Economic Theory (2008) 35: 99–116
DOI 10.1007/s00199-006-0191-6

RESEARCH ARTICLE

Joana Pais

Random matching in the college admissions problem

Received: 6 July 2005 / Revised: 1 December 2006 /
Published online: 12 January 2007

Abstract In the college admissions problem, we consider the incentives confronting agents who face the prospect of being matched by a random stable mechanism. We provide a fairly complete characterization of ordinal equilibria. Namely, every ordinal equilibrium yields a degenerate probability distribution. Furthermore, individual rationality is a necessary and sufficient condition for an equilibrium outcome, while stability is guaranteed in ordinal equilibria where firms act straightforwardly. Finally, we relate equilibrium behavior in random and in deterministic mechanisms.

JEL Classification Number C78

Keywords Matching · College admissions problem · Stability · Random mechanism

1 Introduction

The study of two-sided matching has been mainly devoted to centralized markets. These matching markets work by having each agent of the two sides of the market submit a rank ordered preference list of acceptable matches to a central clearinghouse, which then produces a matching by processing all the preference lists according to some algorithm. Typically, such mechanisms are deterministic in the

I wish to thank an anonymous referee, Flip Klijn, and especially Jordi Massó for their precious comments. I acknowledge the financial support from the Fundação para a Ciência e a Tecnologia under grant SFRH/BD/5308/2001 and from the spanish Ministerio de Ciencia y Tecnología under research grant BEC2002-02130.

J. Pais (✉)
UECE - Research Unit on Complexity and Economics, ISEG/Technical University of Lisbon, Rua Miguel Lupi, 20, 1249-078 Lisboa, Portugal
E-mail: jpais@iseg.utl.pt

sense that the outcome depends on the submitted lists in a way that involves no element of chance. As a consequence, the existing results do not generally allow us to address behavior in many labor markets and other two-sided matching situations where lotteries ultimately determine the outcome. In discrete problems where agents have opposite interests randomization is surely one of the most practical tools to achieve procedural fairness.[1] Hence, equity considerations provide an important justification for the introduction of chance in many instances of centralized matching. On the other hand, lotteries are especially attractive as a means of representing the frictions of a decentralized market. Indeed, in the extremely complex environment of a real life market, decentralized decision making will often lead to uncertain outcomes: the question of who will match with whom depends on the realization of random events—random meetings.

This paper studies a class of matching mechanisms that are random: given agents' behavior, chance determines the final outcome. These mechanisms may be used in centralized markets as a means to promote procedural fairness. Or they may arise in the context of decentralized decision making: starting from an arbitrary matching, agents from the two sides of the market meet bilaterally in a random fashion. We assume that each individual has preferences over the other side of the market and the prospect of being unmatched; however, they are not compelled to behave in a straightforward manner, according to these true preferences. Instead, agents are confronted with a game in which they act in what they perceive to be their own best interest. Hence, upon meeting, the paired agents match if this is consistent with their strategies, and separate otherwise. Since one of the clearest lessons from the study of deterministic procedures is that understanding such incentives is crucial to understand the behavior of the market, the paper is devoted to equilibrium analysis.

Our study was largely motivated by Roth and Vande Vate (1990, 1991). In the context of the marriage problem where matching is one-to-one, Roth and Vande Vate (1990) proved that, starting from an arbitrary matching, the decentralized decision making process of allowing randomly chosen blocking pairs to match will converge to a stable matching with probability one. Under a stable matching no individual or pair of agents has incentives to circumvent the matching. It is argued that such process can be thought of as an approximation to real life dynamics. In the related paper Roth and Vande Vate (1991), strategic considerations are made for the marriage market, focusing on the class of truncation strategies, i.e., strategies that are order-consistent with true preferences, but may regard fewer partners as acceptable. In a one-period game in which every agent states a list of preferences and then a matching stable with respect to those preferences is selected at random, it is shown that all stable matchings can be reached as equilibria in truncations. However, certain unstable matchings can also arise in this way. A multi-period extension is then considered to rule out such undesirable outcomes.

