Department of Economics

and

Center for Research on Economic and Social Theory

Working Paper Series

Strategy-Proofness in Many-to-One Matching Problems

Tayfun Sönmez

October, 1995
Number 95-01

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, MI 48109-1220

Strategy-Proofness in Many-to-One Matching Problems

Tayfun Sönmez*

Department of Economics

University of Michigan,

Ann Arbor, MI 48109-1220

Abstract

We search for ***strategy-proof*** solutions in the context of (many-to-one) matching problems (Gale and Shapley 1962). In this model, whenever the firms can hire as many workers as they want (the capacities are unlimited) the stable set is a singleton. There exists a ***Pareto efficient, individually rational,*** and ***strategy-proof*** matching rule if and only if the capacities are unlimited. Furthermore, whenever the capacities unlimited, the matching rule which selects the unique stable matching is the only matching rule that is ***Pareto efficient***, ***individually rational***, and ***strategy-proof.***

Journal of Economic Literature Classification Numbers: C71, C78, D71, D78

Key Words: Matching Problems, College Admissions Problems, Strategy-Proofness, Implementation, Stable Matching, Core.

*This paper is based on the fourth chapter of my dissertation submitted to the Department of Economics, University of Rochester. I wish to thank William Thomson for his efforts in supervision as well as his useful suggestions. I am grateful to José Alcalde, Salvador Barberà, Alvin Roth, James Schummer, Marilda Sotomayor, an associate editor and particularly to Stephen Ching, Tarık Kara and two anonymous referees for their helpful comments. Tarık Kara also significantly simplified the proof of Proposition 4. All errors are my own responsibility.

1 Introduction

A class of public decision problems that has been extensively analyzed is the class of ***two-sided matching problems.*** (For an exposition of game theoretic modelling and analysis of such problems, see Roth and Sotomayor 1990). In real life applications of these models, many-to-one matching is the most typical case, where one side (of the market) consists of institutions and the other side consists of individuals: colleges admit many students, firms hire many workers, and hospitals employ many interns. On the other hand, students attend one college, workers work for one firm, and interns work for one hospital.

In this paper we deal with many-to-one matching problems[1]. The class of one-to-one matching problems is a subclass of the class of many-to-one matching problems. For simplicity we refer to the institutions side of the market as "firms" and to the individuals side as "workers". Each firm has a preference relation over groups of workers of size at most its capacity, and each worker has a preference relation over firms and being unemployed. The preferences of the workers and the firms are strict. We furthermore assume that the preferences of the firms over groups of workers are ***separable*** (Barberà, Sonnenschein and Zhou 1991): consider two groups of workers such that one of the groups contains the other one and it is larger by one worker. Then the firm prefers the larger group if and only if the additional worker is acceptable by himself. An allocation in this context is a matching of firms and workers.

A matching is ***individually rational*** if no worker prefers being unemployed to his assignment and no firm prefers a subset of its assignment to its assignment. A matching is ***stable*** if it is ***individually rational***, and there does not exist a firm-worker pair such that the worker prefers the firm to his assignment and the firm prefers a subset of its assignment joined with the worker to its assignment.

A ***solution*** is a systematic procedure to select a set of matchings for each matching problem. We refer to solutions as ***matching rules.*** Some examples of matching rules

[1] Also known as ***college admissions problems.*** Gale and Shapley (1962) introduces the college admissions problems. Gale and Sotomayor (1985), Roth (1984,85a,b,86), Roth and Sotomayor (1989) study various aspects of the college admissions problems.

are the *stable rule*, which selects the set of stable matchings, the *Pareto rule*, which selects the set of *Pareto efficient* matchings, and the individually rational *rule*, which selects the set of *individually rational* matchings.

In a public decision problem, agents usually have private information about their own preferences. In most cases it may be unreasonable to expect them to reveal them truthfully. However there may be solutions which give agents the incentive to be truthful independently of the other agents' behavior; this property is known as *strategy-proofness*[2]. Roth (1982a) shows that there is no *strategy-proof* selection from the stable rule in the context of one-to-one matching. Alcalde and Barberà (1994) improve on this result and show that there is no *strategy-proof* matching rule that is *Pareto efficient* and *individually rational* in this context. The class of one-to-one matching problems is a subclass of the class of many-to-one matching problems (where the capacity of each firm is one). Therefore the negative results of Roth (1982a) and Alcalde and Barberà (1994) carry over to many-to-one matching problems.[3]

In this paper we show that there are significant changes in these results if the capacities of the firms are not binding. We consider matching problems with unlimited capacities. That is, each firm can hire as many workers as it wants and thus its capacity is (greater than or) equal to the number of workers. For matching problems with unlimited capacities, there exists a unique stable matching for each matching problem. The stable rule is a *strategy-proof* matching rule and furthermore it is the

[2] *Strategy-proofness* was first analyzed in abstract social choice models where there are few or no restrictions on preferences. Gibbard (1973) and Satterthwaite (1975) show that, under minor conditions *strategy-proofness* is equivalent to dictatorship. In models with more structure (such as economic models) some positive results are available. See for example Barberà, Gül and Stachetti (1993), Barberà and Jackson (1994) Ching (1993,94), Moulin (1980), and Sprumont (1991) for some positive results, Sprumont (1994), Thomson (1994) for surveys on recent results of *strategy-proofness*, and Barberà and Jackson (1995), Barberà and Peleg (1990), Hurwicz (1972), Satterthwaite and Sonnenschein (1981), Tadenuma and Thomson (1991), Zhou (1991) for further negative results.

[3] Kara and Sönmez (1993,94) weakens the incentive requirement and search for Nash implementable matching rules. They show that the stable rule is the minimal Nash implementable solution that is *Pareto efficient* and *individually rational*.

only matching rule that is *Pareto efficient, individually rational,* and *strategy-proof.* Unfortunately it is not *coalitionally strategy-proof*[4] and hence there is no matching rule that is *Pareto efficient, individually rational,* and *coalitionally strategy-proof.* It is important to note that the core correspondence coincides with the stable rule for this model. We also show the independence of the axioms in the characterization result by means of simple examples.

