Discussion Paper No. 1225

Strategic Formation of Coalitions

by

Jinpeng Ma
Northwestern University

June 1996
This version September 10, 1998

Strategic Formation of Coalitions*

Jinpeng Ma[†]

June, 1996
This version September 10, 1998

Abstract

Consider a society with a finite number of players. Each player has personal preferences over coalitions in which he joints. A social outcome is a coalition structure that is defined by a partition of the set of players. We study the strategy proof core and von Neumann and Morgenstern (*vN&M*) solutions.

The roommate problem is a problem in which each coalition contains at most two members. We show that if the core is a singleton, then the core mechanism is coalitionally strategy proof. Since a singleton core defines the largest domain of preferences to admit a mechanism that is strategy proof, individually rational (IR) and Pareto optimal (PO), our result shows that this largest domain is achieved in the roommate problem.

We show in an example that a singleton core is manipulable if coalitions contain more than two members (three, say). We show that if a *vN&M* solution is a singleton, then it is the unique *vN&M* solution and coincides with the core. Moreover the *vN&M* solution mechanism is coalitionally strategy proof in the domain with a singleton *vN&M* solution. In fact the *vN&M* solution is the only mechanism that is strategy proof, IR and PO in the domain.

Journal of Economic Literature Classification Numbers: D40, D71, L10.

Keywords: Strategy Proof, the Core, *vN&M* Solutions, Coalition Structures.

*Dedicated with great admiration, affection and gratitute to Professor Robert J. Aumann on his sixty fifth birthday. I thank the editor, an associate editor, and an anonymous referee for their very helpful comments and suggestions that have greatly improved the exposition of the paper. I also thank Tayfun Sönmez, Jun Wako, Timothy Van Zandt, Lin Zhou, and the conference participants in the Stanford Institute of Theoretical Economics and in the 1996 Summer Meetings of the Econometrics Society in Iowa for helpful comments and suggestions. Part of the work was done when I visited the Center for Rationality at the Hebrew University of Jerusalem as a postdoctoral fellow. I am grateful to the supports from the center. I thank the Kellogg Graduate School of Management at the Northwestern University for the hospitality for my visiting during which this paper is in revision. Certainly any errors are mine.

[†]Department of Economics, Rutgers University, Camden, N.J. 08102. jinpeng@crab.rutgers.edu.

Please Send Proofs to (From September 1 to December 31. 1998):

Jinpeng Ma

MEDS. J.L.Kellogg G.S.M.

Northwestern University. Evanston. IL 60208-2009

Please Send Proofs to (After December 31. 1998):

Jinpeng Ma

Department of Economics

Rutgers University. Camden. NJ 08102

1 Introduction

Consider a society with a finite number of players. Each player has personal preferences over coalitions in which he joints. A social outcome is a coalition structure that is defined by a partition of the set of players. We study the strategy proof core and *vN&M* solutions under this framework.

The *vN&M* solutions and the core are the two very notions in the study of coalition formation in the literature. A solution may be considered "a stable standard of behavior in [a possible social organization]" (von Neumann and Morgenstern (1953, p.436)). Given a profile of preferences, a coalition structure $\mathcal{B}$ dominates the coalition structure $\mathcal{A}$ if there exists a coalition B in $\mathcal{B}$ such that each member in B prefers the coalition B to the coalition in $\mathcal{A}$ in which he joints. It should be noted that the coalition structure $\mathcal{B}$ is not necessarily the one agreed upon by all players in the society. But this definition makes sense because if the coalition B can guarantee each member in it better off by separating from the coalition structure $\mathcal{A}$ no matter what the rest players in the society do, then there is no reason not to believe that the coalition B may well deviate from the coalition structure $\mathcal{A}$ and work alone by themselves to form the coalition B. With this definition of domination, the *vN&M* solutions and the core are defined as usual. A *vN&M* solution is a set of coalition structures such that no element in the solution dominates any other in it and any coalition structure outside the solution is dominated by at least one element in it. Therefore, a solution satisfies both internal and external stability. In contrast, the core contains all coalition structures that are not dominated by any other coalition structures. The core satisfies the internal stability but not necessarily the external stability.

A mechanism is a function from the set of profiles of preferences to the set of coalition structures. Since the underlying true preferences are private information, a mechanism should provide appropriate incentive for each individual player (each coalition) to reveal the truth. This property is known in the literature by (coalitional) strategy proofness. This incentive issue is important for coalition formation. For example the National Resident Matching Program (NRMP) uses a core mechanism to form coalitions between hospitals and physicians. The centralized matching program eventually eliminates the inefficiency and the chaos experienced in the decentralized market (see Roth and Sotomayor (1990) for details and the noted empirical works therein by Professor Al Roth). One important issue in the NRMP is what incentive the hospitals and physicians may have in reporting their ranking lists once a core matching program is implemented. This issue is clearly related to the strategy proof property of a core matching mechanism.

In addition to the strategy proofness a mechanism should have some reasonable properties like

IR and PO. But it is known from Gibbard (1973) and Satterthwaite (1975) that a strategy proof mechanism is dictatorship under the universal domain of preferences (with some minor assumptions). Thus one will have to search under a smaller domain of preferences for a strategy proof mechanism that is also IR and PO. What will be such a domain of preferences? What does such a mechanism look like (if any)?

The literature has provided some hints to the two questions. In the housing market in Shapley and Scarf (1974), Roth (1982b) showed that the core mechanism is strategy proof and Bird (1984) showed that the core mechanism is coalitionally strategy proof. Moreover Ma (1994) showed that the core is the only mechanism that is strategy proof, IR and PO. Further Sonmez (1995) studied a class of generalized matching problems that include both the housing market and the marriage problem in Gale and Shapley (1962). In a generalized matching problem each player is assigned one and only one other individual player or himself under a matching.[1] Sonmez (1995) strenghthened a result in Demange (1987) and showed that the core mechanism is coalitionally strategy proof if it is a singleton and externally stable. Moreover he showed that there exists a mechanism that is strategy proof, IR and PO only if the core is a singleton. Further if such a mechanism exists and the core is not empty, it is the core mechanism itself. Therefore the domain of preferences with a singleton core is the largest domain that admits a mechanism that is strategy proof, IR and PO in the class of generalized matching problems. It also appears that a strategy proof mechanism that is also IR and PO is closely tied up with the core mechanism. This is of interest since the core is a cooperative notion but the three properties are noncooperative in nature. These results invite an interesting question for coalition formation. How robust are the results in the housing market or the generalized matching problems when they apply to the coalition structures? In the coalition structure case players have preferences over coalitions of players and these preferences can be far more complicated than the preferences over individual players in the housing market or the generalized matching problems.

We start with the roommate problem in which each coalition contains at most two members. Here we improve a main result in Sonmez (1995) and show that if the core is a singleton, then the core mechanism is coalitionally strategy proof. Examples also exist such that a singleton core is not externally stable.[2] Our result together with those in Sonmez (1995) shows that a mechanism is strategy proof, IR and PO if and only if it is the core mechanism in the domain defined by a

[1] A matching defined here is not necessarily bilateral since players may form oriented chains like the top trade cycles in the housing market; see Section 3.

