

## **IMPLEMENTATION OF STABLE SOLUTIONS TO MARRIAGE PROBLEMS\***

**José Alcalde\*\***

WP-AD 95-05

---

\* This work was partially supported by Grant PB 92-05909, from the DGICYT of the Spanish Ministry of Educación y Ciencia. This research was initiated while the author was visiting GREMAQ, Université des Sciences Sociales de Toulouse, whose hospitality is gratefully acknowledged. The author is grateful to Luis Corchón, Matthew Jackson, Hervé Moulin, David Pérez Castrillo and Eric Ralph for their comments. The many helpful discussions with Salvador Barberà were specially useful.

\*\* University of Alicante.

Editor: **Instituto Valenciano de  
Investigaciones Económicas, S.A.**  
Primera Edición Marzo 1995.  
ISBN: 84-482-0900-1  
Depósito Legal: V-1200-1995  
Impreso por Copisteria Sanchis, S.L.,  
Quart, 121-bajo, 46008-Valencia.  
Printed in Spain.

# Implementation of Stable Solutions to Marriage Problems

José Alcalde

## Abstract

This paper analyzes the possibility of implementing stable outcomes for marriage markets. Our first result shows a contradiction between the use of stable mechanisms and the hypothesis of agents' behaviour considered in the Nash equilibrium concept.

We analyze the possibility of implementing two sets of stable allocations, by employing two types of mechanisms. The first mechanism is a “now-or-never” choice process that permits us to implement in undominated Nash equilibria the set of all the stable allocations.

The second choice process is the classic algorithm in matching theory, the Gale-Shapley mechanism. A reversal property is observed in such a mechanism when agents act strategically. The use of a mechanism which selects the best solution for one side of the market in the absence of strategic behaviour yields the best stable solution for the agents on the other side under dominance solvability.

Keywords: Matching markets, Implementation, Stability.



## 1. Introduction

In this paper we study the possibility of designing mechanisms to implement stable solutions for marriage markets. A market is given by two finite, disjoint sets of agents and by the preferences of each agent over the individuals on the other side of the market. A one-to-one matching is an assignment of each agent on one side on the market to at most one agent on the other side. Marriage markets are those where outcomes are required to be one-to-one matchings. Among other interesting interpretations, we can think of them as models of the relationship between a set of workers looking for jobs and a set of firms, with one opening each.

A matching is stable if it is individually rational and will not be disrupted by coalitions (or partnerships) formed by one agent from each side of the market. This natural requirement for solutions to market allocation problems can always be satisfied: stable allocations always exist for marriage markets. In addition, stability turns out to be equivalent to coalitional rationality: the set of stable matchings for marriage markets coincides with the core of its associated marriage game. For all these reasons, I am interested in allocational criteria that would select stable matchings for any given market.

I focus on mechanisms, rather than on specific markets. I look for functions that assign a matching to each possible market that can be formed by fixing the agents on each side, and letting their preferences vary. These are called matching mechanisms. We would like to find mechanisms that always select matchings which are stable relative to the preferences of agents. But this dependence of the outcome on the agents' preferences gives rise to incentive problems. Agents operating under a stable matching mechanism will typically be led to misrepresent their preferences, and to manipulate the allocation to their advantage (Alcalde and Barberà [2], Roth [13]). Since this strategic behaviour is unavoidable for general domains of preferences, I try to find mechanisms leading agents to choose strategies that result in stable matchings despite misrepresentation. In other words, I look for mechanisms that implement stable matchings.

The paper presents two results in this direction. First, it introduces a mechanism that will implement the full correspondence of stable matchings in undominated Nash equilibria. This mechanism could be attractive to a designer unwilling to impose a specific stable outcome, but still seeking to guarantee stability as the consequence of the interaction among agents. The second result is in a sense complementary to the first. Suppose now that the designer wants to obtain one specific stable outcome: the one that is most favorable to one of the sides of the market. I exhibit a mechanism that sophisticatedly implements the men's (or the women's) optimal stable matching. Moreover, I get the surprising

result that the designer who favors one side of the market by allowing it to attain its optimum through sophisticated behaviour could in fact claim that the mechanism is intended to benefit just the other side! (if used straightforwardly, that is without strategic misrepresentation).

It is important to stress that the present result cannot be obtained by applying well-known previous results on implementation. One cannot use Maskin's [9] theorem on Nash implementation, or Palfrey and Srivastava's [12] results on implementation by undominated Nash equilibria, because stable social choice correspondences in marriage markets endow each player with veto power. Another sufficient condition for implementability is separability, as used by Jackson et al. [8]: again, it is not satisfied in our context.

An interesting feature on the mechanisms I present is that they constitute reasonable proposals for effective design. Many of the mechanisms suggested in the implementation literature are useful tools to show the possibility of implementing some solution, but does hardly qualify as a set of rules that a designer would be ready to actually propose for practical use. Although it is hard to provide a formal definition for a "reasonable" mechanism, I find several reasons why the ones presented here should qualify. The messages required from agents are simple and refer to their preferences (we could describe the procedures as direct mechanisms). Agents are endowed with a finite number of strategies. This makes the use of the present solution concepts more palatable than they would be under more complex settings.

The existing literature in Matching Theory does not say much about full implementation of stable allocations. Most of the authors have focused on the study of the classical Gale-Shapley mechanism. Roth [14] studies this matching rule and shows the stability of its undominated Nash equilibria. But he does not make any other prediction on which outcomes can be reached. Gale and Sotomayor [7] show that there is a specific strong equilibrium in the Gale-Shapley rule which satisfies a reversal property. A similar result is obtained in Demange and Gale [3] for the assignment problem. But those results do not ensure that this property is satisfied by the outcome of every strong equilibrium. In fact, a result ensuring this property for every strong equilibrium seems to be impossible. This is because the core of every cooperative game can be implemented in strong equilibria by means of a "related" noncooperative game, as proved by Moulin [11].

The result I present relative to the Gale-Shapley mechanism contributes in two ways to a game-theoretical analysis of this rule. First, the equilibrium concept I employ is dominance solvability. This does not require coordination by a set of agents, as in the strong equilibrium case. Second, the equilibrium is unique in outcomes. This result permits us to make some predictions about what would happen when the Gale-Shapley mechanism is used: As Moulin [11] said, "at any fixed profile, dominance solvability is a generalization of strategy-proofness."

