JOURNAL OF ECONOMIC THEORY **68**, 425–439 (1996)
ARTICLE NO. 0024

Nash Implementation of Matching Rules

TARIK KARA AND TAYFUN SÖNMEZ*

Department of Economics, University of Rochester, Rochester, New York 14627

Received December 21, 1993; revised January 23, 1995

We consider the Nash implementation of Pareto optimal and individually rational solutions in the context of matching problems. We show that all such rules are supersolutions of the stable rule. Among these solutions, we show that the "lower bound" stable rule and the "upper bound" Pareto and individually rational rule are Nash implementable. The proofs of these results are by means of a recent technique developed by Danilov [2]. Two corollaries of interest are the stable rule is the minimal implementable solution that is Pareto optimal and individually rational and the stable rule is the minimal Nash implementable extension of any of its subsolutions. *Journal of Economic Literature Classification Numbers*: C78, D78.

1. INTRODUCTION

In a public decision problem, agents usually have private information about their own preferences. In most cases, it may be unreasonable to expect truthful revelation. Nevertheless, it may be possible to reach the allocations selected by a solution in spite of the fact that agents behave strategically, by confronting them with an appropriately constructed "game form". This possibility is the motivation for the *theory of implementation.* A *game form* consists of sets of possible actions for each agent (the strategy spaces) and an *outcome function*, a function which associates with each profile of actions an alternative. Each agent can choose from his strategy space so as to influence the resulting outcome in his favor. The objective is to find a game form such that at equilibrium the desired alternatives are obtained. The equilibrium concept we consider is *Nash equilibrium.* (For expositions of implementation theory see Maskin [10], Moore [11], and

* We thank William Thomson for his enormous efforts in supervision. We are grateful to José Alcalde, Jeffrey Banks, Bhaskar Dutta, Alvin Roth, Rangarajan Sundaram, Manabu Toda, the participants of workshops at Bilkent University and the University of Rochester, and in particular to Stephen Ching and an associate editor for their helpful comments. All errors are our own responsibility.

0022-0531/96 $18.00

Thomson [24]). In this paper we are concerned with the (Nash) implementation of Pareto optimal and individually rational social choice rules in the context of matching problems.

A *matching problem* consists of two finite disjoint sets of agents, and a preference relation for each agent. The two disjoint sets are interpreted as a set of men, denoted by M, and a set of women, denoted by W. Each man has a preference relation over the set of women and staying single. Similarly each woman has a preference relation over the set of men and staying single. A *matching* is a function which assigns each man at most one woman and each woman at most one man such that if a man, m, is assigned to a woman, w, then the woman w is assigned to the man m. In this case m and w are each others' mates. If a man is not assigned to any woman then he stays single. Similarly if a woman is not assigned to any man then she stays single.

A matching is *stable*[1] if no agent prefers staying single to his or her mate (i.e., if it is individually rational), and there does not exist any pair of man and woman who prefer each other to their mates. The game theoretic analysis of matching problems starts with Gale and Shapley [5]. They show that each matching problem has at least one stable matching. They also show that when preferences are strict there is a stable matching which is weakly preferred by all men to any other stable matching. This matching is referred as the *man-optimal stable matching*. There is an analogous stable matching which is referred as the *woman-optimal stable matching*. (For an exposition of game theoretic modeling and analysis of matching problems see Roth and Sotomayor [18].)

In this paper we consider the case where the set of men and the set of women are fixed. Therefore each matching problem is represented by its preference profile. A *matching rule* is a correspondence which assigns to each matching problem (hence each preference profile) a set of matchings. A matching rule is a *social choice rule* (or simply a *rule*) in the context of matching problems. Some of the matching rules of interest are: The *stable rule*, which assigns the set of *stable* matchings to each preference profile; the *man-optimal stable rule*, which assigns the *man-optimal stable* matching; the *individually rational rule*, which assigns the set of *individually rational matchings*; the *Pareto rule*, which assigns the set of *Pareto optimal matchings*; and the *Pareto and individually rational rule*, which assigns the set of *Pareto optimal and individually rational matchings*. Each of these matching rules has some appealing properties.

Maskin [9] shows that a *monotonicity* condition is necessary for a rule to be implementable. This condition requires a rule to be such that: If an

[1] See Demange *et al.* [3], Gale and Sotomayor [6, 7], Knuth [8], and Roth [13, 15] for an extensive analysis of stable matchings.

alternative is recommended by the rule for some preference profile and all agents weakly improve the relative ranking of that alternative, then that alternative should be recommended by the rule for the new preference profile. A rule satisfies *no veto power* if an alternative is recommended by the rule whenever every agent but possibly one ranks that alternative at the top of their preferences. Maskin [9] also shows that *monotonicity* and *no veto power* together are sufficient conditions for implementability. He provides a game form to implement the given rule whenever it is *monotonic* and satisfies *no veto power*.