[2] That is examples exist such that the core is a singleton but the $vN\&M$ solution is not.

singleton core in the roommate problem (also in the marriage problem). Since a singleton core defines the largest domain in which a mechanism is strategy proof, IR and PO, we show that this largest domain is achievable in the roommate problem.

When we turn to the general coalition structures beyond the roommate problem, things become more complicated. We adopt an example from Roth (1985) to show that a singleton core is no longer strategy proof. This example shows that the "size" of coalitions and the preferences over coalitions do matter in the study of strategy proof mechanisms. What will be the answers to the two questions raised above? Since a singleton core is not strategy proof here, the domain that admits a mechanism that is strategy proof, IR and PO will be smaller than that with a singleton core. It turns out that this is the domain with a singleton *vN&M* solution. Indeed we show that if a *vN&M* solution is a singleton, then it is the unique *vN&M* solution and the *vN&M* solution mechanism is coalitionally strategy proof (in the domain with a singleton *vN&M* solution). Further it is shown that the *vN&M* solution is the only mechanism that is strategy proof, IR and PO in the domain.

How about the core? The results about the core in Sonmez (1995) for the generalized matching problems still apply to the coalition structures. That is, there exists a mechanism that is strategy proof, IR and PO only if the core is a singleton for the coalition structures. Moreover if the core is nonempty and such a mechanism exists, it must be the core mechanism. But, as we noted above, this largest domain with a singleton core is no longer achievable for the general coalition structures. In contrast, this largest domain with a singleton core is achievable in the housing market, the marriage problem and the roommate problem each of which is an example of the generalized matching problems. Naturally we may conjecture that this largest domain is achievable for all generalized matching problems. But we do not have a proof for this conjecture at this point.

The rest of the paper is structured as follows. Section 2 introduces the model and definitions. Section 3 studies the roommate problem. Section 4 studies the general coalition structures. Section 5 presents the characterization of the core and the *vN&M* solutions in terms of strategy proofness, IR and PO properties.

2 The Model

Consider a society with a finite number of players $N = \{1, 2, \cdots, n\}$. A social outcome is a coalition structure that is defined by a partition of the set of players N. Let $\mathcal{P}(N)$ denote the set of all coalition structures. A coalition structure in $\mathcal{P}(N)$ may represent a social or an economic structure.

A coalition in a coalition structure may represent an organization. Let

$$N_i = \{S \subset N : i \in S\}$$

denote the set of all coalitions that contain player i. Let R_i denote the preference for player $i \in N$ over the set of all coalitions N_i. Let $\mathcal{R}_i$ denote the class of all preferences R_i for player i. The strict and indifference preferences for player i are denoted by P_i and I_i respectively. Define $\mathcal{R}^n = \Pi_{i=1}^{i=n}\mathcal{R}_i$. A profile of preferences $R \in \mathcal{R}^n$ is a list of preferences $R = (R_1, R_2, \cdots, R_n)$, one from each player. Let $T \subset N$. Denote $R_T = (R_i)_{i \in T}$ and $R_{-T} = R_{N-T}$.

Given a coalition structure $\mathcal{A} \in \mathcal{P}(N)$, let $\mathcal{A}(i)$ denote the coalition in $\mathcal{A}$ that contains player i.

Definition (Domination) *A coalition structure $\mathcal{B} \in \mathcal{P}(N)$ dominates the coalition structure $\mathcal{A}$ under $R \in \mathcal{R}^n$ if there exists a coalition $B \in \mathcal{B}$ such that $BP_i\mathcal{A}(i)$ for all $i \in B$.*

For any subset X of $\mathcal{P}(N)$, define

$$E(X) = \{\mathcal{A} \in \mathcal{P}(N) : \mathcal{A} \text{ is dominated by some } \mathcal{B} \in X\}. \tag{1}$$

$E(X)$ is the set of all coalition structures each of which is dominated by some coalition structure(s) in the set X.

A subset V of $\mathcal{P}(N)$ is called a *stable set* or a *vN&M* solution whenever

$$\begin{aligned} V \cap E(V) &= \emptyset \quad (2)\\ V \cup E(V) &= \mathcal{P}(N). \quad (3) \end{aligned}$$

These two definitions are called internal and external stability, respectively. No coalition structure in a solution V dominates the other in V and any coalition structure outside V is dominated by some coalition structures in V. We use $V(R)$ rather than V to denote a *vN&M* solution and $\mathcal{V}(R)$ to denote the set of all *vN&M* solutions under $R \in \mathcal{R}^n$.

Definition (Weak domination) *A coalition structure $\mathcal{B} \in \mathcal{P}(N)$ weakly dominates the coalition structure $\mathcal{A}$ under $R \in \mathcal{R}^n$ if there exists a coalition $B \in \mathcal{B}$ such that $BR_i\mathcal{A}(i)$ for all $i \in B$ and $BP_i\mathcal{A}(i)$ for some $i \in B$.*

Weak domination and domination may be different. But when preferences over coalitions are strict, they are the same.[3] Henceforth we do not make any distinction between weak domination and domination when preferences over coalitions are strict. Given a profile of preferences $R \in \mathcal{R}^n$, the *core* $C(R)$ consists of all coalition structures that are not dominated under R by any other coalition structures. Note that the core may be empty and the *vN&M* solutions may not exist.

Definition *A coalition structure $\mathcal{B} \in \mathcal{P}(N)$ is PO under $R \in \mathcal{R}^n$ if there is no other coalition structure $\mathcal{A} \in \mathcal{P}(N)$ such that $\mathcal{A}(i)R_i\mathcal{B}(i)$ for all $i \in N$ and $\mathcal{A}(i)P_i\mathcal{B}(i)$ for some $i \in N$. Let $PO(R)$ denote the set of all PO coalition structures under R.*

Definition *A coalition structure $\mathcal{B}$ is IR under $R \in \mathcal{R}^n$ if $\mathcal{B}(i)R_i\{i\}$ for all $i \in N$. Let $IR(R)$ denote the set of all IR coalition structures under R.*

Definition *A mechanism $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ is a map from profiles of preferences $\mathcal{R}^n$ to the set of coalition structures $\mathcal{P}(N)$. A mechanism φ and the underlying true profile $R \in \mathcal{R}^n$ induce a direct revelation game.*

This definition may deserve some comments. Why are we interested in what players may report? One reason is that no matter where preferences come from, only players know in reality the "values" of their coalitions (coalition structures). If the value of a coalition depends on players' underlying preferences or their ideology, then these underlying preferences are not known publicly. Therefore, the consideration of the reported information of the above mechanism is as usual as in the literature of mechanism design such as voting. We see this as a merit in the study of coalition formation over the coalitional form games in some aspects because the incentive issues naturally arise in the current framework. The strategic issues, though important, are often too complicated to be considered in the coalitional form games.

