)$ where $SE_g(R_1,\ldots,R_n)$ is the set of strong equilibria for the game form g with respect to the preferences $(R_1,\ldots,R_n)$.

I should note that if g implements f for strong equilibrium, it does <u>not</u> necessarily implement it for Nash equilibrium. The reason for this apparent anomaly is that g may possess Nash equilibria which are not strong and which, furthermore, do not lead to outcomes in $f(R_1,\ldots,R_n)$. For example, consider the following two person game form, where Player 1 chooses rows as strategies, and Player 2, columns.

	s_2	s_2'
s_1	a	a
s_1'	a	b

This game form implements for strong equilibrium the SCC f* :

$$f^* : \mathcal{R} \times \mathcal{R} \overset{\rightarrow}{\rightarrow} \{a,b\}$$

where $\mathcal{R} = \begin{Bmatrix} a & b \\ b, & a \end{Bmatrix}$ and

where $f^* \begin{pmatrix} a & a \\ b, & b \end{pmatrix} = f^* \begin{pmatrix} a & b \\ b, & a \end{pmatrix} = f^* \begin{pmatrix} b & a \\ a, & b \end{pmatrix} = \{a\}$, and $f^* \begin{pmatrix} b & b \\ a, & a \end{pmatrix} = \{b\}$.

It does <u>not</u> implement f^* for Nash equilibrium, however, because (s_1,s_2) constitutes a non f^*-optimal Nash equilibrium with respect to $\begin{pmatrix} b & b \\ a, & a \end{pmatrix}$.

2. No Veto Power

A principal theorem in Maskin [1] is the assertion that an SCC satisfying no veto power can be implemented for Nash equilibrium iff it is monotonic. In this section I show that the picture is quite different for strong equilibrium. Monotonicity remains a necessary condition (theorem 1), but if the number of alternatives exceeds two, no veto power and implementability become mutually incompatible, at least when preferences are unrestricted.

<u>Theorem 1</u>: If $f : \prod_{j=1}^{n} \mathcal{R}_j \overset{\rightarrow}{\rightarrow} A$ can be implemented for strong equilibrium, then f is monotonic.

<u>Proof</u>: If f is not monotonic then there exist $a \in A$ and $(R_1,\ldots,R_n)$, $(R_1',\ldots,R_n') \in \prod_{j=1}^{n} \mathcal{R}_j$ such that $a \in f(R_1,\ldots,R_n)$ and $[\forall i \ \ \forall b \ \ aR_ib \Rightarrow aR_i'b]$, yet $a \notin f(R_1',\ldots,R_n')$. Now if $g : \prod_{j=1}^{n} S_j \rightarrow A$ implements f for strong equilibrium, there exists $\bar{s} \in \prod_{j=1}^{n} S_j$ such that $\bar{s}$ is a strong equilibrium

for $(R_1,\ldots,R_n)$ and $g(\bar{s}) = a$. But observe that $\bar{s}$ is also a strong equilibrium with respect to $(R_1',\ldots,R_n')$, a contradiction of the definition of implementation.

Q.E.D.

<u>Theorem 2</u>: If $|A| \geq 3$, and if $f : \mathcal{R}_A^n \rightrightarrows A$ satisfies no veto power, it cannot be implemented for strong equilibrium.

<u>Proof</u>: Write $A = \{a(1),\ldots,a(m)\}$. Suppose that $g : \Pi S_j \to A$ implements f for strong equilibrium. Assume first that $m \geq n$. Choose $(R_1,\ldots,R_n) \in \mathcal{R}_A^n$ so that

$$a(1)P_1a(2) \ldots P_1a(n)P_1a(n+1) \ldots P_1a(m)$$

$$a(n)P_2a(1) \ldots P_2a(n-1)P_2a(n+1) \ldots P_2a(m)$$

$$a(n-1)P_3a(n) \ldots P_3a(n-2)P_3a(n+1) \ldots P_3a(m)$$

$$\vdots$$

$$a(2)P_na(3) \ldots P_na(1)P_na(n+1) \ldots P_na(m)$$

Now suppose that s* is a strong equilibrium with respect to $(R_1,\ldots,R_n)$. If $g(s^*) = a(p)$ where $p \in \{1,\ldots,m)$ then,

$$\text{for} \begin{cases} \text{any } i \in \{1,\ldots,n\} & \text{if } p \geq n+1 \\ i = n-p+2 & \text{if } 2 \leq p \leq n \\ i = 1 & \text{if } p = 1 \end{cases}$$