If a rule is not *monotonic* (hence not implementable) it is natural to ask how close it is to being *monotonic*. One appealing way to proceed is to expand the rule such that the resulting rule is *monotonic*. Sen [20] introduces the concept of a "minimal monotonic extension" of a rule and shows that this concept is well defined. (To see some applications of the concept of *minimal monotonic extensions* in the classical domain see Thomson [23].)

Many interesting rules are *monotonic* yet do not satisfy *no veto power*. Therefore Maskin's original result does not apply to such cases. Danilov [2] and Moore and Repullo [12] provide necessary and sufficient conditions for implementation. Danilov's condition is a certain monotonicity property. Moore and Repullo's condition is a condition which states the existence of particular sets. Yet Moore and Repullo do not provide a method to construct these sets. Sjöström [21] introduces an algorithm to check these necessary and sufficient conditions.

In this paper, we are interested in the implementation of Pareto optimal and individually rational matching rules (i.e., we are interested in the implementation of subsolutions of the Pareto and individually rational rule). Some examples of such rules are, the man-optimal stable rule, the stable rule, and the Pareto and individually rational rule itself. Are these rules implementable? Not all of them. In particular the man-optimal stable rule is not implementable. We show that, if a subsolution of the Pareto and individually rational rule is *monotonic*, then it should be a supersolution of the stable rule. That is, it should select all the stable matchings and possibly some other matchings. As *monotonicity* is a necessary condition for implementability, we need to concentrate on the supersolutions of the stable rule. A natural point of departure is of course the stable rule itself.

Even though the stable rule is *monotonic* it does not satisfy *no veto power*. Hence it is not possible to use Maskin's result to determine whether the stable rule is implementable. However, we show that (when there are at least three agents) the stable rule is implementable using Danilov's [2] and Yamato's [26] results. One corollary to these results is that the stable rule is the "minimal" implementable subsolution of the Pareto and individually rational rule in the sense that any subsolution of the Pareto and individually rational rule that is implementable is a supersolution of the stable rule. Another corollary is that the stable rule is the minimal implementable extension of any of its subsolutions. In particular it is the minimal implementable extension of the man-optimal stable rule.

The stable rule has very appealing properties. As already noted, it is Pareto optimal and individually rational. Furthermore it recommends the set of core matchings for each preference profile.[2] Our results further increase the importance of the stable rule.

What about the Pareto and individually rational rule itself? Before answering this question, let us first consider separately the Pareto rule and the individually rational rule. The Pareto rule is *monotonic* and it satisfies *no veto power*. Therefore Maskin's result applies and it can be implemented. The individually rational rule is *monotonic*; yet it does not satisfy *no veto power*. Nevertheless, Alcalde [1] shows that the individually rational rule is implementable and he provides a class of game forms to implement it. In this paper we provide a simple game form that achieves the same purpose.

Monotonicity is preserved under intersection (as long as the intersection is welldefined). Therefore the Pareto and individually rational rule is *monotonic*. Yet it fails to satisfy no veto power. Nevertheless, using Danilov's [2] result once again, we show that (as long as there are at least three agents) it is implementable. In this paper we use Danilov's [2] and Yamato's [26] results to show the implementability of the stable rule and the Pareto and individually rational rule. Nevertheless we have alternative proofs based on the results of Moore and Repullo [12] and Sjöström [21]. To our knowledge, this paper is one of the first papers to present some applications of these (full characterization implementation) results. We believe it provides some support for the importance of these results.

2. Definitions and Notation

We start by defining implementability and related concepts in the general social choice theory framework. We will make use of these concepts in the framework of matching problems.

The set of alternatives is A. The set of agents is $N=\{1, 2, ..., n\}$. For each agent $i \in N$, $\mathscr{R}_i$ is the set of his possible preference relations. Let $\mathscr{R} = \prod_{i \in N} \mathscr{R}_i$. The **lower contour set of R_i at $a \in A$** is $L(a, R_i) = \{b \in A \mid aR_ib\}$.

A **social choice rule** (or simply a **rule**) is a correspondence: $\varphi\colon \mathscr{R} \to A$. A preference profile $\tilde{R}$ is obtained by a **monotonic transformation of R at $a \in A$**, if $L(a, R_i) \subseteq L(a, \tilde{R}_i)$ for all $i \in N$. Let $MT(R, a)$ denote the set of preference profiles which are obtained by a *monotonic* transformation of R

[2] See Roth [16].

at a. A rule, φ, is **monotonic** if for all $R, \tilde{R} \in \mathscr{R}$, and for all $a \in \varphi(R)$, if $\tilde{R} \in MT(R, a)$, then $a \in \varphi(\tilde{R})$. A rule φ satisfies **no veto power** if, for all $i \in N$, and for all $R \in \mathscr{R}$, if $A = L(a, R_j)$ for all $j \in N \setminus \{i\}$, then $a \in \varphi(R)$.