As in Roth and Vande Vate (1991) we assume that random meeting among agents will eventually converge to a stable matching with respect to the chosen strategy profile. Hence, such process induces a lottery exclusively over stable outcomes. However, the present paper extends their contribution in two ways. First,

[1] At least to move towards procedural fairness. A random matching mechanism is procedurally fair whenever the sequence of moves for the agents is drawn from a uniform distribution. See Moulin (1997, 2003).

we take equilibrium analysis further, going beyond the analysis of truncations. A concept of equilibrium based on first-order stochastic dominance is used, given that preferences are ordinal in nature and probability distributions over matchings are to be compared. The notion of ordinal Nash equilibrium guarantees that each agent plays his best response to the others' strategies for every utility representation of the preferences.[2]

Second, the analysis is conducted in the context of the college admissions problem. In this setting, agents belonging to two disjoint sets (henceforth firms and workers) have preferences over the other side of the market; in addition, each firm can employ at most some fixed number of workers, while each worker can fill only one position. Strategic issues in this context have been studied for a deterministic stable matching rule. Roth (1985) shows that no stable matching rule exists that makes it a dominant strategy for all players to report their true preferences. Moreover, he proves that there are equilibrium misrepresentations that generate any individually rational matching with respect to the true preferences.[3] Ma (2002) shows that in order to obtain stability with respect to true preferences, we have to use a refinement of the Nash equilibrium concept and restrict to truncations at the match point (i.e., strategies that preserve the ordering of the true preferences, but rank as unacceptable all the agents that are less preferred than the current match). More precisely, all strong equilibria in truncations at the match point produce stable outcomes. Further, Ma (2002) establishes that every Nash equilibrium profile admits at most one stable matching with respect to the true preferences; if, indeed, such a matching is admitted, it will always be achieved.

In this paper we characterize equilibria arising in the game induced by a random stable matching mechanism, providing simultaneously some results that extend to deterministic mechanisms. First, we show that when ordinal Nash equilibria are considered, a unique matching is obtained as the outcome of the random process. In addition, this outcome is individually rational with respect to the true preferences. Since every individually rational matching for the true preferences can be achieved as an equilibrium outcome, we establish that a matching can be reached at an ordinal Nash equilibrium if and only if it is individually rational for the true preferences. We then turn our attention to equilibria where firms behave straightforwardly. In fact, there are reasons to contemplate truth telling as a salient form of behavior in situations involving uncertainty; further, sophisticated strategic play does not even make sense in settings where firms follow an objective criterion to fill their positions. We prove that, even though workers may not play straightforwardly, stability with respect to the true preferences holds for any matching that results from a play of equilibrium strategies in which firms reveal their true preferences. Conversely, every matching that is stable for the true preferences can be achieved as an equilibrium outcome. In closing, we relate the equilibrium strategy profiles in the games induced by both random and deterministic mechanisms. In particular, for any random stable matching mechanism that always assigns positive probability to two different stable matchings (when they exist), we show that a

[2] This concept was introduced in d'Aspremont and Peleg (1988); it has been used in the context of voting theory in Majumdar and Sen (2004) and in matching markets in Ehlers and Massó (2003), Majumdar (2003), Pais (2004a,b).

[3] For a detailed explanation of these and other results see Roth and Sotomayor (1990), a comprehensive treatment of the matching problem.

strategy profile is an ordinal Nash equilibrium if and only if it has a unique stable matching and it is a Nash equilibrium in the game induced by some deterministic stable mechanism.

We proceed as follows. In Sect. 2 we present the college admissions problem and introduce notation. We describe random matching mechanisms and the equilibrium concept used in Sect. 3. In Sect. 4 we turn our attention to individual decision making. The matching process is modeled as a one-period game and its equilibria are then characterized. In Sect. 5 we briefly discuss equilibria in the context of a sequential game. Some concluding remarks follow in Sect. 6.

2 The model

The agents in the college admissions problem are two finite and disjoint sets, the set $W = \{w_1, \ldots, w_p\}$ of workers and the set $F = \{f_1, \ldots, f_n\}$ of firms. We let $V = W \cup F$ and sometimes refer to a generic agent by v, while w and f represent a generic worker and firm, respectively. Each worker w can work for at most one firm and each firm f has a *quota* q_f, the maximal number of workers it may employ.

