

**THE ROLE PLAYED BY SIMPLE OUTCOMES IN THE COALITION  
FORMATION PROCESS IN MATCHING MARKETS**

by

MARILDA SOTOMAYOR<sup>1</sup>

Department of Economics

Universidade de São Paulo, Cidade Universitária, Av. Prof. Luciano Gualberto 908  
05508-900, São Paulo, SP, Brazil

e-mail: [marildas@usp.br](mailto:marildas@usp.br)

9/2006

**ABSTRACT**

This paper provides a novel way to analyze matching markets, which is close to the dynamic adjustment process by which such markets are presumably cleared. This approach allows to show how the possibility of transfers in the continuous case and the structure of preferences in the discrete case determine core existence and other properties. A clearer view of the connections and differences between the various one-to-one matching models is then obtained by using quite elementary notions and arguments. We show that some property of the core of the two-sided matchings is a necessary and sufficient condition for the non-emptiness of the core of the one-sided markets. A necessary and sufficient condition for the non-emptiness of the strong core of the discrete markets is also obtained.

Keywords: core, stable matching

JEL numbers: C78, D78

---

<sup>1</sup> This paper is partially supported by CNPq-Brazil.

## INTRODUCTION

The matching markets have as primary object the formation of partnerships. When agents from a set  $F$  can only pair with agents from a set  $W$ , with  $F \cup W = N$  and  $F \cap W = \emptyset$ , the matching market is said to have two sides. Two-sided matching markets have been widely studied by many authors since the seminal paper of Gale and Shapley (1962). A more general matching market is obtained when  $F = W = N$ . These are the so-called *one-sided matching markets*.

For a wide class of matching markets, each player is allowed to form one partnership at most. We say that the matching is one-to-one. In some of them, whenever two agents are matched they can share the surplus they produce in any way they want. In some other ones, the payoffs of the partners are pre-set and fixed, and constitute one of the factors that determine the preferences of the agents. Traditionally, the former kind of market has been modeled as a continuous matching market with linear utilities. An *outcome* specifies a matching plus a payoff for each agent. The latter has been modeled as a discrete matching market. An *outcome* only specifies a matching. The discrete two-sided and discrete one-sided matching models are known respectively as Marriage Problem and Roommate Problem since Gale and Shapley (1962). The continuous two-sided matching model is the well known Assignment Game of Shapley and Shubik (1972). The one-sided Assignment game has not yet been studied. These four models are the matching models treated in this paper.

The traditional cooperative game theoretic approach assumes as postulate the existence of an environment where players are permitted to freely communicate, to make binding agreements and preferences over outcomes, as well as the rules of the game are common knowledge. Two fundamental questions can be posed: A. What coalitions are likely to form? B. How will the members of a coalition split their joint payoff?

The answer to these questions is the central issue of games in which the main activity of the players is to form coalitions. It has to do with the choice of an adequate concept of equilibrium. Typically, players do not first join a coalition and then negotiate their payoffs; rather, both coalitions and payoffs are determined simultaneously. Under the assumption above, only those coalitions whose members are engaged into optimal cooperation are expected to occur. That is, in equilibrium, a coalition with a specific

agreement will not emerge unless the members of the coalition believe that more favorable terms cannot be obtained elsewhere.

This intuitive idea of optimal cooperation is, in the matching markets, captured by the notion of setwise-stability or stability for short. An outcome  $X$  is setwise-stable if there is no coalition of agents who by forming new partnerships only among themselves, possibly dissolving some partnerships in  $X$  and possibly keeping other ones, can all obtain a higher payoff than the one given by  $X$ . For the discrete and continuous models treated here the matching is one-to-one. Then, in order to check instabilities in the discrete markets it suffices to verify the existence of pairs of agents, not matched to each other, but who could be better off by becoming partners. In order to check instabilities in the continuous markets we only need to verify the existence of pairs of agents, not matched to each other, but who, by becoming partners, could negotiate a payoff, which is higher than their current payoff. The core and the set of stable outcomes coincide in the four models: a pair of agents causes an instability if and only if it blocks the outcome. Also, the strong core and the set of strongly stable outcomes coincide in the four models. In the continuous models and in the discrete models with strict preferences, the core equals the strong core. The core of the one-sided matching models and the strong core of the Marriage model with non-necessarily strict preferences may be empty.

However, the cooperative game theoretic approach does not provide the essential informations concerning the precise bargaining process. The history underlying a core outcome is not clear and it is a matter for the players to work out for themselves in order to get a specific agreement.

This paper provides us with a new conceptual view of matching markets that can contribute to a better understanding of both dynamics of the coalition formation process underlying a given stable outcome and the coalitional structure of such outcome. The results proved here propitiate a clearer view of the connections and fundamental differences between the four one-to-one matching models. The concepts and results for the one-sided markets are obtained from the corresponding concepts and results for the two-sided markets by making the two sides to coincide with the whole set of players. Which results are affected by the number of sides of the market, and which ones are not, become more evident. The concepts as well as the results for the continuous and discrete

models are presented together, so readers can know the distance of the two kinds of models immediately. This was made possible because each concept for the continuous models, as well as each result proved here for these models, have an analogue in the discrete cases. Furthermore, the proofs presented here for the one-sided and two-sided Assignment games do not make use of the linearity of the utility functions and the nature of the mathematical arguments is combinatorial<sup>2</sup>. Consequently, all proofs are very similar to those for the Marriage and the Roommate markets.

The key concepts are those of *simple outcome* and *Pareto optimal simple outcome*. For a subset  $A$  of players, an outcome is *A-simple* if it is feasible, individually rational and no matched player belonging to  $A$  is part of a blocking pair. That is, if a player in  $A$  is engaged into cooperation then this cooperation is optimal for him/her. For the discrete markets with non necessarily strict preferences we will use the concept of strongly *A-simple* outcome similarly defined. (See Definition 7). The outcome where every one is unmatched is *A-simple* and strongly *A-simple*. Also, every core outcome is *A-simple* and every strong core outcome is strongly *A-simple*.

In the two-sided matching models we will concentrate on *A-simple* outcomes where  $A=N$  or  $A$  is one of the sides of the market - to fix ideas we will choose  $A=W$ . For the one-sided matching models we will consider  $A=N$ . We say that the *W-simple* outcome  $X$  *F-extends* the *W-simple* outcome  $Y$ , with  $X \neq Y$ , if either every *F-agent* weakly prefers  $X$  to  $Y$  and at least one *F-agent* strictly prefers  $X$  to  $Y$ , or (only pertinent when the market is continuous) every *F-agent* gets the same payoff in  $X$  and  $Y$  but  $X$  has less blocking pairs than  $Y$ . A *W-simple* outcome is *Pareto optimal W-simple outcome* if it cannot be *F-extended* to any other *W-simple* outcome. By making  $F=W=N$  we obtain the concept of *N-simple extension* and *Pareto optimal N-simple outcome*. Then, the *N-simple* outcome  $X$  *N-extends* (extends, for short) the *N-simple* outcome  $Y$ , if every agent weakly prefers  $X$  to  $Y$  and at least one agent strictly prefers  $X$  to  $Y$ . The concepts of *strongly N-simple extension* and *strongly W-simple extension* are similarly defined.

---

<sup>2</sup> This fact has already been observed for the two-sided matching markets in Roth and Sotomayor (1996), where a unified model is presented. An analogous formulation of this model can be used to unify the one-sided matching markets.

In outcomes that are not  $N$ -simple, some players who are engaged into cooperation can block, so they are engaged into cooperation which is not optimal for them.

The existence of strong core outcomes of the Marriage and Roommate models with non-necessarily strict preferences has not yet been explored in the literature. The strong core of these models may be empty. (See Example 1 for the Marriage model and Example 3 of Gale and Shapley (1962) for the Roommate model. Also, the core of the one-sided matching models may be empty. (See Example 2 for the continuous case)<sup>3</sup>. Our main finding is the identification of a new necessary and sufficient condition for the non-emptiness of the strong core of the discrete models. A convenient adaptation of this condition constitutes a necessary and sufficient requirement for the non-emptiness of the core of the one-sided Assignment game and the Roommate model with strict preferences.

The existence results can be summarized as follows:

*1. For the one-sided Assignment game and for the Roommate model with strict preferences, the core is non-empty if and only if every unstable and  $N$ -simple outcome can be extended to an  $N$ -simple outcome.*

When preferences are not strict for the Roommate model, this condition is not necessary for the non-emptiness of the core. We show, through an example, that we may have unstable and  $N$ -simple outcomes which cannot have an  $N$ -simple extension even when the core is non-empty. However,

*2. For the Marriage and Roommate models with non-necessarily strict preferences, the strong core is non-empty if and only if every unstable and strongly  $N$ -simple matching can be extended to a strongly  $N$ -simple matching.*

The proofs that the requirements are sufficient use a very simple argument. All of them follow from the fact shown here that Pareto optimal  $N$ -simple outcomes, as well as Pareto optimal strongly  $N$ -simple outcomes, always exist for the four models. If the conditions are satisfied, these outcomes cannot be unstable. (We must point out that a

---

<sup>3</sup> Since Gale and Shapley (1962) the problem of existence of stable matchings for the Roommate model has been the subject of several research articles. In Irvin (1985) an algorithm is presented to find a stable matching when the set of stable matchings is non-empty. Tan (1991) has identified a necessary and sufficient condition, stated in terms of preference restriction, for the existence of stable matchings for the roommate problem. Chung (2000) has identified a condition called *no odd rings* that is sufficient, but not necessary, for the existence of stable matchings for this market.

Pareto optimal  $N$ -simple matching for the Roommate problem with non-strict preferences may be unstable, even when the core is non-empty (see Example 4)). The proof that the conditions are necessary is not so straightforward.

The important application of this condition is that it is satisfied by the unstable  $W$ -simple outcomes of both two-sided matching models treated here:

*For the two-sided Assignment game and for the Marriage model with non-necessarily strict preferences, every unstable and  $W$ -simple outcome can be extended to a  $W$ -simple outcome.*

The basic idea of the proof of this result is the following. For the Marriage market, if a  $W$ -simple matching  $\mu$  is unstable, then we can choose a blocking pair  $(f,w)$  such that by matching the pair and leaving unmatched  $\mu(f)$ , if any, no  $F$ -player is hurt,  $f$  is better off and the matching continues to be  $W$ -simple. The crucial point is that we cannot do that with the Pareto optimal  $W$ -simple matching, which implies that this outcome is stable. For the two-sided Assignment game, if a  $W$ -simple outcome is unstable then it has a  $W$ -simple extension. In fact, otherwise it is Pareto optimal  $W$ -simple outcome. However, this outcome is stable, for otherwise we could choose a blocking pair that yields a chain of new trades that keeps the payoffs of the  $F$ -players and converges to a  $W$ -simple outcome with less blocking pairs, which is a contradiction.

Since Pareto optimal  $W$ -simple outcomes for these models always exist a non-constructive proof of the non-emptiness of the core of the Marriage model and a combinatorial proof of the existence of core outcomes of the two-sided Assignment game are then obtained.<sup>4</sup>

The essence of these results is captured by the following coalition formation process in which  $N$ -simple and  $W$ -simple outcomes play a fundamental role. Starting from the outcome where every player is unmatched, we can gradually increase cooperation by making  $N$ -simple (respectively,  $W$ -simple) Pareto improvements and still staying within  $N$ -simple (respectively,  $W$ -simple) outcomes, until no transaction is able to

---

<sup>4</sup> In Sotomayor (1999) and (2000) the proof of the non-emptiness of the core of respectively, the Marriage market and the one-to-one hybrid market, is obtained directly by showing that Pareto  $W$ -optimal simple outcomes are stable. In these papers the concept of Pareto optimal  $W$ -simple outcome is slightly different from the one presented here. For the sake of completeness these results will be proved again in the present paper. The hybrid market unifies and extends the Assignment game and the Marriage market.

benefit the agents involved, or until the payoff cannot be  $N$ -simple (respectively,  $W$ -simple) anymore. In the first case the core has been reached (respectively, the core has been reached without hurting any  $F$ -agent). In the other case, the core is empty, so increase in payoffs is only available through non-optimal (in the selfish sense) cooperation of some agents.