A **game form** is a pair $\Gamma = (S, h) = (\prod_{i \in N} S_i, h)$, where S_i is agent *i's* **strategy space**, and h: $S \to A$ is an **outcome function**. The pair (Γ, R) defines a **game**. Let $N(\Gamma, R)$ denote the set of pure strategy Nash equilibria for the game (Γ, R). The game form Γ **implements φ (in pure strategy Nash equilibria)**, if $h(N(\Gamma, R)) = \varphi(R)$ for all $R \in \mathscr{R}$.

Maskin [9] shows that *monotonicity* is a necessary condition for implementability. He further shows that *monotonicity* and *no veto power* together are sufficient for implementability.[3]

A **matching problem** is an ordered triplet (M, W, R), where M and W are two non-empty, finite and disjoint sets, $R = (R_i)_{i \in M \cup W}$ is a preference profile such that R_i is a linear order on $M \cup \{i\}$ if $i \in W$ and on $W \cup \{i\}$ if $i \in M$. We call M the set of men and W the set of women. Given a set of men, M, and a set of women, W, a **matching on $M \cup W$**, μ, is a one-to-one function from the set $M \cup W$ onto itself of order two (that is $\mu^2(i) = i$) such that for any $m \in M$ if $\mu(m) \neq m$ then $\mu(m) \in W$ and for any $w \in W$ if $\mu(w) \neq w$ then $\mu(w) \in M$. We refer to $\mu(i)$ as the **mate of *i***. We denote the set of all matchings on $M \cup W$ by $\mathscr{M}$.

In the present context of matching problems, $N = M \cup W$ and A is the set of matchings $\mathscr{M}$. Given a preference relation of a man $m \in M$ we extend his preference, initially defined over the set $W \cup \{m\}$, to the set of matchings $\mathscr{M}$, in the following natural way: m prefers the matching μ to the matching μ' if and only if he prefers his mate under μ to his mate under μ'. We use R_m to represent both. The same can be done for each woman $w \in W$. We also consider the case where M and W are fixed. Therefore a preference profile defines a matching problem.

A matching μ is **blocked by individual *i* under *R*** if $iP_i\mu(i)$, i.e., if individual i prefers to stay single to being matched to $\mu(i)$. Otherwise $\mu(i)$ is said to be **acceptable** to i. The matching μ is **blocked by the pair (*m*, *w*) under *R*** if

$$wP_m\mu(m) \qquad \text{and} \qquad mP_w\mu(w),$$

i.e., the man and the woman, m and w, prefer each other to their mates. The matching μ is **individually rational for *R*** if it is not blocked by any agent $i \in M \cup W$, i.e., each agent is acceptable to his/her mate. We denote the set of all individually rational matchings for R by $\mathscr{I}(R)$. A matching μ is **stable under *R*** if it is not blocked by any individual or any pair. We

[3] See also Saijo [19].

denote the set of all stable matchings under R by $\mathcal{S}(R)$. Gale and Shapley [5] show that for any R there exists a matching $\mu_M(R) \in \mathcal{S}(R)$ such that

$$\forall \mu \in \mathcal{S}(R), \forall m \in M \quad \mu_M(R)(m)\ R_m \mu(m).$$

We call the matching $\mu_M(R)$, **the man-optimal stable matching for the preference profile *R***. The **women-optimal stable matching** for the preference profile R, $\mu_W(R)$ is defined similarly. A matching μ is **Pareto optimal under *R*** if there is no other matching μ' such that $\mu'(i)\ R_i \mu(i)$ for all $i \in M \cup W$ and $\mu'(j)\ P_j \mu(j)$ for some $j \in M \cup W$. A matching μ is **weakly Pareto optimal under *R*** if there is no other matching μ' such that $\mu'(i)\ P_i \mu(i)$ for all $i \in M \cup W$. We denote the set of all Pareto optimal matchings under R by $\mathcal{P}(R)$ and the set of all weakly Pareto optimal matchings under R by $\mathcal{P}^w(R)$.

In the context of matching problems, rules will be referred to as matching rules. An example of such a rule is the matching rule which associates the man-optimal stable matching, $\mu_M(R)$, to each preference profile $R \in \mathcal{R}$. We call this rule the man-optimal stable rule and denote it by μ_M. We define the woman-optimal stable rule analogously and denote it by μ_W. Other examples of matching rules are the stable rule, which associates the set of the stable matchings $\mathcal{S}(R)$, the Pareto rule, which associates the set of Pareto optimal matchings $\mathcal{P}(R)$, the individually rational rule, which associates the set of individually rational matchings $\mathcal{I}(R)$, and the Pareto and individually rational rule, which associates the set of Pareto optimal and individually rational matchings $\mathcal{P}(R) \cap \mathcal{I}(R)$, for each preference profile $R \in \mathcal{R}$. We denote the stable rule by $\mathcal{S}$, the Pareto rule by $\mathcal{P}$, the individually rational rule by $\mathcal{I}$, and the Pareto and individually rational rule by $\mathcal{PI}$.