Definition *A mechanism $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ is IR and PO if $\varphi(R) \in IR(R) \cap PO(R)$ for all profiles $R \in \mathcal{R}^n$.*

Definition *A mechanism φ is strategy proof if for all $R \in \mathcal{R}^n$, all $i \in N$, and all $R'_i \in \mathcal{R}_i$,*

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

[3] By definition, if $\mathcal{B}$ weakly dominates $\mathcal{A}$ under R, then there exists a coalition $B \in \mathcal{B}$ such that $BR_i\mathcal{A}(i)$ for all $i \in B$ and $BP_i\mathcal{A}(i)$ for some $i \in B$. Since preferences over coalitions are strict, it follows that $BP_i\mathcal{A}(i)$ for all $i \in B$.

It is coalitionally strategy proof if all $R \in \mathcal{R}^n$, all $T \subset N$, and all $R'_T \in \mathcal{R}_T$ there exists $i \in T$ such that

$$\varphi(R_{-T}, R_T)(i) R_i \varphi(R_{-T}, R'_T)(i).$$

3 The Roommate Problem

Let $\mathcal{M} \subset \mathcal{P}(N)$ be a subset of $\mathcal{P}(N)$ such that $\mid \mathcal{A}(i) \mid \leq 2$ for all $i \in N$ for every $\mathcal{A} \in \mathcal{M}$. That is each player forms a coalition with at most one other player. Given a coalition structure $\mathcal{A} \in \mathcal{M}$, define $\mu(i) = j$ if $\exists j \in \mathcal{A}(i)$ such that $j \neq i$, and $\mu(i) = i$ otherwise to be the matching obtained from the coalition structure $\mathcal{A}$. A matching μ is bilateral in the sense that $\mu(\mu(i)) = 1$ for all $i \in N$. Note that the matching μ such that $\mu(i) = i$ for all $i \in N$ is in $\mathcal{M}$. Let $\mathcal{A}$ and $\mathcal{B}$ be two coalition structures in $\mathcal{M}$. Let μ and ν be the two matchings obtained from $\mathcal{A}$ and $\mathcal{B}$ respectively. Then players' preferences over matchings are naturally defined by $\mu(i) R_i \nu(i)$ if and only if $\mathcal{A}(i) R_i \mathcal{B}(i)$ for all i. Thus players have restrictive preferences over individual players in the roommate problem. We slightly abuse the notation and also use $\mathcal{M}$ to denote the set of all (bilateral) matchings and R_i to denote the preferences over matchings or individual players.

Given a profile of preferences R, a pair (i, j) blocks a matching μ if $\mu(i) \neq j$ but $j P_i \mu(i)$ and $i P_j \mu(j)$. A matching μ is *stable* if $\mu(i) R_i i$ for all $i \in N$ and it is not blocked by any pair. Let $S(R)$ denote the set of all stable matchings under a profile of preferences R. Unlike the marriage problem, the roommate problem may not have stable matchings, i.e., $S(R)$ may be empty for some (even strict) profile R. A matching λ dominates the other μ via a coalition $T \subset N$ if $\lambda(i) \in T$ for all $i \in T$ and all players in T prefer λ to μ. Given a profile $R \in \mathcal{R}^n$, the core $C(R)$ is a subset of $\mathcal{M}$ each of which is not dominated by any other matching in $\mathcal{M}$ via any coalition.

Theorem 1 *Let $P \in \mathcal{R}^n$ be a profile of strict preferences. Then $S(P) = C(P)$.*

Proof Clearly $C(P) \subset S(P)$. Let μ be a stable matching. Suppose that μ is not in the core. Then μ is dominated by some other matching λ via some coalition T. Since μ is individually rational, it follows that $\mid T \mid \geq 2$. Let $i \in T$ and $j = \lambda(i)$. The fact that λ dominates μ via T implies that $j P_i \mu(i)$ and $i P_j \mu(j)$, since preferences are strict. This shows that μ is not stable. $\square$

Sonmez (1995) introduced a class of generalized matching problems. Recall that N_i is the set of all subsets of N that contain i. A generalized matching problem is a triple (N, S, R), where N is the set of players, $S = (S_1, S_2, \cdots, S_n)$ is a given list of subsets of N such that $S_i \in N_i$ for all $i \in N$, and $R = (R_1, R_2, \cdots, R_n)$ is a profile of preferences. Each S_i may be considered as the set of players that are feasible to player i. Each preference R_i is a linear order on S_i. A matching μ in a generalized matching problem is a map from the set N into itself such that $\mu(i) \in S_i$ for all $i \in N$ and $\mid \mu^{-1}(i) \mid= 1$ for all $i \in N$. Each preference R_i over S_i can be extended to matchings as follows: a player i prefers a matching μ to the other ν if and only if he prefers $\mu(i)$ to $\nu(i)$. The flexibility in the definition of the generalized matching problems provides a uniform framework for the housing market and the roommate problem. For example, the roommate problem is a generalized matching problem by setting $S_i = N$ for all $i \in N$ and defining a matching μ to satisfy the additional bilateral property, i.e., $\mu(\mu(i)) = i$ for all $i \in N$.

Demange (1987) defined a notion of coalitionally strategy proofness for correspondences and showed that the core correspondence is coalitionally strategy proof if it is nonempty and satisfies the external stability (in domination). Sonmez (1995) showed that the core in the class of generalized matching problems is coalitionally strategy proof if the core is a singleton and satisfies the external stability (in weak domination).

Theorem 2 (Sonmez) *Let $\mid C(R) \mid= 1$ and assume that $C(R)$ satisfies the external stability (in weak domination) for all profiles of preferences $R \in \mathcal{R}^n$ in the class of generalized matching problems. Then the core mechanism C is coalitionally strategy proof.*

Our theorem below improves this result in the roommate problem and it shows that the core mechanism is coalitionally strategy proof as long as the core is a singleton, no matter whether the core satisfies the external stability or not. Examples exist such that a singleton core does not satisfy the external stability in the roommate problem (also in the marriage problem).

Theorem 3 *Let $\mid C(P) \mid= 1$ for all profiles of strict preferences $P \in \mathcal{R}^n$. Then the core mechanism C is coalitionally strategy proof in the roommate problem.*

Proof Let $T \subset N$ and $Q_T \in \mathcal{R}_T$ such that $\mid C(P_{-T}, Q_T) \mid= 1$. Let $\mu \in C(P_{-T}, Q_T)$ and $\lambda \in C(P)$. Suppose that $\forall i \in T$,

$$\mu(i) P_i \lambda(i).$$

We first show that μ is IR under P. Suppose not. Then there exists $i \in N$ such that $iP_i\mu(i)$. This implies that $i \in T$. But if $i \in T$, then $\mu(i)P_i\lambda(i)R_ii$, a desired contradiction.

Let K denote all players who prefer μ to λ under P. Then $T \subset K$. The local blocking lemma (see Lemma A3 in Appendix) shows that there exists a pair (i, j) that blocks μ under P such that $i \in N - K$ and $j \in \mu(K)$. Since $j \in \mu(K)$, it follows that $\lambda(j)P_j\mu(j)$, by the stability of λ. Hence $j \notin T$. Thus both i and j are not in T. And then the fact that (i, j) blocks μ under P implies that μ is not stable under (P_{-T}, Q_T). Theorem 1 shows that μ is not in the core $C(P_{-T}, Q_T)$, a contradiction. $\square$

The singleton core is a strong assumption. But Sonmez (1995) showed that as long as the core mechanism is well defined, there exists a strategy proof mechanism that is also IR and PO only if the core is a singleton. Moreover such a mechanism must be the core.

