PARETO-STABILITY CONCEPT: A NATURAL SOLUTION CONCEPT FOR THE ROOMMATE AND THE MARRIAGE MODELS

By

MARILDA SOTOMAYOR[1]

Universidade de São Paulo

Department of Economics, , Cidade Universitária, Av. Prof. Luciano Gualberto 908, 05508-900, São Paulo, SP, Brazil

Brown University

Department of Economics, 64 Waterman Street, Providence, RI, USA

e-mail: marildas@usp.br

10/2007

ABSTRACT

The traditional concept of stability assumes that only strong blockings are allowed to rupture the structure of a matching. This paper introduces a natural solution concept for the Roommate and the Marriage models which can be conveniently adapted for other discrete matching models. The idea is to allow that, under non-necessarily strict preferences, also weak blockings can upset a matching once they come from the grand coalition. More specifically, in a decentralized setting, exchange of partners leading to a (weak) Pareto improvement of a stable matching should be allowed. The concept called here Pareto-stability, requires stability plus Pareto optimality.

We show that the set of Pareto-stable matchings is non-empty when the core is non-empty but it may be smaller than the core. Under strict (respectively, weak) preferences, these outcomes are the Pareto-optimal simple matchings[2] for the roommate and the marriage models (respectively, marriage model). A characterization of the strong core is also provided.

Some characteristic properties of the core of the Marriage market are shown not to depend on the fact that this market has two sides.

Keywords: Pareto-optimal, core, stable matching, Pareto-stable matching, simple matching, Pareto-optimal simple matching

JEL numbers: C78, D78

[1] This paper was partially supported by CNPq-Brazil. The author thanks the hospitality of Brown University where most of this work was written. Many thanks are due to Roberto Serrano for his helpful comments.

[2] The individually rational matchings whose blocking pairs, if any, are formed with unmatched agents are called simple matchings. Pareto-optimal simple matchings are the simple matchings that are Pareto-optimal among all simple matchings.

INTRODUCTION

It is known that the core is contained in the set of Pareto-optimal outcomes in the coalitional games that can be represented in a characteristic function form. This property also holds for the discrete matching models with strict preferences, and for the continuous matching models where the structure of the preferences is given by utility functions that are continuous in the money variable, which varies in some interval of the real line. The explanation is that any core point that has a weak Pareto improvement must have a weak blocking coalition and so the given point cannot be in the strong core. On the other hand, in all these models, the core and the strong core coincide, so there is no core outcome which has a weak Pareto improvement.

When preferences are non-necessarily strict in the discrete matching models, the core may be different from the strong core. An example in this paper shows that we may have core matchings that are not Pareto-optimal for the Roommate model. Another example illustrates this negative result for the Marriage model.

The presence of outcomes in the core that are not Pareto-optimal suggests that the game-theoretical predictions that in a decentralized setting a stable matching will occur should be revised when preferences are non-necessarily strict and the core is non-empty. To figure out the kind of coalitional interaction taking place between the agents allocated according to a core outcome that is not Pareto-optimal, consider some discrete one-to-one matching model, the Roommate model or the Marriage model. Technically speaking, if a stable matching x is not Pareto-optimal then it can be weakly improved by some other stable matching y via some coalition A. Of course the grand coalition weakly blocks x. However coalition A need not be the grand coalition. It includes at least one weak blocking pair of x and weakly blocks x in an unusual way. The players in A are matched among them at x and y, that is, $x(A)=y(A)$; if j belongs to A then $x(j)\neq y(j)$. Furthermore, the players in A who prefer y to x (there is at least one) are not unmatched at y and form weak blocking pairs of x with their mates at y. Then if j belongs to A and $y(j)=k$ (k might be j in case j is unmatched at y), k belongs to A

and either $\{j,k\}$ is a weak blocking pair of x or j and k are indifferent between x and y. Clearly a coalition with a single weak blocking pair is not able to cause a weak Pareto improvement of x. Therefore there may be core points that are Pareto optimal but are not in the strong core.

Coalition A is then a collection of players who are able *to exchange their partners among them* so that no one is worse off and at least one member of the coalition is better off. We argue that in a context in which players can freely communicate with each other and get together in groups, such a coalitional interaction between agents should be allowed and then the prediction should be that A can be formed, and so the matching x should not occur. *The solution concept which captures this idea of equilibrium is that of Pareto–stability.* (See Example 2). We say that an outcome is Pareto-stable if it is stable and Pareto-optimal . As it is proved here, when preferences are strict then Pareto-stability is equivalent to stability. Hence nothing is changed in this case. However, when preferences are non-necessarily strict this concept is a refinement of the stability concept. In fact, we show that the set of Pareto-stable matchings may be a proper subset of the core.