In Section 4 we relax the unlimited capacities assumption. We show that the unlimited capacity condition is "tight" for the existence of a matching rule that is *Pareto efficient, individually rational,* and *strategy-proof.* That is, whenever there is a firm with a smaller capacity than the number of the workers, there is no matching rule that is *Pareto efficient, individually rational,* and *strategy-proof.* Therefore, there exists a matching rule with these properties if and only if the capacities are unlimited. Another corollary is that whenever there is a matching rule satisfying *Pareto efficiency, individual rationality,* and *strategy-proofness*, it is unique.

2 The Model

A **(many-to-one) matching problem** is a four-tuple (F, W, q, R). The first two components are non-empty, finite, and disjoint sets of firms and workers $F = \{f_1, \ldots, f_n\}$ and $W = \{w_1, \ldots, w_m\}$. We assume that there are at least two firms and two workers. The third component is a vector of positive natural numbers $q = (q_{f_1}, \ldots, q_{f_n})$, where q_{f_i} is the capacity of firm $f_i \in F$. The last component $R = (R_i)_{i \in F \cup W}$ is a list of preference relations of firms and workers. Let P_i denote the strict relation associated with the preference relation R_i for all $i \in F \cup W$. We consider the case where F, W and q are fixed and hence each preference profile defines a matching problem.

The preference relation R_w of worker $w \in W$ is a linear order on $\Sigma_w = \{\{f_1\}, \ldots, \{f_n\}, \emptyset\}$. Let $\mathcal{R}_w$ be the class of all such preference relations for worker $w \in W$. The preference relation R_f of firm $f \in F$ is a linear order on $\Sigma_f^q = \{G\}_{\substack{G \subseteq W \\ |G| \leq q_f}}$

[4] A matching rule is *coalitionally strategy-proof* if no group of agents can benefit by jointly misrepresenting their preferences.

and it is **separable** (Barberà, Sonnenschein, and Zhou 1991). That is, for all $G \subseteq W$ such that $|G| < q_f$, for all $w \in W \backslash G$

$$(G \cup \{w\})P_f G \iff \{w\}P_f\emptyset.$$

Let $\mathcal{R}_f$ be the class of all such preferences for firm $f \in F$. Let $\mathcal{R} = \prod_{f\in F} \mathcal{R}_f \times \prod_{w\in W} \mathcal{R}_w$. For all $R \in \mathcal{R}$, $G \subset F \cup W$, we denote the restriction of R to G by R_G, and the set $(F \cup W)\backslash G$ by $-G$. For all $i \in F \cup W$, we denote the set $(F \cup W)\backslash\{i\}$ by $-i$.

The **choice** of a firm f from a group of workers $G \subseteq W$ under the preference R_f is defined as

$$Ch_f(R_f, G) = \{G' \subseteq G \mid G' \in \Sigma_f^q,\ G'R_fG'' \text{ for all } G'' \subseteq G \text{ such that } G'' \in \Sigma_f^q\}.$$

A **matching** μ is a mapping from the set $F \cup W$ into $2^{F\cup W}$ such that:

1. For all $w \in W$, $|\mu(w)| \leq 1$ and $\mu(w) = \emptyset$ if $\mu(w) \not\subseteq F$;
2. For all $f \in F$, $|\mu(f)| \leq q_f$ and $\mu(f) = \emptyset$ if $\mu(f) \not\subseteq W$;
3. For all $(f, w) \in F \times W$, $\mu(w) = \{f\}$ if and only if $w \in \mu(f)$.

For any $i \in F \cup W$, we refer to $\mu(i)$ as the **assignment** of i at μ. We denote the set of all matchings for the triple (F, W, q) by $\mathcal{M}(F, W, q)$.

Given a preference relation R_f of a firm $f \in F$, initially defined over Σ_f^q, we extend it to the set of matchings $\mathcal{M}(F, W, q)$, in the following natural way: f prefers the matching μ to the matching μ' if and only if it prefers its assignment under μ to its assignment under μ'. We slightly abuse the notation and also use R_f to denote this extension. The same can be done for each worker $w \in W$.

A firm $f \in F$ is **acceptable to a worker** $w \in W$ under R if $\{f\}P_w\emptyset$. A worker $w \in W$ is **acceptable to a firm** $f \in F$ under R if $\{w\}P_f\emptyset$. A matching μ is **blocked by a worker** $w \in W$ under R if $\mu(w)$ is not acceptable to w under R. A matching μ is **blocked by a firm** f under R if $\mu(f) \neq Ch_f(R_f, \mu(f))$. Note that this statement is equivalent to the following under separable preferences: A matching μ is blocked by a firm $f \in F$ under R if there is a worker $w \in \mu(f)$ which is not acceptable to f under R. A matching μ is **individually rational** under R if it is not blocked by a worker or a firm under R. We denote the set of all *individually rational* matchings under R by $\mathcal{I}(R)$.

A matching μ is **blocked by a firm-worker pair** $(f, w) \in F \times W$ under R if $\{f\}P_w\mu(w)$ and $\mu(f) \neq Ch_f(R_f, \mu(f) \cup \{w\})$. A matching μ is **stable** under R if it is not blocked by a worker, a firm, or a firm-worker pair. We denote the set of all stable matchings under R by $\mathcal{S}(R)$.