3. Monotonic Matching Rules

It is easy to see that the stable rule is *monotonic*. Let the matching μ be stable under R and let R' be a monotonic transformation of R at μ. Suppose μ is not stable under R'. Then either an agent i or a pair (m, w) blocks μ under R'. If the former holds then the agent i blocks μ under R as the relative ranking of $\mu(i)$ is weakly better under R'. If the latter holds then the pair (m, w) blocks μ under R with the same reasoning. As both cases contradict μ being stable under R, μ should be stable under R' showing that the stable rule is *monotonic*.

A next natural question is whether there are other subsolutions of the Pareto and individually rational rule that are *monotonic*. Theorem 1 concerns *monotonic* subsolutions of the Pareto and individually rational rule.

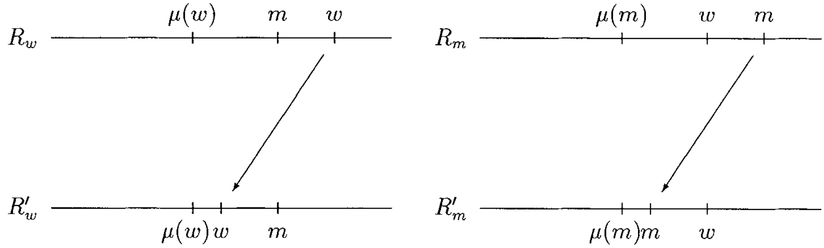

FIGURE 1

It states and proves that any solution that is Pareto efficient, individually rational, and *monotonic* is a supersolution[4] of the stable rule. Therefore, if there is any hope of implementing a Pareto efficient and individually rational rule, it is the stable rule and its supersolutions.

THEOREM 1. *Let $\varphi \subseteq \mathscr{P}\mathscr{I}$ be monotonic. Then $\mathscr{S} \subseteq \varphi$.*[5]

Proof. Let $\varphi \subseteq \mathscr{P}\mathscr{I}$ be *monotonic*. Let $R \in \mathscr{R}$ and $\mu \in \mathscr{S}(R)$. We need to show that $\mu \in \varphi(R)$.

Let $R' \in \mathscr{R}$ be as follows (see Figure 1):

1. For all $m \in M$:
   (a) For all $w, w' \in W$,

$$wP'_m w' \Leftrightarrow wP_m w'.$$

   (b) $\mu(m)\, R'_m m$ and $\nexists w \in W \setminus \{\mu(m)\}$ such that $\mu(m)\, R'_m w R'_m m$.
2. For all $w \in W$:
   (a) For all $m, m' \in M$,

$$mP'_w m' \Leftrightarrow mP_w m'.$$

   (b) $\mu(w)\, R'_w w$ and $\nexists m \in M \setminus \{\mu(w)\}$ such that $\mu(w)\, R'_w m R'_w w$.

Note that $R' \in MT(R, \mu)$ and $\mathscr{S}$ is *monotonic*, therefore $\mu \in \mathscr{S}(R') \subseteq \mathscr{P}\mathscr{I}(R')$. We also have $R \in MT(R', \mu)$.

Let $m \in M$ and $w \in W$ be such that $wP'_m \mu(m)$. Then, since $\mu \in \mathscr{S}(R')$ we have $\mu(w)\, P'_w m$ and hence by construction $(2b)$ we have $wP'_w m$. Therefore

$$\forall \mu' \in \mathscr{I}(R'), \quad \forall m \in M \quad \mu(m)\, R'_m \mu'(m).$$

[4] A rule ψ is a supersolution of the rule φ if $\varphi \subseteq \psi$.

[5] The earlier version of this theorem states that the stable rule is the only subsolution of itself, which is *monotonic*. Comments of an associate editor led us to discover this stronger version.

Similarly

$$\forall \mu' \in \mathcal{I}(R'), \quad \forall w \in W \quad \mu(w)\, R'_w \mu'(w).$$

This, together with $\mu \in \mathcal{P}\mathcal{I}(R')$ implies that $\{\mu\} = \mathcal{P}\mathcal{I}(R')$ and therefore $\varphi(R') = \{\mu\}$. But $R \in MT(R', \mu)$ and φ is *monotonic*, therefore $\mu \in \varphi(R)$.

Theorem 1 states that if we want to implement subsolutions of the Pareto and individually rational rule we need to concentrate on the supersolutions of the stable rule.

One of the matching rules that is often analyzed in the literature is the man-optimal stable rule (woman-optimal stable rule). The counterpart of this rule in the context of many-to-one matching problems is used to match medical interns and hospitals by the National Resident Matching Program in United States.[6] One of the corollaries of Theorem 1 is that the man-optimal stable rule (woman-optimal stable rule) is not *monotonic*.[7] Therefore the man-optimal stable rule is not implementable, and hence it is natural to ask how close it is to being implementable. One appealing operation consists of expanding the rule such that the resulting rule is implementable, and therefore *monotonic*. Among the possible expansions a minimal one is the most desirable. This is the motivation for the following definition from Sen [20]. Given $\varphi: \mathcal{R} \to A$, the **minimal monotonic extension of φ**, $mme(\varphi)$, is defined by

$$mme(\varphi) = \bigcap \{\psi: \mathcal{R} \to A \mid \psi \supseteq \varphi, \text{ where } \psi \text{ is } monotonic\}.$$

Note that $mme(\varphi)$ is well defined as the **feasibility rule**, the rule that assigns the set of all alternatives A for each preference profile, is *monotonic*. We also have $mme(\varphi)$ *monotonic* as *monotonicity* is closed under intersection.