The remaining part of this paper is devoted to investigate the Pareto-stable outcomes. We do that for the well-known Roommate and Marriage models, both introduced by Gale and Shapley in their famous paper of 1962. As it is proved here, Pareto-stable matchings always exist when the core is non-empty. Since the Marriage model is a sub model of the Roommate model, we can concentrate on the last model and extend our positive results to the former model.

We approach these one-to-one matching models by focusing on certain individually rational matchings, whose blocking pairs, if any, are formed with unmatched agents. The matchings that satisfy this property are called *simple*. Thus, simple matchings are individually rational matchings in which none of the matched agents is member of a blocking pair. Simple matchings exist even when stable matchings do not, since the matching where every one is unmatched is simple. Clearly, every stable matching is simple. When preferences are non-necessarily strict we may also focus on strongly simple matchings similarly defined. Of course, simple matchings are strongly simple when preferences are strict.

Under this approach, two other concepts are of fundamental importance in our analysis: *Pareto-optimal simple matching* and *Pareto-optimal strongly simple matching.* A matching is Pareto-optimal simple if it is simple and Pareto-optimal among all simple matchings. Similarly, a matching is Pareto-optimal strongly simple if it is strongly simple and Pareto-optimal among all strongly simple matchings. These outcomes always exist because the set of simple matchings, as well as the set of strongly simple matchings, is non-empty, finite and the preferences are transitive. The main finding of this paper concerns the role played by these outcomes in the characterization of the core outcomes that are Pareto-optimal and of the strong core outcomes that are Pareto-optimal, respectively.

As mentioned before, for the case where preferences are strict and the core is non-empty we show that every stable matching is Pareto-stable. Then the core coincides with the set of Pareto-stable matchings. Of course we may have Pareto-optimal matchings that are unstable. However, if a Pareto-optimal matching is simple then it is stable. Indeed, we need not require that the simple matching be Pareto-optimal . It is enough to be Pareto-optimal only among the simple matchings. In fact, *when preferences are strict and the core is non-empty we characterize the set of Pareto-stable matchings (the core) as the set of Pareto-optimal simple matchings.*

When preferences are non-necessarily strict, this same characterization for the set of Pareto-stable matchings is obtained for the Marriage model. However, as pointed out above, in both models, the set of Pareto-stable matchings may be smaller than the core. That is, under non-necessarily strict preferences, a stable matching need not be Pareto-optimal . On the other hand, the equivalence between the Pareto-stability concept and the concept of Pareto-optimal simple matching does not extend to the Roommate model. In fact, our examples show that we may have Pareto-optimal simple matchings which are not in the core of the Roommate model, even when this set is non-empty. The set of Pareto-stable matchings contains the strong core but may be bigger than this set, even when this set is non-empty. On the other hand, if a Pareto-optimal matching is strongly simple then it is strongly stable. In fact, if the strong core is non-empty we characterize it as the set of Pareto-optimal strongly simple matchings. Then, if all Pareto-optimal simple matchings are unstable, the strong core is empty.

The proofs of our characterization results make use of a technical lemma demonstrated here. This lemma also allows us to prove some characteristic properties of the stable matchings which do not depend on the fact that the market has one or two sides. When preferences are strict these properties are well-known results for the Marriage market. Indeed these properties have some analogue in the one-to-one continuous markets, which indicates that they are even more fundamental than one might expect.

We hope this work has contributed with one step further for a better understanding of the under investigated matching models with non-necessarily strict preferences.

The formal model is described in section 2. The examples in the next section aim to help the understanding of our main results. Section 4 presents two preliminary results. Section 5 is devoted to obtain some characterization of the set of Pareto-stable and of Pareto-strongly stable matchings for the Roommate model and the Marriage model. Section 6 presents some characteristic properties of the two models. Section 7 concludes the paper and presents some related work.

2. DESCRIPTION OF THE MODEL

In this section we will describe the Roommate model and will present the main concepts and definitions for this model. The Marriage model can be regard as a sub model of the Roommate model: Let every man list all the other men as unacceptable and let every woman list all the other women as unacceptable. Then the definitions presented in this section can easily be adapted to the Marriage model.

There is a finite set of players, $N=\{1,2,\ldots,n\}$. Each player is interested in forming at most one partnership with players of N and has complete and transitive preferences over the players in N. Hence, player j's preference can be represented by an ordered list of preferences, $P(j)$, on the set N. 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 $P(j)=i, [m,l], j, \ldots, q$ indicating that j prefers i to m; he/she is indifferent between m and l and prefers any of these two last players to himself/herself; anyone else is *unacceptable* to himself/herself.

