

Ma, Jinpeng

**Working Paper**

## Strategic Formation of Coalitions

Working Paper, No. 1998-10

**Provided in Cooperation with:**

Department of Economics, Rutgers University

*Suggested Citation:* Ma, Jinpeng (1998) : Strategic Formation of Coalitions, Working Paper, No. 1998-10, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/94302>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Strategic Formation of Coalitions\*

Jinpeng Ma<sup>†</sup>

June, 1996

This version March 10, 1998

## Abstract

Consider a society with a finite number of individuals. A coalition structure is a partition of the set of individuals. Each individual has personal preferences over the set of all coalition structures. We study the coalition structure core and the von Neumann and Morgenstern ( $vN\mathcal{M}$ ) solutions.

A roommate problem is a coalition formation problem in which each coalition contains at most two members. We show that as long as the core is single-valued, the core is coalitionally strategy proof. Moreover the core mechanism is characterized by three properties, namely, strategy proofness, individual rationality and Pareto optimality, in the domain with single-valued core. The single-valued core defines the largest domain to obtain a mechanism with the three properties.

We show in an example that the single-valued core is manipulable if coalitions contain more than two members. Nevertheless we show that the single-valued  $vN\mathcal{M}$  solution is coalitionally strategy proof and it is individually rational and Pareto optimal. In fact the  $vN\mathcal{M}$  solution is the only mechanism with the three properties in the domain with single-valued  $vN\mathcal{M}$  solution.

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

**Keywords:** Strategy Proof, the Core,  $vN\mathcal{M}$  Solutions, Individual Rationality, Pareto Optimality, Coalition Structures

---

\*Dedicated with great admiration, affection and gratitude to Professor Robert J. Aumann on his sixty fifth birthday. I thank the editor, an associate editor and an anonymous referee for their 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 and I am grateful to the supports from the center. Certainly any errors are mine.

<sup>†</sup>Department of Economics, Rutgers University, Camden, N.J. 08102. jinpeng@crab.rutgers.edu.

# 1 Introduction

There are numerous organizations or coalitions in a society and an economy. Cartels, firms, families, clubs and political parties are a few examples of coalitions. Many such situations can be formalized and analyzed by means of coalitional form games. Indeed, since von Neumann and Morgenstern (1944) published their seminal book, *The Theory of Games and Economic Behavior*, coalitional form games have been used to study coalition structures in a great number of papers; see Thrall and Lucas (1963), Aumann and Maschler (1964), Aumann and Drèze (1974), Myerson (1977), Drèze and Greenberg (1980), Hart and Kurz (1983), Aumann and Myerson (1988), Greenberg (1990, 1994), and lately, Nouweland (1993), Jackson and Wolinsky (1995) and Qin (1995), among many others<sup>1</sup>.

Individuals may form coalitions for many reasons: monetary benefits, exchange of information, friendships or mutual concerns. When an individual decides in which coalition he likes to joint, he may often concern with the qualitative and quantitative characteristics of that coalition (Drèze and Greenberg (1980)). Therefore, personal preferences over coalitions are important for coalition formation and they can be very complicated. For examples, “personal, family, patriotic, geographical or professional relationship” (Aumann and Drèze (1974)) and “political affinity, ideological similarity and attitudinal proximity” (Kahan and Rapoport (1984)) may be parts of these preferences. This paper explores a framework in one way to treat these preferences as the primitive data in the study of coalition formation. The framework works as follows. Consider a society or an economy with a finite number of players or individuals. A player has personal preferences over the coalitions in which he joints. An outcome is a coalition structure which is a partition of the set of players. Player’s preferences over coalition structures are naturally obtained from his preferences over coalitions.

This framework is closely related to Drèze and Greenberg (1980). Drèze and Greenberg defined individual’s preferences on his consumption bundle and the coalition to which he belongs. They used coalitional form games to study individually stable equilibrium and individually stable contractual equilibrium and the existence properties of their two equilibrium notions. Our framework may be considered as the counterpart of theirs. We are mainly concerned with the  $vNEM$  solutions and the core.

The  $vNEM$  solutions and the core are the two very notions in the study of coalition formation,

---

<sup>1</sup>One should note that networks in Myerson (1977), Aumann and Myerson (1988), Nouweland (1993) and Jackson and Wolinsky (1995) admit richer graphic structures than coalitions.

especially in the coalitional form games. A solution, as it was presented in the first place in von Neumann and Morgenstern (1953, p.436), may be considered “a stable standard of behavior in [a possible social or economic organization]”. The coalitional form games define the solutions and the core in terms of payoffs. In this paper we define the solutions and the core in terms of coalition structures. Given a preference profile, a coalition structure  $\mathcal{A}$  is dominated by another  $\mathcal{B}$  if there exists a coalition  $B$  in  $\mathcal{B}$  such that every member in it prefers coalition structures  $\mathcal{B}$  to  $\mathcal{A}$ . Note that in this definition of domination, coalition structure  $\mathcal{B}$  is not necessary to be the one agreed upon by all individuals. But this definition makes sense because if the coalition  $B$  can guarantee each member better off by separating from previous coalition structure  $\mathcal{A}$  no matter what the rest individuals 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. Therefore, the coalition structure  $\mathcal{A}$  is not so reasonable when such a coalition exists. With this definition of domination, the definitions of the  $vNEM$  solutions and the core become routine. A  $vNEM$  solution is a set of coalition structures such that no element in the solution dominates the 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 (strict) core is defined to satisfy the internal stability only and contains all coalition structures that are not (weakly) dominated by any other coalition structures.

We study the strategic aspects of the core and the  $vNEM$  solutions in the formation of coalitions in the institutional framework. A mechanism is a function from the set of preference profiles to the set of coalition structures. A  $vNEM$  solution (the core) mechanism, if well-defined, is the one that assigns a coalition structure in a  $vNEM$  solution (the core) for each revealed preference profile. To prevent from opportunistic misrepresentation of the true preferences, a mechanism should induce each individual to reveal his truth. This property is known in the literature by strategy proofness. In addition to the strategy proof property, we require a mechanism to satisfy individual rationality and Pareto optimality. An immediate question is what a mechanism with the three properties may look like. The answer to the question is closely related to the core and the  $vNEM$  solution.

