Two Sided Matching with Intrinsic Preferences Over Stated Rankings*

Yair Antler

Tel Aviv University

May 2013

Abstract

We extend the problem of two-sided matching by allowing motives such as shame, pride, embarrassment and insult to enter into agents' considerations. This is accomplished by defining agents' preferences so as to depend on stated rankings (during the matching process). The aim is to derive a matching mechanism such that each matching which is the result of an equilibrium of its induced revelation game is pairwise stable. We then study the Gale Shapley mechanism in this environment and provide an impossibility theorem for individually rational matching mechanisms. We also introduce a sequential variant of the Gale Shapley mechanism that guarantees pairwise stability of matchings generated in equilibria of its induced revelation game.

JEL Classification:

Keywords: Market Design; Behavioral Market Design; Two Sided Matching; Reciprocity; Behavioral Mechanism Design; Psychological Games

*This paper is based on the first chapter of my PhD dissertation which is currently held. I am grateful to Rani Spiegler for his valuable guidance. I thank Eddie Dekel, Erik Eyster, Alex Frug, Jacob Glazer, Tymofiy Mylovnov, Francesco Nava, Alvin Roth, and Ariel Rubinstein for their useful comments. All remaining errors are mine. *Correspondence:* Yair Antler, The Eitan Berglas School of Economics, Tel Aviv University. *E-mail:* yair.an@gmail.com.

1 Introduction

The standard mechanism design literature implicitly assumes that there is a clear distinction between agents' preferences and the mechanism in use. Two exceptions are Glazer and Rubinstein (1998) and Bierbrauer and Netzer (2012) who let agents' preferences depend on actions taken in the mechanism. In this paper, we relax this assumption by extending the standard complete information two-sided matching problem (Roth and Sotomayor (1990), chapter 4). This is done by allowing agents' preferences to depend on the matching process, so that motives such as shame, pride, embarrassment and insult can enter into the agents' considerations.

We assume that actions in the matching process are public knowledge among agents that participate in the mechanism, the social planner and an audience of outside observers. When actions are publicly known, agents may have other motives besides finding the best possible spouse. For example, suppose that *Alice* and *Bob* are two agents that take part in some matching mechanism and that the mechanism matches *Alice* to *Bob*. *Alice* may refuse the match because *Bob* states in public that she is his last choice and as a result she may prefer to remain single over marrying him. Similarly, *Bob* may feel embarrassed if some outside observers (for example *Bob*'s parents) see that his spouse ranks him lower than he ranks her or lower than others rank him.

An example of a matching process in which the existence of an audience affects agents' considerations is the draft choice system in professional sports in the US. In this system, there is a matching process between teams and players who are drafted into the league. Each team in its turn chooses one player who is obligated to play for that team for a few years. Usually teams invest a great deal of resources in scouting for and investigating potential draft choices and know the players' preferences quite well. Consider a player a who declares that he prefers playing for team A or that team A is unacceptable to him. In that case, although team A's manager knows a's true preferences and that his statement is a bluff, the statement may give the manager an additional motive to draft him since team A's fans do not know the player's true preferences and may want him to draft a player who declares that team A is his favorite. Explaining his decision to the team's fans may be costly for the team's manager and therefore the player's declaration is an additional motive for team A's manager to draft him.

Matching mechanisms are a fertile environment for such considerations since they provide a mechanism for direct revelation through which an agent can state his preferences over agents on the other side of the market. Thus, actions are not just abstract messages. We will call this kind of motive a *matching process related motive* (MPRM). These motives may upset the stability of an otherwise stable matching process. For example, if *Bob* and *Alice* are matched and after hearing *Alice*'s stated ranking *Bob* prefers to refuse the match and remain single, then this matching is not stable. On the other hand, some matchings that we consider to be unstable might become stable when we allow for MPRM. For example, if *Bob* and *Alice* are matched, but *Bob* likes *Carry* (who is single and would prefer to be married to *Bob*), then we consider *Bob* and *Carry* to be a blocking pair with regard to the stability of the matching. But if *Bob* changes his preferences between the two women after *Alice* states that he is her top choice, then this matching may be stable.

It will be useful at this point to clarify the notions of stability, preferences and equilibrium in our environment. First, we treat stability as an ex post criterion since a social planner does not want agents to have an incentive to challenge a prescribed match. Thus, we are interested in stability with respect to agents' preferences after the rankings have been announced. Secondly, since agents' preferences depend on stated rankings they form beliefs over the other agents' strategies, which determine their own preferences ex ante (which are intention-based). In equilibrium, the beliefs are correct and the agents' ex ante preferences are identical to their ex post preferences (which are outcome-based).

According to a standard result in the market design literature (Roth (1984b) Gale and Sotomayor (1985)), the matching mechanism induced by the Gale Shapley deferred acceptance algorithm (the GS mechanism) Nash equilibrium with undominated strategies implements the pairwise stable (PS) correspondence (with respect to the true preferences) regardless of which side of the market makes the proposals. We attempt to determine whether a similar result holds in our environment. First, we examine an environment in which only one side of the market contains agents with MPRM. It turns out that the analysis is sensitive to the side of the market that makes the proposals. If the agents on the proposing side of the market do not have MPRM, then each matching that is produced in equilibrium is PS, while if agents on the proposing side of the market have MPRM, then the stability of each matching produced in equilibrium is not guaranteed. On the other hand, the set of matchings produced when agents on the proposing side of the market have MPRM includes the set of matchings produced when agents on the proposing side of the market do not have MPRM.

When both sides of the market contain agents with MPRM the GS mechanism cannot guarantee the stability of each matching resulting from an equilibrium of its induced revelation game. Moreover, we provide an impossibility result which shows that for every simultaneous stable matching mechanism there exists a profile of preferences such that there exists an equilibrium of its induced revelation game that ends in a non-PS matching. In addition, in order to strengthen the theorem, we restrict the domain of MPRM to satisfy some reasonable properties. In that restricted domain, a higher stated ranking serves as a positive externality and the agents prefer being matched over being single since they do not have an outside option. The impossibility result holds even in this restricted domain.

We then turn to the class of sequential matching mechanisms. When considering sequential mechanisms one must be precise about which stated rankings determine an agent's preferences after a deviation. According to the approach we develop, a deviation by one of the first movers may change the second mover's actions, but not the deviator's preferences. We restrict the domain of MPRM to satisfy the positive externality requirement and provide a sequential variant of the GS mechanism such that in its induced game each equilibrium results in a PS matching. In our proposed mechanism, one side of the market moves first and then the other side responds.

The rest of the paper is organized as follows: Section 2 describes the related literature. Section 3 presents the general framework and section 4 presents the analysis of the GS mechanism's induced game. Section 5 covers simultaneous stable mechanisms and section 6 presents the sequential GS mechanism. Section 7 covers sequential matching mechanisms. Section 8 concludes.