The model can then be described by (N,P), where $P=\{P(1),\ldots,P(n)\}$.

Definition 1. *A* ***matching*** x *is a one-to-one correspondence from* N *onto itself of order two (that is,* $x^2(j)=j$*). We refer to* $x(j)$ *as the* ***partner of j at x*** *(even if* $x(j)=j$*).*

If $x(j)=j$ we say that j is **unmatched** at x. Player j prefers matching x to matching y if and only if he/she prefers $x(j)$ to $y(j)$. 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.

Definition 2. *The matching* x *is* ***individually rational*** *if each player is acceptable to his or her partner.*

The key notion is that of stability.

Definition 3. *We say that the pair* (j,k) ***blocks*** *the matching* x *if* $j >_k x(k)$ *and* $k >_j x(j)$*. A matching* x *is* ***stable*** *if it is individually rational and it does not have any blocking pair. If* x *is not stable we say that it is* ***unstable****.*

We say that the pair (j,k) ***weakly blocks*** *the matching* x *if* $j \geq_k x(k)$ *and* $k \geq_j x(j)$*, with strict preference for at least one of the two players. A matching* x *is* ***strongly stable*** *if it is individually rational and it does not have any weak blocking pair.*

It is a matter of verification that a matching is stable (respectively, strongly stable) if and only if it is in the core (respectively, strong core).

Definition 4. *A matching* x *is* ***simple*** *if it is individually rational and no matched agent is part of a blocking pair. The matching* x *is* ***strongly simple*** *if it is individually rational and no matched agent is part of a weak blocking pair.*

Hence, in case (j,k) blocks (respectively, weakly blocks) the simple (respectively, strongly simple) matching x then j and k are unmatched at x.

When preferences are strict *every strongly simple matching is simple and vice versa.* Since the matching at which no partnership is formed is simple and strongly simple, **the set of simple matchings and the set of strongly simple matchings are non-empty.** Clearly, every stable matching is simple and every strongly stable matching is strongly simple.

Definition 5. *Let x and z be individually rational matchings. We say that* ***z extends*** *x or that* ***x has an extension*** *if $z(i) \geq_i x(i)$ for all $i \in N$, with strict preference for at least one player in N. If z is simple (respectively, strongly simple) we say that* ***x has a simple (respectively, strongly simple) extension.***

Therefore, if z extends x then z is a weak Pareto improvement of x. Observe that a weak Pareto improvement of x does not create any new blocking pair.

Definition 6. *Matching x is called* **Pareto-optimal simple matching** *if it is simple and it does not have any simple extension.*

Matching x is called **Pareto-optimal strongly simple matching** *if it is* **strongly simple** *and it does not have any* **strongly simple** *extension.*

Matching x is called **Pareto-optimal** *if it is individually rational and it does not have any extension.*

Matching x is called **Pareto-stable** *if it is stable and Pareto-optimal . It is called* ***Pareto-strongly stable*** *if it is strongly stable and Pareto-optimal .*

Remark 1. Clearly, if a stable matching is Pareto-optimal then it is simple and Pareto-optimal among all simple matchings, so it is Pareto-optimal simple. Also if x is stable and Pareto optimal simple then it is Pareto optimal. In fact, if it was not then there would be some matching y which extends x. By the observation above y must be stable and so it is simple, which contradicts the Pareto optimality of x among all simple matchings. With the same kind of argument we can see that if a matching x is strongly stable then it is Pareto-optimal if and only if it is Pareto-optimal strongly simple. ■

The existence of Pareto-optimal simple and Pareto-optimal strongly simple matchings is guaranteed by the fact that the set of simple matchings and the set of strongly simple matchings are non-empty and finite and preferences are transitive.

3. EXAMPLES

Example 1. (The core of the Roommate model is non-empty and the strong core is empty. Furthermore, some Pareto-optimal simple matching is unstable) Consider the Roommate model where the set of agents are N={1,2,…,6}. The agents' preferences over acceptable partners are given by:

P(1)=2,3,4,1 P(4)=[5,6],4

P(2)=3,1,4,2 P(5)=4,3,5

P(3)=1,5,2,3 P(6)=4,,6

The set of stable matchings is non-empty since matching *y*, such that *y(1)=2, y(3)=5* and *y(4)=6,* is stable. This is the only stable matching for this market. The pair *{5,4}* weakly blocks *y,* so the set of strongly stable matchings is empty. Clearly, *y* is Pareto-optimal , so it is Pareto-stable and Pareto-optimal simple. The fact that *y* is Pareto-optimal is guaranteed by Proposition 2.