We must point out that we are not providing an algorithm to produce core outcomes when they exist. What we are trying to do is to make clearer the understanding of the dynamics of the coalition formation process that leads to a core outcome. It is this process that can be identified, for example, in the deferred-acceptance algorithm of Gale and Shapley for the Marriage model and in the Top Trading Cycles algorithm of Gale for the Housing market of Shapley and Scarf (1974). Each acceptance of a  $W$ -player in the deferred-acceptance algorithm, with the  $W$ -players proposing, yields a  $W$ -simple matching. Also, each cycle formed by the Top Trading Cycles algorithm produces an  $N$ -simple outcome, defined in the straightforward manner for the Housing model (see Sotomayor, 2005). Both algorithms converges to a core outcome after a finite number of steps.

Other properties of the  $N$ -simple outcomes also contribute to a better understanding of the coalition formation process. In the continuous models, those  $N$ -simple Pareto improvements produce a sequence of adjustments in which payoffs are preserved for agents already matched and increase only for those newly matched. That is, once a transaction is done, the agents involved will keep their payoffs (even if they dissolve their current partnerships to enter new ones) at the subsequent adjustments. Consequently, only agents who are not currently trading are able, by trading among them, to increase their payoffs. In addition, at any payoff adjustment, the sum of the values of the trades done by the current trading agents is equal to the sum of the value of the trades done by these agents at any stable outcome, and attains the maximum these agents can obtain by their own.

In the discrete models with strict preferences, those  $N$ -simple Pareto improvements produce a sequence of  $N$ -simple outcomes that keep the trades done and add new ones. Thus, once a transaction is done, it will not be undone at the subsequent adjustments. Only agents who are not currently trading are able, by trading among them,

to be better off. Furthermore, the trading agents at an  $N$ -simple outcome continue to trade only among themselves under any stable outcome.

If at some step of the coalition formation process for the discrete models with strict preferences (respectively, continuous models), some non-trading agent is not part of a blocking pair then he/she will remain unmatched (respectively, gets zero payoff) under every stable outcome .

Finally, meaningful comparative static results can be derived for the two-sided matching markets when agents in some market  $M$  are allocated according to a core outcome  $X$ , and new agents from a given side, say  $W$ , enter the market. Intuitively, before the entrants interacting with the other players they are unmatched and so the resulting outcome  $Y$  is  $W$ -simple for the new market  $M'$ . If  $Y$  is unstable, then it can be extended to a core outcome  $Z$  by a sequence of adjustments in which the trades can be done without hurting any  $F$ -agent and without benefit any  $W$ -agent. Then, every  $F$ -agent weakly prefers  $Z$  to  $X$  and every  $W$ -agent weakly prefers the opposite. Unlike the comparative static results that have been obtained for matching markets, the outcomes in comparison need not be any of the optimal stable outcomes for each side of the market.

This paper is organized as follows. Section 2 is devoted to the one-to-one matching models. Subsection 2.1 gives the conceptual framework and present some preliminary results. Subsection 2.2 presents some characteristic properties of the simple outcomes in connection with stable outcomes. Subsection 2.3 presents the existence results and some new comparative static result for the two-sided matching markets. Section 3 concludes the paper.

## **2. ONE-TO-ONE MATCHING MARKETS.**

In sub-section 2.1, the one-to-one matching models will be described together, aiming to make a parallel between them. The continuous one-sided and the continuous two-sided matching models with linear utilities will be called one-sided and two-sided Assignment games, respectively. The former is a model not yet explored in the literature and the second one is well known and due to Shapley and Shubik (1972). For the discrete case, we will not assume strict preferences. When preferences are strict, the one-sided

matching market is the Roommate model and the two-sided matching market is the well known Marriage market. Both models were introduced in Gale and Shapley (1962).

As discussed in section 1, the concepts for the one-sided markets can be obtained from the corresponding concepts for the two-sided markets by making  $F=W=N$ . The similarity between the results is due to the fact that they all apply when core outcomes and simple outcomes are required to be strong core and strongly simple outcomes, respectively. This requirement is always satisfied by the discrete models when preferences are strict. It is also satisfied by the continuous models, because preferences are continuous and payoffs can be adjusted continuously.

## 2.1 GENERAL FRAMEWORK

There is a finite set of players,  $N=\{1,2,\dots,n\}$ . In the two-sided matching model there are two disjoint sets of players,  $F$ , with  $m$  elements, and  $W$ , with  $p$  elements, such that  $F\cup W=N$ . Each player in  $F$  ( $F$ -agent or  $F$ -player) is interested in forming at most one partnership with players in  $W$  ( $W$ -agent or  $W$ -player) and vice-versa. For the continuous two-sided matching model, associated to each partnership  $(i,j)$ ,  $i\in F$  and  $j\in W$ , there is a nonnegative real number  $a_{ij}$ . A game in coalitional function form with side payments is determined by  $(N,a)$ , with the numbers  $a_{ij}$  being equal to the worth of the coalitions  $\{i,j\}$ . The worth of large coalitions is determined entirely by the worth of the pairwise combinations that the coalition members can form. That is, the coalitional function  $v$  is given by

$$v(S)=0 \text{ if } |S|=1 \text{ or } S\cap F=\emptyset \text{ or } S\cap W=\emptyset,$$

$$v(S)=a_{ij} \text{ if } S=\{i,j\}, \text{ with } i\in F \text{ and } j\in W;$$

$$v(S)=\max\{v(i_1,j_1)+v(i_2,j_2)+\dots+v(i_k,j_k)\}^5 \text{ for arbitrary coalitions } S, \text{ where the maximum is taken over all sets } (i_1,j_1),\dots,(i_k,j_k), i_t\in F\cap S \text{ and } j_t\in W\cap S \text{ for } t=1,\dots,k, \text{ of any } k \text{ distinct pairs in } S.$$

Thus, the rules of the game are that any pair of agents  $(i,j)$ , with  $i\in F$  and  $j\in W$ , can together obtain  $a_{ij}$ , and any larger coalition is valuable only insofar as it can organize itself into such pairs. The members of any coalition may divide among

---

<sup>5</sup> We write  $v(i,j)$  rather than  $v(\{i,j\})$ .

themselves their collective worth in any way they like. We will consider that any player  $h$  can be self-matched and will define  $a_{hh}=0$  for all  $h \in N$ .

This model can then be denoted by  $(F, W, a)$ . It is the two-sided Assignment game. If we make  $F=W=N$  in the description of the two-sided Assignment game above, we obtain a new model that will be called here *one-sided Assignment game*. It will be denoted by  $(N, a)$ . The players of this model will some times be called *N-players* or *N-agents*. Thus, every concept for the one-sided Assignment game is obtained from the same concept for the two-sided Assignment game by making  $F=W=N$ .

The two-sided Assignment game can also be thought of as being a sub-model of the one-sided Assignment game, by allowing every pair of agents from the same side to contribute with a zero payoff.

For the discrete models, each  $F$ -player has complete and transitive preferences over  $W$ - players and vice-versa. These preferences may be such that a player from one of the sides would prefer to remain unpaired rather than to enter a partnership with any other player from the other side. Hence, the preferences of each  $i \in F$  (respectively  $j \in W$ ) can be represented by an ordered list of preferences,  $P(i)$  (respectively  $P(j)$ ), on the set  $W \cup \{i\}$  (respectively  $F \cup \{j\}$ ). Player  $i$  is *acceptable* to player  $j$  if  $j$  does not prefer himself/herself to  $i$ . Player  $j$  is always acceptable to  $j$ . Thus,  $P(j)$  might be of the form  $P(j)=i, [m, l], j, \dots, q$  indicating that  $j$  prefers  $i$  to  $m$ , is indifferent between  $m$  and  $l$ , prefers any of these two last players to himself/herself and anyone else is *unacceptable* to himself/herself.

The model can then be denoted by  $(F, W, P)$ , where  $P$  is the set of all preference lists, one for each player.

If we make  $F=W=N$  we obtain a one-sided matching model, denoted here by  $(N, P)$ . The players of this model will some times be called *N-players* or *N-agents*. Thus, in the rest of this subsection, **every concept for  $(N, P)$  model is the restriction of the same concept for  $(F, W, P)$  model when  $F=W=N$** . When preferences are strict,  $(F, W, P)$  and  $(N, P)$  models are the Marriage model and Roommate model, respectively. The Marriage model is considered a sub-model of the Roommate model where every

agent lists the agents from the same side as unacceptable<sup>6</sup>. We write  $j >_i j'$  to mean  $i$  prefers  $j$  to  $j'$  and  $j \geq_i j'$  to mean  $i$  likes  $j$  at least as well as  $j'$ . A matching is a set of partnerships. It is feasible if every agent form one partnership at most. Formally,

**Definition 1.** For  $(F,W,a)$  and  $(F,W,P)$  a **feasible two-sided matching**  $x$  is a one-to-one correspondence from  $N=F \cup W$  onto itself of order two (that is,  $x^2(h)=h$ ). In addition,  $x(i) \in W \cup \{i\}$  for every  $i \in F$  and  $x(j) \in F \cup \{j\}$  for every  $j \in W$ . We refer to  $x(h)$  as the **partner of  $h$  at  $x$** , for every  $h \in N$ . If  $x(h)=h$  we say that  $h$  is **unmatched** at  $x$ .

When it does not cause any confusion, the term matching will be used without any reference to the number of sides of the market.

In the discrete models, player  $j$  prefers the feasible matching  $x$  to the feasible matching  $y$  if and only if he/she prefers  $x(j)$  to  $y(j)$ . He/she is indifferent between these two outcomes if and only if he/she is indifferent between  $x(j)$  and  $y(j)$ . Therefore, we are assuming that player  $j$  cares about who he/she is matched with, but is not otherwise concerned with the partners of other players. If all players in a set  $A$  prefers the feasible matching  $x$  to the feasible matching  $y$  we write  $x >_A y$ . Similarly we define  $x \geq_A y$ . The matching  $x$  is **individually rational** if each player is acceptable to his or her partner. A payoff vector for  $(F,W,a)$  (respectively,  $(N,a)$ ) is a vector  $u=(t,v) \in \mathbb{R}^m \times \mathbb{R}^p$  (respectively,  $u \in \mathbb{R}^n$ ). The notation  $u$  stands for a payoff vector in  $(F,W,a)$  or  $(N,a)$ . The payoff vector  $u=(t,v) \in \mathbb{R}^m \times \mathbb{R}^p$  (or  $u \in \mathbb{R}^n$ ) is **individually rational** if  $u_i \geq 0$ , for all  $i \in N$ . For our purposes it will be simpler to define:

**Definition 2.** The payoff vector  $u$  is **pairwise-feasible** (feasible, for short) for market  $(F,W,a)$  if there is a two-sided matching  $x$  such that if  $i \in F$  and  $j \in W$  then (a)  $u_i + u_j = a_{ij}$  if  $x(i)=j$  and (b)  $u_i = 0$  if  $x(i)=i$ , and  $u_j = 0$  if  $x(j)=j$ . We say that  $(u,x)$  is a **pairwise-feasible outcome** (feasible outcome, for short) and  $x$  is compatible with  $u$ .