$g(s^*_i, s_{-i}) \neq a(p-1)$ [4/] for all $s_{-i} \in \prod_{j \neq 1} S_j$, because

s^* is an equilibrium and $a(p-1)P_j a(p)$ for all $j \neq i$.

Now consider $\bar{R}_i \in \mathcal{R}_A$ such that $\forall\, a \neq a(p-1)$, $a\bar{P}_i a(p-1)$. Observe that for any $\bar{R}_{-i} \in \mathcal{R}_A^{n-1}$, $a(p-1) \notin f(\bar{R}_i, \bar{R}_{-i})$, because player i can block $a(p-1)$ by playing s^*_i. But this is a contradiction of no veto power. Suppose next that $n > m$. For $i=1,\ldots,m$ choose $(R^i_1,\ldots,R^i_n) \in \mathcal{R}_A^n$ such that

$$a(1)P^i_1 a(2) \ldots a(m-1)P^i_1 a(m)$$

$$a(2)P^i_2 a(3) \ldots a(m)P^i_2 a(1)$$

$$\vdots$$

$$a(m)P^i_m a(1) \ldots a(m)P_2 a(m-1)$$

and $$R^i_1 = R^i_{m+1} = \ldots = R^i_n .$$

Let $\bar{s}^i$ be a strong equilibrium for g with respect to $(R^i_1,\ldots,R^i_n)$. Now, if $g(\bar{s}^i) = a(q) \neq a(i)$, then $\forall\, s^i_{-q} \in \mathcal{R}_A^{n-1}$, $g(\bar{s}^i_q, s^i_{-q}) \neq a(q-1)$ because $a(q-1)P_j a(q)$ $\forall\, j \neq q$. But now choose $\tilde{R}_q \in \mathcal{R}_A$ such that $\forall\, a \neq a(q-1)$ $a\tilde{P}_q a(q-1)$. $\forall\, R_{-q} \in \mathcal{R}_A^n$, $a(q-1) \notin f(R_q, R_{-q})$ because player q can block $a(q-1)$ by playing $\bar{s}^i_q$. This is a violation of no veto power. Thus, $g(\bar{s}^i) = a(i)$. Now, by construction, there exists a coalition C with $|C| = m-1$, such that $\forall j \in C$ $a(i-1)P_j a(i)$. This means that

(1) $\forall\, R_C \in \mathcal{R}_A^{m-1}$ $\exists\, R_{-C} \in \mathcal{R}_A^{n-m+1}$ such that $a(i-1) \notin f(R_C, R_{-C})$.

Furthermore, by the symmetry of the above argument, (1) holds for all i and all coalitions C such that $|C| = m-1$. Now because $m \geq 3$, the cardinality of C is at least 2. From symmetry, we may take $C = \{n-m+2,\ldots,n\}$. Construct $f^* : \mathcal{R}_A^{n-m+2} \rightrightarrows A$ so that

$$\forall (R_1,\ldots,R_{n-m+2}) \in \mathcal{R}_A^{n-m+2} \; f^*(R_1,\ldots,R_{n-m+2}) = f(R_1,\ldots,R_{n-m+2},R_{n-m+2},\ldots,R_{n-m+2})$$

Now f^* satisfies no veto power because f does and because the coalition C has no veto power. Furthermore, taking $S_j^* = S_j$ for $j=1,\ldots,n-m+2$ and $S^*_{n-m+2} = S_{n-m+2} \times \ldots \times S_n$, and defining $g^* : \prod_{j=1}^{n-m+2} S_j \to A$ where $g^*(s_1^*,\ldots,s^*_{n-m+2}) = g(s_1^*,\ldots,s^*_{n-m+2})$, one may easily verify that g^* implements f^* for strong equilibrium. But we have now succeeded in reducing the number of players by m-2. Continuing iteratively, one can reduce the number of players to the number of alternatives. At this point, the argument from the beginning of the proof applies.