We first consider the roommate problem in which each coalition contains at most two members. The roommate problem is a symmetric version of the marriage problem in Gale and Shapley (1962). It is a special instance of the generalized matching problem in Sonmez (1995). Sonmez (1995) strengthened a result in Demange (1987) and showed that the core is coalitionally strategy proof if it is single-valued and satisfies the external stability (in weak domination). We improve Sonmez’s result and show that the single-valued core is always coalitionally strategy proof in the roommate problem. We also give an example to show that the single-valued core in the roommate

and marriage problems may not satisfy the external stability, i.e., the core is singleton but the  $vNEM$  solution is not. Sonmez (1995) also showed that as long as the core is well defined, there exists a mechanism that is strategy proof, individually rational and Pareto optimal only if the core is single-valued. Further it must be the core, as long as the core is nonempty and such a mechanism exists. His result extended those in Roth (1982b), Bird (1984) and Ma (1994) for the housing swapping market (Shapley and Scarf (1974)) to the generalized matching market. Our strategy proof core together with his result provides a characterization of the core, namely, a mechanism is strategy proof, individually rational and Pareto optimal if and only if it is the core in the domain with single-valued core in the roommate and marriage problems. Since the single-valued core defines the largest domain with a mechanism satisfying the three properties, we show that this largest domain is achievable in certain situations.

When we turn to the general coalition structures, things become more complicated. We adopt an example from Roth (1985) to show that the single-valued core is no longer strategy proof when coalition structures are beyond bilateral matchings. Nevertheless we identify a domain that admits at least one mechanism with the three properties. This is the domain with single-valued  $vNEM$  solution. This later domain is smaller than that with single-valued core. We show that the single-valued  $vNEM$  solution is always coalitionally strategy proof. Moreover it is individually rational and Pareto optimal. In fact we show that the  $vNEM$  solution is the only mechanism with the three properties in the domain with single-valued  $vNEM$  solution. We also show in the general coalition structures that there exists a mechanism with the three properties only if the core is single-valued. Moreover it must be the core, as long as the core is nonempty and such a mechanism exists. This is a parallel result to that in Sonmez (1995) for the generalized matching problems.

The rest of the paper is structured as follows. Section 2 introduces the formal model in this paper. Section 3 introduces several additional definitions. Section 4 studies the roommate problem. We extend the decomposition lemma and the blocking lemma in the marriage problem to the roommate problem and show that the single-valued core is always strategy proof. Section 5 studies the general coalition structures. Section 6 presents the characterization of the core and the  $vNEM$  solution in terms of the three properties. Section 7 provides remarks on admissible coalition structures and the networks.

## 2 A Formal Model

Consider a society that consists of a finite number of players,  $N = \{1, 2, \dots, n\}$ . A coalition structure is a partition of the set  $N$  of players. Denote the set of all coalition structures by  $\mathcal{P}(N)$ . 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

$$N_i = \{S \in 2^N : i \in S\}$$

denote the set of all coalitions that contain player  $i$ . Player  $i$  has personal (ordinal) preferences  $R_i$  over coalitions in  $N_i$ .

Given a coalition structure  $\mathcal{A} \in \mathcal{P}(N)$ , we use  $\mathcal{A}(i)$  to denote the coalition in  $\mathcal{A}$  that contains player  $i$ . A player's personal preferences over coalition structures in  $\mathcal{P}(N)$  are naturally derived from his preferences over coalitions, i.e., player  $i$  prefers coalition structures  $\mathcal{A}$  to  $\mathcal{B}$  if and only if he prefers  $\mathcal{A}(i)$  to  $\mathcal{B}(i)$ . We abuse notion slightly and also use  $R_i$  ( $P_i, I_i$ ) to represent player  $i$ 's (strict, indifference) preferences over coalition structures. Denote by  $\mathcal{R}_i$  the set of player  $i$ 's preferences over coalition structures. Define  $\mathcal{R}^n = \prod_{i=1}^n \mathcal{R}_i$ . A preference profile  $R \in \mathcal{R}^n$  is a list of preferences  $R = (R_1, R_2, \dots, R_n)$ , one from each player.

**Definition (Domination)** *Let  $R \in \mathcal{R}^n$  be a profile of preferences. A coalition structure  $\mathcal{B} \in \mathcal{P}(N)$  dominates another  $\mathcal{A}$  if there exists a coalition  $B \in \mathcal{B}$  such that  $\mathcal{B}P_i\mathcal{A}$  for all  $i \in B$ .*

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

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

$E(X)$  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 *vNEM solution* whenever

$$V \cap E(V) = \emptyset \quad (2)$$

$$V \cup E(V) = \mathcal{P}(N). \quad (3)$$

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 vNEM solution and

$\mathcal{V}(R)$  to denote the set of all  $vNEM$  solutions under  $R \in \mathcal{R}^n$ .

**Definition (Weak domination)** Let  $R \in \mathcal{R}^n$  be a profile of preferences. A coalition structure  $\mathcal{B} \in \mathcal{P}(N)$  weakly dominates another coalition structure  $\mathcal{A}$  if there exists a coalition  $B \in \mathcal{B}$  such that  $\mathcal{B}R_i\mathcal{A}$  for all  $i \in B$  and  $\mathcal{B}P_i\mathcal{A}$  for some  $i \in B$ .

Weak domination may not be equivalent to domination. But when preferences are strict, they are the same.<sup>2</sup> Henceforth we assume that preferences are strict and make no distinction between weak domination and domination.

Let  $P \in \mathcal{R}^n$  be a profile of strict preferences, the *core*  $C(P)$  consists of all coalition structures that are not dominated by any other coalition structures. The core may be empty and the  $vNEM$  solutions may not exist. It is known that when the number of agents is at most 4, the  $vNEM$  solutions always exist in the coalitional form games; see Neumann and Morgenstern (1953) and Bondareva, Kulakovskaya and Naumova (1979). Example 1 below shows that some conclusions in the coalitional form games may not apply to the current model. The  $vNEM$  solutions may not exist in a situation with three players  $n = 3$ .