The rest of the paper is organized as follows. Section 2 presents the model

and introduces some standard concepts of implementation. Section 3 presents a preliminary result on the Nash equilibria of revelation games induced by stable matching games. Section 4 introduces a mechanism implementing the full correspondence of stable matchings in undominated Nash equilibria. Section 5 analyzes the sophisticated equilibria of mechanisms whose straightforward result would be to choose the stable matching which is optimal for one side, and proves that this mechanism would actually implement the opposite side's optimal stable matching under this equilibrium concept. Section 6 concludes.

## 2. Preliminary definitions

Let  $M = \{m_1, \dots, m_i, \dots, m_r\}$  be the set of men and  $W = \{w_1, \dots, w_j, \dots, w_s\}$  be the set of women. These sets are kept fixed throughout the paper. Each  $m_i$  in  $M$  is endowed with a complete transitive asymmetric preference relation  $P(m_i)$ , on  $W \cup \{m_i\}$ .  $P(m_i)$  will stand for the preferences of  $m_i$  over all his possible mates, including the possibility of remaining single. Similarly, each  $w_j$  in  $W$  is endowed with a complete transitive asymmetric preference relation  $P(w_j)$  over  $M \cup \{w_j\}$ . We say that  $m_i$  (resp.  $w_j$ ) is acceptable to  $w_j$  under  $P(w_j)$  (resp.  $m_i$ ,  $P(m_i)$ ) if  $m_i P(w_j) w_j$  (resp.  $w_j P(m_i) m_i$ ). An agent  $x$  who is not acceptable to another  $y$  on the opposite side on the market is said unacceptable to her.

We denote by  $\mathcal{P}(x)$  the set of all possible preferences for individual  $x$  in  $M \cup W$ . Elements of  $\tilde{\mathcal{P}} = \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j)$  are called preference profiles. We denote preference profiles by  $\tilde{P}$ . Given a preference profile  $\tilde{P}$ , we denote by  $\tilde{P}/P'(m_i)$  (resp.  $\tilde{P}/P'(w_j)$ ) the profile obtained from  $\tilde{P}$  by changing the preferences of man  $m_i$  (resp. woman  $w_j$ ) from  $P(m_i)$  to  $P'(m_i)$  (resp.  $P(w_j)$ ,  $P'(w_j)$ ), and keeping all other preferences unchanged.

An outcome for a marriage market  $\{M, W; \tilde{P}\}$ , also called a matching, is a function  $\mu$  which assigns to each agent on a side on the market at most one mate who belongs to the opposite side. A matching  $\mu$  in which no agent prefers to be unmatched rather than her mate is said individually rational. A pair whose agents prefer to be matched together rather than to their mates by  $\mu$  is called a blocking pair for this matching. Finally, a matching which is individually rational and does not have any blocking pair is called stable.

Since the set of stable marriage markets coincides with the core of such a "marriage game" for any preference profile, we want to analyze the possibility of implementing stable matchings. That is, we are looking for decentralized mechanisms yielding stable solutions under reasonable assumptions about the strategic behaviour of agents. These rules will be presented as direct mechanisms, in which any agent declares a preference relation, which may differ from its own preferences. The application of the rule will select a matching for our market.

We analyze the possibility of implementing two solution concepts. First, the correspondence of stable matchings; and, second, the M-optimal stable matching (and, by symmetry between both sides on the market, the W-optimal stable matching.) Two rules to be defined will be used to implement these solution concepts, each one by means of an equilibrium concept: undominated Nash equilibria and dominance solvability.

Let  $\Gamma(\varphi) = \left[ \varphi; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P} \right]$  denote a matching game, where  $\varphi$ , the matching rule, is a function that selects a matching for each marriage market  $\{M, W; \tilde{P}'\}$ ;  $\mathcal{P}(x)$  is the strategy space for agent  $x$ ; and  $\{M, W; \tilde{P}\}$  is a marriage market.

A vector of strategies  $\tilde{P}'$  is a Nash equilibrium for  $\Gamma$  when, given the strategies of others, no agent can benefit from using another strategy.

A strategy  $P''(x)$  for agent  $x$  dominates another strategy  $P'(x)$  in the game  $\Gamma(\varphi)$  if the use of  $P''(x)$  always guarantees to him at least as utility as the use of  $P'(x)$  do, and sometimes a higher utility level.  $P'(x)$  is called an undominated strategy for  $x$  when there does not exist any other strategy dominating it. Finally, a strategy which dominates any other strategy is called dominant.

A Nash equilibrium where each agent employs undominated strategies is called an undominated Nash equilibrium (UNE).

Given a matching game  $\Gamma(\varphi) = \left[ \varphi; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P} \right]$ , let  $\Gamma^1(\varphi) = \left[ \varphi; \prod_{i=1}^r \mathcal{P}^1(m_i) \times \prod_{j=1}^s \mathcal{P}^1(w_j); \tilde{P} \right]$  denote the subgame where each agent's strategy set is modified in such a way that dominated strategies are removed from her (original) strategy space. An iterative argument yields to define  $\Gamma^k(\varphi)$  as the subgame where each agent's dominated strategies on  $\Gamma^{k-1}(\varphi)$  have been removed from her initial strategy set. We say that a matching mechanism  $\varphi$  is dominance solvable whenever  $\varphi\left(\prod_{i=1}^r \mathcal{P}^{\bar{k}}(m_i) \times \prod_{j=1}^s \mathcal{P}^{\bar{k}}(w_j)\right)$  is a singleton for some  $\bar{k}$ .

Let  $\Psi$  be a solution concept for marriage markets. We say that mechanism  $\varphi$  implements  $\Psi$  in (undominated) Nash equilibria if (i) for each (undominated) Nash equilibrium,  $\tilde{P}'$ ,  $\varphi(\tilde{P}')$  satisfies property  $\Psi$ , and (ii) for each matching  $\mu$  satisfying this property, there exists an (undominated) Nash equilibrium  $\tilde{P}'$  such that  $\varphi(\tilde{P}') = \mu$ .