Q.E.D.

Theorem 2 is false when the number of alternatives is exactly two, as the following simple example shows. Let n=2, $A = \{a,b\}$ and take $\mathcal{R}_1 = \mathcal{R}_2 = \left(\begin{matrix} a \\ b \end{matrix}, \begin{matrix} b \\ a \end{matrix}\right)$. Let f_{MAJ} be the majority rule SCC. That is, let

$$f_{MAJ}\left(\begin{matrix} a \\ b \end{matrix}, \begin{matrix} a \\ b \end{matrix}, \begin{matrix} a \\ b \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} b \\ a \end{matrix}, \begin{matrix} a \\ b \end{matrix}, \begin{matrix} a \\ b \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} a \\ b \end{matrix}, \begin{matrix} b \\ a \end{matrix}, \begin{matrix} a \\ b \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} a \\ b \end{matrix}, \begin{matrix} a \\ b \end{matrix}, \begin{matrix} b \\ a \end{matrix}\right) = \{a\}$$

$$f_{MAJ}\left(\begin{matrix} b \\ a \end{matrix}, \begin{matrix} b \\ a \end{matrix}, \begin{matrix} b \\ a \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} b \\ a \end{matrix}, \begin{matrix} a \\ b \end{matrix}, \begin{matrix} b \\ a \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} b \\ a \end{matrix}, \begin{matrix} b \\ a \end{matrix}, \begin{matrix} a \\ b \end{matrix}\right) = f_{MAJ}\left(\begin{matrix} a \\ b \end{matrix}, \begin{matrix} b \\ a \end{matrix}, \begin{matrix} b \\ a \end{matrix}\right) = \{b\}$$

The following game form implements f_{MAJ} for strong equilibrium, where player 1 chooses rows, 2, columns, and 3, matrices:

a	a
a	b

a	b
b	b

3. Individual Rationality

No veto power, although appealing, may not always make sense. In some circumstances, one may wish to guarantee players payoffs which leave them no worse off then their initial welfare levels; i.e., one may require the SCC to be individually rational. In such cases, players must be able to veto alternatives which entail net losses. In this section I investigate the set of individually rational SCC's which can be implemented for strong equilibria. It will turn out that this set is a singleton, consisting only of the SCC which for any preference profile, selects all individually rational Pareto optima. I first show that this SCC can indeed be implemented.

Theorem 3: Let $a_o \in A$ be identified as the status quo, and let $f_Q : \mathcal{R}_A^n \rightrightarrows A$ be the SCC such that $\forall (R_1,\ldots,R_n) \in \mathcal{R}_A^n$.

$$f_Q(R_1,\ldots,R_n) = \{a \in A \mid aR_i a_o \ \forall i, \text{ a Pareto optimal with respect to } (R_1,\ldots,R_n)\} .$$

Then f_Q may be implemented.

<u>Proof</u>: The proof is constructive. Suppose that f_Q is the SCC described in the hypotheses. With each $a \in A$ associate a strategy $s(a)$. Take $s_1 = \ldots = s_n = \{s(a) \mid a \in A\}$. Define $g : S_1 \times \ldots \times S_n \to A$ so that

$$g(s_1,\ldots,s_n) = \begin{cases} a, & \text{if } s_1 = s_2 = \ldots = s_n = s(a) \\ a_o, & \text{otherwise} \end{cases}$$

Now if $a \in f(R_1,\ldots,R_n)$, then I claim that $\bar{s} = (\bar{s}_1,\ldots,\bar{s}_n) = (s(a),\ldots,s(a))$ is a strong equilibrium. Clearly $g(\bar{s}) = a$. No coalition C smaller than the grand coalition can improve itself by deviating from $\bar{s}_C$ since deviating yields a_o, which, by individual rationality, no one prefers to a. On the other hand, the grand coalition cannot improve itself because a is Pareto efficient. Therefore $f(R_1,\ldots,R_n) \subseteq SE_g(R_1,\ldots,R_n)$. Now suppose $\bar{s}$ is a strong equilibrium with respect to $(R_1,\ldots,R_n)$. $g(\bar{s})$ is obviously Pareto optimal. Therefore, if $g(\bar{s}) = a_o$, then $a_o \in f(R_1,\ldots,R_n)$. If $g(\bar{s}) = a \neq a_o$, then aR_ia_o for all i since anyone can force the outcome a_o. Therefore $a \in f(R_1,\ldots,R_n)$, and $SE_g(R_1,\ldots,R_n) \subseteq f(R_1,\ldots,R_n)$.