Now, let *x* be the matching that assigns *5* to *4* and leaves unmatched the other players. It is easy to see that *x* is simple and unstable. Also *x* is a Pareto-optimal simple matching, since there is no way to extend *x* to a simple matching. Any arrangement with the unmatched players will have a blocking pair where at least one member is not unmatched. Of course *y* is not Pareto optimal. This matching is not strongly simple, since *{4,6}* is a weak blocking pair and *4* is matched.■

Example 2. (The core and the strong core of the Roommate model are non-empty; a stable matching that is not Pareto-optimal; the set of Pareto-stable matchings is a proper subset of the core)

Consider the Roommate model where the set of agents are N={1,2,…,8}. The agents' preferences over acceptable partners are given by:

P(1)=2 P(5)=8, 6

P(2)=3, 1 P(6)=[3, 5]

P(3)=6, [4,2]   P(7)=4, 8

P(4)=[3, 7]   P(8)= [5, 7]

The matching z where $z(1)=2, z(3)=4, z(5)=6, z(7)=8$ is stable. The matching z is not Pareto-stable since it is not Pareto-optimal. The matching w such that $w(1)=2, w(3)=6, w(5)=8$ and $w(7)=4$ is a weak Pareto improvement of z. It is easy to see that z and w are the only stable matchings and that w is the only matching which is in the strong core. Also, w is Pareto-stable. This fact is implied by Proposition 1. By Remark 1 w is Pareto optimal strongly simple.

There are other Pareto-optimal matchings in this market. For example, the matching z' such that $z'(2)=3, z'(5)=8, z'(4)=7$, 1 and 6 are left unmatched is Pareto-optimal but it is unstable: {3,6} blocks z'. Since 3 is matched then z' is not simple. ■

Example 2. (Continued) (Pareto-stability as a natural solution concept for the Roommate model) The set of agents are N={1,2,…,8}. The agents' preferences over acceptable partners are given by:

P(1)=2   P(5)=8, 6

P(2)=3, 1   P(6)=[3, 5]

P(3)=6, [4,2]   P(7)=4, 8

P(4)=[3, 7]   P(8)= [5, 7]

Consider the stable matching z again, where $z(1)=2, z(3)=4, z(5)=6, z(7)=8$. There are four weak blocking pairs: *{2,3}, {3,6}, {5,8}* and *{7,4}*. Player 2 would like to be with player 3 but player 3 is indifferent between his partner, player 4, and player 2 and player 4 does not want to be with player 2`s partner. The situation is different with the other three weak blocking pairs. By acting together, players $3, 5$ and 7 can be better off by exchanging their partners among them: 3 gets 5's partner, 5 gets 7's partner and 7 gets 3's partner. In fact, it is reasonable to expect that if these players propose this new arrangement to their partners, these ones will accept it, since they are indifferent between their partners under z and their new proposed partners, so they will not be worse off. Moreover, their partners under z agree to leave them and to have different partners.

Matching w such that $w(1)=2$, $w(3)=6$, $w(5)=8$ and $w(7)=4$ will reasonably be the resulting matching of such coalitional interaction.

The reason by which the pairs $\{3,6\}$, $\{5,8\}$ and $\{7,4\}$ can exchange their partners is that they are able to cause a weak Pareto improvement in matching z by matching the agents within each pair to one another. Observe that no weak Pareto improvement of matching z can improve the situation of player 2. In fact, if player 2 gets better off, player 1 gets worse off. ■

As in the Roommate model under non-strict preferences, the strong core of the Marriage model may also be empty and there may be stable matchings that are not Pareto-optimal . These facts are illustrated in Example 3 and Example 4, respectively.

Example 3. (The strong core of the Marriage market may be empty) Consider the Marriage model where the set of agents are $M=\{m_1, m_2\}$ and $W=\{w_1,w_2\}$. Agent m_1 is indifferent between w_1 and w_2; both women prefer m_1 to m_2 and are acceptable to m_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 (m_1,w_1) or by (m_1,w_2), so the strong core is empty. It is a matter of verification that both matchings are Pareto-optimal.■

Example 4. (A stable matching for the Marriage market that is not Pareto-optimal) Consider the Marriage model where the set of agents are $M=\{m_1, m_2\}$ and $W=\{w_1,w_2\}$. Agent m_1 is indifferent between w_1 and w_2; m_2 prefers w_1 to w_2; w_1 is indifferent between m_1 and m_2 and w_2 prefers m_1 to m_2. Both matchings under which no agent is unmatched are stable and are the only stable matchings. The matching x where $x(m_1)=w_1$ and $x(m_2)=w_2$ is not Pareto-optimal and is not in the strong core. It is Pareto improved by matching y where $y(m_1)=w_2$ and $y(m_2)=w_1$. Matching y is in the strong core. ■