2 Related literature

The complete information two-sided matching problem is a standard market design textbook problem (see Roth and Sotomayor (1990), chapter 4). Glazer and Rubinstein (1998) study an implementation problem in which experts have two motives: a public motive to increase the probability that the desirable action will be taken and a private motive to have their recommendation accepted. Bierbrauer and Netzer (2012) introduce intentions-based motives into a mechanism design problem. They show that intentions may fail the revelation principle and that different mechanisms which implement the same social choice function might differ in the kindness sensations that they generate. Gradwohl (2012) studies a complete information implementation problem in which agents have preferences for privacy. Although agents that take part in the mechanism know the true preferences of other participating agents, the social planner or an audience of outside observers do not, such that revealing information may be costly in future interactions with them.

Avery and Levin (2010) give an example of a matching market in which one side of the market contains agents that reciprocate to other agents' preferences, such as schools that take into account the applicants' preferences over schools. Geanakoplos, Pearce and Stacchetti's (1989) (GPS) introduce psychological game theory as another framework that enables the study of reciprocity as a matter of intentions. Rabin (1993) adopts GPS's model and extends it such that an action is perceived as kind or unkind according to the intentions behind it, which depend on the payoff structure.

3 The general framework

The elements of the model are as follows: There are two disjoint sets M and W. $M = \{m_1, ..., m_l\}$ is the set of men and $W = \{w_1, ..., w_n\}$ is the set of women. Each agent $i \in M \cup W$ has a strict linear ordering P_i of the agents on the other side of the market (including the option of being single) such that if $i \in M$ then P_i is a strict ranking of $\{i\} \cup W$, and if $i \in W$ then P_i is a strict ranking of $i \cup M$. We use $P_m = (w, w', m, w'', ...)$ to state that man m ranks woman w first, women w' second, being single third and woman w'' fourth. Whenever iP_ij we say that j is unacceptable to i (for example, w'' is unacceptable to m). Let $P \equiv (P_i)_{i \in M \cup W}$.

We consider direct revelation mechanisms in a complete information environment. Denote the report of each agent i ($i \in M \cup W$) as Q_i and let the set of strategies available for agent i be $\mathcal{Q}_i$; let $Q \equiv (Q_i)_{i \in M \cup W}$, $Q_M \equiv (Q_i)_{i \in M}$, $Q_W \equiv (Q_i)_{i \in W}$, $Q_{-i} \equiv (Q_j)_{j \in M \cup W \setminus \{i\}}$ and $\mathcal{Q} = \times_{i \in M \cup W} \mathcal{Q}_i$. An outcome of the game is a matching: $\nu : M \cup W \rightarrow M \cup W$ such that $w = \nu(m)$ if and only if $m = \nu(w)$ and for all m and w either $\nu(w) \in M$ or $\nu(w) = w$ and either $\nu(m) \in W$ or $\nu(m) = m$. Let $\mathcal{M}$ be the set of matchings and let $g : \mathcal{Q} \rightarrow \mathcal{M}$ be a matching mechanism and for each $i \in M \cup W$ let $g_i(Q)$ be i's spouse given a mechanism g and a strategy profile Q.

Definition 1 *Let $\tilde{\nu}$ be a matching in the marriage market and let S be a profile of strict linear orderings on $M \cup W$. $\tilde{\nu}$ is PS with respect to S if the two following conditions are* *met:*

$(a)\nexists (m \in M, w \in W)$ *such that* $m \neq \tilde{\nu}(w)$, $mS_w\tilde{\nu}(w)$ *and* $wS_m\tilde{\nu}(m)$.

$(b)\nexists i \in M \cup W$ *such that* $iS_i\tilde{\nu}(i)$.

The innovation in this model is that the profile of rankings P does not necessarily represents agents' preferences. For each $I \in \{M, W\}$ and $i \in I$, denote agent i's (strict) preferences as R_i. Agent i's preferences depend on the strategy profile Q_{I^c} and his own ranking P_i. For each $Q \in \mathcal{Q}$ and $Q_{I^c} \in \mathcal{Q}_{I^c}$ such that $Q \supset Q_{I^c}$, let $R_i(Q) \equiv R_i(Q_{I^c}, P_i)$ and let $R(Q) \equiv (R_i(Q))_{i \in M \cup W}$.

Definition 2 *The strategy* Q_i *weakly dominates the strategy* $\hat{Q}_i$ *if* $\nexists Q'_{-i} \in \mathcal{Q}_{-i}$ *such that* $g_i\left(\hat{Q}_i, Q'_{-i}\right) R_i(Q') g_i\left(Q_i, Q'_{-i}\right)$ *and* $\exists \tilde{Q}_{-i} \in \mathcal{Q}_{-i}$ *such that* $g_i\left(Q_i, \tilde{Q}_{-i}\right) R_i\left(\tilde{Q}\right) g_i\left(\hat{Q}_i, \tilde{Q}_{-i}\right)$.

Definition 3 *A Nash equilibrium that does not involve the use of weakly dominated strategies (NEU) of the revelation game* $\langle R, g\rangle$ *is a profile of strategies* Q^* *such that for each* $i \in M \cup W$:

(1) Q_i^* *is not weakly dominated.*

(2) $\nexists Q_i \in \mathcal{Q}_i$ *such that* $g_i\left(Q_i, Q_{-i}^*\right) R_i(Q^*) g_i\left(Q_i^*, Q_{-i}^*\right)$.

4 The GS deferred acceptance algorithm and its induced revelation game

The following description is taken from Roth and Sotomayor (1990).

4.1 The GS deferred acceptance algorithm with men making the proposals

Each man starts by proposing to his favorite woman within his ranking of acceptable women. A woman rejects the proposals of men who are unacceptable to her and if she receives more than one proposal from acceptable men, she rejects all but the one from her most preferred man. Any man whose proposal has not been rejected at this point is engaged. In each step, a man whose proposal has been rejected in the previous step proposes to his most preferred acceptable woman among those who have not rejected him in the previous steps. If he has been rejected by all women that he finds acceptable, then he issues no further proposals. A woman who receives proposals rejects those of men who are unacceptable to her and if she has received more than one proposal from acceptable men, she rejects all but the one from her most preferred man, including any man she may have been engaged to from the previous step. The algorithm stops after any step in which no man has been rejected.

4.2 Analysis of the induced revelation game

We first present three examples that demonstrate the effects of MPRM. Recall that Roth (1984b) shows that in each NEU of the GS game the produced matching is PS with respect to the true preferences regardless of which side of the market makes the proposals. The first example shows that even if only one agent has MPRM, the GS game might have an NEU Q^* that induces a matching $g(Q^*)$ and a profile of preferences $R(Q^*)$ such that $g(Q^*)$ is not PS with respect to $R(Q^*)$.