---

<sup>6</sup> A way to have a more concise theory is to describe the preferences of the players in the discrete models by separable cardinal utility functions. (See Sotomayor (1999))

**Remark 1.** Given a coalition  $S$ , the definition of  $v$  implies that there is some one-sided (respectively, two-sided) matching  $x$  such that  $x(S)=S$  and  $\sum_{i \in S, i \leq x(i)} a_{i,x(i)} = v(S)$  (respectively,  $\sum_{i \in S \cap F} a_{i,x(i)} = v(S)$ ). Furthermore,  $v(S) \geq \sum_{i \in S, i \leq x'(i)} a_{i,x'(i)}$  (respectively,  $v(S) \geq \sum_{i \in S \cap F} a_{i,x'(i)}$ ) for all feasible one-sided (respectively, two-sided) matching  $x'$  such that  $x'(S)=S$ . Then, it follows from Definition 2 that  $\sum_{i \in S} u_i \leq v(S)$  for all  $S \subseteq N$  and for all feasible outcome  $(u, x)$  with  $x(S)=S$ . In particular,  $\sum_{i \in N} u_i \leq v(N)$  for all feasible outcome  $(u, x)$ .

The equilibrium concept is given by that of *setwise-stability* (stability, for short) presented in section 1. Since every player is allowed to form one partnership at most, in order to check instabilities we only need to consider pairs of agents in  $F \times W$ , for the two-sided markets, and in  $N \times N$ , for the one-sided markets. Then the definition of stability is the following:

**Definition 3.** a) For  $(F, W, a)$ , we say that  $i \in F$  and  $j \in W$  **block** the payoff vector  $u$  if  $u_i + u_j < a_{ij}$ . The payoff vector  $u$  is **stable** if it is feasible, individually rational and is not blocked by any pair of agents. If  $x$  is compatible with  $u$  we say that  $(u, x)$  is a **stable outcome**.

b) For  $(F, W, P)$ , we say that  $i \in F$  and  $j \in W$  **block** the two-sided matching  $x$  if  $j \succ_i x(j)$  and  $i \succ_j x(i)$ . A matching  $x$  is **stable** if it is feasible, individually rational and is not blocked by any pair of agents.

If an outcome is not stable we say that it is **unstable**. The individual rationality requirement means that a player always has the option of remaining unmatched. If agents  $i$  and  $j$  form a blocking pair, then it would pay them to break up their present partnership(s) and form a new one together, because this could give them each a higher payoff.

**Definition 4.** For  $(F, W, P)$ , we say that  $i \in F$  and  $j \in W$  **weakly block** the two-sided matching  $x$  if  $j \succeq_i x(i)$  and  $i \succeq_j x(j)$ , with strictly preference for at least one of the players. A matching  $x$  is **strictly stable** if it is not weakly blocked by any pair.

The set of strictly stable payoffs coincides with the strong core, which is the core defined by weak domination. When preferences are strict, the core and the strong core coincide. (See Roth and Sotomayor, 1990).

**Definition 5.** *The payoff vector  $u \in R^n$  is in the **core** of  $(F, W, a)$  if (a)  $\sum_{i \in N} u_i = v(N)$  and (b)  $\sum_{i \in S} u_i \geq v(S)$  for all  $S \subseteq N$ .*

Note that the definition of stable payoff requires the existence of a compatible matching. This is not required for the core definition. However, these two concepts are equivalent for the models we are treating here. In fact,

**Proposition 1.** *a) The set of stable payoffs equals the core of  $(N, a)$  (respectively,  $(F, W, a)$ ). b) The set of stable payoffs equals the core of  $(N, P)$  (respectively,  $(F, W, P)$ ).*

**Proof.** We will prove a). The proof of part b) for  $(F, W, P)$  is given by Theorem 3.3 in Roth and Sotomayor (1990). The proof for  $(N, P)$  can then be obtained by making  $F=W=N$ . Suppose  $u$  is a stable payoff for  $(N, a)$  (respectively,  $(F, W, a)$ ). Then,  $u$  is pairwise-feasible, so

$$\sum_{i \in N} u_i \leq v(N), \quad (1)$$

by Remark 2.1. Given a coalition  $S$ , it follows from Remark 1 that there is some one-sided (respectively, two-sided) matching  $y$  such that  $y(S)=S$  and  $v(S)=\sum_{i \in S, i \leq y(i)} a_{i, y(i)}$  (respectively,  $v(S)=\sum_{i \in S \cap F} a_{i, y(i)}$ ). The stability of  $u$  implies that  $u_i + u_{y(i)} \geq a_{i, y(i)}$  and  $u_i \geq 0$  for all  $i \in S$ , so

$$\sum_{i \in S} u_i \geq v(S) \text{ for all coalition } S. \quad (2)$$

By (1) and (2) it follows that  $\sum_{i \in N} u_i = v(N)$ . Then, (a) and (b) of Definition 5 are satisfied, so  $u$  is in the core of  $(N, a)$  (respectively,  $(F, W, a)$ ).

Now, suppose  $u$  is in the core of  $(N, a)$  (respectively,  $(F, W, a)$ ). Definition 5–(b) implies that  $u_i + u_j \geq v(i, j) = a_{ij}$  for every coalition  $\{i, j\}$  (respectively,  $(i, j) \in FxW$ ) and  $u_i \geq v(i) = 0$  for all  $i \in N$ , so  $u$  does not have any blocking pair and is individually rational. To see that  $u$  is pairwise-feasible **let  $x$  be a one-sided (respectively, two-sided) matching such that  $v(N) = \sum_{i \leq x(i)} a_{ix(i)}$  (respectively,  $v(N) = \sum_{i \in F} a_{i, y(i)}$ ).** Now use

that  $\sum_{i \in N} u_i = v(N)$ ,  $a_{ii} = 0$  and  $u_i + u_{x(i)} \geq a_{i,x(i)}$  for all  $i \in N$ , to get that  $\sum_{i \in N} u_i = \sum_{i \leq x(i)} a_{i,x(i)} = \sum_{i < x(i)} a_{i,x(i)} + \sum_{i=x(i)} a_{i,x(i)} = \sum_{i < x(i)} a_{i,x(i)} \leq \sum_{i < x(i)} (u_i + u_{x(i)}) = \sum_{x(i) \neq i} u_i \leq \sum_{x(i) \neq i} u_i + \sum_{x(i)=i} u_i = \sum_{i \in N} u_i$  (respectively,  $\sum_{i \in N} u_i = \sum_{i \in F} a_{i,x(i)} = \sum_{i \in F, x(i) \neq i} a_{i,x(i)} \leq \sum_{i \in F, x(i) \neq i} (u_i + u_{x(i)}) = \sum_{i \in F, x(i)=j, j \neq i} (u_i + u_j) \leq \sum_{i \in F, x(i)=j, j \neq i} (u_i + u_j) + \sum_{i \in F, x(i)=i} u_i + \sum_{j \in W, x(j)=j} u_j = \sum_{i \in N} u_i$ ). Then, the inequalities cannot be strict. This and the fact that  $u_i + u_{x(i)} \geq a_{i,x(i)}$  for all  $i \in N$  imply that  $u_i + u_{x(i)} = a_{i,x(i)}$  for all  $i \in N$ . Since  $u_i \geq v(i) = 0$  for all  $i \in N$ ,  $u_i = 0$  for all  $i \in N$  such that  $x(i) = i$ . Then,  $x$  is compatible with  $u$ , so  $u$  is pairwise-feasible and hence  $u$  is stable for  $(N, a)$  (respectively,  $(F, W, a)$ ). ■

**Definition 6.** *The two-sided matching  $x$  is optimal if  $\sum_{i \in F} a_{i,x(i)} = v(N)$  (respectively,  $\sum_{i \leq x(i)} a_{i,x(i)} = v(N)$ ).*

Observe that the second part of the proof of Proposition 1 shows that, if  $u$  is in the core of  $(N, a)$  (respectively,  $(F, W, a)$ ) and  $x$  is an optimal one-sided (respectively, two-sided) matching then  $x$  is compatible with  $u$ . Now, if  $u$  is in the core of  $(N, a)$  (respectively,  $(F, W, a)$ ) and is compatible with  $x$  then we must have  $v(N) = \sum_{i \in N} u_i = \sum_{i \leq x(i)} (u_i + u_{x(i)}) = \sum_{i \leq x(i)} a_{i,x(i)}$  (respectively,  $v(N) = \sum_{i \in N} u_i = \sum_{i \in F} (u_i + u_{x(i)}) = \sum_{i \in F} a_{i,x(i)}$ ), so  $x$  is an optimal matching. We have proved that,

**Proposition 2.** *(a) For  $(N, a)$  (respectively,  $(F, W, a)$ ), if  $x$  is an optimal matching, then it is compatible with any stable payoff. (b) If  $(u, x)$  is a stable outcome for  $(N, a)$  (respectively,  $(F, W, a)$ ), then  $x$  is an optimal matching.*

This proposition makes clear why in both continuous models, unlike their discrete version, we can concentrate on the payoffs to the agents rather than on the underlying matching. The key notions in this paper are those of *A-simple outcome* and *Pareto optimal A-simple outcome* given in section 1. To fix ideas we will choose  $A = W$ . Due to the importance of these concepts we present them again.

**Definition 7.** (a) The outcome  $(u,x)$  for  $(F,W,a)$  is **W-simple** if it is feasible, individually rational and no matched  $W$ -agent is part of a blocking pair.

(b) The matching  $x$  for  $(F,W,P)$  is **W-simple** if it is feasible, individually rational and no matched  $W$ -agent is part of a blocking pair. The matching  $x$  is **strongly W-simple** if it is  $W$ -simple and no matched  $W$ -agent is part of a weak blocking pair.

Then, in an  $N$ -simple (respectively, strongly  $N$ -simple) outcome only unmatched agents can block (respectively, weakly block). When there is no confusion, we will refer to  $W$ -simple and  $N$ -simple outcomes as simple outcomes, and to strongly  $W$ -simple and strongly  $N$ -simple outcomes as strongly simple outcomes.

When preferences are strict, every strongly simple matching is simple and vice versa. Since the outcome in which no partnership is formed is simple and strongly simple, **the set of simple outcomes and strongly simple outcomes are non-empty for all models we are treating here.** Clearly, every stable outcome is simple and every strongly stable outcome is strongly simple.

**Definition 8.** (a) In the market  $(F,W,a)$ , we say that  $(t^*,v^*;z)$  **F-extends**  $(t,v;x)$  if either (i)  $t^*_i \geq t_i$  for all  $i \in F$  with strict inequality for at least one player in  $F$  or (ii)  $t=t^*$  and  $(t^*,v^*;z)$  has less blocking pairs than  $(t,v;x)$ . If  $(t,v;x)$  and  $(t^*,v^*;z)$  are  $W$ -simple outcomes we say that  $(t,v;x)$  has a  $W$ -simple extension.

(b) In the market  $(F,W,P)$ , we say that matching  $z$  **F-extends** matching  $x$  if  $z(i) \geq x(i)$  for all  $i \in F$ , with strict inequality for at least one player in  $F$ . If  $x$  and  $z$  are  $W$ -simple matchings (respectively, strongly  $W$ -simple matchings) we say that  $x$  **has a  $W$ -simple (respectively, strongly  $W$ -simple) extension.**

Note that in the  $N$ -simple extensions for the continuous models condition (ii) is never satisfied.