Theorem 4 (Sonmez) *Suppose that $C(R) \neq \emptyset$ for all profiles $R \in \mathcal{R}^n$ and let $\varphi : \mathcal{R}^n \to \mathcal{M}$ be a strategy proof mechanism that is also IR and PO in the generalized matching problems. Then $\mid C(R) \mid = 1$ for all $R \in \mathcal{R}^n$ and $\varphi(R) = C(R)$ for all R.*

Since the roommate problem is belong to the class of generalized matching problems, Theorems 3 and 4 show that the core mechanism is characterized by strategy proofness, individual rationality and Pareto optimality in the domain with a singleton core.

Corollary 5 *A mechanism $\varphi : \mathcal{R}^n \to \mathcal{M}$ is strategy proof, IR and PO if and only if φ is the core mechanism in the domain of preferences with a singleton core in the roommate problem.*

Example 6 (Roth (1982)) *Let $M = \{m_1, m_2, m_3\}$ be the set of men and $W = \{w_1, w_2, w_3\}$ be the set of women. The strict preferences are as follows:*

$$P_{m_1} = (w_2, w_1, w_3, m_1) \qquad P_{w_1} = (m_1, m_2, m_3, w_1)$$
$$P_{m_2} = (w_1, w_2, w_3, m_2) \qquad P_{w_2} = (m_3, m_1, m_2, w_2)$$
$$P_{m_3} = (w_1, w_2, w_3, m_3) \qquad P_{w_3} = (m_1, m_2, m_3, w_3).$$

Without any confusion, we represent these preferences in a manner that is consistent with the marriage problem. Note that every bilateral matching in this example satisfies individual rationality.

There exists a unique stable matching μ in the core $C(P)$

$$\mu = [(m_1, w_1), (m_2, w_3), (m_3, w_2)].$$

The matching λ

$$\lambda = [(m_1, w_2), (m_2, w_3), (m_3, w_1)]$$

is not stable but it is in the $vN\&M$ solution. In fact one can show that this matching market has a unique $vN\&M$ solution that consists of the two matchings $\{\mu, \lambda\}$.[4] Therefore the core $C(P) = \{\mu\}$ does not satisfy the external stability since it does not dominate the matching λ.

But it is known that the core mechanism is coalitionally strategy proof as long as the core is a singleton in the marriage problem: see Dubins and Freedman (1981) and Roth (1982). Hence the external stability is not a necessary condition for a singleton core to be coalitionally strategy proof in the roommate and the marriage problems. It remains open if a singleton core, without the external stability, is coalitionally strategy proof for all generalized matching problems. When we consider coalition structures, examples exist such that a singleton core is not strategy proof in the domain with a singleton core. Thus the results in this section depend on heavily the structures of matchings and the preferences over matchings.

4 Coalition Structures

The set $\mathcal{P}(N)$ admits much richer coalition structures than the set of bilateral matchings $\mathcal{M}$. When we pay attention to $\mathcal{M}$, the singleton core mechanism has the appealing coalitionally strategy proof property as shown in Theorem 3. A natural question is if a singleton core is coalitionally strategy proof in the situation of general coalition structures. We now adopt an example[5] from Roth (1985) to show that this is not the case. Even if the core is a singleton, the core mechanism is manipulable by some player.

Example 7 *Let $N = C \cup S$, where $C = \{C_1, C_2, C_3\}$ and $S = \{S_1, S_2, S_3, S_4\}$. The strict preferences are as follows (as a convention those coalitions that do not appear in the preferences are ranked below 'staying alone'):*

$$P_{C_1} = (\{C_1, S_1, S_2\}, \{C_1, S_2, S_3\}, \{C_1, S_1, S_4\}, \{C_1, S_2, S_4\},$$

[4] A formal proof of this is available from the author.

[5] This example is different from that in Roth (1985) in students' preferences. In the college admissions problem in Roth (1985), students have preferences over individual colleges and they are not concerned with the other students the colleges may be assigned.

$$
\begin{aligned}
& \{C_1, S_3, S_4\}, \{C_1, S_1\}, \{C_1, S_2\}, \{C_1, S_3\}, \{C_1, S_4\}, \{C_1\}) \\
P_{C_2} &= (\{C_2, S_1\}, \{C_2, S_2\}, \{C_2, S_3\}, \{C_2, S_4\}, \{C_2\}) \\
P_{C_3} &= (\{C_3, S_3\}, \{C_3, S_1\}, \{C_3, S_2\}, \{C_3, S_4\}, \{C_3\}) \\
P_{S_1} &= (\{C_3, S_1\}, \{C_1, S_1\}, \{C_2, S_1\}, \{S_1\}) \\
P_{S_2} &= (\{C_2, S_2\}, \{C_1, S_2, S_3\}, \{C_1, S_2, S_4\}, \{C_1, S_2\}, \{C_3, S_2\}, \{S_2\}) \\
P_{S_3} &= (\{C_1, S_3, S_4\}, \{C_1, S_2, S_3\}, \{C_3, S_3\}, \{C_2, S_3\}, \{S_3\}) \\
P_{S_4} &= (\{C_1, S_3, S_4\}, \{C_1, S_2, S_4\}, \{C_2, S_4\}, \{C_3, S_4\}, \{S_4\}).
\end{aligned}
$$

Then the coalition structure $\mathcal{A}$

$$\mathcal{A} = [(C_1: S_3, S_4), (C_2: S_2), (C_3: S_1)]$$

is the unique one in the core $C(P)$. To see this, note that every player S_i must be matched with some C_j in a core coalition structure. For example, if a core coalition structure leaves S_1 single, then it is dominated by $\{C_2, S_1\}$. If a core coalition structure leaves S_3 single, then it is dominated by the coalition $\{C_3, S_3\}$. If a core coalition structure leaves S_2 single, then it must match C_2 with S_1, C_3 with S_3, and leave S_4 single. But then it is dominated by $\{C_1, S_2, S_4\}$. If a core coalition structure μ leaves S_4 single, C_2 and C_3 must match with some player in $\{S_1, S_2, S_3\}$ in μ. This implies that C_1 can only match at most (also at least) one player in $\{S_1, S_2, S_3\}$ in μ. If C_1 is matched with S_1 in μ, then C_3 must be matched with S_3. Then μ is dominated by $\{C_1, S_3, S_4\}$. If C_1 is matched with S_2 in μ, then it is dominated by $\{C_1, S_2, S_4\}$. If C_1 is matched with S_3 in μ, then it is dominated by $\{C_1, S_3, S_4\}$.