Each worker w has a complete, transitive, and strict preference relation P_w over the set $F \cup \{w\}$. For example, the preferences of w on $\{f_1, f_2, f_3, f_4\} \cup \{w\}$ can be represented by $P_w : f_1, f_2, w, f_3, f_4$, indicating that the best firm for w is f_1, his second choice is f_2, and he prefers being unemployed than working for either f_3 or f_4. Each firm f also has a complete, transitive, and strict preference relation P_f over the set $W \cup \{f\}$. For example, the preferences of f on $\{w_1, w_2, w_3, w_4\} \cup \{f\}$ can be represented by $P_f : w_3, w_1, f, w_2, w_4$, indicating that the best worker for f is w_3, its second choice is w_1, and it prefers having a position unfilled to hiring any other worker. A worker is *acceptable* if the firm prefers to employ him rather than having a position unfilled. Formally, the set of acceptable workers for f is $A(P_f) = \{w \in W : w P_f f\}$. Given P_w, we can similarly define an acceptable firm and the set of acceptable firms for w as $A(P_w) = \{f \in F : f P_w w\}$. In the above examples, the set of acceptable workers for f is $A(P_f) = \{w_1, w_3\}$ and the set of acceptable firms for w is $A(P_w) = \{f_1, f_2\}$. We let $P = (P_{f_1}, \ldots, P_{f_n}, P_{w_1}, \ldots, P_{w_p})$ denote the profile of all agents' preferences; we sometimes write it as $P = (P_v, P_{-v})$ where P_{-v} is the set of preferences of all agents other than v. We let $\mathcal{P}_v$ denote the set of all possible preference relations for agent v and let $\mathcal{P} = \prod_{v \in V} \mathcal{P}_v$ be the set of admissible preference profiles. We write $v' P_v v''$ when v' is preferred to v'' under preferences P_v and we say that v *prefers* v' to v''. Since agents will have to compare two potential partners v' and v'' that may actually be the same, we write $v' R_v v''$ to denote that either $v' = v''$ or else $v' P_v v''$. In this case, we say that v likes v' *at least as well as* v''. The set of agents that v likes at least as well as v'' is $U_{P_v}(v'') = \{v' \in V : v' R_v v''\}$.

Each firm with quota greater than one must be able to compare groups of workers. Following Roth (1985), we assume firms' preferences over groups of workers are *responsive* to the preferences over single agents. A preference $\bar{P}_f$ for f over sets of workers is responsive to its preference P_f over single workers if, for all $S \in 2^W$ such that $|S| < q_f$,

1. For all $w, w' \in W \backslash S$, $S \cup \{w\} \bar{P}_f S \cup \{w'\}$ if and only if $w P_f w'$;
2. For all $w \in W \backslash S$, $S \cup \{w\} \bar{P}_f S$ if and only if $w P_f f$;

and for all $S \in 2^W$ such that $|S| > q_f$, $\emptyset \bar{P}_f S$.[4]

Responsive preferences are assumed throughout the paper.

Since firms may have to compare two groups of workers S and S' that may actually be the same, we use $\bar{R}_f$, a responsive extension of R_f. We write $S \bar{R}_f S'$ to denote that either $S = S'$ or else $S \bar{P}_f S'$. We let $U_{\bar{P}_f}(S) = \{S' \in 2^W : S' \bar{R}_f S\}$ denote the set of groups of workers f likes at least as well as S.

An outcome for the college admissions problem (F, W, P) is a *matching* , a mapping μ from the set V into $2^W \cup V$ satisfying the following:

1. For all $w \in W$, either $\mu(w) \in F$ or else $\mu(w) = w$;
2. For all $f \in F$, $|\mu(f)| \le q_f$ and $\mu(f) \in 2^W$;
3. For all $(w, f) \in W \times F$, $\mu(w) = f$ if and only if $w \in \mu(f)$.

Observe that, while a worker may be matched to a firm or to himself—the latter meaning being unmatched— a firm is always matched to a subset of workers and being matched to the empty set stands for being unmatched. We denote the set of all matchings by $\mathcal{M}$.

We can extend preferences over partners to preferences over matchings in the following, natural, way: each worker's preferences over matchings correspond precisely to his preferences over his own assignments at the matchings; similarly, firms' preferences over matchings are tantamount to the preferences over its assignments. For instance, w prefers μ to μ' when $\mu(w) P_w \mu'(w)$, while f prefers μ to μ' if $\mu(f) \bar{P}_f \mu'(f)$.