A corollary to Theorem 1 is that the stable rule is the minimal monotonic extesion of any of its subsolutions. In particular, it is the minimal monotonic extension of the man-optimal stable rule (woman-optimal stable rule).

COROLLARY 1. *Let $\varphi \subseteq \mathcal{S}$ be monotonic. Then $\varphi = \mathcal{S}$.*[8]

Proof. Follows from *monotonicity* of the stable rule and Theorem 1.

Corollary 1 is of independent interest. One may be interested in implementation of only the subsolutions of the stable rule. Consider the

[6] See Roth [14].

[7] Tadenuma [22] shows this independently.

[8] The proof of Lemma 3.2 in Toda [25] can be interpreted as an independent proof of Corollary 1.

hospital-intern market in the United States. Interns can decline a job to which they had been matched by the National Resident Matching Program and arrange instead another job. Similarly hospitals can decline an intern with whom they had been matched in favor of another intern. Such situations can be avoided by implementing subsolutions of the stable rule. In fact in United Kingdom most of the unstable matching rules that are used to match interns and hospitals have been abandoned after a short period, whereas the stable matching rules have been in use in United States and United Kingdom for several decades. (See Roth [14, 17] for extensive analysis of medical intern-hospital markets in United States and United Kingdom.)

4. Nash Implementable Matching Rules

The next natural question is whether there are subsolutions of the Pareto and individually rational rule which are implementable. Due to Theorem 1, we need to concentrate on the supersolutions of the stable rule. A natural point of departure seems to be checking whether the stable rule itself is implementable.

Monotonicity is a necessary condition for implementability. *Monotonicity* together with *no veto power* are sufficient conditions for implementability. The stable rule is *monotonic*. Yet it does not satisfy *no veto power*. Hence we cannot determine whether the stable rule is implementable by appealing to Maskin's original result.

Nevertheless, the next theorem shows that the stable rule is implementable as long as there are at least three agents. We will need further notation and results to prove the theorem. For this, we will return to the general context.

Let $\varphi\colon \mathscr{R} \to A$ and $X \subseteq A$. An alternative $x \in X$ is **essential for agent $\boldsymbol{i \in N}$ in $\boldsymbol{X}$ for $\boldsymbol{\varphi}$** if

$$\exists R \in \mathscr{R} \quad L(x, R_i) \subseteq X \qquad \text{and} \qquad x \in \varphi(R).$$

We denote the set of essential alternatives for agent $i \in N$ in X for φ by $E(\varphi, i, X)$. A rule, φ, is **essentially monotonic** if for all R, $\tilde{R} \in \mathscr{R}$ and for all $a \in \varphi(R)$, if $E(\varphi, i, L(a, R_i)) \subseteq L(a, \tilde{R}_i)$ for all $i \in N$, then $a \in \varphi(\tilde{R})$. Danilov [2] shows that if $|N| \geqslant 3$ then φ is implementable if and only if it is **essentially monotonic**.

Danilov proves his result on a domain where preferences are linear orders on A. Yamato [26] generalizes his result as follows: Let $\mathscr{R}$ be such that, for all $a \in A$, $R \in \mathscr{R}$, $i \in N$, and $b \in L(a, R_i)$, there exists $R' \in \mathscr{R}$ such that $L(b, R'_i) = L(a, R_i)$ and for all $j \neq i$, $L(b, R'_j) = A$. Then, if $|N| \geqslant 3$, a social choice rule φ is implementable if and only if it is *essentially monotonic*.

Our domain does not satisfy Danilov's domain assumption. (This is due to the assumption that agents are indifferent between two matchings which assign them the same mate.) Nevertheless it does satisfy Yamato's domain assumption. Therefore in the following lemma and theorem we will refer to Yamato's result.

LEMMA 1. *For all $R \in \mathcal{R}$, $\mu \in \mathcal{S}(R)$, and $i \in M \cup W$ we have $E(\mathcal{S}, i, L(\mu, R_i)) = L(\mu, R_i)$.*

Proof. Let $R \in \mathcal{R}$, $\mu \in \mathcal{S}(R)$, $i \in M \cup W$.

1. $E(\mathcal{S}, i, L(\mu, R_i)) \subseteq L(\mu, R_i)$: Let $\mu' \in E(\mathcal{S}, i, L(\mu, R_i))$. Then there exists a preference profile $R' \in \mathcal{R}$ such that $L(\mu', R'_i) \subseteq L(\mu, R_i)$. Therefore $\mu' \in L(\mu, R_i)$.