Note that the coalition structure

$$[(C_1: S_3, S_4), (C_2: S_1), (C_3: S_2)]$$

is dominated by $\{C_3, S_1\}$. There are four other coalition structures in which every player is matched (that are also IR for S_1):

$$
\begin{aligned}
\mathcal{B} &= [(C_1: S_2, S_4), (C_2: S_1), (C_3: S_3)] \text{ or } [(C_1: S_2, S_4), (C_2: S_3), (C_3: S_1)] \\
\mathcal{D} &= [(C_1: S_2, S_3), (C_2: S_4), (C_3: S_1)] \\
\mathcal{E} &= [(C_1: S_2, S_3), (C_2: S_1), (C_3: S_4)]
\end{aligned}
$$

They are dominated by $\{C_1, S_2, S_3\}$, $\{C_2, S_2\}$ and $\{C_3, S_1\}$, respectively. This completes the proof that $\mathcal{A}$ is the only element in the core $C(P)$. But the core mechanism is not strategy proof at the profile P. To show this, consider

$$Q_{C_1} = (\{C_1, S_2, S_4\}, \{S_2, C_1\}, \{S_4, C_1\}, \{C_1\}).$$

Then the coalition structure $\mathcal{B}$

$$\mathcal{B} = [(C_1; S_2, S_4), (C_2; S_1), (C_3; S_3)]$$

is the unique one in the core $C(Q_{C_1}, P_{-C_1})$. Hence any core mechanism must assign $\{S_2, S_4\}$ to C_1 at (Q_{C_1}, P_{-C_1}) and $\{S_3, S_4\}$ to C_1 at P. But C_1 prefers $\{C_1, S_2, S_4\}$ to $\{C_1; S_3, S_4\}$. Thus the core mechanism is not strategy proof even if it is a singleton at P. Note also that the core is a singleton at (Q_{C_1}, P_{-C_1}) as well. □

The following is an extension of Theorem 2 from the generalized matching problems to the coalition structures when preferences are strict.

Theorem 8 *Let $| C(P) |= 1$ and assume that $C(P)$ is externally stable for all profiles of strict preferences P in the coalition structures. Then the core mechanism C is coalitionally strategy proof.*

Proof Let $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ be the core mechanism C. Suppose, by the way of contradiction, that there exists a profile of strict preferences (P_{-T}, Q_T) such that $| C(P_{-T}, Q_T) |= 1$ and for all $i \in T$,

$$\varphi(P_{-T}, Q_T)(i) P_i \varphi(P_{-T}, P_T)(i),$$

where $\varphi(P_{-T}, Q_T) \in C(P_{-T}, Q_T)$.

The external stability of $C(P_{-T}, P_T)$ shows that $\exists B \in \varphi(P_{-T}, P_T)$ such that $B P_i \varphi(P_{-T}, Q_T)(i)$ for all $i \in B$. Then it follows from the external stability of $C(P_{-T}, Q_T)$ that $\exists A \in \varphi(P_{-T}, Q_T)$ such that $A P_i \varphi(P_{-T}, P_T)(i)$ for $i \in A - T$ and $A Q_i \varphi(P_{-T}, P_T)(i)$ for $i \in T \cap A$. But, by the assumption that $\varphi(P_{-T}, Q_T)(i) P_i \varphi(P_{-T}, P_T)(i)$ for all $i \in T$, it follows that $A P_i \varphi(P_{-T}, P_T)(i)$ for all $i \in T \cap A$. Therefore $A P_i \varphi(P_{-T}, P_T)(i)$ for all $i \in A$, a desired contradiction, since $\varphi(P_{-T}, P_T)$ is in the core $C(P_{-T}, P_T)$. This completes the proof. □

Clearly a singleton, externally stable core is a singleton *vN&M* solution. The next result shows that the converse of this is also true: A singleton *vN&M* solution is a singleton, externally stable core. In fact a singleton *vN&M* solution is the unique *vN&M* solution.

Theorem 9 *Let $P \in \mathcal{R}^n$ be a profile of strict preferences. Suppose $V(P) = \{\mathcal{A}\}$ for some $V(P) \in \mathcal{V}(P)$. Then $\mathcal{V}(P) = \{V(P)\}$. Moreover $V(P) = C(P)$.*

Proof We first show the claim that $\nexists B \subset N$ such that $BP_i\mathcal{A}(i)$ for all $i \in B$. Suppose, by the way of contradiction, that this is not the case, i.e., there exists a coalition B such that $BP_i\mathcal{A}(i)$ for all $i \in B$. Then we can construct an infinite chain $A_1, A_2, \cdots$, of non-empty, disjoint sets in $\mathcal{A}$, as shown below. But $\mathcal{A}$ is finite, a desired contradiction.

Let $K \in N$. Given $\{A_1, \cdots, A_K\} \subset \mathcal{A}$ such that $\{B, A_1, \cdots, A_K\}$ are disjoint (if $K = 0$, then $\{B\}$ is trivially disjoint). Let $\mathcal{B}_K = \{B, A_1, \cdots, A_K, N-B-A_1-A_2-\cdots-A_K\}$. Since $BP_i\mathcal{A}(i)$ for all $i \in B$, it follows $B \notin \mathcal{A}$. Hence $\mathcal{B}_K \neq \mathcal{A}$. By the external stability of $\mathcal{A}$, there exists $A_{K+1} \in \mathcal{A}$ such that $A_{K+1}P_i\mathcal{B}_K(i)$ for all $i \in A_{K+1}$. For $k = 1, \cdots, K$, since $A_k \in \mathcal{B}_K$ and $A_k \in \mathcal{A}$, it follows that $A_{K+1} \neq A_k$ and $A_{K+1} \cap A_k = \emptyset$. Since $BP_i\mathcal{A}(i)$ for all $i \in B$, the fact that $A_{K+1} \in \mathcal{A}$ mplies that $B \cap A_{K+1} = \emptyset$. Hence, $B, A_1, \cdots, A_{K+1}$ are disjoint and $\{A_1, \cdots, A_{K+1}\} \subset \mathcal{A}$. This completes the proof of the claim.

Now, let $V_1(P), V_2(P) \in \mathcal{V}(P)$ be any two distinct *vN&M* solutions. Suppose $V_1(P) = \{\mathcal{A}\}$. There exists $\mathcal{D}$ in $V_2(P)$ such that $\mathcal{D} \neq \mathcal{A}$ since $V_1(P) \neq V_2(P)$. Moreover $\mathcal{A} \notin V_2(P)$, since $\mathcal{A}$ dominates $\mathcal{D}$, by the assumption $V_1(P) = \{\mathcal{A}\}$. By the external stability of $V_2(P)$, there exists some coalition structure $\mathcal{B}$ that dominates $\mathcal{A}$ via some coalition B, i.e., $\exists B \in \mathcal{B}$ such that $BP_i\mathcal{A}(i)$ for all $i \in B$. But this is a contradiction to our claim above. This shows that $\mathcal{V}(P) = \{V(P)\}$.

Note that a corollary of the claim above shows that the coalition structure $\mathcal{A}$ dominates every other coalition structure (since it is externally stable) and is not dominated by any other coalition structure. Thus it follows that $V(P) = C(P)$. □

Definition *A mechanism $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ is a vN&M solution mechanism if $\varphi(R) \in V(R)$ for some $V(R) \in \mathcal{V}(R)$ for all $R \in \mathcal{R}^n$.*

The above definition may not be well defined since a *vN&M* solution may not exist. This definition will not generate problems though in the context below because we are concerned with the domain where a *vN&M* solution exists. A corollary of Theorems 8 and 9 is as follows.