A matching μ is *individually rational* if, for every $w \in W$, $\mu(w) R_w w$ and if, for every firm f and w in $\mu(f)$, $w P_f f$.[5] A firm f and a worker w are a *blocking pair* for μ if they are not matched under μ but prefer one another to one of their assignments, i.e., $w \notin \mu(f)$ but $f P_w \mu(w)$, $w P_f f$, and either (i) $|\mu(f)| < q_f$ or (ii) if $|\mu(f)| = q_f$ then there exists $w' \in \mu(f)$ such that $w P_f w'$. A matching μ is *stable* if it is individually rational and if there is no blocking pair for μ. Note that the stability of μ depends on preferences over individuals, irrespective of the responsive extension that is being used. We let $IR(P)$ and $S(P)$ denote the set of all individually rational and the set of all stable matchings respectively with respect to a profile P. A firm f and a worker w are *achievable* for each other if f and w are matched under some stable matching.

The proof of existence of stable matchings in Gale and Shapley (1962) is constructed by means of the deferred-acceptance algorithm. For a given preference profile P, proposals are issued by one side of the market accordingly, while the other side merely reacts to such offers by rejecting all but the best in P. In the case that firms make job offers, the algorithm arrives at the firm-optimal stable matching $\mu^F[P]$, with the property that all firms are in agreement that it is the best stable matching. The deferred-acceptance algorithm with workers proposing produces the worker-optimal stable matching $\mu^W[P]$ with corresponding properties. Further, the optimal stable matching for one side of the market is the worst stable matching for every agent on the other side of the market, a result presented in Knuth (1976) but attributed to John Conway.

[4] Note that, while $\bar{P}_f$ is used to compare sets of workers, namely the empty set, P_f compares single workers and f itself, the latter representing having an unfilled position.

[5] By responsiveness, the latter requirement is equivalent to $\mu(f) \bar{R}_f S$, for every $S \subseteq \mu(f)$.

3 Random matching and ordinal Nash equilibria

Many matching markets do not employ centralized procedures. Agents are free to issue offers and make acceptations and rejections as they please and matching is performed over the telephone network, using the mail, or through the Internet. In such environments, randomness determines the order in which agents communicate: it may depend on which telephone call goes through, on the speed of the mail, or on how fast firms react to eventual proposals. When a central clearinghouse does exist, chance is widely used to restore procedural fairness—any deterministic mechanism is bound to favor a subset of the agents involved. In two-sided matching markets, the need for compromise solutions is especially intense given the strong polarization of interests of agents reflected in the structure of the set of stable matchings. Some real life applications of random procedures concern allocation problems as on-campus housing, namely in American universities, or public housing.[6] Student placement mechanisms that assign students to colleges are another example of mechanisms where randomness plays a role, as well as procedures used to match students to optional courses or even children to summer camps.[7] Finally, randomness is present in any matching mechanism where the position in a queue or the order of arrival may influence assignments.

Formally, a *random matching rule* $\tilde{\varphi}$ is a mapping from preference profiles to lotteries over the set of matchings: $\tilde{\varphi} : \mathcal{P} \longrightarrow \triangle\mathcal{M}$. A *random matching* $\tilde{\varphi}[Q]$ is the image of a preference profile Q under a random matching rule, i.e., a lottery over matchings. Throughout the paper, we consider only *random stable matching rules* by restricting the range of random matching rules to the set of lotteries whose supports are subsets of the sets of stable matchings, i.e., we consider $\tilde{\varphi}$ such that, for every Q in $\mathcal{P}$, the support of $\tilde{\varphi}[Q]$, denoted by supp $\tilde{\varphi}[Q]$, is included in $S(Q)$. While $\tilde{\varphi}[Q]$ denotes a lottery over matchings, we let $\tilde{\varphi}_v[Q]$ represent the probability distribution induced over agent v's achievable matches. Whenever the probability distribution $\tilde{\varphi}[Q]$ is degenerate, we abuse the notation slightly by letting $\tilde{\varphi}[Q]$ denote the unique outcome matching; similarly, if the distribution $\tilde{\varphi}_v[Q]$ is degenerate for some agent v, $\tilde{\varphi}_v[Q]$ denotes v's unique match in the random stable matching $\tilde{\varphi}[Q]$. Observe however that in general supp $\tilde{\varphi}[Q]$ is a subset of the set of stable matchings $S(Q)$. In contrast, a *deterministic matching rule* φ is a function from preference profiles to matchings: $\varphi : \mathcal{P} \longrightarrow \mathcal{M}$. We consider only *deterministic stable matching rules* that produce a unique stable matching $\varphi[Q]$ for every profile of preferences Q. In particular, φ^F and φ^W denote the deterministic stable matching rules that yield the firm-optimal $\mu^F[Q]$ and the worker-optimal $\mu^W[Q]$ stable matchings, respectively, for every Q in $\mathcal{P}$. Finally, we let $\varphi_v[Q]$ denote v's partner under the matching $\varphi[Q]$.