Let $A_f(R_f)$ denote the set of acceptable workers for firm $f \in F$ under $R_f \in \mathcal{R}_f$. Formally

$$A_f(R_f) = \{w \in W \mid \{w\}P_f\emptyset\} \qquad \text{for all } f \in F, \text{ for all } R_f \in \mathcal{R}_f.$$

Let $B_f(R_W)$ denote the set of workers for whom firm $f \in F$ is acceptable. Similarly, $B_w(R_F)$ denotes the set of firms for whom worker $w \in W$ is acceptable. Formally

$$\begin{aligned} B_f(R_W) &= \{w \in W \mid \{f\}P_w\emptyset\} \qquad \text{for all } f \in F, \text{ for all } R_W \in \mathcal{R}_W, \\ B_w(R_F) &= \{f \in F \mid \{w\}P_f\emptyset\} \qquad \text{for all } w \in W, \text{ for all } R_F \in \mathcal{R}_F. \end{aligned}$$

Note that $w \in A_f(R_f)$ if and only if $f \in B_w(R_F)$ for all $f \in F$, for all $w \in W$, and for all $R_F \in \mathcal{R}_F$.

A matching μ is **blocked by a coalition** $G \subseteq F \cup W$ via a matching μ' under R if

1. $\mu'(i) \subseteq G$ for all $i \in G$,
2. $\mu'(i)R_i\mu(i)$ for all $i \in G$,
3. $\mu'(j)P_j\mu(j)$ for some $j \in G$.

A matching μ is in the **core** of the matching problem $R \in \mathcal{R}$ if it is not blocked by any coalition via any matching.

A matching μ is **Pareto efficient** under R if there is no other matching μ' such that $\mu'(i)R_i\mu(i)$ for all $i \in F \cup W$ and $\mu'(j)P_j\mu(j)$ for some $j \in F \cup W$. We denote the set of all *Pareto efficient* matchings under R by $\mathcal{P}(R)$.

A **matching rule** is a correspondence $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$. A matching rule φ is **Pareto efficient** if $\varphi(R) \subseteq \mathcal{P}(R)$ for all $R \in \mathcal{R}$. A matching rule φ is **individually rational** if $\varphi(R) \subseteq \mathcal{I}(R)$ for all $R \in \mathcal{R}$. The next property concerns single-valued solutions. A matching rule φ is **strategy-proof** if

$$\varphi(R)(i)R_i\varphi(R_{-i}, R_i')(i) \quad \text{for all } R \in \mathcal{R}, \text{ for all } i \in F \cup W, \text{ for all } R_i' \in \mathcal{R}_i.$$

3 Results under Unlimited Capacities

In this section, we consider the model when $q_f \geq |W|$ for all $f \in F$. That is, each firm can employ as many workers as it wants. We refer to this case as matching problems with **unlimited capacities**. One application for this case is the model where each firm produces one good (not necessarily the same), it is price taker, labor is the only variable factor of production, and the marginal contribution of a worker to a firm is independent of the other workers employed by that firm. Another application is the membership recruitment of American sororities. Here, in most situations the number of available slots at any sorority is more than the number of the applicants. (See Mongell and Roth 1991.) In the next section, we will drop the unlimited capacities assumption.

For this model, the separability of the preferences of the firms implies that the matching μ is blocked by a pair $(f, w) \in F \times W$ under R if and only if

$$\{f\}P_w\mu(w) \quad \text{and} \quad \{w\}P_f\emptyset.$$

3.1 The Stable Rule

In general many-to-one matching problems, in which there is no restriction on the preferences of the agents, the set of stable matchings may be empty. Gale and Shapley (1962), Roth (1984), Roth and Sotomayor (1990) [Chapter 6, Section 1]

identify natural restrictions on preferences for which the set of stable matchings is non-empty. Proposition 1 concerns the set of stable matchings in matching problems with unlimited capacities. It says that with separable preferences not only the stable set is non-empty, but it is a singleton.

Proposition 1: The stable set $\mathcal{S}(R)$ is a singleton for each preference profile $R \in \mathcal{R}$ on the class of matching problems with separable preferences and unlimited capacities.

Proof: Let (F, W, q) be such that $q_{f_i} \geq |W|$ for all $f_i \in F$. Let $R \in \mathcal{R}$. Let $\mu \in \mathcal{M}(F, W, q)$ be such that, for all $w \in W$,

$$\mu(w) = \{f \in B_w(R_F) \mid \{f\}P_w\emptyset \text{ and } \{f\}R_w\{f'\} \text{ for all } f' \in B_w(R_F)\}.$$

As preferences are strict, there exists a unique such matching.

We claim that μ is the unique stable matching under R. First we show that μ is stable under R. We have $\mu(w)R_w\emptyset$ for all $w \in W$. Furthermore $w \in \mu(f)$ only if $w \in A_f(R_f)$ for all $f \in F$. This, with preferences being separable, implies that μ is *individually rational* under R. Next consider any firm-worker pair $(f, w) \in F \times W$ such that $w \notin \mu(f)$. Then, either $w \notin A_f(R_f)$ and therefore $\emptyset P_f\{w\}$, or $\mu(w)P_w\{f\}$. Hence no worker-firm pair blocks μ under R. Therefore μ is stable under R.

Finally we show that μ is the only stable matching under R. Let $\mu' \in \mathcal{M}(F, W, q)$ and $\mu' \neq \mu$. If μ' is not *individually rational*, then μ' is not stable and we are done. Otherwise, there is a firm-worker pair $(f, w) \in F \times W$ such that $w \in A_f(R_f), w \notin \mu'(f)$ and $\{f\}P_w\mu'(w)$. Thus, $(\mu'(f) \cup \{w\})P_f\mu'(f)$ and $\{f\}P_w\mu'(w)$, and therefore the firm-worker pair (f, w) blocks μ' under R. Hence μ' is not stable under R. Q.E.D.

Let $\mathcal{S} : \mathcal{R} \to \mathcal{M}(F, W, q)$ be the matching rule which selects the set of stable matchings for each preference profile $R \in \mathcal{R}$. Due to Proposition 1, $\mathcal{S}$ is a single-valued matching rule on the class of matching problems with unlimited capacities. We call $\mathcal{S}$ the **stable rule**.