Example 1 *Let $M = \{m, m'\}$, $W = \{w, w'\}$ and let $P_m = (w, w', m)$, $P_{m'} = (w', w, m')$, $P_w = (m', m, w)$, $P_{w'} = (m', m, w')$. For each $i \in \{m', w, w'\}$ and $Q \in \mathcal{Q}$, let $R_i(Q) = P_i$ and let R_m be such that for each $j \in W$, $jR_m(Q)\,m$ if and only if jQ_jm. Recall that by Roth (1982) the agents on the side of the market that makes the proposals have a dominant strategy, i.e. stating their true preferences. This is true in our setup for agents that do not have MPRM. In this case, when women make the proposals, they must use their dominant strategies. When men make the proposals, women do not have a dominant strategy. It follows that a profile Q^* in which man m ranks both women as unacceptable and woman w misrepresents her preferences and ranks man m as unacceptable is a NEU in the revelation game that is induced by the GS mechanism when men make the proposals. Let Q^* be: $Q^*_w = (m', w, m)$, $Q^*_{w'} = (m', w', m)$, $Q^*_m = (m, w, w')$, $Q^*_{m'} = (w', w, m')$. Then, $g(Q^*) = (m, (m', w'), w)$ ((m', w') are married to each other and m and w are singles) is a NEU of $\langle g, R\rangle$ and is not PS with respect to $R(Q^*)$ since it is blocked by the pair (m, w).*

Along with Roth (1984b)'s result, Gale and Sotomayor (1985) show that the PS correspondence is NEU implemented by both versions of the GS mechanism. The second example shows that it matters which side of the market makes the proposals in the sense that the set of matchings that can be supported by an NEU in the induced revelation game depends on it. In the example, there is only one matching that can be supported by an NEU in the revelation game induced by the GS mechanism when men make the proposals, while there are two that can be supported by an NEU in the revelation game induced by the GS mechanism when women make the proposals. In the example, all of these matchings are PS with respect to the corresponding equilibrium's profile of preferences, although in general this need not be the case.

Example 2 *Let* $M = \{m, m'\}$, $W = \{w, w', w''\}$ *and let* $P_m = (w, m, w', w'')$, $P_{m'} = (w, w'', w', m')$, $P_w = (m, w, m')$, $P_{w''} = (w'', m, m')$. *For each* $i \in \{m, m', w, w''\}$ *and* $Q \in \mathcal{Q}$ *, let* R_i *be such that* $R_i(Q) = P_i$ *and let*

$$R_{w'}(Q) = \begin{pmatrix} m', w', m & \text{if } Q_{m'} \text{ ranks } w' \text{ at least second} \\ w', m', m & \text{otherwise} \end{pmatrix}.$$

Woman w'*'s preferences are interpreted as follows: She likes man* m' *(and only man* m'*), but she is insulted if she thinks that he ranks her lower than second place. In that case, she reciprocates and prefers to be single. When men make the proposals, then each* $m \in M$ *uses his weakly dominant strategy (Roth, (1982)) and ranks the women exactly as in* P_m*. It follows that the only matching that can be supported by an NEU of the induced revelation game is* $((m, w), m', w', w'')$*. When women make the proposals there are two matchings that can be supported by an NEU in the induced revelation game:* $((m, w), m', w', w'')$ *and* $((m, w), (m', w'), w'')$*.* $((m, w), m', w', w'')$ *is supported by a report profile in which man* m' *reports his true rankings and therefore, in NEU woman* w' *reciprocates such that* $w' Q_{w'} m'$*.* $((m, w), (m', w'), w'')$ *is supported by a ranking profile in which man* m'*'s report is* $Q_{m'} = (w, w', w'', m')$ *such that woman* w' *reciprocates and states* $m' Q_{w'} w'$*. Note that* $((m, w), (m', w'), w'')$ *is at least as good for men and is strictly better for man* m'*.*

Example 3 presents another profile of preferences in which one side of the market contains agents with MPRM. This profile of preferences is closely related to the class of priority mechanisms (see Ergin and Somez (2006)) since it corresponds to the Boston mechanism.

Example 3 *Suppose that for each* $m \in M$ *and* $Q \in \mathcal{Q}$, $R_m(Q) = P_m$ *and that for each* $w \in W, \mathcal{Q} \in Q$, $R_w(Q)$ *is as follows: (1) If* m *ranks* w *higher than* m' *does, then* $m R_w(Q) m'$*; (2) if* m *ranks* w *exactly as high as* m' *does, then* $m R_w(Q) m'$ *if and only if*

$mP_w m'$. This profile of preferences corresponds to the Boston mechanism where women are in the position of the schools. In the old Boston school allocation system, the schools were not strategic, but the allocation system acted as if they were and as if they have lexicographic preferences, where the first criterion is students' stated rankings and the following ones are eligibility criteria. Proposition 1 implies that if schools were strategic and if the students make the proposals, then the revelation game that is induced by the GS mechanism induces only stable matchings. Moreover, the students have an incentive not to act strategically since truth telling is a dominant strategy for them.

We now present results for the GS mechanism when agents on one side of the market may have MPRM but those on the other side do not. The first result shows that in this environment, if the side of the market that makes the proposals is chosen carefully, then the GS mechanism's induced revelation game produces only PS matchings. This result relies on standard arguments in the market design literature.

Proposition 1 *Let $I \in \{M, W\}$. Let g be the GS mechanism with agents on side I of the market making the proposals and for each $i \in I$ and $Q \in Q$ let $R_i(Q) = P_i$. If Q^* is an NEU of the revelation game $\langle g, R\rangle$, then $g(Q^*)$ is PS with respect to $R(Q^*)$. Moreover, the PS correspondence with respect to $R(P)$ is NEU-implemented by $\langle g, R\rangle$.*

Proof. See appendix. ■

It appears that if one's main consideration is PS, then the side of the market with no MPRM should make the proposals. The following result generalizes example 2 and shows that if the side of the market that makes the proposals is the one with MRPM, then the set of PS matchings is weakly larger. For the next result, we place two restrictions on MPRM.

Definition 4 *For each $I \in \{M, W\}$, $i \in I$, $Q_i \in \mathcal{Q}$, and $j \in I^c$, let $B_j^{Qi} = \{k \in I^c \cup i : jQ_ik\}$. We say that the strategy Q_i^* ranks j at least as highly as the strategy Q_i' if $B_j^{Q_i^*} \supseteq B_j^{Q_i'}$.*

Note that this definition is more restrictive than requiring that only the relative stated rankings matter since agents may care about who is ranked below/above them and if they are ranked as acceptable or not.

Condition 1 *(reciprocity) R satisfies reciprocity if $iR_j(Q_i, Q_{-i})\, i'$ implies $iR_j\left(\tilde{Q}_i, Q_{-i}\right) i'$ whenever $\tilde{Q}_i$ ranks j at least as highly as Q_i for each $I \in \{M, W\}$,$i \in I$,$i' \in I \cup \{j\}$, $j \in I^c$, $Q_{-i} \in \mathcal{Q}_{-i}$.*

Condition 2 *(independence) R satisfies independence if for any two strategy profiles $Q, \hat{Q} \in \mathcal{Q}$ such that $Q_i = \hat{Q}_i$ and $Q_{i'} = \hat{Q}_{i'}$ it is true that $iR_j(Q)\,i'$ if $iR_j\left(\hat{Q}\right)i'$ and $jR_j(Q)\,i'$if and only if $jR_j\left(\hat{Q}\right)i'$ for each $I \in \{M, W\}$, $i, i' \in I$ and $j \in I^c$.*

These two conditions make a high stated ranking a positive externality. Note that example 1 exhibits MPRM but does not satisfy *reciprocity* while examples 2 and 3 satisfy both conditions.