Let  $\Psi$  be a solution concept for marriage markets. We say that mechanism  $\varphi$  sophisticatedly implements  $\Psi$  if  $\varphi$  is dominance solvable and  $\Psi \equiv \varphi(\tilde{P}')$ , the (unique) outcome obtained from  $\varphi$ 's dominance solvability.



### 3. Implementing the set of individually rational matchings

This section motivates the need to be specific about the type of equilibrium under which to study the strategic behaviour of agents. When rationality of the economic agents is postulated, it is paradoxical to propose (or to reject) a mechanism on the basis of properties which arise from the use of dominated strategies. The key question is, then, what kind of properties are satisfied when undominated strategies are employed by the agents.

Proposition 3.1 presents a quite negative result: any stable mechanism implements, in Nash equilibria, the correspondence of individually rational matchings. Consider a matching where no worker was employed. It is always individually rational. This is even true if, as in the original Gale and Shapley [6] model, every worker prefers to be employed rather than unemployed and every firm also prefers to cover each position rather than to have any vacancy. Thus Proposition 3.1 shows that Nash equilibria is a very weak equilibrium concept when stable rules are employed.

**Proposition 3.1.** *Let  $\varphi$  be a direct mechanism choosing a stable matching for each marriage market. Then, for each market,  $\varphi$  implements in Nash equilibria the correspondence of individually rational matchings.*

*Proof.* Let  $\{M, W; \tilde{P}\}$  be a marriage market and  $\varphi$  a stable rule. It can be easily checked that a matching which is not individually rational cannot be supported by any Nash equilibrium.

On the other hand, we can see that each individually rational matching can be supported by a Nash equilibrium. Let  $\mu$  be an individually rational matching for  $\{M, W; \tilde{P}\}$ . Define  $\tilde{P}'$  so that for each  $m_i$  in  $M$  and  $w_j$  in  $W$

$$(1) \quad w_j P'(m_i) m_i \iff w_j = \mu(m_i), \quad \text{and}$$

$$(2) \quad m_i P'(w_j) w_j \iff m_i = \mu(w_j)$$

The reader can easily check that  $\mu$  is the only stable matching for  $\{M, W; \tilde{P}'\}$ , and  $\tilde{P}'$  is a Nash equilibrium for the game induced by  $\varphi$ . ■

The negative result presented in Proposition 3.1 is due to the combination of two facts. One is that we always choose stable outcomes and the other is that we require Nash behaviour for all agents. This weakness is disclosed by Example 3.2.

We present a marriage market and analyze a particular type of behaviour on the part of the agents, when the M-optimal stable rule is employed. We study the set of allocations that can be attained as the outcome for some Nash equilibrium in undominated strategies. There is a stable allocation that cannot be obtained as the outcome for any UNE. So, this example shows how the selection of a specific stable mechanism influences the set of allocations it implements.

**Example 3.2.** Let  $\{M, W; \tilde{P}\}$  be a marriage market, in which  $M = \{m_1, m_2\}$ ,  $W = \{w_1, w_2\}$  and agents preferences are defined as follows:

$$P(m_1) = w_1 \ w_2 \ m_1 \qquad P(w_1) = m_2 \ m_1 \ w_1$$

$$P(m_2) = w_2 \ w_1 \ m_2 \qquad P(w_2) = m_1 \ m_2 \ w_2$$

Let  $\varphi^M(\tilde{P}')$  be the mechanism leading to the M-optimal stable matching for  $\{M, W; \tilde{P}'\}$ .<sup>1</sup>

In this example there are two stable matchings, but only one of these can be the outcome associated with any UNE for  $\left[\varphi^M; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$ . We can easily check it with the help of two well-known properties satisfied when this mechanism applies:<sup>2</sup>

- (i) for each  $m_i$  in M, it is a dominant strategy of this game to declare his true preferences,  $P(m_i)$ , and
- (ii) for each  $w_j$  in W, if  $P'(w_j)$  is an undominated strategy, then the  $P(w_j)$ -maximal must be maximal under  $P'(w_j)$ .

Therefore, at any  $\tilde{P}'$  UNE for this game, the payoff should satisfy that  $\varphi^M(\tilde{P}') = \mu'$ , where  $\mu'(m_1) = w_2$  and  $\mu'(m_2) = w_1$ . Because of the stability of  $\mu$  (where  $\mu(m_1) = w_1$  and  $\mu(m_2) = w_2$ ) with respect to  $\{M, W; \tilde{P}\}$ , we conclude that the Gale-Shapley mechanism does not implement the full correspondence of stable matchings.

Notice that this argument applies in a more general context, namely the set of marriage markets satisfying the following two properties:

---

<sup>1</sup>The Gale-Shapley mechanism is presented in an algorithmic way: each man makes a proposal to his most favored woman. Women who receive more than one proposal reject all but their most favored man from among those which have made a proposal. Rejected men then make a proposal to their most favored woman, who accepts it if she has not received better proposals; and so on. This mechanism stops when no man is rejected by the woman he proposes to.

For more details, the interested reader is addressed to the original article by Gale and Shapley [6]. The book by Roth and Sotomayor [15] also contains some examples of the execution of the algorithm.

<sup>2</sup>These properties are deduced in Gale and Sotomayor [7].

- (i) there are, at least, two stable matchings for this market, and
- (ii) there is an individually rational matching, in which any  $w_j$  in  $W$  is matched to her maximal element for her true preferences  $P(w_j)$ .

## 4. Implementing the correspondence of stable matchings

In this section, we introduce a mechanism that implements, for each marriage market, the correspondence of its stable matchings. We know, from Proposition 3.1, that either this mechanism will select unstable matchings, or else we should be using an equilibrium concept stronger than Nash's. Our result is based on both types of relaxation. We present a matching rule which can select unstable allocations. But when agents employ undominated strategies, the Nash equilibrium outcomes will be stable with respect to the agents' true preferences.

The algorithm we propose is similar, in some aspects, to that of Gale and Shapley [6]. Agents on one side on the market make the proposals, and those who are on the other side evaluate them.

The optimal algorithms analyzed by these authors allow players to reject proposals at any step and remain in the market. We propose a "now-or-never" mechanism, where an agent who has received a proposal once cannot receive subsequent proposals. In contrast with the Gale-Shapley optimal algorithms, our mechanism implements the correspondence of stable matchings in undominated Nash equilibria. (See Example 3.2.)