**Definition 9.** (a) For the market  $(F,W,a)$ , the payoff vector  $u$  is **Pareto optimal W-simple** if it is  $W$ -simple and it does not have any  $W$ -simple extension. If  $x$  is compatible with  $u$  we say that  $(u,x)$  is a Pareto optimal  $W$ -simple outcome.

(b) For the market  $(F, W, P)$ , the two-sided matching  $x$  is called **Pareto optimal  $W$ -simple (respectively, strongly  $W$ -simple) matching** if it is  $W$ -simple (respectively, strongly  $W$ -simple) and it does not have any  $W$ -simple (respectively, strongly  $W$ -simple) extension.

Then, a Pareto optimal  $N$ -simple outcome is an  $N$ -simple outcome that is Pareto optimal among all  $N$ -simple outcomes. The existence of Pareto optimal simple payoffs for both continuous models is guaranteed by the fact proved below that **the set of  $W$ -simple payoffs and the set of  $N$ -simple payoffs are non-empty and compact sets of  $\mathbf{R}^n$** . Then, there is some  $W$ -simple payoff  $(t^*, v^*)$  (respectively,  $N$ -simple payoff  $u^*$ ) such that  $\sum_{i \in F} t_i^* \geq \sum_{i \in F} t_i$  (respectively,  $\sum_{j \in N} u_j^* \geq \sum_{j \in N} u_j$ ) for all  $W$ -simple payoffs  $(t, v)$  (respectively, for all  $N$ -simple payoffs  $u$ ). Choose  $(t^*, v^*)$  such that it has the minimum number of blocking pairs among all such outcomes. Clearly,  $(t^*, v^*)$  (respectively,  $u^*$ ) is a Pareto optimal  $W$ -simple payoff (respectively, Pareto optimal  $N$ -simple payoff).

**Proposition 3.** *The set of  $W$ -simple payoffs and the set of  $N$ -simple payoffs are compact sets of  $\mathbf{R}^n$ .*

**Proof.** The set of  $N$ -simple (respectively,  $W$ -simple) payoffs is bounded, since  $0 \leq u_j \leq v(N)$ , for all  $j \in N$  and all simple payoff  $u$ . To see that it is closed, take any sequence  $(u^t)_t$  of  $N$ -simple (respectively,  $W$ -simple) payoffs, with  $u^t \rightarrow u$ , when  $t$  tends to infinity. Since the set of matchings is finite, there is some matching  $x$ , which is compatible with infinitely many terms of the sequence  $u^t$ . We will use the same notation  $(u^t)$  for this subsequence. Then, if  $x(h) = k$ ,  $u_h + u_k = \lim_{t \rightarrow \infty} (u_h^t + u_k^t) = \lim_{t \rightarrow \infty} a_{hk} = a_{hk}$ ; if  $x(h) = h$  then  $u_h = \lim_{t \rightarrow \infty} u_h^t = 0$ . Thus,  $x$  is compatible with  $u$ , so  $(u, x)$  is pairwise-feasible. Now, observe that if  $h \in N$  (respectively,  $h \in W$ ) is matched at  $x$  then  $h$  is not part of a blocking pair of  $(u^t, x)$ . Then,  $u_h + u_k = \lim_{t \rightarrow \infty} (u_h^t + u_k^t) \geq \lim_{t \rightarrow \infty} a_{hk} = a_{hk}$ , so  $h$  is not part of a blocking pair of  $(u, x)$ . Therefore,  $(u, x)$  is  $N$ -simple (respectively,  $W$ -simple), so  $u$  is an  $N$ -simple (respectively,  $W$ -simple) payoff. Hence, the set of  $N$ -simple (respectively,  $W$ -simple) payoffs is bounded and closed, so it is compact. ■

The existence of Pareto optimal simple matchings for the Marriage market and for the Roommate market is guaranteed by the fact that the set of strongly  $W$ -simple

matchings and the set of strongly  $N$ -simple matchings are non-empty and finite and preferences are transitive.

## 2.2 PROPERTIES OF THE SIMPLE OUTCOMES

We are assuming that players preferences in the discrete markets are non-necessarily strict. Under the assumption of strict preferences, all the results for the discrete markets hold if we replace strongly simple and strongly stable by simple and stable, respectively. The notation  $\sigma=(u,x)$  stands for an outcome in  $(N,a)$  or  $(F,W,a)$ . This section presents some characteristic properties of simple outcomes. We will prove the properties of the  $W$ -simple outcomes and strongly  $W$ -simple outcomes. The proofs for the  $N$ -simple outcomes and strongly  $N$ -simple outcomes (even for the two-sided matching markets) follow the same arguments and are obtained by just making  $F=W=N$ . The following proposition is a key lemma that enables us to derive all of our results.

**Lemma 1.** (a) For the market  $(F,W,a)$ , let  $\sigma=(t,v,x)$  be a  $W$ -simple outcome and let  $\tau=(t',v',y)$  be a stable outcome. Let  $T=\{j \in W; x(j) \neq j\}$ ,  $M_\sigma(F)=\{i \in F; t_i > t'_i\}$  and  $M_\tau(W)=\{j \in T; v'_j > v_j\}$ . Then  $x(M_\sigma(F))=y(M_\sigma(F))=M_\tau(W)$ .

(b) For the markets  $(N,a)$  and  $(F,W,a)$ , let  $\sigma=(u,x)$  be an  $N$ -simple outcome and let  $\tau=(u',y)$  be a stable outcome. Let  $T=\{j \in N; x(j) \neq j\}$ ,  $M_\sigma(N)=\{j \in N; u_j > u'_j\}$  and  $M_\tau(N)=\{j \in T; u'_j > u_j\}$ . Then  $x(M_\sigma(N))=y(M_\sigma(N))=M_\tau(N)$ .

(c) For the market  $(F,W,P)$ , let  $x$  be a strongly  $W$ -simple matching and let  $y$  be a strongly stable matching. Let  $T=\{j \in W; x(j) \neq j\}$ . Let  $M_x(F)=\{i \in F; x(i) >_i y(i)\}$ ,  $M_y(W)=\{j \in T; y(j) >_j x(j)\}$ . Then  $x(M_x(F))=y(M_x(F))=M_y(W)$ .

(d) For the markets  $(N,P)$  and  $(F,W,P)$ , let  $x$  be a strongly  $N$ -simple matching and let  $y$  be a strongly stable matching. Let  $T=\{j \in N; x(j) \neq j\}$ . Let  $M_x(N)=\{j \in N; x(j) >_j y(j)\}$ ,  $M_y(N)=\{j \in T; y(j) >_j x(j)\}$ . Then  $x(M_x(N))=y(M_x(N))=M_y(N)$ .

**Proof.** For part (a), all  $i$  in  $M_\sigma(F)$  are matched under  $x$ , since  $t_i > t'_i \geq 0$ . Analogously, all  $j$  in  $M_\tau(W)$  are matched under  $y$ , since  $v'_j > v_j \geq 0$ . If  $i$  is in  $M_\sigma(F)$  then  $j=x(i)$  is in  $M_\tau(W)$ , for if not

$$a_{ij} = t_i + v_j > t'_i + v'_j,$$

which implies that  $(i,j)$  blocks  $(t',v',y)$ , contradiction. On the other hand, if  $j$  is in  $M_\tau(W)$  then  $i=y(j)$  is in  $M_\sigma(F)$ , for if not

$$a_{ij} = t'_i + v'_j > t_i + v_j,$$

which implies that  $(i,j)$  blocks  $(t,v,x)$ . However,  $j$  is in  $T$ , so  $j$  is matched at  $x$ , which contradicts the fact that  $(t,v,x)$  is simple. Therefore,  $x(M_\sigma(F)) \subseteq M_\tau(W)$  and  $y(M_\tau(W)) \subseteq M_\sigma(F)$ . Since  $x$  and  $y$  are one-to-one and  $M_\sigma(F)$  and  $M_\tau(W)$  are finite, the conclusion follows. Hence the proof of part (a) is complete.

For part (c), all  $i$  in  $M_x(F)$  are matched under  $x$ , since  $x(i) >_i y(i) \geq_i i$ . Analogously, all  $j$  in  $M_y(W)$  are matched under  $y$ , since  $y(j) >_j x(j) \geq_j j$ . If  $i$  is in  $M_x(F)$  then  $j=x(i)$  is in  $M_y(W)$ , for if not  $i=x(j) \geq_j y(j)$ , so  $y$  will be weakly blocked by  $i$  and  $j$ , which contradicts the assumption that  $y$  is strongly stable. On the other hand, if  $j$  is in  $M_y(W)$  then  $i=y(j)$  is in  $M_x(F)$ , for if not  $j=y(i) \geq_i x(i)$ , so  $x$  will be weakly blocked by  $i$  and  $j$ . However,  $j$  is in  $T$ , so  $j$  is matched at  $x$ , which contradicts the fact that  $x$  is simple. Therefore,  $x(M_x(F)) \subseteq M_y(W)$  and  $y(M_y(W)) \subseteq M_x(F)$ . Since  $x$  and  $y$  are one-to-one and  $M_x(F)$  and  $M_y(W)$  are finite, the conclusion follows. Hence the proof of part (c) is complete. ■

The following immediate consequence implies that, if both outcomes are stable for  $(F,W,a)$  or  $(F,W,P)$  (respectively,  $(N,a)$  or  $(N,P)$ ) then  $x$  and  $y$  also map the set of  $F$ -agents (respectively,  $N$ -agents) who prefer the first outcome to the second outcome onto the set of  $W$ -agents (respectively,  $N$ -agents) who prefer the opposite. That is,

**Corollary 1.** (a) ) Let  $\sigma=(t,v,x)$  and  $\tau=(t',v',y)$  be stable outcomes for the market  $(F,W,a)$ . Let  $M_\sigma(F)=\{i \in F; t_i > t'_i\}$  and  $M_\tau(W)=\{j \in W; v'_j > v_j\}$ . Analogously define  $M_\sigma(W)$  and  $M_\tau(F)$ . Then  $x(M_\sigma(F))=y(M_\sigma(F))=M_\tau(W)$  and  $x(M_\tau(F))=y(M_\tau(F))=M_\sigma(W)$ .<sup>7</sup>

(b) Let  $\sigma=(u,x)$  and  $\tau=(u',y)$  be stable outcomes for the market  $(N,a)$ . Let  $M_\sigma(N)=\{j \in N; u_j > u'_j\}$  and  $M_\tau(N)=\{j \in N; u'_j > u_j\}$ . Then  $x(M_\sigma(N))=y(M_\sigma(N))=M_\tau(N)$ .

<sup>7</sup> This property is due to Demange and Gale (1985). It can also be seen in Roth and Sotomayor (1990).

(c) Let  $x$  and  $y$  be strongly stable matchings for  $(F, W, P)$ . Let  $M_x(F) = \{i \in F; x(i) \succ y(i)\}$  and  $M_y(W) = \{j \in W; y(j) \succ_j x(j)\}$ . Analogously define  $M_x(W)$  and  $M_y(F)$ . Then  $x(M_x(F)) = y(M_x(F)) = M_y(W)$ .<sup>8</sup>

(d) Let  $x$  and  $y$  be strongly stable matchings for  $(N, P)$ . Let  $M_x(N) = \{j \in N; x(j) \succ_j y(j)\}$  and  $M_y(N) = \{j \in N; y(j) \succ_j x(j)\}$ . Then  $x(M_x(F)) = y(M_x(F)) = M_y(W)$ .