Proposition 2 *Let R satisfy reciprocity and independence. Suppose that $\exists I \in \{M, W\}$ such that for each $i \in I$, $R_i(Q) = P_i$ for each $Q \in \mathcal{Q}$. Let g be the GS mechanism with agents on the I side of the market making the proposals and let g' be the GS mechanism with agents on the I^c side of the market making the proposals. If Q^* is an NEU of the revelation game $\langle g, R\rangle$, then there exists a profile of strategies $\bar{Q}$ such that:*

(1) $\bar{Q}$ is an NEU of the revelation game $\langle g', R\rangle$.
(2) $g'(\bar{Q}) = g(Q^*)$.
(3) $g'(\bar{Q})$ is PS with respect to $R\left(\bar{Q}\right)$.

Proof. See appendix. ■

When both sides of the market might contain agents with MPRM, weakly dominant strategies need not exist and the set of PS matchings may depend on the strategy profile. The next section provides some negative results. Not only that the GS mechanism can not guarantee pairwise stability in every NEU of its induced revelation game, there exists no other simultaneous stable matching mechanism that can guarantee pairwise stability in every NEU of its induced revelation game, even if we restrict the MPRM to satisfy some plausible conditions.

5 Simultaneous stable matching mechanisms

We now focus on stable matching mechanisms, which have received a great deal of attention in market design theory. Most matching mechanisms in use today are stable ones, including the NRMP matching mechanism (Roth, 1984a) and the current Boston public schools allocation system (Abdulkadiroglu et al., 2005 and 2006).

Definition 5 *A stable matching mechanism is a matching mechanism g such that for each $Q \in \mathcal{Q}$, $g(Q)$ is PS with respect to Q.*

First, we show what cannot be achieved given our framework. To do so, we provide an impossibility theorem.

Proposition 3 *For every simultaneous stable matching mechanism g, there exists a profile of preferences R and a profile of rankings Q^* such that Q^* is an NEU of the revelation game $\langle g, R\rangle$ and $g(Q^*)$ is not PS with respect to $R(Q^*)$.*

Proof. The proof follows directly from the proof of proposition 7. ■

To strengthen the theorem, we restrict the domain of MPRM to satisfy *independence*, *reciprocity* and *no outside option*. *no outside option* corresponds to a plausible scenario in which workers and firms do not have an outside option, such that each pair of firm and worker is mutually acceptable. Technically, this condition limits our ability to use the individual rationality (IR) property of stable matching mechanisms.

Condition 3 *(no outside option) We say that R satisfies no outside option if for each $I \in \{M, W\}$, $i \in I$, $j \in I^c$, $Q \in \mathcal{Q}$ it is true that $iR_j(Q)\,j$ and $jR_i(Q)\,i$.*

Proposition 4 *For every simultaneous stable matching mechanism g, there exists a profile of preferences R that satisfies ndependence, reciprocity and no outside option and a profile of rankings Q^* such that Q^* is an NEU of the revelation game $\langle g, R\rangle$ and $g(Q^*)$ is not PS with respect to $R(Q^*)$.*

Proof. See appendix. ■

6 The Sequential Gale Shapley mechanism

In the previous section, we claimed that it is impossible to find a simultaneous stable matching mechanism such that every equilibrium of its induced revelation game induces a PS matching (with respect to the produced profile of preferences). In this section, we

introduce the sequential Gale Shapley (SGS) mechanism, a sequential variant of the GS mechanism, and claim that it produces only PS matchings in equilibria of its induced revelation game in a restricted domain of MPRM. The SGS mechanism works as follows:

- Each woman $w \in W$ submits a ranking Q_w of $w \cup M$.

- Q_W becomes common knowledge among the men.

- Each man $m \in M$ submits a ranking Q_m of $m \cup W$.

- The GS algorithm with men making the proposals is used.

Since the SGS mechanism induces a sequential revelation game, one must define an appropriate solution concept. We use the Sub-Game Perfect Nash equilibrium that does not involve the use of weakly dominated strategies in any sub-game (SGPU). For the men, the MPRM are outcome-based. After each man $m \in M$ sees Q_W , his preferences will be $R_m(Q_W)$. As for women, a question arises as to whether preferences should be outcome-based or intentions-based. This distinction is meaningless in a static game since the two consolidate in NEU. We follow GPS (1989) and Rabin (1993) by considering preferences that are intentions-based (a short discussion appears in the next section) such that a woman that deviates from her equilibrium's strategy updates her beliefs on men's succeeding actions but not her preferences.

Definition 6 *Let g be the SGS mechanism. A SGPU of the revelation game $\langle g, R\rangle$ is a profile of strategies $(Q_W^*, Q_M^*(Q_W))$ such that:*

(1) For each $m \in M$, $Q_m^*(Q_W)$ is not weakly dominated and for each $Q_W \in \mathcal{Q}_W$ $\nexists Q_m^{**}(Q_W)$ such that:

$$g_m\left(Q_m^{**}(Q_W), Q_{M\backslash\{m\}}^*(Q_W), Q_W\right) R_m\left(Q_M^*(Q_W), Q_W\right) g_m\left(Q_m^*(Q_W), Q_{M\backslash\{m\}}^*(Q_W), Q_W\right)$$

(2)For each $w \in W$ Q_w^* is not weakly dominated and $\nexists Q_w^{**}$ such that $g_w\left(Q_w^{**}, Q_{W\backslash\{w\}}^*, Q_M^*(Q_W)\right)$ $R_w\left(Q_W^*, Q_M^*(Q_W)\right)$ $g_w\left(Q_W^*, Q_M^*(Q_W)\right)$

The following is an existence result and to prove it we restrict ourselves to preferences that satisfy *reciprocity* and *independence*.

Proposition 5 *Let R satisfy independence and reciprocity. Then, a SGPU of the revelation game that is induced by the SGS mechanism exists.*

Proof. See appendix. ■

We now present the main positive result of this paper. We continue to restrict ourselves to MRPM that satisfies *recuprocity* and *independence* (in proposition 7, we show that this restricted domain is necessary for this proof). We show that each of the SGS mechanism-induced revelation game's SGPUs ends in a PS matching (with respect to the equilibrium's profile of preferences). The idea underlying the proof is that the second movers (i.e. men) know their preferences when it is their turn to take an action. Thus, they have a weakly dominant strategy (or strategies) that they must use in SGPU. Using *reciprocity* and *independence*, the coordination problem is solved since the men respond to the women's rankings.

Proposition 6 *Let R satisfy independence and reciprocity. Let g be the SGS mechanism. Let $(Q_W^*, Q_M^*(Q_W))$ be an SGPU of the revelation game $\langle g, R \rangle$. Then, $g(Q_W^*, Q_M^*(Q_W))$ is PS with respect to $R(Q_W^*, Q_M^*(Q_W))$.*

Proof. See appendix. ■

This mechanism enables us to obtain a result similar to Roth (1984b)'s, whereby a small change in the mechanism that is generally in use enables us to overcome the coordination problem that is generated by MPRM. In what follows, we expand the scope of the model. We define a sequential matching mechanism and show why our restriction of the domain of MRPM is needed.