Corollary 10 *Assume that $|\ V(P)\ | = 1$ for some $V(P) \in \mathcal{V}(P)$ for all profiles of strict preferences P. Then the vN&M solution mechanism is coalitionally strategy proof.*

Proof Theorem 9 shows that a singleton *vN&M* solution is the unique *vN&M* solution. Moreover it is a singleton, externally stable core. Thus it follows from Theorem 8 that the *vN&M* solution mechanism is coalitionally strategy proof in the domain with a singleton *vN&M* solution. □

Example 7 and Theorem 8 show that the external stability is necessary for a singleton core to be strategy proof in coalition structures. In contrast the external stability is not necessary in the housing market, the marriage problem and the roommate problem each of which is an example of the generalized matching problems. But it remains unclear if the external stability is necessary for all generalized matching problems.

5 *vN&M* Solutions versus the Core

In this section we show an analogy to Theorem 4 and Corollary 5 for the general coalition structures. Since a conclusion in the generalized matching models may not apply to the coalition structures, we present a formal proof of them.

The results presented here are closely related to those for the housing market in Shapley and Scarf (1974) in which each agent owns a house and consumes at most one house. An allocation in the housing market is a permutation of all houses. Roth and Postlewaite (1977) showed that the core in the housing market is a singleton and it satisfies the external stability (both defined in weak domination); Roth (1982b) showed that the core is individually strategy proof; Bird (1984) showed that the core is coalitionally strategy proof; and Ma (1994) showed that the core is the only mechanism that is strategy proof, IR and PO. Sonmez (1995) generalized these results to the class of generalized matching problems.

In what follows we prove that there exists a strategy proof mechanism that is also IR and PO only if the core is a singleton for the general coalition structures. Moreover it must be the core, as long as the core is nonempty and such a mechanism exists. The main idea in the proof is to construct the profile Q of preferences in the use of the induction approach. This idea has been used before in Ma (1994) and Sonmez (1995). The proof in Sonmez (1995) depends on the structure of a matching μ. The proof in Ma (1994) depends on the housing market. Our proof below follows Ma (1994) closely.

Let $\mathcal{A}$ and $\mathcal{B}$ be any two coalition structures in $\mathcal{P}(N)$, define

$$J(\mathcal{A}, \mathcal{B}; P) = \{j \in N : \mathcal{A}(j) P_j \mathcal{B}(j)\}$$

to be the set of all players j who prefer coalition $\mathcal{A}(j)$ to coalition $\mathcal{B}(j)$ in which he joints. Therefore, the three sets, $J(\mathcal{A}, \mathcal{B}; P)$, $J(\mathcal{B}, \mathcal{A}; P)$ and $N - J(\mathcal{A}, \mathcal{B}; P) - J(\mathcal{B}, \mathcal{A}; P)$, form a partition of N.

Lemma 11 *Let P be a profile of strict preferences. Also let $\mathcal{A}, \mathcal{B} \in PO(P)$ be two PO coalition structures such that $\mathcal{A} \neq \mathcal{B}$. Then $J(\mathcal{A}, \mathcal{B}; P) \neq \emptyset$ and $J(\mathcal{B}, \mathcal{A}; P) \neq \emptyset$.*

Proof If $J(\mathcal{A}, \mathcal{B}; P) = \emptyset$, then $\mathcal{B}(i) R_i \mathcal{A}(i)$ for all $i \in N$. Since $\mathcal{B} \neq \mathcal{A}$, it follows that there exists at least one $i \in N$ such that $\mathcal{B}(i) P_i \mathcal{A}(i)$, since preferences are strict. This implies that $\mathcal{A}$ is not in $PO(P)$. Symmetrically we also have $J(\mathcal{B}, \mathcal{A}; P) \neq \emptyset$. □

Lemma 12 *Let P be a profile of strict preferences. Let $\mathcal{A} \in C(P)$ be a core coalition structure and $\mathcal{B} \in IR(P) \cap PO(P)$ be an IR and PO coalition structure. If $\mathcal{B} \neq \mathcal{A}$, then $\exists j \in J(\mathcal{A}, \mathcal{B}; P)$ such that $\mathcal{A}(j) P_j \mathcal{B}(j) P_j \{j\}$.*

Proof By definition $\mathcal{A}(j) P_j \mathcal{B}(j)$ for all $j \in J(\mathcal{A}, \mathcal{B}; P)$. Lemma 11 shows that $J(\mathcal{A}, \mathcal{B}; P) \neq \emptyset$. If the lemma is false, then there exists no player $j \in J(\mathcal{A}, \mathcal{B}; P)$ such that $\mathcal{A}(j) P_j \mathcal{B}(j) P_j \{j\}$. This implies that $\{j\} R_j \mathcal{B}(j)$ for all $j \in J(\mathcal{A}, \mathcal{B}; P)$. Since $\mathcal{B}$ is IR under P, it follows that $\mathcal{B}(j) R_j \{j\}$ for all $j \in J(\mathcal{A}, \mathcal{B}; P)$. Thus $\mathcal{B}(j) I_j \{j\}$ for all $j \in J(\mathcal{A}, \mathcal{B}; P)$. Since preferences are strict, it follows that $\mathcal{B}(j) = \{j\}$ for all $j \in J(\mathcal{A}, \mathcal{B}; P)$. That is all players in $J(\mathcal{A}, \mathcal{B}; P)$ stay alone in $\mathcal{B}$. Lemma 11 shows that $J(\mathcal{B}, \mathcal{A}; P) \neq \emptyset$. Thus $N - J(\mathcal{A}, \mathcal{B}; P)$ forms a coalition that weakly dominates the coalition structure $\mathcal{A}$. This is a contradiction to the assumption that $\mathcal{A}$ is in the core. □

Theorem 13 *Suppose that $C(P) \neq \emptyset$ for all profiles of strict preferences $P \in \mathcal{R}^n$ and let $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ be a mechanism that is strategy proof, IR and PO. Then $\mid C(P) \mid = 1$ for all $P \in \mathcal{R}^n$ and $\varphi(P) = C(P)$ for all P.*

Proof Let $P \in \mathcal{R}^n$ such that $C(P) \neq \emptyset$. Let $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ be a mechanism that is strategy proof, IR and PO. We first show that $\mid C(P) \mid \leq 1$ for all $P \in \mathcal{R}^n$.

Let $\mathcal{D} \in C(P)$ be a core coalition structure. Construct a profile of strict preferences Q as follows: $\forall i \in N$,

$$Q_i = \begin{cases} (\overbrace{\cdots\cdots\cdots\cdots,\mathcal{D}(i),}^{\text{truncation of } P_i \text{ up to } \mathcal{D}(i)} \quad \{i\},\cdots) & \text{if } \mathcal{D}(i) \neq \{i\} \\ P_i & \text{otherwise} \end{cases}$$

In the construction of the preference Q_i when player i does not 'stay alone' in the core coalition structure $\mathcal{D}$, we move $\{i\}$ right after the coalition $\mathcal{D}(i)$ in which he joints. The ranking order after $\{i\}$ in Q_i is arbitrary.