**Example 1**  $N = \{1, 2, 3\}$ .  $\mathcal{P}(N) = \{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5\}$ , where

$$\begin{aligned} \mathcal{A}_1 &= \{\{1\}, \{2\}, \{3\}\} & \mathcal{A}_2 &= \{\{1\}, \{2, 3\}\} & \mathcal{A}_3 &= \{\{1, 2\}, \{3\}\} \\ \mathcal{A}_4 &= \{\{1, 3\}, \{2\}\} & \mathcal{A}_5 &= \{\{1, 2, 3\}\} \end{aligned}$$

The preferences are as follows

$$\begin{aligned} P_1 &= (\{1, 2\}, \{1, 3\}, \{1, 2, 3\}, \{1\}) & P_2 &= (\{2, 3\}, \{1, 2\}, \{1, 2, 3\}, \{2\}) \\ P_3 &= (\{1, 3\}, \{2, 3\}, \{1, 2, 3\}, \{3\}). \end{aligned}$$

Then a  $vNEM$  solution does not exist. Note that  $\mathcal{A}_2$  is dominated by  $\mathcal{A}_4$  via  $\{1, 3\}$ ,  $\mathcal{A}_3$  is dominated by  $\mathcal{A}_2$  via  $\{2, 3\}$  and  $\mathcal{A}_4$  is dominated by  $\mathcal{A}_3$  via  $\{1, 2\}$ . The internal stability requires that only one among these three is available for a  $vNEM$  solution. But then it is impossible to find a set of coalition structures to satisfy the external stability.  $\square$

---

<sup>2</sup>By definition, if  $\mathcal{B}$  weakly dominates  $\mathcal{A}$ , then there exists a coalition  $B \in \mathcal{B}$  such that  $\mathcal{B}R_i\mathcal{A}$  for all  $i \in B$ . By strict preferences, we must have  $\mathcal{B}P_i\mathcal{A}$  for all  $i \in B$ .

### 3 Some Definitions

In what follows we provide some additional definitions.

**Definition** Let  $P \in \mathcal{R}^n$  be a profile of strict preferences. A coalition structure  $\mathcal{B} \in \mathcal{P}(N)$  is Pareto optimal under  $P$  if  $\mathcal{B}$  is not dominated by any  $\mathcal{A} \in \mathcal{P}(N)$  via the grand coalition  $N$ . Denote by  $PO(P)$  the set of all Pareto optimal coalition structures under  $P$ .

If a player  $i$  stays single in a coalition structure  $\mathcal{S} = (\{i\}, \dots)$ , this coalition structure is denoted by  $\mathcal{S}_i$ . Note that there are an equivalent class of coalitions like  $\mathcal{S}_i$  under  $R_i$ . Without confusion, we use  $\mathcal{S}_i$  to represent this class or an element in it.

**Definition** Let  $P \in \mathcal{R}^n$  be a profile of strict preferences. A coalition structure  $\mathcal{B}$  is individually rational if  $\mathcal{B} R_i \mathcal{S}_i$  for all  $i \in N$ . Denote by  $IR(P)$  the set of all individually rational coalition structures under  $P$ .

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

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 formations 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 consider in the coalitional form games.

**Definition** A mechanism  $\varphi : \mathcal{R}^n \rightarrow \mathcal{P}(N)$  satisfies the individual rationality and Pareto optimality if  $\varphi(P) \in IR(P) \cap PO(P)$  for all profiles  $P \in \mathcal{R}^n$ .

An additional key property of a mechanism is the strategy proofness. A strategy proof mechanism prevents agents from misrepresenting their true preferences toward to their own interests. Let  $T \subset N$ . Denote  $P_T = (P_i)_{i \in T}$  and  $P_{-T} = P_{N-T}$ .

**Definition** A strict profile  $Q \in \mathcal{R}^n$  is a dominant strategy equilibrium of the game  $(\varphi, P)$  if

$$\varphi(Q_{-i}, Q_i) P_i \varphi(Q_{-i}, Q'_i) \quad (4)$$

for all  $Q_{-i}, Q'_i$  and all  $i \in N$ .

**Definition** A mechanism  $\varphi : \mathcal{R}^n \rightarrow \mathcal{P}(N)$  is (individually) coalitionally strategy proof if each underlying true profile  $P$  is a dominant strategy equilibrium of the game  $(\varphi, P)$  for all coalitions  $T \subset N$  (with size  $|T| = 1$ ).

## 4 The Roommate Problem

Let  $\mathcal{M} \subset \mathcal{P}(N)$  be a subset of  $\mathcal{P}(N)$  such that  $|\mathcal{A}(i)| \leq 2$  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 an element  $\mathcal{A} \in \mathcal{M}$ , define  $\mu(i) = j$  if  $j \in \mathcal{A}(i)$  and  $\mu(i) = i$  otherwise to be the matching corresponding to the coalition structure  $\mathcal{A}$ . A matching  $\mu$  is bilateral in the sense that  $\mu(\mu(i)) = i$  for all  $i \in N$ . We assume that the matching  $\mu$  such that  $\mu(i) = i$  for all  $i \in N$  is in  $\mathcal{M}$ . This is the noted roommate problem, a symmetric version of the marriage problem (Gale and Shapley (1962)).

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

**Theorem 2** 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  $\mu$  be a stable matching. Suppose that  $\mu$  is not in the core. Then  $\mu$  is dominated by some other matching  $\lambda$  via some coalition  $T$ . Since  $\mu$  is individually rational, it follows that  $|T| \geq 2$ . Let  $i \in T$  and  $j = \lambda(i)$ . The fact that  $\lambda$  dominates  $\mu$  via  $T$  implies

that  $jP_i\mu(i)$  and  $iP_j\mu(j)$ , since preferences are strict. This shows that  $\mu$  is not stable.  $\square$

Sonmez (1995) considered a generalized matching problem that unifies the roommate problem (thus the marriage problem) and the housing swapping problem in Shapley and Scarf (1974). He defined a matching  $\mu$  to be a map from the set  $N$  into itself such that  $\mu(i) \in N_i$  and  $|\mu^{-1}(i)| = 1$  for all  $i \in N$ . This matching  $\mu$  is not necessarily bilateral as in the roommate and the marriage problems. Clearly the roommate and the marriage problems are examples of the generalized matching problem.

Demange (1987) defined a notion of coalitionally strategy proofness for correspondences and showed that the core is coalitionally strategy proof if it is nonempty and satisfies the external stability (in domination). Sonmez (1995) showed that the core mechanism in the generalized matching problem is coalitionally strategy proof if the core is single-valued and satisfies the external stability (in weak domination).

**Theorem 3** (Sonmez) *Let  $|C(R)| = 1$  and assume that  $C(R)$  satisfies the external stability (in weak domination) for all profiles  $R \in \mathcal{R}^n$  of preferences in the generalized matching problem. Then the core  $C$  is coalitionally strategy proof.*

Our theorem below improves this result for the roommate problem and it shows that the core is coalitionally strategy proof as long as the core is single-valued, no matter whether the core satisfies the external stability or not.