The existence of two sides in the Marriage market causes fundamental differences between the two models, starting with the fact that the core is always non-empty for the

Marriage model, under any kind of preferences, and the optimal stable matchings for each side of the market always exist when preferences are strict. Under strict preferences, if these two optimal stable matchings coincide, the core is a singleton. However, when preferences need not be strict, Example 5 shows a situation where the man-optimal stable matching coincides with the woman-optimal stable matching, but the core is not a singleton.

Example 5. (The man-optimal stable matching coincides with the woman-optimal stable matching but the core of the Marriage market is not a singleton)
Consider Example 4 again. Matching y is clearly optimal for the men and for the women but matching x is also in the core. ■

4. SOME PRELIMINARY RESULTS

As the examples in section 2 illustrate, Pareto-optimal matchings, Pareto-optimal strongly simple matchings and Pareto-optimal simple matchings need not be stable. When preferences need not be strict Example 2 shows a situation in which a stable matching is not Pareto-optimal . Proposition 1 below implies that if preferences are strict then every stable matching is Pareto-optimal . This follows from the fact that in this case the strong core and the core coincide.

Proposition 1. *The strong core of the Roommate model (N,P) is a subset of the set of Pareto-optimal matchings.*

Proof. Let x be a strongly stable matching. Suppose x is not Pareto-optimal . Then there is some individually rational matching y such that $y(i)\geq_i x(i)$ for all $i \in N$ with strict preferences for at least one agent. Then let j be such an agent. The fact that $y(j)>_j x(j)$ implies that $y(j)=k$ for some $k \neq j$ and $y(j)\neq x(j)$. By hypothesis we have that $j=y(k)\geq_k x(k)$. Then $\{j,k\}$ weakly blocks x, which is a contradiction. ■

Hence every matching in the strong core is Pareto-optimal . However, how to guarantee the existence of Pareto-stable matchings when the strong core is empty? In Example 1 the strong core of the Roommate model is empty and in Example 3 the strong core of the Marriage model is also empty. In despite of this the set of Pareto-stable matchings is non-empty in both examples. The existence of Pareto-stable matchings in these examples is guaranteed by Proposition 2 below.

Proposition 2. *Suppose the core of the Roommate model* (N,P) *is non-empty. Then the set of Pareto-stable matchings is non-empty.*

Proof. If agents' preferences are strict then we are done, because in this case the core is a subset of the set of Pareto-optimal matchings by Proposition 1. Then suppose the preferences are non-necessarily strict. Take any stable matching x. Suppose by way of contradiction that there is no Pareto-stable matching. Then x is not Pareto-optimal and so there must exist some matching y_1 which extends x, so $y_1 \neq x$. The matching y_1 is stable since any pair of agents who blocks it also blocks x, which contradicts the stability of x. (In fact, if $\{j,k\}$ blocks y_1 then $k >_j y_1(j)$ and $j >_k y_1(k)$. Then, $k >_j x(j)$ and $j >_k x(k)$). Then y_1 is not Pareto-optimal and so there must exist some matching y_2 which extends y_1, so $y_2 \neq y_1$. Clearly, y_2 is also an extension of x, so $y_2 \neq x$. Again, matching y_2 is stable since it does not create any blocking pair, so it is not Pareto-optimal , and so on. Thus, we have a sequence of stable matchings $x, y_1, y_2, \ldots, y_t, \ldots$ where every term of the sequence is an extension of the previous ones. Therefore all these matchings are distinct, from which follows that the sequence does not cycle and so it is infinite, which is absurd. Hence, there always exists a Pareto-stable matching. ■

5. CHARACTERIZATION OF THE PARETO-STABLE AND OF THE PARETO-STRONGLY STABLE MATCHINGS

It follows from Proposition 1 and Remark 1 that *every stable matching is Pareto-optimal simple when preferences are strict.* It turns out that the Pareto-optimal simple matchings are stable and are exactly the Pareto-optimal matchings that are in the core when this set is non-empty and preferences are strict. More generally, we can characterize the Pareto-optimal matchings which are in the strong core under any kind of preferences, strict or non-strict, as *the set of Pareto-optimal strongly simple matchings.*