7 Sequential matching mechanisms

A sequential matching mechanism $g = (a, S)$ includes a matching algorithm $a : \mathcal{Q} \to \mathcal{M}$ and a partition $S = (s_1, ..., s_k)$ of $M \cup W$ such that $1 \leq k \leq |M \cup W|$. Each cell $s_l \in S$ represents a stage in g such that each $i \in s_l$ submits a ranking $Q_i\left(Q_j\right)_{j \in \{s_1,...,s_{l-1}\}}$ in stage l. At the end of each stage l, each Q_i such that $i \in s_l$ becomes common knowledge among all agents. For each $i \in s_l$, let $\bar{b}_i$ be agent i's beliefs over $\{Q_j : j \neq i, j \in s_m, \text{ and } m \geq l\}$ in stage l, conditional on the events that took place in stages $1, ..., l-1$ and on i's strategy and let $\underline{b}_i = \left(Q_j\right)_{j \in \{s_1,...,s_{l-1}\}}$. Once Q is revealed, a predetermined matching algorithm a is applied and a matching $g(Q)$ is produced.

Since agents' preferences depend on their beliefs, no single set of payoffs adequately describes the strategic situation. This phenomenon induces a psychological game (GPS (1989)). GPS defined a solution concept that demands that only initial beliefs enter into the agents' payoffs. We could prove propositions 5,6 and 7 with GPS's solution concept but we think that it is not suitable to the strategic situations that we describe in this work since in these situations it is reasonable to think that agents' preferences change when their beliefs change. We define a solution concept that allows later movers' beliefs based preferences to change due to actions of preceding movers. As in section 6, we assume that agents' preferences are intention based so that an agent that deviates changes his beliefs over the actions of his successors but his preferences are left unchanged due to his deviation[1].

Definition 7 *A Psychological Sub-Game Perfect Nash Equilibrium that does not involve the use of weakly dominated strategies in any sub-game (PSGPU) of the psychological game* $\langle g, R \rangle$ *is a profile of strategies* $\left(Q_i^*\left(\underline{b}_i\right)\right)_{i \in M \cup W}$ *such that:*
(1) For each $j \in W \cup M$ $Q_j^*\left(b_j\right)$ *is not weakly dominated.*
(2) For each $j \in W \cup M$: $\nexists Q_j\left(b_j\right)$ *such that* $g\left(Q_j\left(b_j\right), \left(Q_i^*\left(\underline{b}_i\right)\right)_{i \in M \cup W/\{j\}}\right) R_j\left(\underline{b}_j, \bar{b}_j\right)$ $g\left(Q_j^*\left(b_j\right), \left(Q_i^*\left(\underline{b}_i\right)\right)_{i \in M \cup W/\{j\}}\right)$ *for each possible* $\underline{b}_j$.

The following proposition is a negative result which states that we must restrict ourselves to a smaller domain of MRPM if we wish to find an IR matching mechanism such that each of its induced revelation game's equilibria ends in a PS matching, which will justify the restricted MPRM domain in proposition 6. Note that proposition 3 is an immediate corollary of this proposition.

Definition 8 *An IR matching mechanism is a matching mechanism* g *such that for each* $Q \in \mathcal{Q}$, $\nexists i \in M \cup W$ *such that* $iQ_i g_i(Q)$.

IR is interpreted as a veto property. An IR matching mechanism enables agent A to veto agent B, such that regardless of the other agents' actions, the mechanism will never match A to B. IR matching mechanisms include the class of stable matching mechanisms. We show that there exists a market and a profile of preferences for which no IR matching mechanism can guarantee that in the psychological game that is induced by the mechanism, each PSGPU results in a PS matching (with respect to the equilibrium's profile of preferences). We then generalize the proof to a market of any size. The proof is a generalization of example 1.

[1] Another possibility is to use a solution concept in which only updated beliefs enter into the agent's payoffs a la Battigalli and Dufwenberg (2009). In our opinion, using updated beliefs is not suitable to the strategic situations described here since it allows an agent that moves earlier to determine his own preferences by choosing different actions which have effect on the later movers' rankings.

Proposition 7 *For every IR matching mechanism $g = (a, S)$, there exists a profile of preferences R and a profile of strategies Q^* that forms a PSGPU of the psychological game $\langle g, R \rangle$ and $g(Q^*)$ is not PS with respect to $R(Q^*)$.*

Proof. See appendix. ∎

8 Concluding remarks

This paper is the first attempt to incorporate MPRM into matching theory. A particularly restrictive issue in matching markets involves coordination problems and MPRM exacerbates those problems. It is shown that sequential mechanisms can solve these coordination problems, while simultaneous mechanisms cannot. This is in contrast to the standard model, in which excluding dominated strategies is sufficient to solve the coordination problem.

A possible direction for future research would be to compare mechanisms that use complete rankings to those that use partial rankings. Partial rankings may solve some of the coordination problems, although in some cases may enlarge the set of possible PS matchings. It would be useful to find the minimal exposition of rankings that is needed to achieve PS when agents have preferences for privacy.

9 Appendix - proofs

Proposition 1 *Let $I \in \{M, W\}$. Let g be the GS mechanism with agents on side I of the market making the proposals and for each $i \in I$ and $Q \in \mathcal{Q}$ let $R_i(Q) = P_i$. If Q^* is an NEU of the revelation game $\langle g, R \rangle$, then $g(Q^*)$ is PS with respect to $R(Q^*)$. Moreover, the PS correspondence with respect to $R(P)$ is NEU-implemented by $\langle g, R \rangle$.*

Proof. Without loss of generality, let $I = M$. Since men do not have MPRM, one can apply Roth's (1982) result and therefore in every NEU men must state their true preferences. Assume by negation that Q^* is an NEU of the revelation game $\langle g, R \rangle$ and that $g(Q^*)$ is not PS with respect to $R(Q^*)$. $g(Q^*)$ cannot be blocked by any individual $i \in M \cup W$ since in that case i can profitably deviate to $Q'_i = (i, ...)$. If $g(Q^*)$ is blocked by a pair (m', w'), then since m' uses his dominant strategy, it must be that w' can deviate to $Q_{w'} = (m', ...)$ and profit, a contradiction of Q^* being an NEU of $\langle g, R \rangle$. Since $Q_M = P_M$ in every NEU and $R_M(Q) = P_M \ \forall Q \in \mathcal{Q}$ it must be that if Q is an NEU of $\langle g, R \rangle$, then $R(Q) = R(P)$.

In the opposite direction, we assume that μ is a PS matching with respect to $R(P)$ and refer to $\mu(i)$ as i's spouse under μ. We wish to show that there exists a profile Q' such that Q' is an NEU of $\langle g, R \rangle$ and $g(Q') = \mu$. The first step of the proof is to construct a profile of strategies Q^* (which may include the use of dominated strategies by women) that is a Nash equilibrium (NE) of $\langle g, R \rangle$, such that $g(Q^*) = \mu$. Let $Q^*_M = P_M$ and let Q^*_W be such that for each $w \in W$:

$$Q^*_w = \left\{ \begin{array}{ll} \mu(w), w, ... & \text{if } \mu(w) \neq w \\ w, ... & \text{if } \mu(w) = w \end{array} \right\}$$

First, it is straightforward to show that $g(Q^*) = \mu$. Second, assume by negation that Q^* is not an NE of $\langle g, R \rangle$. Men use their dominant strategy and therefore no man can deviate and profit. If some woman, say woman w, deviates and as a result marries man $m' \neq \mu(w)$, then by the algorithm it must be that m' proposes to her before he makes a proposal to $\mu(m')$. Therefore, it must be that $w P_{m'} \mu(m')$, such that (m', w) block μ, which violates the stability of μ with respect to $R(P)$.