**Theorem 4** *Let  $|C(P)| = 1$  for all profiles  $P \in \mathcal{R}^n$  of strict preferences. Then the core  $C$  is coalitionally strategy proof in the roommate problem.*

The single-valued core is a strong assumption. But Sonmez (1995) showed in the generalized matching problem that as long as the core mechanism is well defined, there exists a strategy proof mechanism that is also individually rational and Pareto optimal only if the core is single-valued. Moreover it must be the core. Applying his result to the roommate problem, one obtains that

**Theorem 5** (Sonmez) *Suppose that  $C(P) \neq \emptyset$  for all  $P \in \mathcal{R}^n$  and let  $\varphi : \mathcal{R}^n \rightarrow \mathcal{M}$  be a strategy proof mechanism that also satisfies individual rationality and Pareto optimality. Then  $|C(P)| \leq 1$  for all  $P \in \mathcal{R}^n$ . Moreover  $\varphi(P) = C(P)$  for all  $P$  with  $C(P) \neq \emptyset$ .*

Theorems 4 and 5 show that in the roommate problem the core mechanism is characterized by three properties of strategy proofness, individual rationality and Pareto optimality in the domain with a single-valued core.

**Corollary 6** *A mechanism  $\varphi : \mathcal{R}^n \rightarrow \mathcal{M}$  is strategy proof, individually rational and Pareto optimal if and only if  $\varphi = C$  in the roommate problem in the domain with single-valued core, i.e., the domain such that  $|C(P)| = 1$  for all  $P \in \mathcal{R}^n$ .*

In what follows we will show Theorem 4. The proof depends on the extensions of the decomposition and the blocking lemmas in the marriage problem to the roommate problem.

Given a profile  $P$  of strict preferences, let  $\mu$  and  $\lambda$  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  $\mu$  to  $\lambda$ . 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  $\mu$  and  $\lambda$  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 7** (Decomposition lemma) *Let  $P \in \mathcal{R}^n$  be a profile of strict preferences. Then for any two core matchings  $\mu$  and  $\lambda$  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  $\lambda$  is a stable matching (Theorem 2). 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  $\mu$  and  $\lambda$  are one-to-one, the conclusion follows because both  $J(\mu, \lambda)$  and  $J(\lambda, \mu)$  are finite.  $\square$

An immediate result of Lemma 7 is that

**Corollary 8** *With strict preferences a player who is single under a core matching remains single under every core matching.*

**Lemma 9** (Local blocking lemma) *Let  $P \in \mathcal{R}^n$  be a profile of strict preferences such that  $|C(P)| = 1$ . Let  $T \subset N$  and let  $Q_T \in \mathcal{R}^{|T|}$  such that  $|C(P_{-T}, Q_T)| = 1$ , 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  $K$  denote all agents who prefer  $\mu$  to  $\lambda$ . Suppose that  $T \subset K$  and  $\mu$  is individually rational under  $P$ . Then there exists a pair  $(i, j)$  that blocks  $\mu$  under  $P$  such that  $i \in N - K$  and  $j \in \mu(K)$ .*

The matching  $\mu$  in Lemma 9 can be any individually rational one in the blocking lemma in the marriage problem; see Lemma 3.5 in Roth and Sotomayor (1990). The local blocking lemma above only applies to the special matching  $\mu$ , not any one. The proof below 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  $\mu$ , it is possible that  $\mu(K) = \lambda(K)$ . Due to the assumption on  $\mu$ , this case has been excluded in the local blocking lemma above.

**Proof** 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  $\lambda$  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 blocking pair of  $\mu$ .

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  $\lambda$  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  $\lambda$  restricted to  $M \cup W$  is still a stable matching in the market  $(M, W, P')$ , since any pair that blocks  $\lambda$  under  $P'$  would also block it under  $P$ . Let  $\mu_M$  be the M-optimal stable matching for the market  $(M, W, P')$ . Thus

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

If  $\mu_M = \lambda$ , then Theorem 2.27 in Roth and Sotomayor (1990) shows that no individually rational matching  $\mu$  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

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

Since no player in  $W$  is single under the matching  $\lambda$ , it follows from Corollary 8 that no player in  $W$  is single under  $\mu_M$ . Hence, by the construction of  $P'$ , we must have that

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

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

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

contradicting  $\lambda$  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  $\mu$  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  $\mu_M$  in the market  $(M, W, P')$  contradicting  $\mu_M$  is stable under  $P'$ .

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

(a) Suppose that  $j \in M$  and  $i \in N - M - W$ . Then, by (\*) and the definition of  $\nu$ ,

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

contradicting  $\lambda$  is stable under  $P$ .

(b) 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  $\nu$ ,  $M$  and  $W$ , we have

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

contradicting  $\mu$  is stable under  $(P_{-T}, Q_T)$ .

(c) Suppose that both  $i$  and  $j$  are in  $N - M - W$ . Then, by the definition of  $\nu$ , we have

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

contradicting  $\lambda$  is stable under  $P$ . This completes the proof.  $\square$

With the help of Lemma 9, we now show Theorem 4.

**Proof of Theorem 4** Let  $T \subset N$  and  $Q_T \in \mathcal{R}^{|T|}$  such that  $|C(P_{-T}, Q_T)| = 1$ . 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  $\mu$  is individually rational. Suppose not. Then there exists  $i \in N$  such that  $iP_i\mu(i)$ . This implies that  $i \in T$ . But then  $\mu(i)P_i\lambda(i)R_i i$ , if  $i \in T$ , a desired contradiction.

Let  $K$  denote all players who prefer  $\mu$  to  $\lambda$ . Then  $T \subset K$ . The local blocking lemma (Lemma 9) shows that there exists a pair  $(i, j)$  that blocks  $\mu$  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  $\lambda$ . Hence  $j \notin T$ . Thus both  $i$  and  $j$  are not in  $T$ . The fact that  $(i, j)$  blocks  $\mu$  implies that  $\mu$  is not stable under  $(P_{-T}, Q_T)$ . Theorem 2 shows that  $\mu$  is not in the core  $C(P_{-T}, Q_T)$ , a contradiction.  $\square$

In the close of this section we use an example from Roth (1982) to show that the single-valued core may not satisfy the external stability in the roommate and marriage problems; as assumed in Theorem 3. Yet the single-valued core is coalitionally strategy proof.