2. $L(\mu, R_i) \subseteq E(\mathcal{S}, i, L(\mu, R_i))$: Without loss generality we assume $i = m \in M$. Let $\mu' \in L(\mu, R_m)$. Let $R' \in \mathcal{R}$ be such that:

1. (a) $\mu'(m) \, R'_m m$
   (b) $\forall w \in W, \; w R'_m \mu'(m)$.
2. $\forall m' \in M \backslash \{m\}$,
   (a) $\mu'(m') \, R'_{m'} m'$
   (b) $\forall w \in W \backslash \{\mu'(m')\}, \; m' R'_{m'} w$.
3. $\forall w \in W$,
   (a) $\mu'(w) \, R'_w w$
   (b) $\forall m' \in M \backslash \{\mu'(w)\}, \; w R'_w m'$.

We have $\mu' \in \mathcal{S}(R')$ and for all $\tilde{\mu} \in L(\mu', R'_m)$, $\tilde{\mu}(m) \in \{\mu'(m), m\}$. Therefore $\mu(m) \, R_m \tilde{\mu}(m)$ or equivalently $\tilde{\mu} \in L(\mu, R_m)$. Therefore $L(\mu', R'_m) \subseteq L(\mu, R_m)$. We also have $\mu' \in S(R')$, and hence $\mu' \in E(\mathcal{S}, m, L(m, R_m))$.

THEOREM 2. *Let $|M \cup W| \geqslant 3$. Then the stable rule, $\mathcal{S}$, is implementable.*

Proof. Lemma 1 with *monotonicity* of the stable rule implies the stable rule is *essentially monotonic*. Therefore the stable rule is implementable by Yamato [26].

Note that Theorem 2 and Theorem 1 together imply that the stable rule is the minimal implementable subsolution of the Pareto and individually rational rule. The next corollary states that (as long as there are three or more agents) the stable rule is the minimal implementable extension of any of its subsolutions.

COROLLARY 2. *Let* $|M \cup W| \geqslant 3$, *and* $\varphi \subseteq \mathcal{S}$ *be implementable. Then* $\varphi = \mathcal{S}$.

Proof. Follows from Theorem 2 and Corollary 1.

The next step is implementing the stable rule, whenever we have at least three agents. As we have satisfied the necessary and sufficient conditions noted in Danilov [2], the game form introduced in that paper can be used to implement the stable rule.

Moore and Repullo [12], Dutta and Sen [4], and Sjöström [21] provided necessary and sufficient condition, for implementability whenever there are two agents. Using their conditions it is straightforward to show that the stable rule is not implementable whenever $|M| = |W| = 1$. Nevertheless, we give a direct proof of this result.

PROPOSITION 1. *The stable rule is not implementable when* $|M| = |W| = 1$.

Proof. Let $M = \{m\}$ and $W = \{w\}$. Then $\mathcal{M} = \{\mu_1, \mu_2\}$ where μ_1 and μ_2 are such that $\mu_1(m) = w$ and $\mu_2(m) = m$. Let the preference relations R^1_m, R^2_m, R^1_w, and R^2_w be such that

$$w R^1_m m \quad m R^1_w w$$

$$m R^2_m w \quad w R^2_w m.$$

Note that $\mathcal{R} = \{(R^i_m, R^j_w)\colon i, j \in \{1, 2\}\}$.

Assume that a game form $\Gamma = (S, g)$ implements $\mathcal{S}$. Then there exists strategy profiles $s^{i,j} \in S$, $i, j \in \{1, 2\}$ such that $s^{i,j} \in N(\Gamma, R^i_m, R^j_w)$ and

$$g(s^{i,j}) = \begin{cases} \mu_1 & \text{if} \quad i = j = 1 \\ \mu_2 & \text{otherwise.} \end{cases}$$

Since $s^{1,2} \in N(\Gamma, R^1_m, R^2_w)$ we have

$$\forall s_m \in S_m \quad m = g(s^{1,2})(m)\; R^1_m\; g(s_m, s^{1,2}_w)(m).$$

Therefore $g(s_m, s^{1,2}_w)(m) = m$ for any $s_m \in S_m$. Similarly, since $s^{2,1} \in N(\Gamma, R^2_m, R^1_w)$ we have

$$\forall s_w \in S_w \quad w = g(s^{2,1})(w)\; R^1_w\; g(s_w, s^{2,1}_m)(w).$$

Therefore $g(s^{2,1}_m, s_w)(w) = w$ for any $s_w \in S_w$. Thus

$$\forall s_m \in S_m \quad m = g(s^{2,1}_m, s^{1,2}_w)(m)\; R^1_m\; g(s_m, s^{1,2}_w)(m) = m$$

$$\forall s_w \in S_w \quad w = g(s^{2,1}_m, s^{1,2}_w)(w)\; R^1_w\; g(s^{2,1}_m, s_w)(w) = w.$$

Thus $(s_m^{2,1}, s_w^{1,2}) \in N(\Gamma, R_m^1, R_w^1)$. But we have $g(s_m^{2,1}, s_w^{1,2}) = \mu_2 \notin \mathscr{S}(R_m^1, R_w^1)$, contradicting the assumption that Γ implements $\mathscr{S}$. Hence $\mathscr{S}$ is not implementable.