Q.E.D.

Next I show that f_Q is the only implementable SCC which is individually rational.

<u>Theorem 4</u>: If $f : \mathcal{R}_A \times \ldots \times \mathcal{R}_A \rightrightarrows A$ is individually rational and implementable for strong equilibrium, then $f = f_Q$.

<u>Proof</u>: Suppose not. Then there exists an implementable $f \neq f_Q$. This implies that there exist $(R_1,\ldots,R_n) \in \mathcal{R}_A^n$ and $a \in A$ such that a is Pareto optimal for $(R_1,\ldots,R_n)$, $aR_i a_o$ for all i, and yet $a \notin f(R_1,\ldots,R_n)$. Let $g : \Pi S_j \to A$ implement f for strong equilibrium. $\forall\ \bar{s} \in \Pi S_j$ such that $g(\bar{s}) = a$, there exists a coalition C such that $\exists\ b \in A\ \exists\ s_C \in \prod_{j\in C} S_j$ for which

(1) $\forall j \in C\ bP_j a$;

(2) $g(s_C, \bar{s}_{-C}) = b$.

Choose $(\bar{R}_1,\ldots,\bar{R}_n) \in \mathcal{R}_A^n$ such that

(3) $(\bar{R}_1,\ldots,\bar{R}_n) : A \setminus (\{a_o\} \cup D) = (R_1,\ldots,R_n) : A \setminus (\{a_o\} \cup D)$ 5/

where $D = \{b \mid bI_i a \text{ for all } i\}$.

(4) $\forall b\ aP_i b \Rightarrow a_o \bar{P}_i b_i$

(5) $\forall i\ a\bar{P}_i b$ for all $b \in D$.

Suppose that $\bar{s} \in SE_g(\bar{R}_1,\ldots,\bar{R}_n)$. If $g(\bar{s}) = a$, then from the above argument there exist a coalition C, alternative b, and deviation $s_C \in \prod_{j\in C} S_j$ such that (1) and (2) are satisfied, a contradiction of $\bar{s}$'s being an equilibrium. Suppose $g(\bar{s}) = b \neq a$, $b \notin D$. Because a is Pareto optimal for $(R_1,\ldots,R_n)$, there exists $i \in \{1,\ldots,n\}$ such that aP_ib. From (4), $a_o\bar{P}_ib$. Therefore b is not an individually rational outcome, contradicting the hypotheses on f. If $g(\bar{s}) \in D$, then from (5), $g(\bar{s})$ is a Pareto inefficient outcome for $(\bar{R}_1,\ldots,\bar{R}_n)$, and therefore $\bar{s}$ cannot be a strong equilibrium. Thus, in all cases, contradictions arise. So $SE_g(\bar{R}_1,\ldots,\bar{R}_n) = \phi$.

Footnotes

1. Throughout, P_i and I_i shall denote, respectively, the strong and indifference relations associated with R_i.

2. $g(s_i, \bar{s}_{-i}) = g(\bar{s}_1,\ldots,\bar{s}_{i-1}, s_i, \bar{s}_{i+1},\ldots,\bar{s}_n)$.

3. $g(s_C, \bar{s}_{-C}) = g(s)$ where $s_i = \begin{cases} s_i, & i \in C \\ \bar{s}_i, & i \notin C \end{cases}$

4. If $p = 1$, let $a(p-1) = a(n)$.

5. The notation $R : T$ denotes "R restricted to the set T."

Reference

[1] Maskin, E. S. , "Nash Equilibrium and Welfare Optimality," forthcoming in <u>Mathematics of Operations Research</u>.