**Example 10** (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:

$$\begin{aligned} P_{m_1} &= (w_2, w_1, w_3, m_1) & P_{w_1} &= (m_1, m_2, m_3, w_1) \\ P_{m_2} &= (w_1, w_2, w_3, m_2) & P_{w_2} &= (m_3, m_1, m_2, w_2) \\ P_{m_3} &= (w_1, w_2, w_3, m_3) & P_{w_3} &= (m_1, m_2, m_3, w_3). \end{aligned}$$

Without any confusion, we represent these preferences in a manner that is consistent with the marriage problem. Note that every matching in this example satisfies individual rationality. There exists a unique stable matching  $\mu$  in the core  $C(P)$

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

The matching  $\lambda$

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

is not stable but it is in the  $vNEM$  solution. In fact one can show that this matching market has a unique  $vNEM$  solution that consists of the two matchings  $\{\mu, \lambda\}$ .<sup>3</sup> Therefore the core  $C(P) = \{\mu\}$  does not satisfy the external stability since it does not dominate the matching  $\lambda$ .

But it is known that the core mechanism is coalitionally strategy proof as long as the core is single-valued in the marriage problem; see Dubins and Freedman (1981) and Roth (1982). Hence the external stability is not a necessary condition for the single-valued core to be coalitionally strategy proof in the roommate and the marriage problems. It remains open if the single-valued core in the generalized matching problem in Sonmez (1995) is coalitionally strategy proof. In the general coalition structure framework below we have an example to show that the single-valued core is not strategy proof. Thus the results in this section depend on heavily the structures of matchings.

## 5 Coalition Structures

The set  $\mathcal{P}(N)$  admits much richer coalition structures than the set  $\mathcal{M}$  of matchings. When we pay attentions to  $\mathcal{M}$ , the single-valued core mechanism has the appealing coalitionally strategy proof property as shown in Theorem 4. A natural question is if the single-valued core is coalitionally strategy proof for the general situations with coalition structures. We now adopt an example<sup>4</sup> in Roth (1985) to show that this is not the case. Even if the core is single-valued, the core mechanism is manipulable by some individual toward to his own interests.

**Example 11** *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:*

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

<sup>3</sup>A formal proof of this is available from the author.

<sup>4</sup>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.

$$P_{S_4} = (\{C_1, S_3, S_4\}, \{C_1, S_2, S_4\}, \{C_2, S_4\}, \{C_3, S_4\}, \{S_4\}).$$

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  $\mu$  leaves  $S_4$  single,  $C_2$  and  $C_3$  must match with some player in  $\{S_1, S_2, S_3\}$  in  $\mu$ . This implies that  $C_1$  can only match at most (also at least) one player in  $\{S_1, S_2, S_3\}$  in  $\mu$ . If  $C_1$  is matched with  $S_1$  in  $\mu$ , then  $C_3$  must be matched with  $S_3$ . Then  $\mu$  is dominated by  $\{C_1, S_3, S_4\}$ . If  $C_1$  is matched with  $S_2$  in  $\mu$ , then it is dominated by  $\{C_1, S_2, S_4\}$ . If  $C_1$  is matched with  $S_3$  in  $\mu$ , then it is dominated by  $\{C_1, S_3, S_4\}$ .

There are four other possibilities in which every player is matched:

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

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

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

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 unique one 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  $(Q_{C_1}, P_{-C_1})$ . But  $C_1$  prefers  $\mathcal{B}$  to  $\mathcal{A}$ . Thus the core is not strategy proof even if it is single-valued, unlike Theorem 4.  $\square$

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

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. Our next result shows that the  $vN\&M$  solution as a mechanism is always coalitionally strategy proof as long as it is well defined and single-valued. First we need a lemma and a corollary.

**Lemma 12** 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  $\exists B \subset N$  such that  $BP_i\mathcal{A}(i)$  for all  $i \in B$ .

**Proof** By the way of contradiction. Suppose that  $V(P) = \{\mathcal{A}\}$  for some  $vN\&M$  solution  $V(P) \in \mathcal{V}(P)$ . Suppose on the contrary that there exists a coalition  $B$  such that  $BP_i\mathcal{A}(i)$  for all  $i \in B$ . Then we can construct an infinite chain  $A_1, A_2, \dots$ , of non-empty, disjoint sets in  $\mathcal{A}$ , as shown below. But  $\mathcal{A}$  is finite, a desired contradiction.

Let  $K \in N$ . Given  $\{A_1, \dots, A_K\} \subset \mathcal{A}$  such that  $\{B, A_1, \dots, A_K\}$  are disjoint (if  $K = 0$ , then  $\{B\}$  is trivially disjoint). Let  $\mathcal{B}_K = \{B, A_1, \dots, A_K, N - B - A_1 - A_2 - \dots - A_K\}$ . 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 external stability, 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  $k = 1, \dots, K$ , 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  $A_{K+1} \in \mathcal{A}$  implies that  $B \cap A_{K+1} = \emptyset$ . Hence,  $B, A_1, \dots, A_{K+1}$  are disjoint and  $\{A_1, \dots, A_{K+1}\} \subset \mathcal{A}$ . This completes the proof of the lemma.  $\square$

**Corollary 13** Let  $P \in \mathcal{R}^n$  be a profile of strict preferences such that  $V(P) = \{\mathcal{A}\}$  for some  $V(P) \in \mathcal{V}(P)$ . Then  $\mathcal{A}$  dominates every other coalition structure and is not dominated by any other coalition structure.

Next we need to relate the  $vN\&M$  solution to the core and answer the question when there exists a unique  $vN\&M$  solution.

**Theorem 14** Let  $P \in \mathcal{R}^n$ . Suppose  $V(P) = \{\mathcal{A}\}$  for some  $V(P) \in \mathcal{V}(P)$ . Then  $\mathcal{V}(P) = \{V(P)\}$  and  $V(P) = \{\mathcal{A}\}$ . Moreover  $V(P) = C(P)$ .