Theorem 1. *Suppose the strong core of the Roommate model (N,P) is non-empty. Then, the set of Pareto-strongly stable matchings equals the set of Pareto-optimal strongly simple matchings.*

If preferences are strict, Theorem 1 implies that the Pareto-stable matchings are exactly the Pareto-optimal simple matchings. When preferences are non-necessarily strict this result is not true for the Roommate model (recall Example 1). Nevertheless a characterization of the set of Pareto-stable matchings as the set of Pareto-optimal simple matchings for the Marriage market under any kind of preferences is provided by Theorem 2 below:

Theorem 2. *Consider the Marriage model (M,W,P). The set of Pareto-stable matchings equals the set of Pareto-optimal simple matchings.*

In order to prove these theorems we need some technical results which will be given in sub-section 5.1.

5.1 TECHNICAL RESULTS

The following result is a powerful lemma that enables us to derive all our theorems.

Lemma 1. *Consider the Roommate model (N,P). Let* x *be a strongly simple matching and let* y *be a strongly stable matching. Let* $T=\{j\in N;\ x(j)\neq j\}$, $M_x=\{j\in N;\ x(j) >_j y(j)\}$ *and* $M_y=\{j\in T;\ y(j) >_j x(j)\}$. *Then* $x(M_x)= y(M_x)=M_y$ *and* $x(M_y)=y(M_y)=M_x$.

Proof. All j in M_x are matched under x, since $x(j)>_j y(j) \geq_j j$. Analogously, all j in M_y are matched under y, since $y(j)>_j x(j)\geq_j j$. If j is in M_x then $k=x(j)$ is in M_y, for if not $j=x(k) \geq_k y(k)$, so y will be weakly blocked by j and k, which contradicts the assumption that y is strongly stable. On the other hand, if k is in M_y then $j=y(k)$ is in M_x, for if not $k=y(j) \geq_j x(j)$, so x will be weakly blocked by j and k. However, k is in T, so k is matched under x, which contradicts the fact that x is simple. Therefore,

$x(M_x)\subseteq M_y$ and $y(M_y)\subseteq M_x$. Since x and y are one-to-one and M_x and M_y are finite sets, the conclusion follows. Hence the proof is complete.■

That is, if x is a strongly simple matching and y is a strongly stable matching, then both x and y map the set of agents who prefer x to y onto the set of agents who prefer y to x and are matched at x.

Suppose to fix ideas that the preferences are strict. Given a simple matching x, which is not in the core, it is always possible to obtain a new matching z by doing the following: Keep the partnerships formed under x, if any, and add some new partnerships. Of course, these new partnerships are formed with blocking pairs of x. Proposition 3 asserts that matching z can be constructed so that it is stable.

Proposition 3 *Consider the Roommate model (N,P). Let x be a strongly simple matching that is not strongly stable. If the set of strongly stable matchings is non-empty then there exists a strongly stable matching z that extends x.*

Proof. Let y be a strongly stable matching. Using the notation of Lemma 1, set $S\equiv M_x\cup M_y$. We claim that all of S are matched among them under x and y. In fact, let $j\in S$. If $j\in M_x$ then Lemma 1 implies that $x(j)\in M_y$ and $y(j)\in M_y$, so $x(j)\neq j$, $y(j)\neq j$, $x(j)\in S$ and $y(j)\in S$. With similar argument, if $j\in M_y$ then $x(j)\in S$ and $y(j)\in S$. 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. We are going to show that z is strongly stable. In fact, that z is individually rational it is immediate from the individual rationality of x and y. To see that z does not have any weak blocking pair consider any pair $\{j,k\}$. The fact that x is strongly simple and y is strongly stable implies that $\{j,k\}$ does not weakly block z in the cases where $\{j,k\}\subseteq S$ and $\{j,k\}\subseteq N\text{-}S$. Then, without loss of generality, suppose $k\in S$ and $j\in N\text{-}S$. If $\{j,k\}$ weakly blocks z then $j\geq_k z(k)=x(k)$ and $k\geq_j z(j)=y(j)\geq_j x(j)$, 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 simple. Hence, in any case, $\{j,k\}$ does not weakly block z, so z is strongly stable. It remains to show that z extends x. We have that $z(j)=x(j)$ for every $j\in S$ and $z(j)=y(j)\geq_j x(j)$ for every $j\notin S$. Furthermore, $x\neq z$ due to the fact that x is not strongly stable and z is. Hence, $z(j)\geq_j x(j)$

for every *j*, with strict preference for at least one $j \in N$. Then, z extends x and the proof is complete.■

When preferences are strict, Proposition 3 can be restated as follows:
Let x be an unstable and simple matching for the Roommate model. If the set of stable matchings is non-empty, then there exists a stable matching z that extends x.
We can now prove Theorem 1.