Remark 1: Proposition 1 still holds if the capacities of all firms are greater than or equal to the number of workers minus one. However as long as there is a firm, say firm f, which has a capacity less than the number of workers minus one, this is no longer true. In this case we can construct an example in which the top two choices of the firm f differ by (at least) two workers and the firm f is the only acceptable firm to any worker. In this example the matching which assigns the firm f its top choice and the emptyset to all other firms, and the matching which assigns the firm f its second choice and the emptyset to all other firms are both stable. Therefore on the class of matching problems with separable preferences the capacities of all firms being greater than or equal to the number of workers minus one is a necessary and sufficient condition for the stable rule to be single-valued.

Remark 2: Roth and Sotomayor (1990) [Theorem 6.7] shows that on the class of matching problems with substitutable preferences[5] (which is a larger class than matching problems with separable preferences) there is a stable matching which is preferred to any other stable matching by all firms. They refer to this matching as the firm-optimal stable matching and construct an algoritm (which is a variant of the deferred acceptance algoritm of Gale and Shapley) to obtain it. They also construct an algoritm to obtain an analagous stable matching, the worker-optimal stable matching, which is preferred to any other stable matching by all workers. One can obtain an alternative proof of Proposition 1 by observing that these two algoritms lead to the same matching when the preferences are separable and the capacities are unlimited.

3.2 Existence and Characterization

Roth (1982a) shows that in the context of one-to-one matching problems (also known as marriage problems) there is no matching rule which always selects a stable matching and is *strategy-proof.* Alcalde and Barberà (1994), improve on this result and show that there is no matching rule that is *Pareto efficient*, *individually rational*

[5]A firm f has substitutable preferences if whenever its preferred set of employees from a group G includes the workers w and w', its preferred set of employees from the group $G\backslash\{w'\}$ also includes w. Formally the preference relation R_f is substitutable if

$$\text{For all } G \subseteq W \text{ with } w, w' \in Ch_f(R_f, G), \quad w \in Ch_f(R_f, G\backslash w').$$

and *strategy-proof.* These impossibilities extend to many-to-one matching problems. However there are significant changes in these results when firms can hire as many workers as they want and their preferences are separable.

Proposition 2: The stable rule is *strategy-proof* on the class of matching problems with separable preferences and unlimited capacities.

Proof: Let (F, W, q) be such that $q_{f_i} \geq |W|$ for all $f_i \in F$. Let $R \in \mathcal{R}$, $\mu = \mathcal{S}(R)$,[6] and $w \in W$. We have $\mu(w) R_w \emptyset$ and $\mu(w) R_w \{f\}$ for all $f \in B_w(R_F)$. Furthermore $\mathcal{S}(R_{-w}, R'_w)(w) \subseteq B_w(R_F)$ for all $R'_w \in \mathcal{R}_w$. Therefore

$$\mathcal{S}(R)(w) = \mu(w) R_w \mathcal{S}(R_{-w}, R'_w)(w) \qquad \text{for all } R'_w \in \mathcal{R}_w$$

and hence no worker can benefit by unilateral deviation.

Let $f \in F$. Let $w \in W\backslash\mu(f)$ be such that $\{w\} P_f \emptyset$. Then we have $\mu(w) P_w \{f\}$. We have two cases to consider.

Case 1: $\mu(w) = \emptyset$ Let $R'_f \in \mathcal{R}_f$ and $\mu' \in \mathcal{M}(F, W, q)$ be such that $w \in \mu'(f)$. Then w blocks μ' under (R_{-f}, R'_f) and therefore $\mu' \notin \mathcal{S}(R_{-f}, R'_f)$.

Case 2: $\mu(w) \in F$

Let $R'_f \in \mathcal{R}_f$ and $\mu' \in \mathcal{M}(F, W, q)$ be such that $w \in \mu'(f)$. We have $\mu(w) P_w \{f\}$, $w \in A_{\mu(w)}(R_{\mu(w)})$, and therefore (f, w) blocks μ' under (R_{-f}, R'_f). Thus $\mu' \notin \mathcal{S}(R_{-f}, R'_f)$.

This, together with the individual rationality of the stable rule, implies that

$$\mathcal{S}(R)(f) = \mu(f) R_f \mathcal{S}(R_{-f}, R'_f)(f) \qquad \text{for all } R'_f \in \mathcal{R}_f$$

and hence no firm can benefit by unilateral deviation either. Q.E.D.

Next we show that the stable rule is the only matching rule that is *Pareto efficient*, *individually rational*, and *strategy-proof* on this domain.

Theorem 1: The stable rule is the only matching rule that is *Pareto efficient*, *individually rational*, and *strategy-proof* on the class of matching problems with separable preferences and unlimited capacities.

[6]We write $\mu = \mathcal{S}(R)$ instead of $\{\mu\} = \mathcal{S}(R)$.

Proof: Let (F, W, q) be such that $q_{f_i} \geq |W|$ for all $f_i \in F$. *Strategy-proofness* of the stable rule follows from Proposition 2, and its *individual rationality* follows from the definition of stability. Roth and Sotomayor (1990) [Chapter 6] show *Pareto efficiency* for a wider class, the matching problems with substitutable preferences and unlimited capacities. Nevertheless we prove *Pareto efficiency* for the sake of completeness.