**Definition 4.1.** *The mechanism  $\varphi^{IR}$  associates the outcome of the following algorithm to each marriage market  $\{M, W; \tilde{P}\}$ :<sup>3</sup>*

*Each agent  $m_i$  in  $M$  makes a proposal to his maximal element under  $P(m_i)$ . Each agent  $w_j$  in  $W$  who receives any proposal is matched with her  $P(w_j)$ -maximal element on the set of agents who propose to her and herself. Matched agents are removed from the market. Each man remaining in the market make a proposal to his most preferred mate between those remaining in the market. Women who receive any proposal evaluate it as in the previous step and leave the market matched with their mate, and so on. This algorithm stops when no woman receives any proposal. Agents who are still unmatched, if any, are selfmatched.*

---

<sup>3</sup>Alcalde [1] presents a more detailed version of this algorithm and illustrates the workings of this mechanism by means of an example.

**Remark 1.** Notice that, in general, when  $\varphi^{IR}$  applies to  $\{M, W; \tilde{P}\}$ , we do not obtain stable matchings for this market. In fact, efficiency cannot be guaranteed either.

The remainder of the section analyzes some properties for the game form with the payoff function generated by our algorithm. Propositions 4.2 and 4.5 characterize the set of undominated strategies for the game  $\left[\varphi^{IR}; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$ . It is a dominant strategy for each agent in  $W$  to declare her true preferences. Moreover, for each  $m_i$  in  $M$ , a strategy  $P'(m_i)$  is dominated for this game if, and only if, the set of acceptable individuals for this strategy is the same as for his true preferences  $P(m_i)$ .<sup>4</sup>

**Proposition 4.2.** At game  $\left[\varphi^{IR}; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$ , it is a dominant strategy, for each  $w_j$  in  $W$ , to state her true preferences,  $P(w_j)$ .

The proof of this proposition is straightforward. Observe that in the unique step in which  $w_j$  is matched, her mate is the maximal for the stated preferences among the proposals she received.

Lemmas 4.3 and 4.4 will be useful to prove Proposition 4.5.

**Lemma 4.3.** Let  $m_i \in M$ .  $P'(m_i)$  is a dominated strategy for  $m_i$  at game  $\left[\varphi^{IR}; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$  if there is some  $w_j$  in  $W$  which is acceptable for  $m_i$  under  $P'(m_i)$  but unacceptable under  $P(m_i)$ .

**Lemma 4.4.** Let  $m_i \in M$ .  $P'(m_i)$  is a dominated strategy for  $m_i$  at game  $\left[\varphi^{IR}; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$  if there is some  $w_j$  in  $W$  which is acceptable for  $m_i$  under  $P(m_i)$  but unacceptable under  $P'(m_i)$ .

**Proposition 4.5.** Suppose that  $\#M \geq 2$ . At game  $\left[\varphi^{IR}; \prod_{i=1}^r \mathcal{P}(m_i) \times \prod_{j=1}^s \mathcal{P}(w_j); \tilde{P}\right]$ ,  $P'(m_i)$  is an undominated strategy for  $m_i \in M$  if, and only if, the sets of acceptable women for  $m_i$  under  $P(m_i)$  and  $P'(m_i)$  coincide:

$$\forall w_j \in W, \quad w_j P'(m_i) m_i \iff w_j P(m_i) m_i \quad (4.1)$$

---

<sup>4</sup>Throughout this paper, we will denote by  $P(x)$  the true preferences of agent  $x$ , and by  $P'(x)$ ,  $P^2(x)$ , etc., we mean alternative preferences for agent  $x$  other than  $P(x)$ .

*Proof.* The necessity condition follows directly from Lemmas 4.3 and 4.4 above. To prove sufficiency, consider two preferences for  $m_i$ ,  $P'(m_i)$  and  $P^2(m_i)$ . Suppose that  $P'(m_i)$  satisfies Condition (4.1) and is dominated by  $P^2(m_i)$ . Without loss of generality, suppose that  $P^2(m_i)$  also satisfies (4.1). One can easily check that the set of  $m_i$ 's acceptable women under  $P(m_i)$  has to have at least two elements.

Let  $w' = \operatorname{argmax}_{\{w \in W\}} P^2(m_i)$  and suppose that agents other than  $m$  play strategies  $P'(x)$ , satisfying:

- (i)  $w' P'(w') m_i$ ,
- (ii)  $\exists m'$  in  $M \setminus \{m_i\}$  s.t.  $\forall w \in W \setminus \{w'\} m_i P'(w) w P'(w) m'$ ,
- (iii)  $w P'(m') m'$  and  $\{w'' \in W : w'' P^2(m_i) w\} = \{w'' \in W : w'' P'(m_i) w\} \cup \{w'\}$  for every  $w \in W \setminus w'$  such that  $w P(m_i) m_i$ , and
- (iv)  $\forall m'' \in M \setminus \{m_i, m'\}$ ,  $m'' P'(m'') w$  for all  $w$  in  $W$ .

Let  $\mu' = \varphi^{IR}(\tilde{P}')$ , and  $\mu^2 = \varphi^{IR}(\tilde{P}'/P^2(m_i))$ . In this case  $\mu^2(m_i) = m_i$ . Since  $P'(m_i) \neq P^2(m_i)$ , and  $\{w \in W : w P'(m_i) m_i\} = \{w \in W : w P^2(m_i) m_i\}$  there will be  $w \in W$  such that

$$\{w'' \in W : w'' P'(m_i) w\} = \{w'' \in W : w'' P'(m') w\} \cup \{w'\} \quad (4.2)$$

Then  $\mu'(m_i) = \max_{P'(m_i)} \{w \in W : w \text{ satisfies (4.2)}\}$ . This fact contradicts the domination of  $P'(m_i)$  by  $P^2(m_i)$ . ■

**Remark 2.** When  $M$  is a singleton  $\varphi^{IR}$  coincides with  $\varphi^M$ . So,  $P(m_i)$  should be a dominant strategy for  $m_i$ , as Dubins and Freedman [4] proved.

We can now introduce the main result for this section.