What about the Pareto and individually rational rule itself? Before answering this question, let us first consider the Pareto rule and the individually rational rule. The Pareto rule is *monotonic*. It also satisfies *no veto power*. Therefore Maskin's original result applies. The Pareto rule is implementable.[9] The individually rational rule is also *monotonic*. Yet it does not satisfy *no veto power*.[10] Therefore Maskin's original result does not apply. Alcalde [1] shows in a constructive way that the individually rational rule is implementable. His game forms are *preference revelation* games, that is, games in which each agent announces a preference relation. Our contribution here is to propose a simple game form to implement the individually rational rule in which each agent announces a mate.

Let $\Gamma = (S, h)$ be as follows: $S_m = W \cup \{m\}$ for all $m \in M$ and $S_w = M \cup \{w\}$ for all $w \in W$; h: $S \to \mathscr{M}$ is defined as

$$\forall s \in S, \quad \forall i \in M \cup W \qquad h(s)(i) = \begin{cases} s_i & \text{if} \quad s_{s_i} = i \\ i & \text{otherwise.} \end{cases}$$

Claim. The game form Γ implements the individually rational rule. (See the Appendix for a proof.)

The Pareto and individually rational rule is *monotonic* since both the Pareto and individually rational rules are *monotonic*. However the Pareto and individually rational rule does not satisfy *no veto power*, so that once again Maskin's original result does not apply. Nevertheless, the Pareto and individually rational rule is implementable as long as there are at least three agents, and the proof is similar to the proof of Theorem 2 (the only modification is replacing $\mathscr{S}$ with $\mathscr{PI}$ in the proof). The Danilov game form with $\varphi = \mathscr{PI}$ implements the Pareto and individually rational rule. The Pareto and individually rational rule is not implementable whenever $|M| = |W| = 1$.[11]

We summarize some of these results in Table I.

[9] Let the weak Pareto rule, $\mathscr{P}^w$ be the matching rule which associates the set of weak Pareto optimal matchings for each preference profile $R \in \mathscr{R}$. Similar results holds if the weak Pareto rule is replaced by the Pareto rule.

[10] This is due to the fact that, even if all agents but one rank a matching as one of their top choices, this matching will not be selected by the individually rational rule unless the remaining agent weakly prefers his/her mate to staying single.

[11] Similar results hold if the weak Pareto and individually rational rule is replaced by the Pareto and individually rational rule.

TABLE I

	No veto power	Monotonicity	Minimal monotonic extension	Nash implementable	Notes
$\mathscr{S}$	No	Yes	Itself	Yes	We use Danilov's results to show that $\mathscr{S}$ is implementable
μ_M	No	No	$\mathscr{S}$	No	Follows from Maskin's original result
$\varphi \subsetneqq \mathscr{S}$	No	No	$\mathscr{S}$	No	Follows from Maskin's original result
$\mathscr{P}$	Yes	Yes	Itself	Yes	Follows from Maskin's original result
$\mathscr{P}^w$	Yes	Yes	Itself	Yes	Follows from Maskin's original result
$\mathscr{I}$	No	Yes	Itself	Yes	We provide a simple game form to implement $\mathscr{I}$
$\mathscr{P}\mathscr{I}$	No	Yes	Itself	Yes	We use Danilov's result to show that $\mathscr{P}\mathscr{I}$ is implementable
$\mathscr{P}^w\mathscr{I}$	No	Yes	Itself	Yes	We use Danilov's result to show that $\mathscr{P}^w\mathscr{I}$ is implementable

Appendix

Here we propose a simple game form which implements the individually rational rule.

Let $\Gamma = (S, h)$, where $S = \prod_{i \in M \cup W} S_i$ is such that

$$\forall m \in M \quad S_m = W \cup \{m\} \qquad \text{and} \qquad \forall w \in W \quad S_w = M \cup \{w\}$$

and let $h\colon S \to \mathscr{M}$ be defined as

$$\forall s \in S, \quad \forall i \in M \cup W \qquad h(s)(i) = \begin{cases} s_i & \text{if} \quad s_{s_i} = i \\ i & \text{otherwise.} \end{cases}$$

Claim. The game form Γ implements the individually rational rule.

Proof.