**Proof** 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  $\mathcal{B}$  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 Lemma 12.  $V(P) = C(P)$  follows from Corollary 13.  $\square$

**Theorem 15** *Let  $|V(P)| = 1$  for all profiles  $P \in \mathcal{R}^n$  of strict preferences. Then the  $vN\&M$  solution  $V$  is coalitionally strategy proof.*

**First Proof of Theorem 15** Let  $\varphi = V$  be the  $vN\&M$  solution mechanism. By the way of contradiction suppose that there exists a profile  $(P_{-T}, Q_T)$  such that

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

Let  $\mathcal{A} = \varphi(P_{-T}, Q_T)$  and  $\mathcal{B} = \varphi(P_{-T}, P_T)$ . Then  $\mathcal{A} \notin V(P_{-T}, P_T)$  and  $\mathcal{B} \notin V(P_{-T}, Q_T)$ . The external stability shows that  $\exists B \in \mathcal{B}$  such that  $\mathcal{B}P_i\mathcal{A}$  for all  $i \in B$ . Since  $\mathcal{B} \notin V(P_{-T}, Q_T)$ , by the external stability again,  $\exists A \in \mathcal{A}$  such that  $\mathcal{A}P_i\mathcal{B}$  for  $i \in A - T$  and  $\mathcal{A}Q_i\mathcal{B}$  for  $i \in T \cap A$ . But, by the assumption that  $\mathcal{A}P_i\mathcal{B}$  for all  $i \in T$ , it follows that  $\mathcal{A}P_i\mathcal{B}$  for all  $i \in T \cap A$ . Therefore,  $\mathcal{A}P_i\mathcal{B}$  for all  $i \in A$  contradicting Corollary 13. This completes the proof.  $\square$

Next we provide a second proof of Theorem 15 by means of the core and the ideas in Demange (1987) and Sonmez (1995).

**Second Proof of Theorem 15** Again, by the way of contradiction, suppose that there exists a profile  $P_{-T}$  and  $Q_T$  such that

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

Since  $\varphi(P_{-T}, Q_T)$  is not in  $V(P)$ , there exists a coalition  $S$  such that  $\varphi(P_{-T}, Q_T)$  is dominated by  $\varphi(P_{-T}, P_T)$  via the coalition  $S$ . That is

$$\varphi(P_{-T}, P_T)P_i\varphi(P_{-T}, Q_T), \forall i \in S.$$

Thus  $S \cap T = \emptyset$ . Since

$$\varphi(P_{-T}, P_T)(i) \in S, \forall i \in S$$

it follows that  $\varphi(P_{-T}, P_T)$  dominates  $\varphi(P_{-T}, Q_T)$  under  $(P_{-T}, Q_T)$  via  $S$ . Since  $S \cap T = \emptyset$ , it follows that  $\varphi(P_{-T}, Q_T)$  is not in the core  $C(P_{-T}, Q_T)$  contradicting Theorem 14.  $\square$

If a core is singleton and satisfies the external stability, then it must be a singleton  $vNEM$  solution. Theorem 14 shows that it will be the unique  $vNEM$  solution. Further Theorem 14 shows that the converse is also true. If there exists a singleton  $vNEM$  solution, then it must coincide with the core. Thus the core must be singleton and externally stable. Sonmez (1995) showed that the singleton core with external stability is coalitionally strategy proof. Thus Theorem 15 is a corollary of the result in Sonmez (1995) and Theorem 14. Or the result in Sonmez (1995) is a corollary of Theorems 14 and 15. Either way, Lemma 12 is essential.

## 6 $vNEM$ Solutions versus the Core

In this section we show an analogy to Theorem 5 and Corollary 6 for the general coalition structures. Since a conclusion in the generalized matching model may not apply to the coalition structures, we present a formal proof of them for the coalition structures.

The results presented here are closely related to those for the housing swapping market in Shapley and Scarf (1974) in which each agent owns a house and consumes at most one house. An allocation is a permutation of all houses. Roth and Postlewaite (1977) showed that the core in the housing swapping market is single-valued 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. Ma (1994) showed that the core mechanism is the only mechanism that satisfies strategy proofness, individual rationality and Pareto optimality. Sonmez (1995) generalized these results to the generalized matching market.

In what follows we prove that there exists a strategy proof mechanism that is also individually rational and Pareto optimal only if the core is single-valued for the situations with 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  $Q$  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 structures of the matching  $\mu$ . The proof in Ma (1994) is for the housing swapping market. Our proof below follows Ma (1994) closely. Before the proof, we need two simple but useful lemmas.

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

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

to be the set of players who prefer  $\mathcal{A}$  to  $\mathcal{B}$ . 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 16** *If  $\mathcal{A}$  and  $\mathcal{B}$  are Pareto optimal under  $P$  and  $\mathcal{A} \neq \mathcal{B}$ , then  $J(\mathcal{A}, \mathcal{B}; P) \neq \emptyset$ .*

**Proof** If  $J(\mathcal{A}, \mathcal{B}; P) = \emptyset$ , then  $\mathcal{B}R_i\mathcal{A}$  for all  $i \in N$ . Since  $\mathcal{B} \neq \mathcal{A}$ , it follows that there exists at least one  $i \in N$  such that  $\mathcal{B}P_i\mathcal{A}$ . This implies that  $\mathcal{A}$  is not Pareto optimal.  $\square$

**Lemma 17** *Let  $\mathcal{A} \in C(P)$  be a core coalition structure and  $\mathcal{B} \in IR(P) \cap PO(P)$  be an individually rational and Pareto optimal coalition structure. If  $\mathcal{B} \neq \mathcal{A}$ , then  $\exists j \in J(\mathcal{A}, \mathcal{B}; P)$  such that  $\mathcal{A}P_j\mathcal{B}P_jS_j$ .*

**Proof** Suppose on the contrary that  $\mathcal{B}(j) = \{j\}$  for all  $j \in J(\mathcal{A}, \mathcal{B}; P)$ .<sup>5</sup> That is all agents in  $J(\mathcal{A}, \mathcal{B}; P)$  remain single in  $\mathcal{B}$ . Lemma 16 shows that  $J(\mathcal{A}, \mathcal{B}; P) \neq \emptyset$ . Thus  $N - J(\mathcal{A}, \mathcal{B}; P)$  forms a coalition that weakly dominates  $\mathcal{A}$  under the coalition structure  $\mathcal{B}$ . By strict preferences, there exists a coalition  $B \in \mathcal{B}$  such that  $\mathcal{B}$  dominates  $\mathcal{A}$  via  $B$ . In fact  $J(\mathcal{B}, \mathcal{A}; P)$  is such a coalition. This is a contradiction to the assumption that  $\mathcal{A}$  is in the core.  $\square$

**Theorem 18** *Suppose that  $C(P) \neq \emptyset$  for all profiles  $P \in \mathcal{R}^n$  of strict preferences and let  $\varphi : \mathcal{R}^n \rightarrow \mathcal{P}(N)$  be a strategy proof mechanism that also satisfies individual rationality and Pareto optimality. Then  $|C(P)| \leq 1$  for all  $P \in \mathcal{R}^n$ . Moreover  $\varphi(P) = C(P)$  for all  $P$  if  $C(P) \neq \emptyset$ .*

**Proof** Let  $P \in \mathcal{R}^n$  such that  $C(P) \neq \emptyset$ . Let  $\varphi : \mathcal{R}^n \rightarrow \mathcal{P}(N)$  be a mechanism that is strategy proof, individually rational and Pareto optimal. We first show that  $|C(P)| \leq 1$  for all  $P \in \mathcal{R}^n$ .