For the next step of the proof, let W^d be the set of women that use dominated strategies under Q^*. If $W^d = \varnothing$, then the proof is completed. If not, then there must be a woman $\hat{w} \in W^d$. Change woman $\hat{w}$'s strategy from $Q^*_{\hat{w}}$ to $Q'_{\hat{w}}$, which dominates $Q^*_{\hat{w}}$ and is not dominated by any other strategy. First, one must show that the profile of strategies $Q'_{\hat{w}}, Q^*_{-\hat{w}}$ induces μ. Woman $\hat{w}$ uses a strategy that dominates $Q^*_{\hat{w}}$ and therefore cannot marry a man who is inferior to $\mu(\hat{w})$ according to $R_{\hat{w}}(Q^*)$. Since Q^* is a NE, neither can she marry a man who is superior to $\mu(\hat{w})$ according to $R_{\hat{w}}(Q^*)$. It follows that under the new profile of strategies $\hat{w}$ marries $\mu(\hat{w})$. By the algorithm, since $\hat{w}$ does not reject nor keeps any additional men, the set of men who propose to each of the other women is the same as in Q^* and therefore μ must be the induced matching. Since the set of men that eventually propose to each of the women is the same in Q^* and in $Q'_{\hat{w}}, Q^*_{-\hat{w}}$ if a woman $w \in W \setminus \{\hat{w}\}$ can deviate and profit in $Q'_{\hat{w}}, Q^*_{-\hat{w}}$ then that profitable deviation exists also in Q^*. However, Q^* is an NE of $\langle g, R \rangle$, a contradiction. The process is repeated if necessary.

■

Proposition 2 *Let R satisfy reciprocity and independence. Suppose that $\exists I \in \{M, W\}$ such that for each $i \in I$, $R_i(Q) = P_i$ for each $Q \in \mathcal{Q}$. Let g be the GS mechanism with agents on the I side of the market making the proposals and let g' be the GS mechanism with agents on the I^c side of the market making the proposals. If Q^* is an NEU of the revelation game $\langle g, R\rangle$, then there exists a profile of strategies $\bar{Q}$ such that:*

(1) $\bar{Q}$ is an NEU of the revelation game $\langle g', R\rangle$.
(2) $g'(\bar{Q}) = g(Q^*)$.
(3) $g'(\bar{Q})$ is PS with respect to $R\left(\bar{Q}\right)$.

Proof. Without loss of generality, assume that $I = M$. The proof consists of four steps. First, we construct a profile of (perhaps weakly dominated) strategies Q' such that $g'(Q') = g(Q^*)$. The second step is to show that $g'(Q')$ is a NE of $\langle g', R\rangle$. The third step identifies which strategies might dominate any $Q'_w \in Q'$. The fourth step shows that if there exists a weakly dominated strategy $Q'_w \in Q'$, then given the third step, one can change Q' to a profile of weakly undominated strategies $\bar{Q}$ such that Q''s properties are preserved.
Step 1: For each $w \in W$, let $Q'_w = R_w(Q')$. For each $m \in M$, let Q'_m rank the women exactly as P_m does, with the following changes:

(1)If $g_m(Q^*) \neq m$, then if she is ranked in second place (or lower) in P_m, then Q'_m ranks $g_m(Q^*)$ in second place and m in third place.
(2)If $g_m(Q^*) \neq m$ and she is ranked in first place according to P_m, them Q'_m ranks m in second place.
(3)If $g_m(Q^*) = m$, then Q_m ranks m exactly as P_m does or in second place, whichever is higher.

We proceed by showing that $g'(Q') = g(Q^*)$. Assume that this is not the case. Then it must be that there exists a woman $w \in W$ such that $g'_w(Q') \neq g_w(Q^*)$. First, we show that there exists no woman $w \in W$ such that $g_w(Q^*) = w$ and $g'_w(Q') \in M$. Assume that there exists such a woman and denote her as w. By the construction of Q', it must be that $g'_w(Q')$ ranks w first in $P_{g'_w(Q')}$. Since Q^* is a NEU of $\langle g, R\rangle$, it must be that

$g'_w(Q')$ carries out his dominant strategy and ranks w first in $Q^*_{g'_w(Q')}$ as well. By construction of Q', $g'_w(Q')R_w(Q')\,w$. By *reciprocity* and *independence*, also $g'_w(Q')R_w(Q^*)\,w$, a contradiction of Q^* being an NEU of $\langle g, R\rangle$.

Secondly, we claim that each $Q'_w \in \{Q_w \in Q'_W : g_w(Q^*) \neq w\}$ is such that $g_w(Q^*)\,Q'_w m$ for any man $m \in \{m' \in \{M/\{g_{w'}(Q^*)\}\} : w'P_{m'}i$ for each $i \in W \cup m\}$. Assume that this is not the case. Then there exists a man $m \in M$ such that $m \neq g_w(Q^*)$, $mR_w(Q')\,g_w(Q^*)$ and $wP_m i$ for each $i \in W \cup m$. By *reciprocity* and *independence*, also $mR_w(Q^*)\,g_w(Q^*)$, a contradiction of Q^* being PS. From these two claims and the construction of Q', it follows that no woman w exists such that $g'_w(Q') \neq g_w(Q^*)$.

Step 2: We proceed by showing that Q' is an NE of $\langle g', R\rangle$. Clearly, each $Q'_w \in Q'$ is a best response to Q'_{-w}. If $Q'_m \in Q'$ is not a best response to Q'_{-m}, then there exists a strategy $\tilde{Q}_m$ such that $g'_m\left(\tilde{Q}_m, Q'_{-m}\right)P_m g'_m(Q')$. Denote $g'_m\left(\tilde{Q}_m, Q'_{-m}\right)$ as w^*. Q^* is a NEU of $\langle g, R\rangle$ and therefore $g(Q^*)$ is PS with respect to $R(Q^*)$ and by *reciprocity* and *independence* if $mR_{w^*}(Q')\,g_{w^*}(Q^*)$, then $mR_{w^*}(Q^*)\,g_{w^*}(Q^*)$, a contradiction of $g(Q^*)$ being PS with respect to $R(Q^*)$.