An important property that is characteristic of both two-sided Assignment game and Marriage model with strict preferences, states that there is an opposition of interests between the two sides of the market along the whole core<sup>9</sup>:

*If  $\sigma$  and  $\tau$  are stable outcomes for  $(F, W, a)$  (respectively,  $(F, W, P)$ ), then  $\sigma_i \geq \tau_i$ , (respectively,  $\sigma \geq_i \tau$ ) for every  $i \in F$  if and only if  $\tau_j \geq \sigma_j$  (respectively,  $\tau \geq_j \sigma$ ) for every  $j \in W$ .*

This property is a direct consequence of Corollary 1. It is shared with many matching models and it is crucial for the proof of the lattice property of the core of  $(F, W, a)$  and  $(F, W, P)$ . Due to the one-sidedness of  $(N, a)$  and  $(N, P)$ , this kind of opposition of interests does not hold for these markets. Nevertheless, an immediate consequence of Lemma 1 reflects an opposition of interests between the players involved in a partnership regarding two stable outcomes (respectively, strongly stable matchings) for the one-sided markets. It also implies that if some player gets different payoffs (respectively, mates) under the two stable outcomes (respectively, strongly stable outcomes) then he/she is matched at both outcomes and so are his/her mates. Formally,

**Property 1.** (a) Let  $\sigma = (u, x)$  and  $\tau = (u', y)$  be stable outcomes for  $(N, a)$  or  $(F, W, a)$ . If  $u_j > u'_j$  then  $j \neq x(j) = k$ , for some  $k$ , and  $j \neq y(j) = h$ , for some  $h$ . Furthermore,  $u'_k > u_k$  and  $u'_h > u_h$ .

(b) Let  $x$  and  $y$  be strongly stable matchings for the  $(N, P)$  or  $(F, W, P)$ , with non-necessarily strict preferences. If  $j$  prefers  $x$  to  $y$  then  $j \neq x(j) = k$ , for some  $k$ , and  $j \neq y(j) = h$ , for some  $h \neq k$ . Furthermore, both  $k$  and  $h$  prefer  $y$  to  $x$ .

<sup>8</sup> The case of strict preferences was proved by Knuth (1976) and it is also a corollary of the Decomposition Lemma of Gale and Sotomayor (1985a,b).

<sup>9</sup> This property was proved by Shapley and Shubik (1972) for the two-sided Assignment game and by Knuth (1976) for the Marriage market.

**Proof.** (a) Since  $j$  is in  $M_\sigma(N)$ , Lemma 1-b implies that  $j$  is matched under  $x$  and  $y$  to an agent in  $M_\tau(N)$ . Analogous argument applies to (b). ■

Lemma 1 also concurs to the following result, which concerns some set of players who are indifferent between stable outcomes and simple outcomes:

**Property 2.** (a) *If an  $F$ -player (respectively,  $W$ -player) is unmatched under a stable outcome for the two-sided Assignment game, then he/she gets a zero payoff under any  $W$ -simple (respectively,  $F$ -simple) outcome.*

(b) *If a player is unmatched under a stable outcome for  $(N,a)$  or  $(F,W,a)$ , then he/she gets a zero payoff under any  $N$ -simple outcome.*

(c) *Let  $x$  and  $y$  be a strongly  $W$ -simple matching and a stable matching, respectively, for  $(F,W,P)$ . If  $i \in F$  is unmatched under  $y$  then  $i$  is indifferent between  $x(i)$  and being unmatched. If preferences are strict and  $i \in F$  is unmatched under  $y$  then  $i$  is unmatched under  $x$ .*

(d) *Let  $x$  and  $y$  be a strongly  $N$ -simple matching and a stable matching, respectively, for  $(N,P)$  or  $(F,W,P)$ . If  $h \in N$  is unmatched under  $y$  then  $h$  is indifferent between  $x(h)$  and being unmatched. If preferences are strict and  $i \in N$  is unmatched under  $y$  then  $h$  is unmatched under  $x$ .*

**Proof.** For part (a), suppose  $i \in F$  is unmatched under the stable outcome  $\tau=(t',v';y)$  but has a positive payoff under the  $W$ -simple outcome  $\sigma=(t,v;x)$ . Define  $M_\sigma(F)$  as in Lemma 1. Then  $i \in M_\sigma(F)$ , so Lemma 1 implies that  $i$  is matched under  $y$ , which is a contradiction. The other assertion follows dually.

(c) If  $x(i)=i$  then we are done. Otherwise, we cannot have  $x(i) >_i y(i)=i$ , because if so  $i \in M_x(F)$ . Then, by Lemma 1,  $y(i) \in M_y(W)$ , so there would be some  $j \in W$  such that  $y(i)=j$ , which gives the necessary contradiction. Hence,  $x(i) \approx_i y(i)=i$  if preferences are non-strict and  $x(i)=i$ , otherwise. ■

Since every stable outcome is  $N$ -simple and  $W$ -simple, we can obtain Property 3 below.

**Property 3.** (a) In the market  $(F,W,a)$  (respectively,  $(N,a)$ ) every unmatched player under a stable outcome gets zero payoff under any stable outcome<sup>10</sup>.

(b) Let  $x$  and  $y$  be strongly stable matchings for the market  $(F,W,P)$  (respectively,  $(N,P)$ ). If  $h \in N$  is unmatched under  $y$  then  $h$  is indifferent between  $x(h)$  and being unmatched.

(c) In the market  $(F,W,P)$  (respectively,  $(N,P)$ ) with strict preferences, the set of unmatched players under a stable matching is the same in all stable matchings<sup>11</sup>.

Properties 4 and 5 concerns the coalition structure of unstable simple outcomes. Property 4 is an immediate consequence of Lemma 1. It asserts that, in the continuous models, given an  $N$ -simple outcome and a stable outcome, **the trading agents at the  $N$ -simple outcome who are not indifferent between the two outcomes, trade among themselves at the stable outcome.** Similar result holds for the discrete models if we replace the  $N$ -simple and the stable outcomes by strongly  $N$ -simple and strongly stable matchings, respectively. If preferences are strict then **trading agents at a simple matching always make their transactions under a stable matching within the same pool.** Formally,

**Property 4.** Following the notations of Lemma 1, (a) let  $\sigma=(u,x)$  be an  $N$ - simple payoff and let  $\tau=(u',y)$  be in the core of  $(N,a)$  or  $(F,W,a)$ . Let  $S= M_{\sigma}(N) \cup M_{\tau}(N)$ . Then, if  $j \in S$ , then  $y(j) \neq j$  and  $y(j) \in S$ ;

(b) let  $x$  be a strongly  $N$ - simple matching and let  $y$  be a strongly stable matching for  $(N,P)$  or  $(F,W,P)$ . Let  $S= M_x(N) \cup M_y(N)$ . If  $j \in S$  then  $y(j) \neq j$  and  $y(j) \in S$ .

(c) Let  $x$  be a simple matching and let  $y$  be a stable matching for  $(N,P)$  or  $(F,W,P)$ . Suppose preferences are strict. Let  $T=\{j \in N; x(j) \neq j\}$ . If  $j \in T$  then  $y(j) \neq j$  and  $y(j) \in T$ .

The following property is a sort of converse of Property 2:

<sup>10</sup> This property was originally proved in Demange and Gale (1985) for a two-sided matching market with continuous utilities.

<sup>11</sup> The result for the Marriage market is due to Gale and Sotomayor (1985a).

**Property 5** (a) Let  $\sigma$  be a  $W$ -simple (respectively,  $N$ -simple) outcome for  $(F, W, a)$  (respectively,  $(N, a)$ ). If  $j \in W$  (respectively,  $j \in N$ ) is unmatched under  $\sigma$  and it is not part of any blocking pair then  $j$  gets zero payoff under every stable payoff.

(b) Let  $x$  be a strongly  $W$ -simple (respectively,  $N$ -simple) matching for  $(F, W, P)$  (respectively,  $(N, P)$ ). If  $j \in W$  (respectively,  $j \in N$ ) is unmatched at  $x$  and it is not part of a blocking pair  $j$  is indifferent between being unmatched and any strongly stable matching.

(c) Let  $x$  be a  $W$ -simple (respectively,  $N$ -simple) matching for  $(F, W, P)$  (respectively,  $(N, P)$ ) with strict preferences. If  $j \in W$  (respectively,  $j \in N$ ) is unmatched at  $x$  and it is not part of a blocking pair then  $j$  is unmatched at every stable matching.

**Proof.** (a) Consider the model  $(F, W, a)$ . If  $\sigma = (t, v; x)$  is stable for  $(F, W, a)$ , the result follows from Property 2. Then suppose  $\sigma$  is unstable. Let  $\tau = (t', v'; y)$  be a stable outcome for  $(F, W, a)$ . If  $v'_j > 0$  then  $y(j) = i$ , for some  $i \neq j$ . Define  $M_\sigma(F)$  and  $M_\tau(W)$  as in Lemma 1. The fact that  $j$  is unmatched at  $x$  implies that  $j \notin M_\tau(W)$ . Since  $\{i, j\}$  cannot block  $\sigma$  by hypothesis, we cannot have  $t_i \leq t'_i$ , because then  $t_i + v_j < t'_i + v'_j = a_{ij}$ , contradiction. On the other hand, if  $t_i > t'_i$  then Lemma 1 would imply that  $i \in M_\sigma(F)$  and so  $j = y(i) \in M_\tau(W)$ , contradiction. Hence,  $v'_j = 0$ . Now, make  $F = W = N$  in the proof above to get the conclusion for model  $(N, a)$ .

(b) Consider the model  $(F, W, P)$ . If  $x$  is strongly stable the result follows from Proposition 4. Then suppose  $x$  is not strongly stable. Let  $y$  be a strongly stable matching. If  $y(j) = j$  then we are done. Otherwise, we cannot have  $y(j) >_j x(j) = j$ , because if so  $j \in M_y(W)$ . Then, by Lemma 1,  $x(j) \in M_x(F)$ , so there would be some  $k \in F$  such that  $x(j) = k$ , which gives the necessary contradiction. Now, make  $F = W = N$  in the proof above to get the conclusion for model  $(N, P)$ .