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

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

---

<sup>5</sup>By definition,  $\mathcal{A}P_j\mathcal{B}$  for all  $j \in J(\mathcal{A}, \mathcal{B}; P)$ . If the lemma is false, it must be the case that  $S_jR_j\mathcal{B}$  for all  $j \in J(\mathcal{A}, \mathcal{B}; P)$ . Since  $\mathcal{B}$  is individually rational, it must be the case that  $\mathcal{B}(j)I_jS_j$  for all  $j \in J(\mathcal{A}, \mathcal{B}; P)$ .

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

**Proof** Suppose there exists  $\mathcal{A} \in C(Q)$  or  $\mathcal{A} \in IR(Q) \cap PO(Q)$  such that  $\mathcal{A} \neq \mathcal{D}$ . Then Lemma 17 shows that there exists  $j \in J(\mathcal{D}, \mathcal{A}; Q)$  such that  $\mathcal{D}Q_j \mathcal{A}Q_j S_j$ . But there exists no  $\mathcal{A}$  between  $\mathcal{D}$  and  $S_j$  under  $Q_j$ .  $\square$

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

**Proof** We use the induction on the size  $|T|$  of the coalition  $T$  to show this. When  $|T| = 0$ , 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  $|T| \leq k$ .

Suppose, by the way of contradiction, that  $\varphi(P_T, Q_{-T}) \neq \mathcal{D}$  for some  $T \subset N$  such that  $|T| = k + 1$ . Let  $Q' = (P_T, Q_{-T})$ . Then Lemma 17 shows that there exists  $j \in J(\mathcal{D}, \varphi(Q'); Q')$  such that

$$\mathcal{D}Q'_j \varphi(Q') Q'_j S_j. \quad (5)$$

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

$$\mathcal{D}Q_j \varphi(Q') Q_j S_j.$$

But there exists no coalition structure  $\varphi(Q')$  between  $\mathcal{D}$  and  $S_j$  under  $Q_j$ . This shows that  $j \in T$ .

If  $j \in T$ , then it follows from (5) that

$$\mathcal{D}P_j \varphi(Q'). \quad (6)$$

It follows from the induction hypothesis that

$$\varphi(Q_j, Q'_{-j}) = \mathcal{D}. \quad (7)$$

Thus it follows from (6) and (7) that

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

contradicting  $\varphi$  is strategy proof.  $\square$

Since  $\mathcal{D}$  is arbitrarily chosen, it follows that  $|C(P)| \leq 1$  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  $|C(P)| = 1$ . This completes the proof.  $\square$

**Corollary 19** *A mechanism  $\varphi : \mathcal{R}^n \rightarrow \mathcal{P}(N)$  is strategy proof, individually rational and Pareto optimal if, and only if  $\varphi$  is the  $vN\&M$  solution mechanism  $V$  in the domain with single-valued  $vN\&M$  solution.*

**Proof** Theorem 15 shows that the  $vN\&M$  solution  $V$  is coalitionally strategy proof. Theorem 14 shows that  $V$  coincides with the core  $C$ . Therefore  $V$  is a mechanism that is strategy proof, individually rational and Pareto optimal. The “only if” part follows from Theorems 18 and 14.  $\square$

These results are related to the question what may be the largest domain for a mechanism with the three properties, namely, strategy proofness, individual rationality and Pareto optimality. The largest domain admitting such a mechanism is that with the single-valued core (Theorem 18). We also show that this largest domain is reached in the roommate and marriage problems (Theorem 4). But it cannot be reached for general coalition structures since the single-valued core may not be strategy proof (Example 11). Nevertheless the domain with the single-valued  $vN\&M$  solution, though relatively smaller than that with single-valued core, admits at least one mechanism that satisfies the three properties, since we show that the  $vN\&M$  solution is coalitionally strategy proof and it is individually rational and Pareto optimal (Theorems 14 and 15). Then Corollary 19 shows that the  $vN\&M$  solution is the only strategy proof, individually rational and Pareto optimal mechanism in this smaller domain.

## 7 Two Remarks

**Remark 1** In practice whether a coalition structure forms depends on not only players’ personal preferences but also those factors beyond preferences. On one hand, there are many factors that encourage coalitions to form. On the other hand, there are some factors that may prevent coalition formation. For example, cultural differences, the law and communication costs may limit coalition formation to certain degree. In Section 2, we define players’ preferences over the set of all coalition structures. In this definition, we implicitly exclude the impacts of those factors beyond preferences on coalition formation and assume that all coalition structures are admissible. This frictionless

assumption may be unrealistic when those factors such as the law present. To embody these factors into account, we can consider admissible coalition structures. Precisely, let  $\mathcal{B}^f(N) \subset \mathcal{P}(N)$  such that  $(\{i\}_{i \in N}) \in \mathcal{B}^f(N)$  be an admissible subset of all coalition structures. Therefore, culture and the law will at least give individuals the right to not joint in any coalition. We may then restrict players' preferences to  $\mathcal{B}^f(N)$ . And then all results in this paper apply to  $\mathcal{B}^f(N)$ .

**Remark 2** A network is a subset of the set of all subsets  $2^N$  with size 2. A link between  $i$  and  $j$  is the subset  $\{i, j\}$ . A network is complete if each player in it is linked with each other. A complete subnetwork is equivalent to a coalition in a coalition structure. Typically a network of the players  $N$  admits a more general graphic structure than a coalition structure, because a coalition structure often consists of several complete subnetworks. Due to these differences, the results obtained for the coalition structures may not apply to the networks.