In a matching market (F, W, P), we consider the game induced by a random stable matching rule $\tilde{\varphi}$ in which agents are each faced with the decision of what strategies to act on. As a first approach, we examine a one-period game where the strategy space of player v in the game is the set of all possible preference lists $\mathcal{P}_v$. Given the true preference ordering P_v, each player v may eventually reveal a different order Q_v over the players on the other side of the market, and then a matching μ stable with respect to the stated preferences Q is selected at random among all

[6] See Abdulkadiroglu and Sönmez (1999).

[7] See Abdulkadiroglu and Sönmez (2003) for a description of student assignment mechanisms.

the potential matchings, i.e., the elements of supp $\tilde{\varphi}[Q]$. To be precise, we consider the mechanism $(\mathcal{P}, \tilde{\varphi})$, where $\mathcal{P}$ is the set of admissible strategy profiles and $\tilde{\varphi}$ is a random stable matching rule; we refer to $(\mathcal{P}, \tilde{\varphi})$ as a random stable matching mechanism. Once the preferences P of the agents are specified, the above mechanism induces the game $(\mathcal{P}, \tilde{\varphi}, P)$. Analogously, $(\mathcal{P}, \varphi)$ is a deterministic stable matching mechanism that induces the game $(\mathcal{P}, \varphi, P)$. In Sect. 5, we discuss an extension of the obtained results to a more complex setting where agents' strategy spaces are broader.

In the game $(\mathcal{P}, \tilde{\varphi}, P)$, agents compare probability distributions over matchings when deciding which strategic course to take. Since preferences are ordinal, there is no natural utility representation of these preferences for expected utility calculations. It follows that to address strategic questions we need to develop ideas about what constitutes a "best decision" to be taken by an agent. With this purpose in mind, let $\hat{Q}$ be a strategy profile and consider $w \in W$. Let $\tilde{\varphi}_w[\hat{Q}](S)$ be the probability that w obtains a partner in $S \subset F \cup \{w\}$ when the profile $\hat{Q}$ is used in the game $(\mathcal{P}, \tilde{\varphi}, P)$; in particular, let $\tilde{\varphi}_w[\hat{Q}](U_{P_w}(v))$ be the probability that w is matched to a partner at least as good as v when the profile $\hat{Q}$ is used in $(\mathcal{P}, \tilde{\varphi}, P)$. Given a random stable matching rule $\tilde{\varphi}$ and given $\hat{Q}_{-w}$, we say that the strategy Q_w *stochastically P_w-dominates Q'_w* if, for all $v \in F \cup \{w\}$, $\tilde{\varphi}_w[Q_w, \hat{Q}_{-w}](U_{P_w}(v)) \geq \tilde{\varphi}_w[Q'_w, \hat{Q}_{-w}](U_{P_w}(v))$. This means that, for all $v \in F \cup \{w\}$, the probability of w being assigned to v or to a strictly preferred agent is higher under $\tilde{\varphi}_w[Q_w, \hat{Q}_{-w}]$ than under $\tilde{\varphi}_w[Q'_w, \hat{Q}_{-w}]$. Similarly, given $\tilde{\varphi}$ and given $\hat{Q}_{-f}$, we say that the strategy Q_f *stochastically P_f-dominates Q'_f* if, for all $S \in 2^W$ and for every responsive extension $\bar{P}_f$ of P_f, we have $\tilde{\varphi}_f[Q_f, \hat{Q}_{-f}](U_{\bar{P}_f}(S)) \geq \tilde{\varphi}_f[Q'_f, \hat{Q}_{-f}](U_{\bar{P}_f}(S))$. This means that f is not able to increase the probability of obtaining any set of workers S' (with whom it may end up matched) and all sets ranked higher than S' in its list of preferences $\bar{P}_f$, when using Q'_f instead of Q_f. Hence, if we consider the problem that agent v faces given the strategy choices $\hat{Q}_{-v}$ of the other players, a particular strategy choice Q_v may be preferred if it stochastically dominates every other alternative strategy. This provides the basis for the solution concept we will adopt throughout the paper.