**Theorem 4.6.** *The mechanism  $\varphi^{IR}$  implements the correspondence of stable matchings in undominated Nash equilibria for each marriage market.*

*Proof.* When  $M$  is a singleton, because of Proposition 4.2 and Remark 2, the proof is trivial. Let's consider  $\#M \geq 2$ .

Let  $\tilde{P}'$  be an undominated Nash equilibrium for this mechanism. Because of Propositions 4.2 and 4.5,  $\varphi^{IR}(\tilde{P}')$  should be individually rational for  $\{M, W; \tilde{P}'\}$ . Then, if  $\varphi^{IR}(\tilde{P}')$  is unstable for this market, there should be a pair  $(m_i, w_j)$  blocking it. Since  $P'(w_j) = P(w_j)$ ,  $m_j$  can improve his utility by playing a strategy for which  $w_j$  is the maximal element, which contradicts that  $\tilde{P}'$  is an undominated Nash equilibrium.

On the other hand, let  $\mu$  be a stable matching for  $\{M, W; \tilde{P}'\}$ . Consider the strategy profile  $\tilde{P}'$  satisfying:

- (i)  $P'(w_j) = P(w_j)$  for all  $w_j$  in  $W$ ,
- (ii) for each  $m_i$  in  $M$ , the sets of acceptable women under  $P'(w_j)$  and  $P(w_j)$  coincide, and
- (iii) for each  $m_i$  in  $M$  such that  $\mu(m_i) \neq m_i$ ,  $\mu(m_i)$  is the  $P'(m_i)$ -maximal element.

It can straightforwardly be verified that  $\tilde{P}'$  is a Nash equilibrium for  $\varphi^{IR}$  which satisfies that  $\varphi^{IR}(\tilde{P}') = \mu$ . To conclude, remember that  $P'(x)$  is an undominated strategy for any agent  $x$  because of Propositions 4.2 and 4.5. ■

## 5. Reversal properties of the Gale-Shapley mechanism

Example 3.2 shows a reversal property satisfied by the Gale-Shapley mechanism, when agents behaviour is strategic. If the men's (resp. women's) optimal rule applies, the agents' strategic behaviour can lead us to obtain the women's (resp. men's) optimal stable allocation. Following our example, we argued that this previous argument can be straightforwardly extended to a family of marriage markets. A natural question arises: When such mechanisms apply, could we expect similar properties in any market? More precisely, does there exist a natural extension of the individual behaviour used in this example allowing us to generalize this property? This section provides affirmative answers to these questions.

In Section 2, we presented dominance solvability as a natural extension to the use of undominated strategies by the agents. Such a "generalization" builds on the assumption that all agents share the same beliefs. The main requirement of dominance solvability is the uniqueness of the outcome when dominated strategies are iteratively removed.

In this section we show that there exist dominance solvable mechanisms selecting stable outcomes. In particular, we present a mechanism which sophisticatedly implements the men's optimal matching as its outcome. The mechanism we analyze is very simple. Each agent,  $x$ , in the market declares her preferences,  $P'(x)$ , and the women's optimal algorithm applies to the market  $\{M, W; \tilde{P}'\}$ . Observe that preference stated by the agents can be different to their true preferences.

Our results strengthen those presented in Thomson [16] about the reversal of asymmetric allocations when agents' behaviour is strategic. Results by Demange and Gale [3], and by Gale and Sotomayor [7] are similar.  $W$ -optimal stable matching can be obtained as the outcome of a family of strong equilibria for the  $M$ -optimal stable matching game for each marriage market.

An important difference between these results and those I present is the equilibrium concept employed. The strong equilibria they studied require a very high

level of communication among the agents on one side on the market. Dominance solvability, the solution concept I employ, does not require such a collusive behaviour: this is a direct consequence of the uniqueness of the equilibria outcomes. (See, for example, Moulin [10].) There is another qualitative difference: our result ensures that this reversion always occurs. This cannot be predicted from Demange and Gale [3] result, nor from that by Gale and Sotomayor [7].

**Theorem 5.1.** *Let  $\varphi^W$  be the mechanism selecting the  $W$ -stable matching for  $\{M, W; \tilde{P}'\}$ , for each preference profile  $\tilde{P}'$ . Then,*

- (i)  $\varphi^W$  is dominance solvable, and
- (ii) for each marriage market  $\{M, W; \tilde{P}\}$ ,  $\varphi^W$  sophisticatedly implements its  $M$ -optimal stable outcome.

The proof of this result is presented, in a constructive way, by Lemmas 5.2 to 5.5. The first one is a classical result in this literature, proved in Dubins and Freedman [4]. Lemma 5.3 presents a condition that must be satisfied by the preferences stated by agents in  $M$  in order to be undominated strategies for the game related to  $\varphi^W$ . Lemma 5.4 generalizes it and introduces the “iterative argument” of dominance solvability. Finally, Lemma 5.5 establishes part (ii) of our Theorem 5.1.

**Lemma 5.2. (Dubins and Freedman [4].)**

*When mechanism  $\varphi^W$  applies, it is a dominant strategy for any agent in  $W$  to declare her true preferences.*

**Lemma 5.3. (Roth [14].)**

*Any strategy  $P'(m_i)$  is dominated for agent  $m_i$  in the game induced by  $\varphi^W$  if the maximal under  $P(m_i)$  and  $P'(m_i)$  do not coincide.*

The set of undominated strategies for this mechanism,  $\tilde{\mathcal{P}}^1$  is the set of preference profiles such that

- (i) for all  $w_j$  in  $W$ ,  $P'(w_j) = P(w_j)$ , and,
- (ii) maximal elements for  $P'(m_i)$  and  $P(m_i)$  coincide for each  $m_i$  in  $M$ .