**Example 20** Let  $N = \{1, 2, 3, 4, 5\}$ . The preferences are as follows

$$\begin{aligned} P_1 &= (\{1, 2, 5\}, \{1, 2\}, \{1, 5\}, \{1\}) & P_2 &= (\{1, 2, 3\}, \{2, 3\}, \{1, 2\}, \{2\}) \\ P_3 &= (\{2, 3, 4\}, \{3, 4\}, \{2, 3\}, \{3\}) & P_4 &= (\{3, 4, 5\}, \{4, 5\}, \{3, 4\}, \{4\}) \\ P_5 &= (\{1, 4, 5\}, \{1, 5\}, \{4, 5\}, \{5\}). \end{aligned}$$

There are five (nontrivial) coalition structures that satisfy individual rationality.

$$\begin{aligned} \mathcal{B}_1 &= \{\{1, 2\}, \{3, 4\}, \{5\}\} & \mathcal{B}_2 &= \{\{1, 2\}, \{4, 5\}, \{3\}\} \\ \mathcal{B}_3 &= \{\{1, 5\}, \{2, 3\}, \{4\}\} & \mathcal{B}_4 &= \{\{2, 3\}, \{4, 5\}, \{1\}\} \\ \mathcal{B}_5 &= \{\{1, 5\}, \{3, 4\}, \{2\}\} \end{aligned}$$

These coalition structures are dominated by the coalitions  $\{4, 5\}$ ,  $\{2, 3\}$ ,  $\{3, 4\}$ ,  $\{1, 5\}$  and  $\{1, 2\}$ , respectively. Hence, the coalition structure core is empty.

But at the preference profile above the network core is nonempty and there exists a unique network  $vNEM$  solution. Indeed, the network core and the network  $vNEM$  solution coincide and consist of the network  $\{12, 23, 34, 45, 51\}$ , in which 1 is linked with 2 and 5, 2 with 1 and 3, 3 with 2 and 4, 4 with 3 and 5, and 5 with 1 and 4.

## References

- [1] AUMANN, R. J. and J.H. DRÈZE (1974): “Cooperative Games with Coalition Structures,” *International Journal of Game Theory*, 3, 217-237.
- [2] AUMANN, R.J. and M. MACHLER (1964): The Bargaining Set for Cooperative Games, in “Advances in Game Theory”, eds., M. Dresher, L.S.Shapley and A.W. Tucker, Princeton University Press, Princeton.
- [3] AUMANN, R.J. and R. MYERSON (1988): Endogenous Formation of Links between Players and of Coalitions: an application of the Shapley value, in “The Shapley Value: Essays in Honor of Lloyd Shapley”, ed., A. Roth, Cambridge University Press, Cambridge.
- [4] BIRD, C. (1984): “Group Incentive Compatibility in a Market with Indivisible Goods,” *Economic Letter*, 14, 309-313.
- [5] BONDAREVA, O.N., T.E. KULAKOVSKAYA and N.I. NAUMOVA (1979): “The Solution of the Four-Person Games,” *Leningrad University Vestnik* 6, 104-105.
- [6] DEMANGE, G. (1987): “Nonmanipulable Cores,” *Econometrica*, 55, 1057-1074.
- [7] DRÈZE, J.H. and J. GREENBERG (1980): “Hedonic Coalitions: Optimality and Stability,” *Econometrica*, 48, 987-1003.
- [8] DUBINS, L.E. and D.A. FREEDMAN (1981): “Machiavelli and the Gale-Shapley Algorithm,” *American Mathematical Monthly* 88, 485-94.
- [9] HART, S. and M. KURZ (1983): “Endogenous Formation of Coalitions,” *Econometrica*, 51, 1047-1064.
- [10] GALE, D. and L. SHAPLEY (1962): “College Admissions and the Stability of Marriage,” *American Mathematical Monthly*, 69, 9-15.
- [11] GREENBERG, J. (1990): The Theory of Social Situations: An Alternative Game Theoretic Approach, Cambridge University Press.
- [12] GREENBERG, J. (1994): Coalitional Structures, in “Handbook of Game Theory, V. 2” edited by R. AUMANN and S. HART, Elsevier Science Publisher.

- [13] JACKSON, M and A. WOLINSKY (1995): A Strategic Model of Social and economic Networks, Northwestern University.
- [14] KAHAN, J. and A. RAPOPORT (1984): Theories of Coalition Formation, Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey.
- [15] LUCAS, W. (1992): Von Neumann-Morgenstern Stable Sets, in “Handbook of Game Theory (I), with economic applications” edited by R.J. Aumann and S. Hart, Elsevier Science Publishers, Amsterdam.
- [16] MA, J. P. (1994): “Strategy Proofness and the Strict Core in a Market with Indivisibilities,” *International Journal of Game Theory*, 23, 75-83.
- [17] MYERSON, R. (1977): “Graphs and Cooperation in Games,” *Mathematics of Operations Research*, 2, 225-229.
- [18] NEUMANN, J. VON and O. MORGENSTERN (1944, 1953): Theory of Games and Economic Behavior. Princeton University Press, Princeton, New Jersey.
- [19] NOUWELAND, A.V.D. (1993): Games and Graphs in Economic Situations, Ph.D. Thesis, Tilburg University.
- [20] QIN, C-Z (1995): Endogenous Formation of Cooperation Structures, University of California at Santa Barbera.
- [21] ROTH, A. (1982): “The Economics of Matching: Stability and Incentives,” *Mathematics of Operation Research*, 7, 617-628.
- [22] ROTH, A. (1982b): “Incentive Compatibility in a Market with Indivisible Goods,” *Economic Letters*, 9, 127-132.
- [23] ROTH, A. (1985): “The College Admissions Problem is not Equivalent to the Marriage Problem,” *Journal of economic Theory*, 36, 277-88.
- [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.E. and M. SOTOMAYOR (1990): Two-sided Matching: a Study in Game-theoretic Modeling and Analysis, Cambridge University Press, Cambridge.

- [26] SHAPLEY, L. and H. E. SCARF (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 23-28.
- [27] SÖNMEZ, T. (1995): Strategy-proofness and Singleton Cores in Generalized Matching Problems, University of Rochester.
- [28] THRALL, R. and W. LUCAS (1963): "n-Person Games in Partition Function Form," *Naval Research Logistics Quarterly*, 10, 281-298.