Let $R \in \mathcal{R}$ and $\mu = \mathcal{S}(R)$. Let $\mu' \in \mathcal{M}(F, W, q)$ be such that $\mu'(i) P_i \mu(i)$ for some $i \in F \cup W$. To prove *Pareto efficiency* we need to show that $\mu(j) P_j \mu'(j)$ for some $j \in F \cup W$. We have two cases to consider:

Case 1: $i \in F$ We have $\mu = \mathcal{S}(R)$ and $\mu'(i) P_i \mu(i)$, therefore there exists a worker $w \in W$ such that $\{w\} P_i \emptyset$, $w \notin \mu(i)$, and $\{w\} \in \mu'(i)$. This implies $w \in A_i(R_i)$ and thus $\mu(w) R_w \{i\} = \mu'(w)$. But $w \notin \mu(i)$ and preferences are strict, therefore $\mu(w) P_w \mu'(w)$.

Case 2: $i \in W$

We have $\mu = \mathcal{S}(R)$, and $\mu'(i) P_i \mu(i)$, therefore $\mu'(i) \subseteq F$ and $i \notin A_{\mu'(i)}(R_{\mu'(i)})$. Let $\mu'(i) = \{f\}$. If $\mu(f) P_f \mu'(f)$, then we are done. If $\mu'(f) R_f \mu(f)$, then there exists a worker $w \in W$ such that $w P_f \emptyset$, $w \notin \mu(f)$ and $w \in \mu'(f)$. This implies $w \in A_f(R_f)$ and thus $\mu(w) R_w \{f\} = \mu'(w)$. But $w \notin \mu(f)$ and preferences are strict, therefore $\mu(w) P_w \mu'(w)$.

Conversely let $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$ be *Pareto efficient, individually rational*, and *strategy-proof.* Suppose $\varphi \neq \mathcal{S}$. Then there exists a preference profile $R \in \mathcal{R}$ such that $\varphi(R) \neq \mu$ where $\mu = \mathcal{S}(R)$. This, together with $\varphi(R) \subseteq \mathcal{I}(R)$, implies that there is a worker $w \in W$ such that

$$\mu(w) P_w \varphi(R)(w) \quad \text{and} \quad w \in A_{\mu(w)}(R_{\mu(w)}).$$

Let $R'_w \in \mathcal{R}_w$ be as follows:

1. $\mu(w) P'_w \emptyset$,
2. $\emptyset P'_w \{f\}$ for all $f \in F \backslash \{\mu(w)\}$.

Since φ is *individually rational*, we either have $\varphi(R_{-w}, R'_w)(w) = \emptyset$ or $\varphi(R_{-w}, R'_w) = \mu(w)$. But $\mu(w) P'_w \emptyset$, $w \in A_{\mu(w)}(R_{\mu(w)})$, and φ is *Pareto efficient*, therefore

$\varphi(R_{-w}, R'_w)(w) = \mu(w)$. That is, worker w benefits from announcing the false preference R'_w, contradicting *strategy-proofness.* Q.E.D.

3.3 Independence of the Axioms

The following examples establishes the independence of the axioms in Theorem 1. Consider the matching problems with unlimited capacities.

Example 1: Let $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$ be such that, for all $R \in \mathcal{R}$, for all $i \in F \cup W$, $\varphi(R)(i) = \emptyset$.

It is easy to see that the matching rule φ is *individually rational*, and *strategy-proof*, yet it is not *Pareto efficient.*

Example 2: Let $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$ be such that, for all $R \in \mathcal{R}$, for all $w \in W$,

$$\varphi(R)(w) = \{f \in F \mid \{f\} R_w \{f'\} \text{ for all } f' \in F \text{ and } \{f\} R_w \emptyset\}$$

It is easy to see that the matching rule φ is *Pareto efficient*, and *strategy-proof*, yet it is not *individually rational.*

Example 3: Let $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$ be such that, for all $R \in \mathcal{R}$

$$\begin{aligned}
\varphi(R)(f_1) &= Ch_{f_1}(R_{f_1}, B_{f_1}(R_W)) \\
\varphi(R)(f_2) &= Ch_{f_2}(R_{f_2}, B_{f_2}(R_W) \backslash \varphi(R)(f_1)) \\
\vdots \quad &\vdots \qquad\qquad \vdots \\
\varphi(R)(f_j) &= Ch_{f_j}(R_{f_j}, B_{f_j}(R_W) \backslash \cup_{k=1}^{j-1} \varphi(R)(f_k)) \\
\vdots \quad &\vdots \qquad\qquad \vdots \\
\varphi(R)(f_n) &= Ch_{f_n}(R_{f_n}, B_{f_n}(R_W) \backslash \cup_{k=1}^{n-1} \varphi(R)(f_k))
\end{aligned}$$

It is easy to see that the matching rule φ is *individually rational* and yet it is not *strategy-proof.* The following sketch shows that it is also *Pareto efficient*: As the matching $\varphi(R)$ is *individually rational* any matching that Pareto dominates it should also be *individually rational.* Suppose there is such a matching μ. Under the matching $\varphi(R)$ firm 1 gets its top choice among those it can get at any *individually rational* matching and therefore $\mu(f_1) = \varphi(R)(f_1)$. Given this, with a similar argument we can show that $\mu(f_2) = \varphi(R)(f_2)$ and so on. This eventually leads to $\mu = \varphi(R)$ contradicting that μ Pareto dominates $\varphi(R)$.

3.4 An Impossibility Result

A matching rule is *strategy-proof* if no agent can benefit by unilaterally misrepresenting his preferences. However, even if a matching rule is *strategy-proof*, a group of agents can benefit by jointly misrepresenting their preferences.

A matching rule $\varphi : \mathcal{R} \to \mathcal{M}(F, W, q)$ is **coalitionally strategy-proof** if for all $G \subseteq F \cup W$ and for all $R'_G \in \prod_{f \in F \cap G} \mathcal{R}_f \times \prod_{w \in W \cap G} \mathcal{R}_w$, there is an agent $i \in G$ such that

$$\varphi(R)(i) R_i \varphi(R_{-G}, R'_G)(i).$$

Proposition 3: The stable rule is not *coalitionally strategy-proof* on the class of matching problems with separable preferences and unlimited capacities.