Step 1: *$C(Q) = \{\mathcal{D}\}$ and $IR(Q) \cap PO(Q) = \{\mathcal{D}\}$.*

Proof Since any core coalition structure is IR and PO, it is sufficient to show that $IR(Q) \cap PO(Q) = \{\mathcal{D}\}$. Suppose that there exists $\mathcal{A} \in IR(Q) \cap PO(Q)$ such that $\mathcal{A} \neq \mathcal{D}$. Lemma 11 shows that $J(\mathcal{D}, \mathcal{A}; Q) \neq \emptyset$. Then Lemma 12 shows that there exists $j \in J(\mathcal{D}, \mathcal{A}; Q)$ such that $\mathcal{D}(j) Q_j \mathcal{A}(j) Q_j \{j\}$. But there exists no $\mathcal{A}(j)$ between $\mathcal{D}(j)$ and $\{j\}$ under Q_j, by the construction of Q_j. □

Step 2: *$\varphi(P_T, Q_{-T}) = \mathcal{D}$ for all $T \subset N$.*

Proof We use the induction on the size $\mid T \mid$ of the coalition T to show this. When $\mid T \mid = 0$, Step 1 shows that $\varphi(Q) = \mathcal{D}$, since $\varphi(Q) \in IR(Q) \cap PO(Q)$. Now assume that

$$\varphi(P_T, Q_{-T}) = \mathcal{D}$$

for all $T \subset N$ such that $\mid T \mid \leq k$.

Suppose, by the way of contradiction, that $\varphi(P_T, Q_{-T}) \neq \mathcal{D}$ for some $T \subset N$ such that $\mid T \mid = k+1$. Let $Q' = (P_T, Q_{-T})$. Since $\mathcal{D} \in C(P)$, it follows that $\mathcal{D} \in C(Q')$ by the construction of the profile Q (every coalition that dominates $\mathcal{D}$ under Q' dominates $\mathcal{D}$ under P as well). Then Lemma 12 shows that there exists $j \in J(\mathcal{D}, \varphi(Q'); Q')$ such that

$$\mathcal{D}(j) Q'_j \varphi(Q')(j) Q'_j \{j\}. \tag{4}$$

If $j \in N - T$, then $Q'_j = Q_j$. We obtain from the above that

$$\mathcal{D}(j) Q_j \varphi(Q')(j) Q_j \mathcal{S}_j.$$

But there exists no coalition $\varphi(Q')(j)$ between $\mathcal{D}(j)$ and $\{j\}$ under Q_j, by the construction of Q_j. This shows that $j \in T$.

If $j \in T$, then it follows from (4) that

$$\mathcal{D}(j) P_j \varphi(Q')(j). \tag{5}$$

It follows from the induction hypothesis that

$$\varphi(Q_j, Q'_{-j}) = \mathcal{D}. \tag{6}$$

Thus it follows from (5) and (6) that

$$\varphi(Q_j, Q'_{-j})(j) P_j \varphi(P_j, Q'_{-j})(j)$$

contradicting φ is strategy proof. □

Since $\mathcal{D}$ is arbitrarily chosen, it follows that $\mid C(P) \mid \leq 1$ for all $P \in \mathcal{R}^n$. Moreover it follows from Step 2 that $\varphi(P) = C(P)$ for all $P \in \mathcal{R}^n$ such that $\mid C(P) \mid = 1$. This completes the proof. □

Corollary 14 *A mechanism $\varphi : \mathcal{R}^n \to \mathcal{P}(N)$ is strategy proof, IR and PO if, and only if φ is the vN&M solution mechanism V in the domain with a singleton vN&M solution.*

Proof Corollary 10 shows that the *vN&M* solution V is coalitionally strategy proof. Theorem 8 shows that V coincides with the core C. Therefore V is a mechanism that is strategy proof, IR and PO. The "only if" part follows from Theorems 9 and 13. □

These results are related to the question what may be the largest domain to admit a mechanism that is strategy proof, IR and PO. The largest domain that admits such a mechanism turns out to be that with a singleton core both in the generalized matching problems and in the general coalition structures (Theorems 4 and 13). We also show that this largest domain is reached in the roommate problem (Corollary 5). But this largest domain cannot be reached in the general coalition structures since a singleton core may not be strategy proof (Example 7). Nevertheless the domain with a singleton *vN&M* solution, though relatively smaller than that with a singleton core, admits at least one mechanism that is strategy proof, IR and PO, since we show that the *vN&M* solution is coalitionally strategy proof and it is IR and PO in the domain (Theorem 9 and Corollary 10). Then Corollary 14 shows that the *vN&M* solution is the only mechanism that is strategy proof. IR and PO in this smaller domain.

Appendix

This appendix is to prove the local blocking lemma used in the proof of Theorem 3. The local blocking lemma is an extension of the blocking lemma in the marriage problem to the roommate problem. We need to prove the decomposition lemma first.

Given a profile of strict preferences P, let μ and λ be two core matchings in $C(P)$. Define

$$J(\mu, \lambda) = \{i \in N : \mu(i) P_i \lambda(i)\}$$

to be the set of all agents who prefer μ to λ under P. The next lemma, an analogy to the decomposition lemma in the marriage problem (see Corollary 2.21 in Roth and Sotomayor (1990)), shows that both μ and λ define isomorphisms between $J(\mu, \lambda)$ and $J(\lambda, \mu)$. The proof follows from that of Lemma 2.20 in Roth and Sotomayor (1990).

Lemma A1 (Decomposition lemma) *Let $P \in \mathcal{R}^n$ be a profile of strict preferences. Then for any two core matchings μ and λ in $C(P)$,*

$$J(\mu, \lambda) \xleftrightarrow{\mu} J(\lambda, \mu) \xleftrightarrow{\lambda} J(\mu, \lambda).$$

Proof Let $i \in J(\mu, \lambda)$ and $j = \mu(i)$. Since $j = \mu(i) P_i \lambda(i) R_i i$, it follows that $j \neq i$. Moreover $\lambda(j) P_j \mu(j)$. Otherwise $j = \mu(i) P_i \lambda(i)$ and $i = \mu(j) P_j \lambda(j)$ contradicting λ is a stable matching (Theorem 1). This shows that $j \in J(\lambda, \mu)$. Thus $\mu(J(\mu, \lambda)) \subset J(\lambda, \mu)$. It also follows that $\lambda(J(\lambda, \mu)) \subset J(\mu, \lambda)$. Since μ and λ are one-to-one, the conclusion follows because both $J(\mu, \lambda)$ and $J(\lambda, \mu)$ are finite. $\square$

An immediate result of Lemma A1 is that

Corollary A2 *With strict preferences a player who is single under a core matching remains single under every core matching in the roommate problem.*