Step 3: Assume that there exists a $Q'_w \in Q'_W$ that is weakly dominated. If Q'_w is weakly dominated by another strategy $\hat{Q}_w$, then for any two men m and m' such that $mQ'_w m'$ and $m'\hat{Q}_w m$, we argue that it must be that if $wQ''_{m'}m'$ then for any $Q \supset Q''_{m'}$, $m'R_w(Q)\,m$. Assume that this is not the case. Then, by *reciprocity* and *independence*, it must be that there exists a profile of strategies $\tilde{Q}_{-w}$ such that $w\tilde{Q}_{m'}m'$, $w\tilde{Q}_m m$, $mR_w\left(\tilde{Q}\right)m'$, and for each $m^* \in M\backslash\{m, m'\}$ $\tilde{Q}_{m^*}$ is such that $m^*\tilde{Q}_{m^*}w$, and for each $w^* \in W\backslash\{w\}$ $\tilde{Q}_{w^*}$ is such that for each $i \in \{m, m'\}$, $w^*\tilde{Q}_{w^*}i$. Clearly, Q'_w does better than $\hat{Q}_w$ against $\tilde{Q}_{-w}$ and therefore cannot be weakly dominated by it.

Step 4: The fourth step of the proof is to change Q' (if necessary) to a profile $\bar{Q}$ such that for each $i \in M \cup W$, $\bar{Q}$ is not weakly dominated, while $\bar{Q}$ preserves Q''s properties. For each $i \in M$, Gale and Sotomayor's (1985) result applies and therefore Q'_m is not weakly dominated. As for women, assume that for some $w \in W$ Q'_w is weakly dominated by another strategy. Now change w's strategy to another strategy $\bar{Q}_w$ which is not weakly dominated by any other strategy. By step 3, the only possible change is that w proposes to some other men (who reject her) before proposing to $g_w(Q^*)$. Repeat this process if necessary and create $\bar{Q}$.

The last part of the proof is to show that $g'\left(\bar{Q}\right)$ is PS with respect to $R\left(\bar{Q}\right)$ and that $\bar{Q}$ is a NE of $\langle g', R\rangle$. Assume by negation that it is not PS with respect to $R\left(\bar{Q}\right)$. Clearly, if $g'\left(\bar{Q}\right)$ is blocked by an individual, then it is not an NEU of $\langle g', R\rangle$. It follows that $g'\left(\bar{Q}\right)$ is blocked by some pair (m, w). By *reciprocity* and *independence*, if $mR_w\left(\bar{Q}\right)g_w(Q^*)$, then $mR_w(Q^*)\,g_w(Q^*)$, and since $wP_m g_m(Q^*)$, this is a contradiction of $g(Q^*)$ being PS with respect to $R(Q^*)$.

It is straightforward to see that no woman $w \in W$ can gain by a deviation. If a man $m \in M$ can gain from a deviation and marry w^*, then $w^*P_m m$ and by step 3 $mR_{w^*}(Q^*)\,g_{w^*}(Q^*)$, this is a contradiction of $g(Q^*)$ being PS with respect to $R(Q^*)$.
■

Proposition 4 *For every simultaneous stable matching mechanism g, there exists a profile of preferences R that satisfies independence, reciprocity and no outside option and a profile of rankings Q^* such that Q^* is an NEU of the revelation game $\langle g, R\rangle$ and $g(Q^*)$ is not PS with respect to $R(Q^*)$*

Proof. Let $M = \{m, m', m''\}$, $W = \{w, w', w''\}$ and let:

$$R_{m'}(Q) = \begin{pmatrix} w, w', w'' & \text{if } Q_{w'} = (m', ...) \\ w, w'', w' & \text{otherwise} \end{pmatrix}$$

$$R_{m''}(Q) = \begin{pmatrix} w, w'', w' & \text{if } Q_{w''} = (m'', ...) \\ w, w', w'' & \text{otherwise} \end{pmatrix}$$

$$R_{w'}(Q) = \begin{pmatrix} m, m', m'' & \text{if } Q_{m'} = (w', ...) \\ m, m'', m' & \text{otherwise} \end{pmatrix}$$

$$R_{w''}(Q) = \begin{pmatrix} m, m'', m' & \text{if } Q_{m''} = (w'', ...) \\ m, m', m'' & \text{otherwise} \end{pmatrix}$$

For each $Q \in \mathcal{Q}$, let $R_m(Q) = (w, w', w'')$ and $R_w(Q) = (m, m', m'')$. Clearly, the agents' preferences satisfy *reciprocity*, *no outside option* and *independence*. Consider the following profile of strategies: $Q^*_m = (w, ...)$, $Q^*_{m'} = (w, w', w'')$, $Q^*_{m''} = (w, w'', w')$, $Q^*_w = (m, ...)$, $Q^*_{w'} = (m, m', m'')$, $Q^*_{w''} = (m, m'', m')$. g is a stable matching mechanism and therefore $g(Q^*)$ must be PS with respect to Q^*. It follows that $g(Q^*) = ((m, w), (m', w'), (m'', w''))$. We proceed by showing that Q^* does not involve the use of dominated strategies.

For each $i \in M \cup W$, ranking i's top option first is not dominated since it does strictly better than any other strategy that does not rank i's top option first against a profile of strategies in which i is everyone else's top option. For each $j \in \{m', m''\}$, ranking $g_j(Q^*)$ second does better than any other strategy against a profile of strategies in which m and w rank each other as a top option and each $i \in \{w', w''\}$ ranks j as her top option. The same argument applies for each $l \in \{w', w''\}$. For each $k \in \{m, w\}$, let $Q_k^* = (k^c, ...)$ be any undominated strategy. It follows that Q^* does not involve the use of weakly dominated strategies.

To see that Q^* is an NEU of the revelation game $\langle g, R \rangle$, note that since g is a stable matching mechanism and therefore for each $i \in \{m', m'', w', w''\}$, $g_i\left(Q_i', Q_{-i}^*\right) \in \{g_i(Q^*), i\}$ for any strategy $Q_i' \in \mathcal{Q}_i$, m and w cannot gain by deviating since both attain their first choice. It follows that Q^* is an NEU of $\langle g, R \rangle$ and is blocked by (m', w''). To generalize the proof to a market of any size, add agents such that for each "new" agent $i \in I \in \{M, W\}$, i is j's last option for each $j \in I^c$ and $Q \in \mathcal{Q}$. Let Q_j^* rank i as j's last option and let Q_i^* be any undominated strategy such that Q_i^*,Q_{-i}^* is an NEU of $\langle g, R \rangle$ and is blocked by (m', w''). ■

Proposition 5 *Let R satisfy independence and reciprocity. Then, a SGPU of the revelation game that is induced by the SGS mechanism exists.*

Proof. We construct a profile of strategies and show that it must be a SGPU of the SGS mechanism's induced revelation game. For each $i \in I \in \{M, W\}$ let T_i be agent i's strict liner ordering on $I^c \cup \{i\}$ given that each $j \in I^c$ reports i as first place and let $T = (T_i)_{i \in M \cup W}$. Let μ be a matching that is PS with respect to T. For each $w \in W$ such that $\mu(w) \neq w$ let $Q_w^* = (\mu(w), w, ...)$ and for each $m \in M$ such that $\mu(m) \neq m$ let

$$Q_m^*(Q_W) = \left\{ \begin{array}{cc} \mu(m), m, ... & \text{if } Q_W = (Q_w^*)_{w \in W} \\ R_m(Q_W) & \text{otherwise} \end{array} \right\}$$