(c) It is immediate from (b) and the fact that every simple matching is strongly simple and every stable matching is strongly stable when preferences are strict. ■

### 2.3 EXISTENCE OF STABLE OUTCOMES

As discussed in section 1, the simple outcomes of the two-sided matching markets have a special property, stated in Theorem 1, which fails to hold for the  $N$ -simple outcomes of the one-sided matching markets. This property allows us to obtain an

existence proof of core outcomes for both two-sided matching models. To prove this property for the two-sided Assignment game we need some more terminology and the lemma below. Let  $(t, v; x)$  be a simple outcome for  $(F, W, a)$ . Let  $j \in W$  and  $x(j) = j$ . We say that  $i \in F$  is a  $j$ 's *favorite blocking partner* if  $a_{ij} - t_i > 0$  and  $a_{ij} - t_i \geq a_{hj} - t_h$  for all  $h \in F$ . That is, among all players  $i \in F$  such that  $(i, j)$  blocks  $(t, v; x)$ , player  $i$  gives to  $j$  the highest surplus. For every  $j \in W$  define:

$$D_j(t, v; x) \equiv \{i \in F; x(i) \neq j \text{ and } t_i + v_j = a_{ij}\} \cup \{j \text{'s favorite blocking partners}\}.$$

**Lemma 2.** *Let  $(t, v; x)$  be a Pareto optimal  $W$ -simple outcome. Construct a graph whose vertices are  $F \cup W$ . There are two kinds of arcs. If  $x(i) = j$  there is an arc from  $i$  to  $j$ ; if  $i \in D_j(t, v; x)$  there is an arc from  $j$  to  $i$ . Suppose there exists some unmatched  $W$ -agent  $j$  with  $D_j(t, v; x) \neq \emptyset$ . Let  $i_1 \in D_j(t, v; x)$ . Then, there is an oriented path which starts from  $j$ , has  $i_1$  as its second node and reaches an unmatched  $F$ -agent or a  $W$ -agent with payoff zero.*

**Proof.** Suppose there is no such path, and denote by  $S_F$  and  $S_W$  the sets of  $F$ -players and  $W$ -players, respectively, that can be reached from  $i_1$ . Then  $j \notin S_W$ , all of  $S_F$  are matched and  $v_k > 0$  for all  $k \in S_W$ , so all of  $S_W$  are also matched. Furthermore,  $S_F$  is non-empty since  $i_1 \in S_F$  and  $S_W$  is non-empty because all of  $S_F$  are matched to those of  $S_W$  at  $x$ . If  $i \notin S_F$ , then there is no  $k \in S_W$  such that  $x(i) = k$  or  $i \in D_k(t, v; x)$ . Then, for all  $k \in S_W$  and all  $i \notin S_F$ , we must have that  $x(i) \neq k$  and  $i \notin D_k(t, v; x)$ . Furthermore  $(i, k)$  is not a blocking pair, for all  $k \in S_W$  and all  $i \notin S_F$ , because  $k$  is matched and  $(t, v; x)$  is  $W$ -simple. Using the definition of  $D_k(t, v; x)$  it follows that  $v_k > a_{ik} - t_i$  for all  $k \in S_W$  and all  $i \notin S_F$ . Then, since  $v_k > 0$ , there is some  $\lambda > 0$ , sufficiently small, so that  $v_k - \lambda > 0$  and  $v_k - \lambda > a_{ik} - t_i$ , for all  $k \in S_W$  and all  $i \notin S_F$ . Hence, if we decrease  $v_k$  and increase  $t_i$  by  $\lambda > 0$ , for all  $(i, k) \in S_F \times S_W$ , the resulting outcome is still  $W$ -simple (it is enough to see that no pair  $(i, j) \in (F - S_F) \times S_W$  blocks the new outcome). However, all of  $S_F$  are better off, while the players in  $F - S_F$  are indifferent between both outcomes, which contradicts the Pareto optimality of  $(t, v; x)$ . ■

**Theorem 1.** (a) *Every unstable and  $W$ -simple outcome for the two-sided Assignment game has a  $W$ -simple extension.* (b) *Every unstable and  $W$ -simple matching for the Marriage model with non-necessarily strict preferences has a  $W$ -simple extension.*

**Proof.** To prove (a), let  $(t, v; x)$  be some unstable and  $W$ -simple outcome. Suppose by way of contradiction that  $(t, v; x)$  does not have any  $W$ -simple extension. Then,  $(t, v; x)$  is Pareto optimal  $W$ -simple outcome. Since  $(t, v; x)$  is unstable, there is a blocking pair  $(i_1, j)$ , where  $j$  is unmatched at  $x$ . Choose  $i_1$  such that  $i_1$  is some  $j$ 's favorite blocking partner. Then,  $i_1 \in D_j(t, v; x)$ . Lemma 2 guarantees the existence of an oriented path  $c = (j = j_1, i_1, j_2, \dots)$  which starts from  $j$ , has  $i_1$  as its second node, reaches an unmatched  $F$ -agent  $i_s$  or a  $W$ -agent  $w$  with payoff zero. The basic idea is to construct a  $W$ -simple extension of  $(t, v; x)$  by rematching the nodes of the path, which gives the desired contradiction. Then,  $c = (j = j_1, i_1, j_2, \dots, j_s, i_s)$  or  $c = (j = j_1, i_1, j_2, \dots, j_s, i_s, w)$ . Without loss of generality we can suppose that all elements in the path, until  $w$  are distinct. Construct the matching  $x'$  that matches  $j$  to  $i_1$ ,  $j_2$  to  $i_2, \dots, j_s$  to  $i_s$  and leaves  $w$  unmatched if  $w$  is in the path, and that otherwise agrees with  $x$  on every player who is not in the path. Now, give the payoff  $v'_j = a_{i_1 j} - t_{i_1} > 0$  to  $j$  and keep unchanged the payoffs of the rest of the players. We claim that the resulting outcome  $(t, v'; x')$  is simple. That it is feasible is immediate from the following facts: (a)  $j$  is matched to  $i_1$  and  $t_{i_1} + v'_j = a_{i_1 j}$ ; (b) for all  $t = 2, \dots, s$ ,  $j_t$  is matched to  $i_t$ ,  $i_t \in D_{j_t}(t, v; x)$ , so  $t_{i_t} + v'_{j_t} = t_{i_t} + v_{j_t} = a_{i_t j_t}$ . That there is no new blocking pairs follows from the fact that  $j$  is matched to a favorite blocking partner and the payoffs of the players other than  $j$  did not change. However, we have decreased by one the number of blocking pairs, so  $(t, v'; x')$  is a  $W$ -simple extension of  $(t, v; x)$ , contradiction.

To prove (b), let  $x$  be an unstable  $W$ -simple matching for  $(F, W, P)$ . Then, there is a blocking pair  $(i, j) \in F \times W$  where  $j$  is unmatched. Choose  $i$  so that  $i$  is the  $j$ 's favorite blocking partner. Now construct the matching  $y$  which matches  $i$  to  $j$ , leaves  $x(i)$  unmatched, in case  $x(i) \in W$ , and keep the same partners for the other players. Clearly,  $y$  is an extension of  $x$ . It is a matter of verification that  $y$  is simple, since  $j$  does not belong to any blocking pair of  $y$ . ■

It is intuitive that, if every unstable and  $W$ -simple outcome has a  $W$ -simple extension in the Marriage model, then the core can be reached from any unstable and  $W$ -simple matching in a finite number of  $W$ -simple extensions. However, it is not true that any sequence of unstable and  $W$ -simple outcomes converges to a core outcome in the continuous case. It turns out that the fact that the core of this model is non-empty follows from a subtle argument, which also applies to the discrete case: any Pareto optimal  $W$ -simple outcome is stable. Furthermore, any unstable and  $W$ -simple outcome has a core extension. See Theorem 2 and Proposition 4 below.

**Theorem 2.** *The core of the two-sided Assignment game, as well as the core of the Marriage model with non-necessarily strict preferences is non-empty.*

**Proof.** In fact, otherwise a Pareto optimal  $W$ -simple outcome would be unstable, so it could be extended to a  $W$ -simple outcome by Theorem 1, which is a contradiction. ■

**Proposition 4.** (a) *Every unstable and  $W$ -simple outcome for  $(F, W, a)$  can be extended to a stable outcome.* (b) *Every unstable and  $W$ -simple matching for  $(F, W, P)$  with non-necessarily strict preferences can be extended to a stable matching.*

**Proof.** To prove (a), let  $(t', v'; x')$  be an unstable and simple outcome for  $(F, W, a)$ . Set  $A \equiv \{t'' \in \mathbb{R}^m; (t'', v''; x'') \text{ is a } W\text{-simple outcome for some } v'' \in \mathbb{R}^p \text{ and some matching } x'' \text{ and } t''_i \geq t'_i \text{ for all } i \in F\}$ .

We have that  $A \neq \emptyset$ , since  $t' \in A$ . It is easy to see that  $A$  is a compact subset of  $\mathbb{R}^m$ , so there is some  $W$ -simple outcome  $(t^*, v^*; x^*)$  such that  $t^* \in A$  and

$$\sum_{i \in F} t^*_i \geq \sum_{i \in F} t''_i \text{ for all } t'' \in A. \quad (1)$$

Without loss of generality we can choose  $v^*$  and  $x^*$  so that  $(t^*, v^*; x^*)$  has the minimum number of blocking pairs among all  $W$ -simple outcomes that are compatible with  $t^*$ . We are going to show that  $(t^*, v^*; x^*)$  does not have any  $W$ -simple extension. If this is established, Theorem 2 then implies that  $(t^*, v^*; x^*)$  is in the core. Then, if there was some  $W$ -simple outcome  $(t'', v''; x'')$  that extended  $(t^*, v^*; x^*)$ , we should have that  $t''_i \geq t^*_i \geq t'_i$  for all  $i \in F$ , so  $t'' \in A$ , so  $\sum_{i \in F} t''_i \geq \sum_{i \in F} t^*_i$  and so  $t'' = t^*$  by (1). Then, by Definition 8  $(t'', v''; x'')$  has less blocking pairs than  $(t^*, v^*; x^*)$ , which contradicts the definition of  $(t^*, v^*; x^*)$ . Hence,  $(t^*, v^*; x^*)$  is in the core and  $t^*_i \geq t'_i$  for all  $i \in F$ . If

$t^*_i > t'_i$  for some  $i \in F$  we are done. Otherwise, observe that  $(t', v'; x')$  is not in the core, so it has more blocking pairs than  $(t^*, v^*; x^*)$ . Hence, in any case  $(t^*, v^*; x^*)$  extends  $(t', v'; x')$ .

To prove (b), let  $x'$  be an unstable and  $W$ -simple outcome for  $(F, W, P)$ . We are going to show that there is some stable matching  $x^*$  that extends  $x'$ . Theorem 1 implies that  $x'$  can be extended to some  $W$ -simple matching conveniently constructed. Call this matching  $x^1$ . Then,  $x^1(i) \geq_i x'(i)$  for every  $i \in F$ , with strict preference for at least one player in  $F$ , so  $x' \neq x^1$ . If  $x^1$  is stable make  $x^1 = x^*$  and we are done. Otherwise, use again Theorem 1 to obtain some other  $W$ -simple matching, say  $x^2$ , that extends  $x^1$ , so  $x^2(i) \geq_i x^1(i) \geq_i x'(i)$  for all  $i \in F$ . By construction of  $x^1$  and  $x^2$  we have that  $x^2 \neq x^1$ ,  $x^2 \neq x'$ . If  $x^2$  is stable make  $x^2 = x^*$  and we are done, and so on. Since the number of matchings is finite, this sequence of distinct unstable and  $W$ -simple matchings is finite, so it will reach some  $W$ -simple matching  $x^s$  such that  $x^s(i) \geq_i x^{s-1}(i) \geq_i \dots \geq_i x^1(i) \geq_i x'(i)$  for every  $i \in F$ , and  $x^s$  cannot be extended to another  $W$ -simple matching. Then  $x^s$  is stable, so make  $x^s = x^*$  and we have the desired outcome. ■

Example 1 illustrates that unlike the core, the strong core of the Marriage market may be empty.

**Example 1. (The strong core of the Marriage market may be empty)** The set of agents are  $F = \{f_1, f_2\}$  and  $W = \{w_1, w_2\}$ . Agent  $f_1$  is indifferent between  $w_1$  and  $w_2$ ; both  $W$ -agents prefer  $f_1$  to  $f_2$  and are acceptable to  $f_2$ . Both matchings under which no agent is unmatched are stable and are the only stable matchings. However, both of them are weakly blocked by  $(f_1, w_1)$  or by  $(f_1, w_2)$ , so the strong core is empty. ■

Gale and Shapley (1962) show that the core and strong core for the Roommate model may be empty:

**Example 2. (Gale and Shapely)** Consider the Roommate model given by  $N = \{a, b, c, d\}$ ,  $P(a) = b, c, d$ ;  $P(b) = c, a, d$ ;  $P(c) = a, b, d$ ;  $P(d)$  is arbitrary. In every feasible matching where

no one is unmatched the partner of player  $d$  is part of a blocking pair. In the other cases, there are at least two players who form a blocking pair. ■

The following example shows that the core for  $(N,a)$  may also be empty.