1. $\mathscr{I}(R) \subseteq h(N(\Gamma, R))$: Let $\mu \in \mathscr{I}(R)$. Let $s^* \in S$ be such that for any $i \in M \cup W$, $s_i^* = \mu(i)$. Clearly we have $h(s^*) = \mu$. Let $i \in M \cup W$ and $s_i \in S_i \setminus \{s_i^*\}$. Since $s^*_{s_i^*} = i$ and $s_i \neq s_i^*$ we have $s_{s_i} \neq i$. Therefore $h(s_i, s^*_{-i})(i) = i$. Since $\mu \in \mathscr{I}(R)$ we have

$$\forall i \in N, \quad \forall s_i \in S_i \qquad h(s^*)(i) = \mu(i) \; R_i \, i = h(s_i, s^*_{-i})(i).$$

So $s^* \in N(\Gamma, R)$. Therefore we have $\mu \in h(N(\Gamma, R))$, and hence

$$\mathscr{I}(R) \subseteq h(N(\Gamma, R)).$$

2. $h(N(\Gamma, R)) \subseteq \mathscr{I}(R)$: Let $\mu \in h(N(\Gamma, R))$. Since $\mu \in h(N(\Gamma, R))$ there is a strategy profile $s^* \in S$ such that $h(s^*) = \mu$ and $s^* \in N(\Gamma, R)$. Note that by announcing $s_i = i$, agent i can stay single. Since s^* is a Nash equilibrium we have

$$\forall i \in M \cup W \qquad \mu(i) = h(s^*)(i) \; R_i h(i, s^*_{-i})(i) = i.$$

Thus $\mu \in \mathscr{I}(R)$. Hence $h(N(\Gamma, R)) \subseteq \mathscr{I}(R)$.

References

1. J. Alcalde, Implementation of stable solutions to marriage problems, mimeo, Universitat Autònoma de Barcelona, 1993, *J. Econ. Theory*, in press.
2. V. Danilov, Implementation via Nash equilibria, *Econometrica* **60** (1992), 43–56.
3. G. Demange, D. Gale, and M. Sotomayor, A further note on the stable matching problem, *Discrete Appl. Math.* **16** (1987), 217–222.
4. B. Dutta and A. Sen, A necessary and sufficient condition of two-person Nash implementation, *Rev. Econ. Stud.* **58** (1991), 121–128.
5. D. Gale and L. Shapley, College admissions and the stability of marriage, *Amer. Math. Monthly* **69** (1962), 9–15.
6. D. Gale and M. Sotomayor, Ms. Machiavelli and the stable matching problem, *Amer. Math. Monthly* **92** (1985), 261–268.
7. D. Gale and M. Sotomayor, Some remarks on the stable matching problem, *Discrete Appl. Math.* **11** (1985), 223–232.
8. D. Knuth, "Marriages stables," Les Presses de l'Université de Montréal, Montréal, 1976.
9. E. Maskin, Nash equilibrium and welfare optimality, mimeo, MIT, 1977.
10. E. Maskin, The theory of implementation in Nash equilibrium: A survey, "Social Goals and Social Organization: Volume in Memory of Elisha Pazner" (L. Hurwicz, D. Schmeidler, and H. Sonnenschein, Eds.), Cambridge Univ. Press, London/New York, 1985.
11. J. Moore, Implementation in environments with complete information, "Advances in Economic Theory" (J.-J. Laffont, Ed.), Cambridge Univ. Press, London/New York, 1992.
12. J. Moore and R. Repullo, Nash implementation: A full characterization, *Econometrica* **58** (1990), 1083–1099.

13. A. E. ROTH, The economics of matching: Stability and incentives, *Math. Oper. Res.* **7** (1982), 617–628.
14. A. E. ROTH, The evolution of the labor market for medical interns and residents: A case study in game theory, *J. Polit. Econ.* **92** (1984), 991–1016.
15. A. E. ROTH, Misrepresentation and stability in marriage problems, *J. Econ. Theory* **34** (1984), 383–387.
16. A. E. ROTH, Common and conflicting interests in two-sided matching markets, *Europ. Econ. Rev.* **27** (1985), 75–96.
17. A. E. ROTH, A natural experiment in the organization of entry-level labor markets: Regional markets for near physicians and surgeons in the United Kingdom, *Amer. Econ. Rev.* **81** (1991), 415–440.
18. A. E. ROTH AND M. SOTOMAYOR, "Two-sided Matching: A Study in Game Theoretic Modeling and Analysis," Cambridge Univ. Press, London/New York, 1990.
19. T. SAIJO, Strategy space reduction in Maskin's theorem: Sufficient conditions for Nash implementation, *Econometrica* **56** (1988), 693–700.
20. A. SEN, Approximate implementation of non-dictatorial social choice functions, mimeo, Princeton University, 1987.
21. T. SJŐSTRŐM, On the necessary and sufficient conditions for Nash implementation, *Soc. Choice Welfare* **8** (1991), 333–340.
22. K. TADENUMA, Maskin monotonicity in matching problems, mimeo, University of Rochester, 1993.
23. W. THOMSON, Minimal monotonic extensions, mimeo, University of Rochester, 1992.
24. W. THOMSON, Concepts of implementation, mimeo, University of Rochester, 1993.
25. M. TODA, Another characterization of the core of two-sided matching problems, Discussion Paper 9303, Tokyo Keizai University, 1993.
26. T. YAMATO, On Nash implementation of social choice correspondences, *Games Econ. Behav.* **4** (1992), 484–992.