Proof: Let $F = \{f_1, f_2\}$, $W = \{w_1, w_2\}$, and $R \in \mathcal{R}$ be such that

$$\{f_1\} P_{w_1} \{f_2\} P_{w_1} \emptyset \qquad \{w_1, w_2\} P_{f_1} \{w_2\} P_{f_1} \{w_1\} P_{f_1} \emptyset$$
$$\{f_2\} P_{w_2} \{f_1\} P_{w_2} \emptyset \qquad \{w_1, w_2\} P_{f_2} \{w_1\} P_{f_2} \{w_2\} P_{f_2} \emptyset$$

We have $\mathcal{S}(R)(f_1) = \{w_1\}$, $\mathcal{S}(R)(f_2) = \{w_2\}$. Let $R'_{f_1} \in \mathcal{R}_{f_1}, R'_{f_2} \in \mathcal{R}_{f_2}$ be such that

$$\{w_2\} P'_{f_1} \emptyset P'_{f_1} \{w_1, w_2\} P'_{f_1} \{w_1\}$$
$$\{w_1\} P'_{f_2} \emptyset P'_{f_2} \{w_1, w_2\} P'_{f_2} \{w_2\}$$

Then $\mathcal{S}(R_{\{w_1,w_2\}}, R'_{\{f_1,f_2\}})(f_1) = \{w_2\}$, $\mathcal{S}(R_{\{w_1,w_2\}}, R'_{\{f_1,f_2\}})(f_2) = \{w_1\}$. Therefore, firms f_1 and f_2 gain by jointly manipulating their preferences. Q.E.D.

Corollary 1: There is no matching rule that is *Pareto efficient*, *individually rational*, and *coalitionally strategy-proof* on the class of matching problems with separable preferences and unlimited capacities. *Proof:* Follows from Theorem 1 and Proposition 3.

Roth (1985a) introduces the class of responsive preferences, a subclass of the class of separable preferences: A firm's preferences over groups of workers are responsive if it is separable, and for any two assignments that differ in only one worker, the firm prefers the assignment containing the more preferred worker as a group by himself. Formally, the preference relation R_f is **responsive** if

1. For all $G \subseteq W$ with $|G| < q_f$, for all $w \in W \backslash G$, $(G \cup \{w\}) P_f G$ if and only if $\{w\} P_f \emptyset$,

2. For all $G \subseteq W$, with $|G| < q_f$, for all $w, w' \in W \backslash G$, $(G \cup \{w\}) P_f (G \cup \{w'\})$ if and only if $\{w\} P_f \{w'\}$.

Note that the preferences of the firms in the proof of Proposition 3 satisfy this stronger requirement.

Remark 3: Roth (1985a) constructs a matching problem in which there exists an *individually rational* matching which is strictly preferred to any stable matching by all firms. The proof of Proposition 3 makes use of a similar construction: Consider the preference profile R. The matching $\mathcal{S}(R_{\{w_1,w_2\}}, R'_{\{f_1,f_2\}})$ is *individually rational* and it is strictly preferred to the unique stable matching $\mathcal{S}(R)$ by both firms.

Remark 4: Note that each worker's assignment is independent of the other workers' preferences under the stable rule. Therefore no group of workers can benefit by jointly manipulating their preferences under the stable rule. Furthermore, if a group of agents $G \subset F \cup W$ can gain by jointly manipulating, then the group of firms $G \backslash W$ can also gain by jointly manipulating.

3.5 Relations with the Core

The core correspondence is one of the most important solutions of cooperative game theory. Roth (1985b) shows that when the preferences of the firms are responsive, the stable set coincides with the core for each matching problem. Roth and Sotomayor (1990) [Proposition 6.4] extend this result to the case of substitutable preferences and unlimited capacities. (Recall that separability implies substitutability.) Therefore all our results concerning the stable rule applies to the core correspondence.

All these results focus our attention to a related model, the "housing market" (Shapley and Scarf 1974). In that model, each individual owns one indivisible good, and has preferences over the goods held by all the agents in the economy. An allocation is a matching of agents with goods. Roth and Postlewaite (1977) show that when preferences are strict, there is a unique core allocation for each problem. Roth (1982b) shows that the core correspondence is *strategy-proof* and Ma (1994) shows that it is the only rule that is *Pareto efficient*, *individually rational*, and *strategy-proof*. Both in housing markets and matching problems with unlimited capacities, the core correspondence is single-valued, it is *strategy-proof*; and it is the only solution that is *Pareto efficient*, *individually rational*, and *strategy-proof*.

4 Impossibility with Limited Capacities

Alcalde and Barberà (1994) show that there is no matching rule that is *Pareto efficient*, *individually rational*, and *strategy-proof* in the context of one-to-one matching problems. In the next proposition we show that this result extends to the model in which firms cannot hire as many workers as they want.

Proposition 4: Consider the matching problems with separable preferences. Let (F, W, q) be such that $q_{f_i} < |W|$ for some $f_i \in F$. Then, there is no matching rule that is *Pareto efficient*, *individually rational* and *strategy-proof*.