In a recursive way we define, for each  $k \geq 2$ ,  $\tilde{\mathcal{P}}^k$  as the set of profiles in  $\tilde{\mathcal{P}}^{k-1}$  such that, for each  $m_i$  in  $M$  Condition (k-1) holds:

$$\forall w_j \in W, w_j \text{ } P'(m_i) \max_{P(m_i)} \mathcal{W}^{k-1}(m_i) \text{ implies } w_j \notin \mathcal{W}^{k-1}(m_i) \quad (\text{k-1})$$

where  $\mathcal{W}^{k-1}(m_i) = \{x \in W \cup \{m_i\} : \varphi_{m_i}^W(\tilde{P}') = x \text{ for some } \tilde{P}' \text{ in } \tilde{\mathcal{P}}^{k-1}\}$

**Lemma 5.4.** *If  $P'(m_i)$  is an undominated strategy for agent  $m_i$  in the game  $[\varphi^W; \tilde{\mathcal{P}}^{k-1}; \tilde{P}]$  then  $P'(m_i)$  would satisfy Condition (k):*

$$\forall w_j \in W, w_j P'(m_i) \max_{P(m_i)} \mathcal{W}^k(m_i) \text{ implies } w_j \notin \mathcal{W}^k(m_i) \quad (\text{k})$$

*Proof.* To prove the lemma, suppose that  $P'(m_i)$  does not satisfy Condition (k). Then, there should be  $w_j$  in  $\mathcal{W}^k(m_i)$ , such that

- (i)  $m_i P'(w_j) w_j$ , and,
- (ii)  $w^k P(m_i) w_j$  and  $w_j P'(m_i) w^k$ , where  $w^k = \max_{P(m_i)} \mathcal{W}^k(m_i)$ .

. Let  $P^2(m_j)$  be a strategy for  $m$  satisfying Condition (k-1), such that

- (a) for each  $w^2, w^3$  in  $\mathcal{W}^k(m_i) \setminus \{w^k\}$   $w^2 P^2(m_i) w^3 \iff w^2 P'(m_i) w^3$
- (b) for each  $w$  in  $\mathcal{W}^k(m_i) \setminus \{w^k\}$   $w^k P^2(m_i) w$

the reader can check that  $P^2(m_j)$  dominates  $P'(m_i)$ . ■

By the constructive definition of  $\tilde{\mathcal{P}}^k$ , we know that, for any  $1 \leq k' \leq k$ ,  $\tilde{\mathcal{P}}^{k'} \subseteq \tilde{\mathcal{P}}^k$ . Moreover, if we denote by  $\mu_M(m_i)$  the M-optimal mate for  $m_i$ , there should be a value  $k(m_i)$  such that

- (i)  $w^k = \mu_M(m_i)$  for all  $k \geq k(m_i)$ , and,
- (ii)  $w^{k(m_i)-1} \neq \mu_M(m_i)$ .

This result is a direct consequence of the latticial structure of the set of stable matchings.

Let  $k_0 = \max_{m_i \in M} k(m_i)$ . Each preference profile  $\tilde{P}'$  in  $\tilde{\mathcal{P}}^{k_0}$  should satisfy:

- (i) the M-optimal stable matchings for  $\{M, W; \tilde{P}\}$  and  $\{M, W; \tilde{P}'\}$  coincide,
- (ii)  $P'(w_j) = P(w_j)$  for each  $w_j$  in  $W$ , and,
- (iii) For each  $m_i$  in  $M$ , if  $\mu_M(m_i) P(m_i) w_j$  and  $w_j P'(m_i) \mu_M(m_i)$ , then there is no matching  $\mu$ , stable for  $\{M, W; \tilde{P}\}$  such that  $\mu(m_i) = w_j$ .

We can easily check that  $\tilde{\mathcal{P}}^{k_0}$  is non-empty. In fact,  $\tilde{P}$ , the profile of true preferences for the agents, belongs to  $\tilde{\mathcal{P}}^{k_0}$ .

**Lemma 5.5.** *When mechanism  $\varphi^W$  applies,  $P'(m_i)$  is an undominated strategy for agent  $m_i$  satisfying condition  $(k_0)$  if, and only if, for every preference profile  $\tilde{P}'$  in  $\tilde{\mathcal{P}}^{k_0}$ , and for each matching  $\mu$ , stable for  $\{M, W; \tilde{P}'\}$ ,  $\mu(m_i) = \mu_M(m_i)$ , the M-optimal stable matching for  $\{M, W; \tilde{P}\}$ .*



A proof of this lemma can be straightforwardly made by using the latticial structure of the stable matchings.

The proof for Theorem 5.1 follows from Lemma 5.5, concluding dominance solvability for  $\varphi^W$ .

## 6. Final comments

In this paper we analyzed some possible escapes to the negative results by Alcalde and Barberà [2] and Roth [13] for truthful revelation in matching markets. The potentially harmful effects of manipulation in stable mechanisms have been studied, and we can show that some of its most obvious negative consequences (like the lack of Pareto optimality or stability) may or may not occur, depending on the type of equilibrium concept and the mechanism under use. We studied the strategic behaviour of agents faced with two different mechanisms.

First, we provided a mechanism that is similar to the classical Gale-Shapley algorithm, but in which no deferred acceptance is allowed. The second mechanism we analyze is the Gale-Shapley matching algorithm, whose strategic properties have been studied by different authors for alternative equilibrium concepts.

Requiring the use of undominated strategies by the agents guarantees us the stability of the equilibrium outcomes for both mechanisms. In fact, the results of Theorems 4.6 and 5.1 remain also valid when the equilibrium concept we employ is that of Farquharson's [5] sophisticated equilibria.

Since we can state the validity of Theorems 4.6 and 5.1 for the case of sophisticated equilibria, we can now compare the mechanisms introduced in the paper. In the Gale-Shapley mechanism, the agents on one side of the market (those who receive the proposals) have the possibility of receiving proposals at any step of the algorithm. The analysis of strategic behaviour by the agents leads to a unique sophisticated equilibrium outcome. On the other hand, in the  $\varphi^{IR}$  mechanism once an offer is made, there is no chance to receive further proposals in later stages of the procedure. This leads to the possibility of several outcomes being supported as sophisticated equilibria.

This last observation allows us to compare both mechanisms from a "fairness" point of view. Indeed, consider a decision maker who must decide on a mechanism to be used for matching decisions in some job market. Suppose that this agent must select between the Gale-Shapley mechanism and the  $\varphi^{IR}$  rule. The use of the first mechanism implies a clear support to one side of the market and, consequently an unfavorable treatment to the other side. When  $\varphi^{IR}$  is employed this problem disappears.