Lemma A3 (Local blocking lemma) *Let $P \in \mathcal{R}^n$ be a profile of strict preferences such that $\mid C(P) \mid = 1$. Let $T \subset N$ and let $Q_T \in \mathcal{R}_T$ such that $\mid C(P_{-T}, Q_T) \mid = 1$, where (P_{-T}, Q_T) is a profile of strict preferences. Let $\lambda \in C(P)$ and $\mu \in C(P_{-T}, Q_T)$. Let $K = J(\mu, \lambda)$. Suppose that $T \subset K$ and μ is IR under P. Then there exists a pair (i, j) that blocks μ under P such that $i \in N - K$ and $j \in \mu(K)$.*

Proof The proof follows closely from the original idea in the second proof of the blocking lemma in Roth and Sotomayor (1990. pp.57-58) for the marriage problem. In the marriage problem with any individually rational matching μ. it is possible that $\mu(K) = \lambda(K)$. Because of the assumption on the matching μ in the local blocking lemma. this case is excluded. In what follows we will show that $\mu(K) \neq \lambda(K)$. If this is the case. then let $k \in K$ and $\mu(k) = j$ such that $j \in \mu(K) - \lambda(K)$. Hence $j = \mu(k) P_k \lambda(k)$. Since λ is stable under P. it follows that $\lambda(j) P_j \mu(j) = k$. Let $i = \lambda(j)$. Then $i \notin K$ since $j \notin \lambda(K)$. Hence $j = \lambda(i) P_i \mu(i)$. Thus (i, j) is the desired blocking pair of μ under P.

We now show that $\mu(K) \neq \lambda(K)$. Suppose on the contrary that $\mu(K) = \lambda(K)$. Since $\mu(k) P_k \lambda(k)$ for all $k \in K$ and λ is stable for P. it follows $\lambda(j) P_j \mu(j)$ for all $j \in \mu(K)$.

We now define a marriage problem (M, W, P') such that $M = K$ and $W = \mu(K)$. The preferences P' are defined as follows. For all $m \in M$. P'_m is the same as P_m restricted to $W \cup \{m\}$. For all $w \in W$. P'_w is the same as P_w restricted to $M \cup \{w\}$ except that w is now ranked just below $\mu(w)$. Note that λ restricted to $M \cup W$ is still a stable matching in the market (M, W, P'). since any pair that blocks λ under P' would also block it under P. Let μ_M be the M-optimal stable matching for the market (M, W, P'). Thus

$$(*) \qquad \mu_M(m) R_m \lambda(m) \ \forall m \in M.$$

If $\mu_M = \lambda$. then Theorem 2.27 in Roth and Sotomayor (1990) shows that no individually rational matching μ restricted to $M \cup W$ exists such that $\mu(m) P_m \lambda(m)$ for all $m \in M$ contradicting to the definition of M. Thus it follows that

$$(**) \qquad \mu_M(m) P_m \lambda(m) \ \text{ for some } m \in M.$$

Since no player in W is single under the matching λ. it follows from Corollary A2 that no player in W is single under μ_M. Hence. by the construction of P'. we must have that

$$(***) \qquad \mu_M(w) R_w \mu(w) \ \forall w \in W.$$

Define ν on N such that $\nu = \mu_M$ on $M \cup W$ and $\nu = \lambda$ on $N - M - W$. Since $\nu \neq \lambda$. by (**). ν is not stable under P. So let (i, j) be a blocking pair of ν. First we show that we cannot have $\{i, j\}$ in M or $\{i, j\}$ in W. Suppose $\{i, j\}$ in M. Then $(*)$ yields

$$j P_i \mu_M(i) R_i \lambda(i) \text{ and } i P_j \mu_M(j) R_j \lambda(j)$$

contradicting λ is stable under P. Suppose that $\{i, j\}$ is in W. Hence $\{i, j\}$ is not in T since $T \subset K$. Then $(***)$ shows that

$$jP_i\mu_M(i)R_i\mu(i) \text{ and } iP_j\mu_M(j)R_j\mu(j)$$

contradicting μ is stable under (P_{-T}, Q_T).

Second we show that $\{i, j\}$ is not in $M \cup W$. Suppose not. Then, without loss of generality, let $j \in M$ and $i \in W$. Hence

$$jP_i\mu_M(i) \text{ and } iP_j\mu_M(j).$$

Then the construction of P' implies that (i, j) also blocks μ_M in the market (M, W, P') contradicting μ_M is stable under P'.

This leaves three possibilities: $j \in M$ and $i \in N - M - W$, $j \in W$ and $i \in N - M - W$, and $i, j \in N - M - W$. We will show that all of them will lead to a contradiction.

First, suppose that $j \in M$ and $i \in N - M - W$. Then, by $(*)$ and the definition of ν,

$$iP_j\nu(j) = \mu_M(j)R_j\lambda(j) \text{ and } jP_i\nu(i) = \lambda(i)$$

contradicting λ is stable under P.

Second, suppose $j \in W$ and $i \in N - M - W$. Then both i and j are not in T. Thus, by $(***)$ and the definition of ν, M and W, we have

$$iP_j\nu(j)R_j\mu(j) \text{ and } jP_i\nu(i)R_i\mu(i)$$

contradicting μ is stable under (P_{-T}, Q_T).

Finally suppose that both i and j are in $N - M - W$. Then, by the definition of ν, we have

$$iP_j\lambda(j) \text{ and } jP_i\lambda(i)$$

contradicting λ is stable under P. This completes the proof. □

References

[1] BIRD, C. (1984): "Group Incentive Compatibility in a Market with Indivisible Goods," *Economic Letter*, 14, 309-313.

[2] DEMANGE, G. (1987): "Nonmanipulable Cores," *Econometrica*, 55, 1057-1074.

[3] DUBINS, L.E. and D.A. FREEDMAN (1981): "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly* **88**, 485-94.

[4] GALE, D. and L. SHAPLEY (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9-15.

[5] GIBBARD, A. (1973): "Manipulation of Voting Schemes: a general result," *Econometrica*, 41, 587-602.

[6] MA, J. P. (1994): "Strategy Proofness and the Strict Core in a Market with Indivisibilities," *International Journal of Game Theory*, 23, 75-83.

[7] NEUMANN, J. VON and O. MORGENSTERN (1944, 1953): Theory of Games and Economic Behavior, Princeton University Press, Princeton, New Jersey.

[8] ROTH, A. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operation Research*, 7, 617-628.

[9] ROTH, A. (1982b): "Incentive Compatibility in a Market with Indivisible Goods," *Economic Letters*, 9, 127-132.

[10] ROTH, A. (1985): "The College Admissions Problem is not Equivalent to the Marriage Problem," *Journal of Economic Theory*, 36, 277-88.

[11] ROTH, A. and A. POSTLEWAITE (1977): "Weak versus Strong Domination in a Market with Indivisible Goods," *Journal of Mathematical Economics*, 4, 131-137.

[12] ROTH, A.E. and M. SOTOMAYOR (1990): Two-sided Matching: a Study in Game-theoretic Modeling and Analysis, Cambridge University Press, Cambridge.

[13] SATTERTHWAITE, M. (1975): "Strategy-proofness and Arrow's Conditions: existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, 10, 187-217.

[14] SHAPLEY, L. and H. E. SCARF (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 23-28.

[15] SÖNMEZ, T. (1995): Strategy-proofness and Singleton Cores in Generalized Matching Problems, University of Rochester.