Proof: Assume that there exists a firm $f_i \in F$ with $q_{f_i} < |W|$. Let $i = 1$ without loss of generality. Let $W_1 \subset W \backslash \{w_1, w_2\}$ be such that $|W_1| = q_{f_1} - 1$. Let $R \in \mathcal{R}$ be such that

$$\begin{array}{ll}
W_1 \cup \{w_1\} P_{f_1} W_1 \cup \{w_2\} P_{f_1} G P_{f_1} \emptyset & \text{for all } G \in \Sigma^q_{f_1} \backslash \{W_1 \cup \{w_1\}, W_1 \cup \{w_2\}, \emptyset\}. \\
\{w_2\} P_{f_2} \{w_1\} P_{f_2} \emptyset P_{f_2} \{w\} & \text{for all } w \in W \backslash \{w_1, w_2\}, \\
\emptyset P_f G & \text{for all } f \in F \backslash \{f_1, f_2\}, \text{ for all } G \subseteq W, \\
\{f_2\} P_{w_1} \{f_1\} P_{w_1} \emptyset P_{w_1} \{f\} & \text{for all } f \in F \backslash \{f_1, f_2\}, \\
\{f_1\} P_{w_2} \{f_2\} P_{w_2} \emptyset P_{w_2} \{f\} & \text{for all } f \in F \backslash \{f_1, f_2\}, \\
\{f_1\} P_w \emptyset P_w \{f\} & \text{for all } w \in W_1, \text{ for all } f \in F \backslash \{f_1\}, \\
\emptyset P_w \{f\} & \text{for all } w \in W \backslash (W_1 \cup \{w_1, w_2\}), \text{ for all } f \in F,
\end{array}$$

and R_{f_1}, R_{f_2} are responsive.

Let $R' \in \mathcal{R}$ be such that

$$\begin{array}{ll}
W_1 \cup \{w_1\} P'_{f_1} G P'_{f_1} \emptyset P'_{f_1} \{w_2\} & \text{for all } G \in \Sigma^q_{f_1} \backslash \{W_1 \cup \{w_1\}, \{w_2\}, \emptyset\}, \\
\{w_2\} P'_{f_2} \emptyset P'_{f_2} \{w\} & \text{for all } w \in W \backslash \{w_2\}, \\
\{f_2\} P'_{w_1} \emptyset P'_{w_1} \{f\} & \text{for all } f \in F \backslash \{f_2\}, \\
\{f_1\} P'_{w_2} \emptyset P'_{w_2} \{f\} & \text{for all } f \in F \backslash \{f_1\}, \\
R'_i = R_i & \text{for all } i \in (F \cup W) \backslash \{f_1, f_2, w_1, w_2\},
\end{array}$$

and R'_{f_1}, R'_{f_2} are responsive.

Let φ be *Pareto efficient*, *individually rational* and *strategy-proof*.

We have

$$\mathcal{PI}(R_{-\{w_1,w_2\}}, R'_{\{w_1,w_2\}}) = \mu = \begin{pmatrix} f_1 & f_2 \\ W_1 \cup \{w_2\} & \{w_1\} \end{pmatrix}$$

and therefore $\varphi(R_{-\{w_1,w_2\}}, R'_{\{w_1,w_2\}}) = \mu$. Moreover

$$\varphi(R_{-w_2}, R'_{w_2})(w_1) R_{w_1} \varphi(R_{-\{w_1,w_2\}}, R'_{\{w_1,w_2\}})(w_1) = \{f_2\}$$

by *strategy-proofness* and therefore $\varphi(R_{-w_2}, R'_{w_2})(w_1) = \{f_2\}$. This together with *Pareto efficiency* and *individual rationality* of φ implies that

$$\varphi(R_{-w_2}, R'_{w_2}) = \mu$$

Similarly

$$\varphi(R)(w_2) R_{w_2} \varphi(R_{-w_2}, R'_{w_2})(w_2) = \{f_1\}$$

by *strategy-proofness* and therefore $\varphi(R)(w_2) = \{f_1\}$. This together with *Pareto efficiency* and *individual rationality* of φ implies that

$$\varphi(R) = \mu$$

We also have

$$\mathcal{PI}(R_{-\{f_1,f_2\}}, R'_{\{f_1,f_2\}}) = \nu = \begin{pmatrix} f_1 & f_2 \\ W_1 \cup \{w_1\} & \{w_2\} \end{pmatrix}$$

and therefore $\varphi(R_{-\{f_1,f_2\}}, R'_{\{f_1,f_2\}}) = \nu$. Moreover

$$\varphi(R_{-f_2}, R'_{f_2})(f_1) R_{f_1} \varphi(R_{-\{f_1,f_2\}}, R'_{\{f_1,f_2\}})(f_1) = W_1 \cup \{w_1\}$$

by *strategy-proofness* and therefore $\varphi(R_{-f_2}, R'_{f_2})(f_1) = W_1 \cup \{w_1\}$. This together with *Pareto efficiency* and *individual rationality* of φ implies that

$$\varphi(R_{-f_2}, R'_{f_2}) = \nu$$

Similarly

$$\varphi(R)(f_2) R_{f_2} \varphi(R_{-f_2}, R'_{f_2})(f_2) = \{w_2\}$$

by *strategy-proofness* and therefore $w_2 \in \varphi(R)(f_2)$ contradicting $\varphi(R) = \mu$. Q.E.D.

Now we are ready to characterize the class of matching problems that admit *Pareto efficient, individually rational,* and *strategy-proof* matching rules whenever the firms' preferences are separable.

Theorem 2: Consider the matching problems with separable preferences. There exists a *Pareto efficient, individually rational,* and *strategy-proof* matching rule if and only if the capacities are unlimited.

Proof: Follows from Proposition 2 and Proposition 4.

Remark 5: When the preferences are responsive they are also separable. Therefore all our positive results extend to the matching problems with responsive preferences. Furthermore, in the proofs of the negative results (Propositions 3 and 4) firms' preferences are responsive. Finally in the uniqueness part of Theorem 1, we constructed the preferences such that all preferences are responsive as long as the initial preferences has that property. Hence all our results extend to matching problems with responsive preferences.

4.1 More on the Core

The impossibility theorem we obtain in this section focus our attention to a related work, Sönmez (1995). Motivated by the major differences in the results concerning *strategy-proofness* in two closely related models, the housing market and one-to-one matching problems, Sönmez (1995) introduces the generalized matching problems

and studies *strategy-proofness* in this context. Generalized matching problems include both the housing market and one-to-one matching problems, as well as the roommate problems (Gale and Shapley 1962). He shows that on the subclasses with a well-defined core correspondence, there exists a rule that is *Pareto efficient, individually rational,* and *strategy-proof* only if the core correspondence is single-valued. Moreover if there is such a rule, it is the core correspondence. Note that these results are also valid in the context of (many-to-one) matching problems as the core correspondence is single-valued whenever the capacities are unlimited (and the preferences of the firms are separable).