Proof of Theorem 1. By Remark 1, if a matching x is Pareto-optimal and strongly stable then it is a Pareto-optimal strongly simple matching. In the other direction, by Proposition 3, if a matching x is Pareto-optimal strongly simple then it is in the strong core, for otherwise it has a strongly simple extension, which is a contradiction. By Proposition 1 we have that x is Pareto-optimal , so x is Pareto-strongly stable. ■

Example 1 illustrates that when preferences need not be strict we may have unstable and simple matchings for the Roommate model which do not have a simple extension. Unlike the Roommate model, every unstable and simple matching for the Marriage model can be extended to a simple matching under any kind of preferences. This is due to the fact that the core of any Marriage market, under any kind of preferences, is non-empty.

Proposition 4. *Consider the Marriage market (M,W,P). Let x be an unstable and simple matching. Then x can be extended to a simple matching.*
Proof. Since the core of the Marriage model is always non-empty, if preferences are strict then the result follows from Proposition 3. When preferences are non-necessarily strict, let A be the set of unmatched agents at x. Let y be some stable matching for the Marriage market restricted to A (it exists because every Marriage market has a non-empty core). Now construct the matching z which agrees with y on A and with x for the remaining agents. This matching is well defined and it is clearly individually rational. We claim that z is stable. In fact, to see that z does not have any blocking pair, take any pair $(m,w) \in MxW$. The fact that x is simple and y is stable implies that (m,w) does not

block z in the cases where $(m,w) \subseteq N\text{-}A$ and $(m,w) \subseteq A$. Then, without loss of generality, suppose $m \in N\text{-}A$ and $w \in A$. If (m,w) blocks z then $w >_m z(m)=x(m)$ and $m >_w z(w)=y(w) \geq_w x(w)$, so (m,w) blocks x. However, m is matched at x, which contradicts the fact that x is simple. Hence, in any case, (m,w) does not block z, so z is stable. It remains to show that z extends x. We have that $z(j)=x(j)$ for every $j \in N\text{-}A$ and $z(j)=y(j) \geq_j x(j)$ for every $j \in A$. Furthermore, $x \neq z$ due to the fact that x is not stable and z is. Hence, $z(j) \geq_j x(j)$ for every j, with strictly inequality for at least one $j \in N$. Then, z extends x and the proof is complete.■

We now can present a proof for Theorem 2.

Proof of Theorem 2. By Remark 1, if a matching x is Pareto-optimal and stable then it is a Pareto-optimal simple matching. In the other direction, by Proposition 4, if a matching x is Pareto-optimal simple then it is in the core, for otherwise it has a simple extension, which is a contradiction. By Remark 1 we have that x is Pareto-optimal , so x is Pareto- stable.■

6. PROPERTIES OF THE STABLE MATCHINGS

In this section we present some new properties of the stable matchings. Although these properties constitute well-known results for the Marriage-market when preferences are strict, they are also new for this market when preferences need not be strict.

The following result reflects an opposition of interests between the players involved in a partnership regarding two strongly stable matchings:

Theorem 3. *Let x and y be strongly stable matchings for the Roommate model (N,P). If j prefers x to y then $k=x(j) \neq j$, for some k, and $h=y(j) \neq j$, for some $h \neq k$. Furthermore, both k and h prefer y to x.*

Proof. If $x(j) >_j y(j)$ then $j \in M_x$. It follows from Lemma 1 that j is matched under x and under y and both mates belong to M_y. ■

Since the set of stable matchings equals the set of strongly stable matchings when preferences are strict, Theorem 3 holds, in this case, if x and y are stable matchings.

The following corollary asserts that there is an opposition of interests between the two sides of the Marriage market along the whole set of strongly stable matchings.[3]

Corollary 1. *Consider the Marriage market (M,W,P). Let x and y be strongly stable matchings. Then, all men like x at least as well as y if and only if all women like y at least as well as x. That is, $x>_M y$ if and only if $y>_W x$.*

Proof. Suppose $x>_M y$. Then, no man prefers y to x. If there is some woman w such that $x(w)>_w y(w)$ then $x(w)=m$ for some $m\in M$ and m prefers y to x by Theorem 3, contradiction. Hence, no woman prefers x to y and so $y>_W x$. Similarly we show the other direction.■

It follows from Lemma 1 that, if preferences are strict, we can regard the set of trading agents at a simple outcome as a sort of stable coalition in the sense that trading agents at a simple matching always make their transactions under a stable matching within the same pool. In particular, the set of matched agents under a stable matching is the same under any stable matching. If the preferences are non-necessarily strict, given two strongly stable matchings, x and y, the trading agents at x who are not indifferent between the two outcomes, trade among themselves at y. Consequently, an unmatched agent under y either is also unmatched under x or is indifferent between his/her mate under x and being unmatched. Formally,

Theorem 4. *Let x and y be strongly stable matchings for the Roommate model (N,P). If $j\in N$ is unmatched under y then j is indifferent between $x(j)$ and being unmatched.*[4]

When preferences are strict we can state:

[3] Knuth (1976) proved this result for the case where preferences are strict.