There seems to be an inverse relation between the size of the set of stable outcomes supported by sophisticated equilibria and the number of steps in which an agent who receives proposals can reject them and remain in the market for

a mechanism à la Gale-Shapley. With some restriction in this ability to receive offers, as in the  $\varphi^{IR}$  mechanism, the size of the set of outcomes supported by sophisticated equilibria increases.

## References

- [1] J. Alcalde, *Matching: Stability and Incentives*, Ph.D. Dissertation, Universitat Autònoma de Barcelona, Barcelona, 1994.
- [2] J. Alcalde and S. Barberà, Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems, *Economic Theory* **4** (1994), 417-435.
- [3] G. Demange and D. Gale, The Strategy Structure of Two-Sided Matching Markets, *Econometrica* **53** (1985), 873-888.
- [4] L. Dubins and D.A. Freedman, Machiavelli and the Gale-Shapley Algorithm, *American Mathematical Monthly* **88** (1981), 485-494.
- [5] R. Farquharson, "Theory of Voting," Yale University Press, New Haven, 1969.
- [6] D. Gale and L.S. Shapley, College Admissions and the Stability of Marriage, *American Mathematical Monthly* **69** (1962), 9-15.
- [7] D. Gale and M. Sotomayor, Ms. Machiavelli and the Stable Matching Problem, *American Mathematical Monthly* **92** (1985), 261-268.
- [8] M.O. Jackson, T.R. Palfrey and S. Srivastava, Undominated Nash Implementation in Bounded Mechanisms, *Games and Economic Behaviour* **6** (1994), 474-501.
- [9] E. Maskin, *Nash Equilibrium and Welfare Optimality*, mimeo, M.I.T., 1977.
- [10] H. Moulin, Dominance Solvable Voting Schemes, *Econometrica* **47** (1979), 1337-1351.
- [11] H. Moulin, "The Strategy of Social Choice," North-Holland Publishing Company, Amsterdam, 1983.
- [12] T.R. Palfrey and S. Srivastava, Nash Implementation Using Undominated Strategies, *Econometrica* **59** (1991), 479-501.
- [13] A.E. Roth, The Economics of Matching: Stability and Incentives, *Mathematics of Operations Research* **7** (1982), 617-628.
- [14] A.E. Roth, Misrepresentation and Stability in the Marriage Problem, *Journal of Economic Theory* **34** (1984), 383-387.

- [15] A.E. Roth and M. Sotomayor, "Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis," Econometric Society Monograph Series, Cambridge University Press, New York, 1990.
- [16] W. Thomson, Reversal of Asymmetries of Allocation Mechanisms under Manipulation, *Economic Letters* **21** (1986), 227-230.
- [17] L. Zhou, Stable Matchings and Equilibrium Outcomes of the Gale-Shapley's Algorithm for the Marriage Problem, *Economic Letters* **36** (1991), 25-29.

## PUBLISHED ISSUES

- WP-AD 90-01 "Vector Mappings with Diagonal Images"  
C. Herrero, A. Villar. December 1990.
- WP-AD 90-02 "Langrangean Conditions for General Optimization Problems with Applications to Consumer Problems"  
J.M. Gutierrez, C. Herrero. December 1990.
- WP-AD 90-03 "Doubly Implementing the Ratio Correspondence with a 'Natural' Mechanism"  
L.C. Corchón, S. Wilkie. December 1990.
- WP-AD 90-04 "Monopoly Experimentation"  
L. Samuelson, L.S. Mirman, A. Urbano. December 1990.
- WP-AD 90-05 "Monopolistic Competition: Equilibrium and Optimality"  
L.C. Corchón. December 1990.
- WP-AD 91-01 "A Characterization of Acyclic Preferences on Countable Sets"  
C. Herrero, B. Subiza. May 1991.
- WP-AD 91-02 "First-Best, Second-Best and Principal-Agent Problems"  
J. Lopez-Cuñat, J.A. Silva. May 1991.
- WP-AD 91-03 "Market Equilibrium with Nonconvex Technologies"  
A. Villar. May 1991.
- WP-AD 91-04 "A Note on Tax Evasion"  
L.C. Corchón. June 1991.
- WP-AD 91-05 "Oligopolistic Competition Among Groups"  
L.C. Corchón. June 1991.
- WP-AD 91-06 "Mixed Pricing in Oligopoly with Consumer Switching Costs"  
A.J. Padilla. June 1991.
- WP-AD 91-07 "Duopoly Experimentation: Cournot and Bertrand Competition"  
M.D. Alepuz, A. Urbano. December 1991.
- WP-AD 91-08 "Competition and Culture in the Evolution of Economic Behavior: A Simple Example"  
F. Vega-Redondo. December 1991.
- WP-AD 91-09 "Fixed Price and Quality Signals"  
L.C. Corchón. December 1991.
- WP-AD 91-10 "Technological Change and Market Structure: An Evolutionary Approach"  
F. Vega-Redondo. December 1991.
- WP-AD 91-11 "A 'Classical' General Equilibrium Model"  
A. Villar. December 1991.
- WP-AD 91-12 "Robust Implementation under Alternative Information Structures"  
L.C. Corchón, I. Ortuño. December 1991.