**Example 3 (The core of the one-sided Assignment game may be empty)** Consider  $N=\{1,2,3\}$  and  $a_{ij}=1$  for all  $\{i,j\}\subseteq N$ . For every feasible payoff  $u$  there will exist two players  $i$  and  $j$  such that  $u_i+u_j<1$ . Hence, the core is empty. ■

As discussed in section 1, the condition that

*every  $N$ -simple outcome has an  $N$ -simple extension* (\*)

turns out to be a sufficient condition for the non-emptiness of the core of the one-sided matching markets and it does not require any assumption about the strictness or non-strictness of the preferences of the players in the discrete model. However, this condition is not always satisfied by these models. In Examples 2 and 3 the outcome where everyone is unmatched is the only  $N$ -simple outcome, so it cannot be extended to another  $N$ -simple outcome. Also the property that

*every strongly  $N$ -simple outcome has a strongly  $N$ -simple extension* (\*\*)

is not always satisfied, unless the model is the Marriage market with strict preferences or the two-sided Assignment game. In fact, in the Marriage market of Example 1 the matching that matches  $f_1$  to  $w_1$  and leaves  $f_2$  and  $w_2$  unmatched is unstable and strongly  $N$ -simple. It is a matter of verification that this outcome cannot be extended to a strongly  $N$ -simple matching. Theorem 3 shows that (\*) is also a necessary condition for the non-emptiness of the core of the one-sided Assignment game and of the Roommate market with strict preferences. In addition, (\*\*) is a necessary and sufficient condition for the non-emptiness of the strong core of both discrete matching models with non-necessarily strict preferences. Formally,

**Theorem 3.** (a) *The core of the one-sided Assignment game is non-empty if and only if every unstable and  $N$ -simple outcome can be extended to an  $N$ -simple outcome.*

(b) *The strong core of the Roommate model with non-necessarily strict preferences is non-empty if and only if every unstable and strongly  $N$ -simple matching can be extended to a strongly  $N$ -simple matching.*

(c) *The strong core of the Marriage market is non-empty if and only if every unstable and strongly  $N$ -simple matching can be extended to a strongly  $N$ -simple matching.*

(d) *The core of the Roommate model with strict preferences is non-empty if and only if every unstable and  $N$ -simple matching can be extended to an  $N$ -simple matching.*

**Proof.** The proof that the condition is sufficient is trivial and does not require strict preferences for the discrete models. It follows from the fact that if every simple (respectively, strongly simple) unstable outcome has a simple (respectively, strongly simple) extension then the Pareto optimal simple (respectively, strongly simple) outcomes are in the core. To prove that the condition is necessary for  $(N,a)$ , let  $\sigma=(u,x)$  be some unstable and simple outcome for  $(N,a)$ . Let  $\tau=(w,y)$  be any stable outcome. Define  $M_\sigma(N)$  and  $M_\tau(N)$  as in Lemma 1. Set  $T\equiv\{j\in N; x(j)\neq j\}$  and  $M_0\equiv\{j\in T; u_j=w_{jj}\}$ . Then  $T=M_\sigma(N)\cup M_\tau(N)\cup M_0$ . Now construct the outcome  $(u^*,z)$  as follows:  $z(j)=x(j)$  and  $u^*_j=u_j$  if  $j\in M_\sigma(N)\cup M_\tau(N)$ ;  $z(j)=y(j)$  and  $u^*_j=w_j$  otherwise. It follows from Lemma 1 that all of  $M_\sigma(N)$  are matched to  $M_\tau(N)$  under  $y$  and vice versa. Then,  $z$  is a feasible matching. It is clear that,

$$u^*_j=u_j \text{ for all } j\in T; \quad (1)$$

$$\text{and } u^*_j\geq u_j \text{ for all } j\in N. \quad (2)$$

We claim that  $(u^*,z)$  is stable. That  $(u^*,z)$  is pairwise-feasible and individually rational is immediate from the pairwise-feasibility and individual rationality of  $(u,x)$  and  $(w,y)$ . Thus, it remains to show that  $(u^*,z)$  does not have any blocking pair. The fact that  $(u,x)$  is simple and  $(w,y)$  is stable implies that  $\{j,k\}$  does not block  $(u^*,z)$  in the following cases:  $\{j,k\}\subseteq M_\sigma(N)\cup M_\tau(N)$  and  $\{j,k\}\subseteq N-[M_\sigma(N)\cup M_\tau(N)]$ . Then, without loss of generality, suppose  $j\in M_u(N)\cup M_w(N)$  and  $k\in N-[M_u(N)\cup M_w(N)]$ . If  $\{j,k\}$  blocked  $(u^*,z)$ , we should have  $a_{jk}> u^*_j+u^*_k\geq u_j+u_k$ , where in the last inequality we used (2). In this case  $(j,k)$  would block  $(u,x)$ . However,  $j\in M_\sigma(N)\cup M_\tau(N)$ , so  $j$  is matched at  $x$ , which contradicts the fact that  $(u,x)$  is simple. Hence, in any case,  $(u^*,z)$  does not have any blocking pair, so it is stable.

To see that  $(u^*, z)$  extends  $(u, x)$ , observe that  $u^*_i \geq u_i$  for all  $i \in N$ , by (2). The fact that  $(u^*, z)$  is stable,  $(u, x)$  is unstable and  $u^*_j = u_j$  for all  $j \in T$  implies that there is some  $j \notin T$  such that  $u^*_j > u_j$ . Hence the proof of (a) is complete.

To prove that the condition in (b) is necessary for the non-emptiness of the strong core of the Roommate model with non-necessarily strict preferences, let  $y$  be a strongly stable matching and let  $x$  be an unstable and strongly  $N$ -simple matching for  $(N, P)$ . Using the notation of Lemma 1, let  $S = M_x(N) \cup M_y(N)$ . It follows from Property 4 that all of  $S$  are matched among them under  $y$ . Then, we can construct the matching  $z$  as follows:  $z(j) = x(j)$  if  $j \in S$ ;  $z(j) = y(j)$  otherwise. It is clear that  $z$  is well defined and  $z(j) \succeq x(j)$  for every  $j$ , with strictly inequality for at least one  $j \in N$  (if  $j \notin S$  then  $z(j) = y(j) \succeq x(j)$  and  $z \neq x$  due to the fact that  $x \neq y$ , since  $x$  is not strongly stable). Then,  $z$  extends  $x$ . It remains to show that  $z$  is strongly  $N$ -simple for  $(N, P)$ . We will show that indeed  $z$  is strongly stable and so the desired result follows. That  $z$  is individually rational is immediate from the individual rationality of  $x$  and  $y$ . To see that  $z$  does not have any weak blocking pair, take any pair  $\{j, k\}$ . The fact that  $x$  is strongly  $N$ -simple and  $y$  is strongly stable implies that  $(j, k)$  does not weakly block  $z$  in the cases where  $\{j, k\} \subseteq S$  and  $\{j, k\} \subseteq N-S$ . Then, without loss of generality, suppose  $k \in S$  and  $j \in N-S$ . If  $(j, k)$  weakly blocks  $z$  then  $j \succeq_k z(k) = x(k)$  and  $k \succeq_j z(j) = y(j) \succeq_j x(j)$ , with strict preference for at least one of the agents, so  $(j, k)$  weakly blocks  $x$ . However,  $k$  is matched at  $x$ , which contradicts the fact that  $x$  is strongly  $N$ -simple. Hence, in any case,  $\{j, k\}$  does not weakly block  $z$ , so  $z$  is strongly stable and so it is strongly  $N$ -simple.

The proof of part (c) follows from (b) and the fact that the Marriage market is a sub-model of the Roommate model. The proof that the condition is necessary for the Roommate model with strict preferences is similar to the proof for the case with non-strict preferences, making the convenient changes. Hence, the proof is complete. ■

Theorem 3-b implies that in the Roommate model with non-strict preferences, a Pareto optimal strongly  $N$ -simple matching is in the strong core when this set is non-empty. The following example shows that when preferences need not be strict, Pareto optimal  $N$ -simple matchings for the Roommate model may be unstable even when the core is non-empty. See the following example.

**Example 4. (The core is non-empty but the Pareto optimal simple matching is unstable)** The set of agents are  $N=\{1,2,\dots,6\}$ . The agents' preferences over acceptable partners are given by:

$$\begin{array}{ll} P(1)=2,3,4,1 & P(4)=1,2,3,5,6,4 \\ P(2)=3,1,4,2 & P(5)=6,3,5 \\ P(3)=5,1,2,4,3 & P(6)=[4,5],6 \end{array}$$

Consider the matching  $x$  such that  $x(5)=6$  and the other players are unmatched. It is easy to see that  $x$  is simple and unstable. There is no way to extend  $x$  to a simple matching, because any arrangement with the unmatched players will have a blocking pair where at least one member is not unmatched. Matching  $x$  is clearly a Pareto optimal simple matching. However, the core is non-empty, since matching  $y$ , such that  $y(1)=2$ ,  $y(3)=5$ ,  $y(4)=6$  is stable. This is the only stable matching for this market. ■

From the proof of Theorem 3 it can be concluded that when the core of  $(N,a)$  is non-empty, the matched players, under an  $N$ -simple outcome, keep their payoffs under some stable payoff. The following proposition asserts that the sum of the payoffs of the matched agents under an  $N$ - simple outcome is the same under every stable outcome.

**Proposition 5.** *Let  $(u,x)$  be an  $N$ -simple outcome and let  $(w,y)$  be in the core of  $(N,a)$  or  $(F,W,a)$ . Let  $T=\{j \in N; x(j) \neq j\}$ . Then,  $\sum_{i \in T} u_i = \sum_{i \in T} w_i = v(T)$ .*

**Proof.** It follows from Remark 1 and from the fact that  $T$  does not block  $(u,x)$  that  $\sum_{i \in T} u_i = v(T)$ . Define the stable outcome  $(u^*,z)$  as in the proof of Theorem 3-a. Then,  $v(N) = \sum_{i \in N} u^*_i = \sum_{i \in T} u_i + \sum_{i \in N-T} w_i = v(T) + \sum_{i \in N-T} w_i \leq \sum_{i \in T} w_i + \sum_{i \in N-T} w_i = \sum_{i \in N} w_i = v(N)$ , so  $v(T) = \sum_{i \in T} w_i$ , where in the inequality was used that  $T$  does not block  $(w,y)$ . Hence,  $\sum_{i \in T} u_i = \sum_{i \in T} w_i$ , and the proof is complete. ■

**Corollary 2 .** *Let  $(u,x)$  be an  $N$ - simple outcome and let  $(w,y)$  be in the core of  $(N,a)$ . Let  $T(x)=\{j \in N; j \text{ is not part of any blocking pair}\}$ . Then,  $\sum_{i \in T(x)} u_i = \sum_{i \in T(x)} w_i$ .*

**Proof.** Set  $M_0 \equiv \{j \in T(x); x(j) = j\}$ . Then,  $T(x) = T \cup M_0$ , where  $T = \{j \in N; x(j) \neq j\}$ . If  $j \in M_0$  then  $u_j = 0$ . By Property 5  $w_j = 0$ . Now use Proposition 5 to get the result. ■

Two-sided matching markets yields statics results about the effects of increased competition that has no parallel in the one-sided case. Theorem 4 implies that when any finite subset of  $W$ -players (respectively,  $F$ -players) is added to the pool of the two-sided Assignment game or of the Marriage market, and the players are allocated according to some stable outcome, then there is some outcome in the core of the new market under which no  $F$ -player (respectively,  $W$ -player) is worse off and no  $W$ -player (respectively,  $F$ -player) is better off.