[4] This result was proved for the Marriage market with strict preferences by Gale and Sotomayor (1985a). A different proof is provided by McVitie and Wilson (1970) for the particular case where all men and women are mutually acceptable.

Let x and y be stable matchings for the Roommate model (N,P). If $j \in N$ is unmatched under y then j is unmatched under x.

7. FINAL REMARKS AND RELATED WORK

This paper explores the following simple idea. Basically we see four sets, each a superset of the next: The weak Pareto frontier, the core, the set of Pareto-stable outcomes and the strong core. The core is always a subset of the set of weak Pareto-optimal outcomes, understood as those allocations for which the grand coalition cannot strongly block, i.e., the grand coalition does not cause a strict Pareto improvement. We proved that when preferences are strict the three last sets coincide. With indifferences, these three sets open up and allow for this intermediate solution concept between the strong core and the core.

From the technical point of view, the characterization of Pareto-stable matchings provided here in terms of Pareto simple matchings informs us about the role of unmatched versus matched agents. On the conceptual point of view, we argued that in a decentralized setting, recontracts between pairs of agents already allocated according to a stable matching leading to a (weak) Pareto improvement of the original matching should be allowed. This justifies Pareto-stability.

Some of the results already proved for the Marriage model with strict preferences were proved here to hold for the Roommate model. If the preferences are non-strict these properties are true when restricted to the strong core. Several other results can be proved for the Marriage model with non-strict preferences. For instance, when preferences over potential partners are non-necessarily strict in the Marriage model we can represent them by separable cardinal utility functions. This way we have two partial orders defined in the set of stable payoffs (not matchings) as usual. We can then prove that the set of strongly stable payoffs for the Marriage model is a complete lattice. The proof follows the lines of the proof of Theorem 2 of Sotomayor (2000), conveniently adapted. Then we have the existence of a unique man-optimal and a unique woman-optimal strongly stable payoffs. (Of course we may not have uniqueness of the man-optimal strongly stable matching and of the woman-optimal strongly stable matching).

A version of the concept of simple matching as an individually rational matching where the woman involved in a blocking pair is always single was introduced in Sotomayor (1996) for the Marriage market. There, a non-constructive proof of the non-emptiness of the set of stable matchings is presented. The proof consists in demonstrating that any Pareto-optimal simple matching for the men must be stable. The same result is obtained by replacing women for men and vice-versa.

Similar concepts were introduced in Sotomayor (1999) for the discrete many-to-many matching market with substitutable and non-strict preferences and in Sotomayor (2000) for the continuous Assignment game of Shapley and Shubik and for the unified two-sided matching model of Eriksson and Karlander (2000). For all these two-sided matching models, a non-constructive and simple proof of the non-emptiness of the set of pairwise-stable outcomes has been obtained. The argument of these proofs is that the Pareto-optimal simple outcome for one of the sides must be pairwise-stable.

Recently, Sotomayor (2005) has introduced the concept of simple allocation for the one-sided market (not matching market) of Shapley and Scarf (1974). There, a non-constructive proof of the non-emptiness of the core has been obtained by proving that every simple allocation which is not in the core has a simple extension. Then every Pareto-optimal simple allocation must be in the core.

REFERENCES.

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) “Some remarks on the stable matching problem” *Discrete Applied Mathematics,* 11, 223-32.

Knuth, D. E., (1976) “Marriages Stables” *Montreal: Les Presses de l’Université de Montreal.*

McVitie, D. G. and Wilson L. B. (1970) “ Stable Marriage assignments for unequal sets”*,* *BIT,* 10, 295-309.

Roth, A. and Sotomayor, M (1990) “Two-sided matching. A study in game-theoretic modeling and analysis” *Eonometric Society Monographs,* 2nd ed., Vol. 18, Cambridge University Press.

Shapley, L. and Scarf, H. (1974) “On cores and indivisibility” *Journal of Mathematical Economics,* 1, 23-8.

Sotomayor, M. (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.