- WP-AD 92-01 "Inspections in Models of Adverse Selection"  
I. Ortuño. May 1992.
- WP-AD 92-02 "A Note on the Equal-Loss Principle for Bargaining Problems"  
C. Herrero, M.C. Marco. May 1992.
- WP-AD 92-03 "Numerical Representation of Partial Orderings"  
C. Herrero, B. Subiza. July 1992.
- WP-AD 92-04 "Differentiability of the Value Function in Stochastic Models"  
A.M. Gallego. July 1992.
- WP-AD 92-05 "Individually Rational Equal Loss Principle for Bargaining Problems"  
C. Herrero, M.C. Marco. November 1992.
- WP-AD 92-06 "On the Non-Cooperative Foundations of Cooperative Bargaining"  
L.C. Corchón, K. Ritzberger. November 1992.
- WP-AD 92-07 "Maximal Elements of Non Necessarily Acyclic Binary Relations"  
J.E. Peris, B. Subiza. December 1992.
- WP-AD 92-08 "Non-Bayesian Learning Under Imprecise Perceptions"  
F. Vega-Redondo. December 1992.
- WP-AD 92-09 "Distribution of Income and Aggregation of Demand"  
F. Marhuenda. December 1992.
- WP-AD 92-10 "Multilevel Evolution in Games"  
J. Canals, F. Vega-Redondo. December 1992.
- WP-AD 93-01 "Introspection and Equilibrium Selection in 2x2 Matrix Games"  
G. Olcina, A. Urbano. May 1993.
- WP-AD 93-02 "Credible Implementation"  
B. Chakravorti, L. Corchón, S. Wilkie. May 1993.
- WP-AD 93-03 "A Characterization of the Extended Claim-Egalitarian Solution"  
M.C. Marco. May 1993.
- WP-AD 93-04 "Industrial Dynamics, Path-Dependence and Technological Change"  
F. Vega-Redondo. July 1993.
- WP-AD 93-05 "Shaping Long-Run Expectations in Problems of Coordination"  
F. Vega-Redondo. July 1993.
- WP-AD 93-06 "On the Generic Impossibility of Truthful Behavior: A Simple Approach"  
C. Beviá, L.C. Corchón. July 1993.
- WP-AD 93-07 "Cournot Oligopoly with 'Almost' Identical Convex Costs"  
N.S. Kukushkin. July 1993.
- WP-AD 93-08 "Comparative Statics for Market Games: The Strong Concavity Case"  
L.C. Corchón. July 1993.

- WP-AD 93-09 "Numerical Representation of Acyclic Preferences"  
B. Subiza. October 1993.
- WP-AD 93-10 "Dual Approaches to Utility"  
M. Browning. October 1993.
- WP-AD 93-11 "On the Evolution of Cooperation in General Games of Common Interest"  
F. Vega-Redondo. December 1993.
- WP-AD 93-12 "Divisionalization in Markets with Heterogeneous Goods"  
M. González-Maestre. December 1993.
- WP-AD 93-13 "Endogenous Reference Points and the Adjusted Proportional Solution for Bargaining Problems with Claims"  
C. Herrero. December 1993.
- WP-AD 94-01 "Equal Split Guarantee Solution in Economies with Indivisible Goods Consistency and Population Monotonicity"  
C. Beviá. March 1994.
- WP-AD 94-02 "Expectations, Drift and Volatility in Evolutionary Games"  
F. Vega-Redondo. March 1994.
- WP-AD 94-03 "Expectations, Institutions and Growth"  
F. Vega-Redondo. March 1994.
- WP-AD 94-04 "A Demand Function for Pseudotransitive Preferences"  
J.E. Peris, B. Subiza. March 1994.
- WP-AD 94-05 "Fair Allocation in a General Model with Indivisible Goods"  
C. Beviá. May 1994.
- WP-AD 94-06 "Honesty Versus Progressiveness in Income Tax Enforcement Problems"  
F. Marhuenda, I. Ortuño-Ortín. May 1994.
- WP-AD 94-07 "Existence and Efficiency of Equilibrium in Economies with Increasing Returns to Scale: An Exposition"  
A. Villar. May 1994.
- WP-AD 94-08 "Stability of Mixed Equilibria in Interactions Between Two Populations"  
A. Vasin. May 1994.
- WP-AD 94-09 "Imperfectly Competitive Markets, Trade Unions and Inflation: Do Imperfectly Competitive Markets Transmit More Inflation Than Perfectly Competitive Ones? A Theoretical Appraisal"  
L. Corchón. June 1994.
- WP-AD 94-10 "On the Competitive Effects of Divisionalization"  
L. Corchón, M. González-Maestre. June 1994.
- WP-AD 94-11 "Efficient Solutions for Bargaining Problems with Claims"  
M.C. Marco-Gil. June 1994.
- WP-AD 94-12 "Existence and Optimality of Social Equilibrium with Many Convex and Nonconvex Firms"  
A. Villar. July 1994.

- WP-AD 94-13 "Revealed Preference Axioms for Rational Choice on Nonfinite Sets"  
J.E. Peris, M.C. Sánchez, B. Subiza. July 1994.
- WP-AD 94-14 "Market Learning and Price-Dispersion"  
M.D. Alepuz, A. Urbano. July 1994.
- WP-AD 94-15 "Bargaining with Reference Points - Bargaining with Claims: Egalitarian Solutions Reexamined"  
C. Herrero. September 1994.
- WP-AD 94-16 "The Importance of Fixed Costs in the Design of Trade Policies: An Exercise in the Theory of Second Best"  
L. Corchón, M. González-Maestre. September 1994.
- WP-AD 94-17 "Computers, Productivity and Market Structure"  
L. Corchón, S. Wilkie. October 1994.
- WP-AD 94-18 "Fiscal Policy Restrictions in a Monetary System: The Case of Spain"  
M.I. Escobedo, I. Mauleón. December 1994.
- WP-AD 94-19 "Pareto Optimal Improvements for Sunspots: The Golden Rule as a Target for Stabilization"  
S.K. Chattopadhyay. December 1994.
- WP-AD 95-01 "Cost Monotonic Mechanisms"  
M. Ginés, F. Marhuenda. March 1995.
- WP-AD 95-02 "Implementation of the Walrasian Correspondence by Market Games"  
L. Corchón, S. Wilkie. March 1995.
- WP-AD 95-03 "Terms-of-Trade and the Current Account: A Two-Country/Two-Sector Growth Model"  
M.D. Guilló. March 1995.
- WP-AD 95-04 "Exchange-Proofness or Divorce-Proofness? Stability in One-Sided Matching Markets"  
J. Alcalde. March 1995.
- WP-AD 95-05 "Implementation of Stable Solutions to Marriage Problems"  
J. Alcalde. March 1995.
- WP-AD 95-06 "Capabilities and Utilities"  
C. Herrero. March 1995.
- WP-AD 95-07 "Rational Choice on Nonfinite Sets by Means of Expansion-Contraction Axioms"  
M.C. Sánchez. March 1995.
- WP-AD 95-08 "Veto in Fixed Agenda Social Choice Correspondences"  
M.C. Sánchez, J.E. Peris. March 1995.
- WP-AD 95-09 "Temporary Equilibrium Dynamics with Bayesian Learning"  
S. Chatterji. March 1995.