For each $m \in M$ such that $\mu(m) = m$ let $Q_m^*(Q_W) = R_m(Q_W)$ and for each $w \in W$ such that $\mu(w) = w$ let Q_w^* be any undominated strategy. Men play a weakly dominant strategy so it is left to verify that women's strategies consist of an SGPU along with the men's strategies. If woman w can deviate and marry $i \in M$ instead of $g_w(Q_W^*, Q_M^*(Q_W))$ then by the construction of $Q_W^*, Q_M^*(Q_W)$, *reciprocity* and *independence* if w ranks i first, i ranks her better than $g_i(Q_W^*, Q_M^*(Q_W))$ but this contradicts the stability of μ with respect to T since by *reciprocity* w and i block μ. if $wR_w(Q_W^*, Q_M^*(Q_W))\ \mu(w)$ then it is again a contradiction to the stability of μ with respect to T. To complete the proof note that since men can condition on women strategies it is straightforward to see that Q_W^* does not include weakly dominated strategies. ■

Proposition 6 *Let R satisfy independence and reciprocity, let g be the SGS mechanism and let $(Q_W^*, Q_M^*(Q_W))$ be an SGPU of the revelation game $\langle g, R\rangle$. Then, $g(Q_W^*, Q_M^*(Q_W))$ is PS with respect to $R(Q_W^*, Q_M^*(Q_W))$.*

Proof. Assume that $g(Q_W^*, Q_M^*(Q_W))$ is not PS with respect to $R(Q_W^*, Q_M^*(Q_W))$. It must be that $g(Q_W^*, Q_M^*(Q_W))$ is blocked by some individual or by a pair. If it is blocked by individual $i \in M \cup W$, then i can do better since g is an IR matching mechanism. Therefore, $g(Q_W^*, Q_M^*(Q_W))$ must be blocked by a pair (m, w). We proceed by describing men's equilibrium behavior in the second stage of g. The strategy $Q_m(Q_W) = R_m(Q_W)$ is a weakly dominant strategy for each $m \in M$. Any other strategy $Q'_m(Q_W)$ that is undominated for m must produce the same spouse as $R_m(Q_W)$ for any profile of strategies $Q_W, Q_{M\setminus\{m\}}(Q_W)$.

Suppose that $(Q_W^*, Q_M^*(Q_W))$ is an SGPU of $\langle g, R\rangle$ and that some pair (m, w) blocks $g(Q_W^*, Q_M^*(Q_W))$. It follows that $wR_m(Q_W^*, Q_M^*(Q_W))\, g_m(Q_W^*, Q_M^*(Q_W))$ and mR_w $(Q_W^*, Q_M^*(Q_W))\, g_w(Q_W^*, Q_M^*(Q_W))$. Since $wR_m(Q_W^*, Q_M^*(Q_W))\, g_m(Q_W^*, Q_M^*(Q_W))$, then either m proposes to w before proposing to $g_m(Q_W^*, Q_M^*(Q_W))$ or that wQ_w^*m. In both cases, consider a deviation of w to the strategy $Q'_w = (m, w, ...)$. By *reciprocity*, $wR_m\left(Q'_w, Q_{-w}^*, Q_M^*(Q_W)\right) g_m(Q_W^*, Q_M^*(Q_W))$. For each $i \in M$, by *independence*, i does not change his preference for women in $W\setminus\{w\}$. One can assume without loss of generality that i uses $Q_i(Q_W) = R_i(Q_W)$ since any other undominated strategy results in the same spouse for him. It follows that m must propose to w before he proposes to $g_m(Q_W^*, Q_M^*(Q_W))$, which makes her deviation profitable, a contradiction of $(Q_W^*, Q_M^*(Q_W))$ being an SGPU of $\langle g, R\rangle$. ■

Proposition 7 *For every IR matching mechanism $g = (a, S)$, there exist a profile of preferences R and a profile of strategies Q^* that form a PSGPU of the psychological game $\langle g, R\rangle$ and $g(Q^*)$ is not PS with respect to $R(Q^*)$.*

Proof. Let $g = (a, S)$ be an IR matching mechanism and let $M = \{m\}$ and $W = \{w\}$. Assume that the MPRM are such that:

$$R_m(Q) = \left\{ \begin{array}{ll} w, m & \text{if } wQ_w m \\ m, w & \text{otherwise} \end{array} \right\}$$

$$R_w(Q) = \left\{ \begin{array}{ll} m, w & \text{if } mQ_m w \\ w, m & \text{otherwise} \end{array} \right\}$$

Let $m \in s_k$ and $w \in s_l$. Without loss of generality, let $l < k$ so that w moves before m. Let $Q_w^* = (w, m)$ and let $Q_m^*(Q_w) = R_m(Q_w)$. Since g is an IR matching mechanism, any deviation by w or m does not change $g(Q_w^*, Q_m^*(Q_w))$. Any other strategy $Q_w \in \mathcal{Q}_w$ cannot dominate Q_w^* since by the IR property of the mechanism for each $i \in w, m$ the only case in which i can be matched to i^c is when it is strictly inferior for i to do so. $g(Q_w^*, Q_m^*(Q_w))$ is blocked by the pair (m, w). To generalize the proof to a market of any size, add agents such that if i is added to the market, then $R_i(Q) = (i, ...)$ for each $Q \in \mathcal{Q}$. ■

References

Abdulkadiroglu, A., P. A. Pathak, A. E. Roth, and T. Sonmez (2005): "The Boston Public School Match," *American Economic Review*, 95(2), 368–371.

——— (2006): "Changing the Boston School Choice Mechanism," (639).

Avery, C., and J. Levin (2010): "Early Admissions at Selective Colleges," *American Economic Review*, 100(5), 2125–56.

Battigalli, P., and M. Dufwenberg (2009): "Dynamic psychological games," *Journal of Economic Theory*, 144(1), 1–35.

Bierbrauer, F., and N. Netzer (2012): "Mechanism design and intentions," ECON - Working Papers econwp066, Department of Economics - University of Zurich.

Dubins, L. E., and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley Algorithm," *The American Mathematical Monthly*, 88(7), pp. 485–494.

Ergin, H., and T. Sonmez (2006): "Games of school choice under the Boston mechanism," *Journal of Public Economics*, 90(1-2), 215–237.

Gale, D., and M. Sotomayor (1985): "Some remarks on the stable matching problem," *Discrete Applied Mathematics*, 11, 223–232.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): "Psychological games and sequential rationality," *Games and Economic Behavior*, 1(1), 60–79.

Glazer, J., and A. Rubinstein (1998): "Motives and Implementation: On the Design of Mechanisms to Elicit Opinions," *Journal of Economic Theory*, 79(2), 157–173.

Gradwohl, R. (2012): "Privacy in Imlementation," *working paper*.

Rabin, M. (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83(5), 1281–1302.

Roth, A. E. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7(4), pp. 617–628.

——— (1984a): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92(6), 991–1016.

——— (1984b): "Misrepresentation and stability in the marriage problem," *Journal of Economic Theory*.

——— (2008): "Deferred acceptance algorithms: history, theory, practice, and open questions," *International Journal of Game Theory*, 36(3), 537–569.

Roth, A. E., and M. Sotomayor (1990): *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press.