**Theorem 4.** (a) Suppose  $k \in W$  is added to  $M=(F,W,a)$ . Let  $(t,v;x)$  be some stable outcome for  $M$ . Then there is some stable outcome  $(t^*,v^*;x^*)$  for the new market  $M^*$ , such that  $t^*_i \geq t_i$  for every  $i \in F$  and  $v^*_j \leq v_j$  for every  $j \in W$ .

(b). Suppose  $k \in W$  is added to  $M=(F,W,P)$  with strict (respectively, non-necessarily strict) preferences. Let  $x$  be some stable (respectively, strongly stable) matching for  $M$ . Then there is some stable matching  $x^*$  for the new market  $M^*$ , such that  $x^*(i) \geq_i x(i)$  for every  $i \in F$  and  $x(j) \geq_j x^*(j)$  for every  $j \in W$ .

**Proof.** To prove (a), let  $(t',v';x')$  be the outcome that agrees with  $(t,v;x)$  on  $F \cup W$  and leaves  $k$  unmatched with payoff  $v'=0$ . This outcome is  $W$ -simple for  $M^*$ , because any blocking pair, if any, contains  $k$ . If  $(t',v';x')$  is stable we are done. Otherwise, Proposition 4 a) implies the existence of a core outcome  $(t^*,v^*;x^*)$  that extends  $(t',v';x')$ , so  $t^*_i \geq t_i$  for all  $i \in F$ . To show that this core outcome satisfies the other assertion, let  $j \in W$ . Suppose  $v^*_j > v_j$ . Then there exists some  $i \in F$  such that  $x^*(j)=i$ . Using that  $t^*_i \geq t_i$  it follows that  $a_{ij}=t^*_i+v^*_j > t_i+v_j$ , so  $(i,j)$  blocks  $(t,v;x)$ , which is a contradiction. Hence  $v_j \geq v^*_j$  for all  $j \in W$  and we have proved (a).

The proof of part (b) follows the line of argument used in the proof of part (a). Get the contradiction that  $(i,j)$  blocks  $x$ , if  $x$  is stable and weakly blocks  $x$  if  $x$  is strongly stable. ■

**Remark 2.** For the Marriage model, the comparative static result peaks any core outcome and compares it with the limit point of the sequence of  $W$ -simple outcomes for the augmented market, generated by the process described above, starting with the given

outcome. For the two-sided Assignment game, the outcome  $(t^*, v^*; x^*)$  constructed in the proof of Proposition 4 is clearly the  $F$ -optimal stable outcome for  $(F, W, a)$ . However, the optimality of this outcome is not required for the proof of Theorem 4. That is, if  $(t^*, v^*; x^*)$  is any stable outcome that extends  $(t', v'; x')$ , then we still have that  $t^*_i \geq t'_i$  for every  $i \in F$  and  $v^*_j \leq v'_j$  for every  $j \in W$ . Now observe that the set of stable payoffs  $(t'', v'')$  such that  $t'' \geq t'$  is a non-empty sub-lattice of the lattice of the stable payoffs, under the partial order defined by the preferences of the  $F$ -players. Therefore, unless this sub-lattice is a singleton, there always exists some stable outcome in  $M^*$ , other than the  $F$ -optimal stable outcome, that extends  $(t', v'; x')$ . ■

### 3. CONCLUDING REMARKS.

The general idea to concentrate attention on the simple outcomes has been overlooked in the literature.<sup>12</sup> This paper uses the concepts of simple outcome and Pareto optimal simple outcome to provide us with a new conceptual view of one-to-one matching models, which allows for better understanding of the coalition formation process. The novelty here is that the gains in insight with this approach allowed the identification of a new necessary and sufficient condition for the non-emptiness of the strong core of the Marriage market and Roommate market with non-necessarily strict preferences, as well as a new necessary and sufficient condition for the non-emptiness of the core of the one-sided Assignment game and Roommate model with strict preferences. This condition is more readily interpretable economically than the requirement, for instance, that the game is balanced. The unified analysis of the one-sided and the two-sided markets made it easier to access the usefulness of the framework of simple outcomes. We could demonstrate, for example, that the sufficient condition for the non-emptiness of the core is always satisfied by the two-sided markets.

The main characteristic of simple outcomes is that, when the core is non-empty, unstable  $N$ -simple outcomes are endowed with a sort of *internal stability* in the following

---

<sup>12</sup> Similar concepts were introduced in several two-sided matching markets: in Sotomayor (1996), for the Marriage market; in Sotomayor (1999) for the discrete many-to-many matching market with substitutable and non-strict preferences and for the College Admission model, and in Sotomayor (2000), for the Assignment game of Shapley and Shubik and the unified two-sided matching model of Eriksson and Karlander (2000). Recently, Sotomayor (2005) has introduced the concept of simple allocation for the one-sided market (not matching market) of Shapley and Scarf (1974). For all these models, a non-constructive existence proof is provided.

sense: The set of players who are not part of any blocking coalition of an  $N$ -simple outcome, keep their payoffs in some stable outcome in the continuous models and keep their mates in some stable matching in the discrete models with strict preferences. Also, for the continuous models, the total payoff gained by the matched players under an  $N$ -simple outcome is the same under every stable outcome. For the discrete models, the matched players under an  $N$ -simple matching are matched among themselves under any stable matching.

Bondareva (1963) and Shapley (1967) proved that the core of a transferable utility game is non-empty if and only if the game is balanced. Thus, for the transferable utility games considered here, our condition is equivalent to that of balancedness. It turns out that this equivalence is not accidental; rather it persists for all TU games. This result was proved in a companion paper (Sotomayor, 2006). The hindrance for the use of Lemma 1 in a general TU game is that a feasible outcome not always is supported by a matching. The observation that the intuition behind the concept of  $N$ -simple outcome is not related to a matching led us to a convenient adaptation of this concept, consistent with the concept used for the matching models.

The possibility of obtaining a stable matching for the discrete models by starting from an arbitrary matching and successively satisfying blocking pairs has been subject of several papers, motivated by an open problem posed by Knuth (1976): *Let  $\mu$  be an unstable matching for the Marriage model  $(F, W, P)$ . Is there a sequence of matchings  $\mu_1, \dots, \mu_k$ , such that  $\mu = \mu_1$ ,  $\mu_k$  is stable and for each  $i=1, \dots, k-1$ , there is a blocking pair  $(f_i, w_i) \in F \times W$  for  $\mu_i$  such that  $\mu_{i+1}$  is obtained from  $\mu_i$  by matching  $f_i$  to  $w_i$ , their mates,  $\mu_i(f_i)$  and  $\mu_i(w_i)$ , if any, to each other and all other agents to the same mates at  $\mu_i$ ?* If we leave  $\mu_i(f_i)$  and  $\mu_i(w_i)$  unmatched the problem is easily solved by using a version of the deferred-acceptance algorithm of Gale and Shapley. This approach is treated by Roth and Vande Vate (1990), giving origin to several other papers: Abeledo and Rothblum (1995), Tan and Hsueh (1995), Blum, Roth and Rothblum (1997), Blum and Rothblum (2002), Ma (1996) and Cechlárova (2002) for the Marriage model; Diamantoudi *et al.* (2004), for the Roommate model.

## REFERENCES.

- Abeledo, H. and Isaak, G. (1991), "A characterization of graphs that ensure the existence of stable matchings", *Mathematical Social Sciences* vol. 22, Issue 1, 93-96.
- Abeledo, H. and Rothblum, U. G. (1995) "Paths to marriage stability", *Discrete Applied Mathematics*, 63.
- Blum, Y., Roth, A. and Rothblum, U. G. (1997) "Vacancy chains and equilibration in senior-level labor markets", *Journal of Economic Theory*, 76.
- Blum, Y., Rothblum, U. G. (2002) "Timing is everything" and marital bliss., *Journal of Economic Theory*, 103.
- Cechlárová, K. (2002) "Randomized matching mechanism revisited", mimeo.
- Chung, Kim-Sau (2000) "On the existence of stable roommate matchings", *Games and Economic Behavior* 33, 206-230.
- Demange, G. and Gale, D. (1985) "The strategy structure of two-sided matching markets", *Econometrica*, 53, 873-888.
- Diamantoudi, E., Miyagawa, E. and Xue, L. (2004) "Random paths to stability in the roommate problem", *Games and Economic Behavior*, 48.
- Eriksson, K. and Karlander, J. (2000) "Stable matching in a common generalization of the marriage and assignment models", *Discrete Mathematics* 217, 135-156.
- Gale, D. and Shapley, L. (1962) "College admissions and the stability of marriage", *American Mathematical Monthly* 69, 9-15.
- Gale, D. and Sotomayor M. (1985-a) "Some remarks on the stable matching problem" *Discrete Applied Mathematics*, 11, 223-32.
- Gale, D. and Sotomayor M. (1985-b) "Ms. Machiavelli and the stable matching problem", *American Mathematical Monthly*, 92, 261-268.
- Gusfield, D. (1988) "The structure of the stable roommate problem: efficient representation and enumeration of all stable assignments", *SIAM Journal on Computing*, 17, 742-69.
- Irving, R. W. (1985) "An efficient algorithm of the stable room-mates problem", *Journal of Algorithms*, 6, 577-95.
- Knuth (1976)
- McVitie, D. G. and Wilson L. B. (1970) "Stable marriage assignments for unequal sets", *BIT*, 10, 295-309.

- Roth, A. (1984) "The evolution of the labor market for medical interns and residents; a case study in game theory", *Journal of Political Economy*, 92, 991-1016.
- Roth, A. and Sotomayor, M (1990) "Two-sided matching. A study in game-theoretic modeling and analysis" *Econometric Society Monographs*, 2<sup>nd</sup> ed., Vol. 18, Cambridge University Press.
- Roth, A. and Vande Vate, J. H. (1990) "Random paths to stability in two-sided matching", *Econometrica*, 58.
- Shapley, L. and Scarf, H. (1974) "On cores and indivisibility" *Journal of Mathematical Economics*, 1, 23-8.
- Shapley, L. and Shubik, M. (1972) "The assignment game I: The core", *Journal of Game Theory* 1, 111-130.
- Sotomayor, M. (2006) "On simple outcomes and cores of the coalitional games", mimeo.
- \_\_\_\_\_ (2005) "An elementary non-constructive proof of the non-emptiness of the core of the housing-market of Shapley and Scarf" *Mathematical Social Sciences*, to appear.
- \_\_\_\_\_ (2000) "Existence of stable outcomes and the lattice property for a unified matching market", *Mathematical social Science* 39 , 119-132.
- \_\_\_\_\_ (1999) "Three remarks on the many-to-many stable matching problem", *Mathematical Social Sciences* 38, 55-70
- \_\_\_\_\_ (1996) "A non-constructive elementary proof of the existence of stable marriages", *Games and Economic Behavior* 13, 135-137.
- Tan, J. J. M. (1991) "A necessary and sufficient condition for the existence of a complete stable matching", *J. of Algorithms*, vol. 12, 154-178.
- Tan, J. J. M., Hsueh, Y.C. (1995) " A generalization of the stable matching problem" *Discrete Applied Mathematics*, 59.