5 Concluding Remarks

As we need the unlimited capacities assumption to obtain a positive result, the practical relevance of our results are likely to be rather thin. On the other hand, our results suggest that in the context of many-to-one matching problems it is possible to achieve *strategy-proofness* together with *Pareto efficiency* and *individual rationality* only by the means of the core correspondence.[7] We believe this result provides a link between cooperative game theory and non-cooperative game theory, giving important non-cooperative support to the core correspondence.

[7] Note that there are subclasses of matching problems where the core correspondence is single-valued, yet it is not *strategy-proof*. For example this is the case on the subclasses where each firm has a capacity equal to the number of the workers minus one.

References

[1] Alcalde, J. and S. Barberà, 1994, Top dominance and the possibility of strategy-proof stable solutions to matching problems, Economic Theory 4, 417-435.

[2] Barberà, S., F. Gül, E. Stachetti, 1993, Generalized median voter schemes and committees, Journal of Economic Theory 61, 262-289.

[3] Barberà, S. and M. Jackson, 1994, A characterization of strategy-proof social choice functions for economies with pure public goods, Social Choice and Welfare 11, 241-252.

[4] Barberà, S. and M. Jackson, 1995, Strategy-proof exchange, Econometrica 63, 51-87.

[5] Barberà, S. and B. Peleg, 1990, Strategy-proof voting schemes with continuous preferences, Social Choice and Welfare 7, 31-38.

[6] Barberà, S., H. Sonnenschein and L. Zhou, 1991, Voting by committees, Econometrica 59, 595-609.

[7] Ching, S., 1993, Strategy-proofness and median voters, University of Rochester mimeo.

[8] Ching, S., 1994, An alternative characterization of the uniform rule, Social Choice and Welfare 11, 131-136.

[9] Gale, D. and L. Shapley, 1962, College admissions and the stability of marriage, American Mathematical Monthly 69, 9-15.

[10] Gale, D. and M. Sotomayor, 1985, Some remarks on the stable matching problem, Discrete Applied Mathematics 11, 223-232.

[11] Gibbard, A., 1973, Manipulation of voting schemes: a general result, Econometrica 41, 587-601.

[12] Hurwicz, L., 1972, On informationally decentralized systems, Chapter 14 in C.B. Mc Guire and R. Radner, eds., Decision and organisation, (North-Holland, Amsterdam) 297-236.

[13] Kara, T. and T. Sönmez, 1993, Nash implementation of matching rules, University of Rochester mimeo, forthcoming in Journal of Economic Theory.

[14] Kara, T. and T. Sönmez, 1994, Implementation of college admission rules, University of Rochester mimeo, forthcoming in Economic Theory.

[15] Ma, J., 1994, Strategy-proofness and the strict core in a market with indivisibilities, International Journal of Game Theory 23, 75-83.

[16] Mongell, S. and A. Roth, 1991, Sorority rush as a two-sided matching mechanism, American Economic Review 81, 441-464.

[17] Moulin, H., 1980, On strategy-proofness and single peakedness, Public Choice 35, 437-455.

[18] Roth, A., 1982a, The economics of matching: stability and incentives, Mathematics of Operations Research 7, 617-628.

[19] Roth, A., 1982b, Incentive compatability in a market with indivisibilities, Economics Letters 9, 127-132.

[20] Roth, A., 1984, The evolution of the labor market for medical interns and residents: a case study in game theory, Journal of Political Economy 92, 991-1016.

[21] Roth, A., 1985a, The college admissions problem is not equivalent to the marriage problem, Journal of Economic Theory 36, 277-288.

[22] Roth, A., 1985b, Common and conflicting interests in two-sided matching markets, European Economic Review 27, 75-96.

[23] Roth, A., 1986, On the allocation of residents to rural hospitals: a general property of two-sided matching markets, Econometrica 54, 425-427.

[24] Roth, A. and A. Postlewaite, 1977, Weak versus strong domination in a market with indivisible goods, Journal of Mathematical Economics 4, 131-137.

[25] Roth, A. and M. Sotomayor, 1989, The college admissions problem revisited, Econometrica 57, 559-570.

[26] Roth, A. and M. Sotomayor, 1990, Two-Sided matching. A study in game-theoretic modeling and analysis (Cambridge University Press, London/New York).

[27] Satterthwaite, M.A., 1975, Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions, Journal of Economic Theory 10, 187-216.

[28] Satterthwaite, M.A. and H. Sonnenschein, 1981, Strategy-proof allocation mechanisms at differentiable points, Review of Economic Studies 48, 587-597.

[29] Shapley, L. and H. Scarf, 1974, On cores and indivisibility, Journal of Mathematical Economics 1, 23-28.

[30] Sönmez, T., 1995, Strategy-proofness and singleton cores in generalized matching problems, University of Michigan working paper.

[31] Sprumont, Y., 1991, The division problem with single-peaked preferences: a characterization of the uniform rule, Econometrica 59, 509-519.

[32] Sprumont, Y., 1994, Strategyproof collective choice in economic and political environments, Canadian Journal of Economics 28, 68-107.

[33] Tadenuma, K. and W. Thomson, 1991, Games of fair division, University of Rochester mimeo, forthcoming in Games and Economic Behavior.

[34] Thomson, W., 1994, Strategy-proof allocation rules, University of Rochester mimeo.

[35] Zhou, L., 1991, Impossibility of strategy-proof mechanisms in economies with pure public good, Review of Economic Studies 58, 107-119.