Definition 1 The profile of strategies Q is an *ordinal Nash equilibrium* (*ON equilibrium*) in the game $(\mathcal{P}, \tilde{\varphi}, P)$ if, given Q_{-v}, Q_v stochastically P_v-dominates every alternative strategy Q'_v for every agent v.

It follows from the above definition that Q is an ordinal Nash equilibrium when no agent v can gain in expected utility terms by unilaterally deviating from Q_v, no matter what utility function is used to represent its true preferences. We will then be concerned in finding a profile of strategies Q that is a Nash equilibrium for every utility representation of agents' preferences.

4 Equilibrium analysis

We now turn to characterize ordinal Nash equilibria in the game induced by a random stable mechanism. Proposition 1 asserts that no ordinal equilibrium supports more than one stable matching. Using the decentralized interpretation, we can say that the outcome in equilibrium is immune to the order in which agents meet when players behave strategically, even though truth revealing often leads to a lottery over matchings. Agents manipulate to protect themselves against uncertainty.

Proposition 1 *Let Q be an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. Then, a single matching is obtained with probability one.*

Proof By contradiction, assume that Q is an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$ and $|\text{supp } \tilde{\varphi}[Q]| \geq 2$. Then, there exists a worker $w \in W$ and matchings μ, $\hat{\mu} \in \text{supp } \tilde{\varphi}[Q]$ such that $\mu(w) \neq \hat{\mu}(w)$. Let $\mu'(w)$ be the best match among all given by the elements of supp $\tilde{\varphi}[Q]$, i.e., $\mu'(w) R_w \mu(w)$, for all $\mu \in \text{supp } \tilde{\varphi}[Q]$. Let Q'_w be such that $A(Q'_w) = \{\mu'(w)\}$ and let $Q' = (Q'_w, Q_{-w})$. Note that μ' is stable for Q and, once w changes his strategy, it remains stable for Q' (it remains individually rational and no blocking pairs emerge). Further, since the set of matched agents is the same under every stable matching, w is matched to $\mu'(w)$ under every matching in $S(Q')$. Then, $1 = \tilde{\varphi}_w[Q'](U_{P_w}(\mu'(w))) > \tilde{\varphi}_w[Q](U_{P_w}(\mu'(w)))$ and Q_w does not stochastically P_w-dominate Q'_w. It follows that Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. $\square$

As a consequence, in the particular case that the random matching rule always assigns positive probability to at least two different matchings (if such matchings exist), the set of stable matchings of each ordinal Nash equilibrium is a singleton. In general, however, the set of stable matchings of an ordinal Nash equilibrium may contain several elements. As proved in Ma (2002) for a deterministic stable matching rule, the random stable rule then chooses the matching that is unanimously preferred among all the stable matchings with respect to the submitted profile.

Lemma 1 *Let Q be an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. Then, for any matching $\mu \in S(Q)$,*

1. *$\tilde{\varphi}_w[Q] R_w \mu(w)$ for every $w \in W$ and*
2. *$\tilde{\varphi}_f[Q] \bar{R}_f \mu(f)$ for every $f \in F$ and every responsive extension $\bar{R}_f$ of R_f.*

Proof By Proposition 1, $\tilde{\varphi}[Q]$ is degenerate. The result then follows from Lemma 6 in Ma (2002). $\square$

For illustration, consider the following example.

Example 1 Let $F = \{f_1, f_2\}$, $W = \{w_1, w_2\}$, and $q_{f_1} = q_{f_2} = 1$. Suppose that the true preferences are as follows:

$$P_{w_1} : f_1,\ f_2,\ w_1 P_{f_1} : w_1,\ w_2,\ f_1$$
$$P_{w_2} : f_2,\ f_1,\ w_2 P_{f_2} : w_2,\ w_1,\ f_2.$$

Let $Q_{w_1} : f_2, f_1, w_1$ and $Q_{w_2} : f_1, f_2, w_2$ and note that the preference profile $Q = (Q_{w_1}, Q_{w_2}, P_F)$ is an ordinal Nash equilibrium in $(\mathcal{P}, \varphi^F, P)$, the game induced by the mechanism that yields the firm-optimal stable matching. Now let $\tilde{\varphi}$ be a random matching rule that assigns probability 0.5 to both the worker-optimal and firm-optimal stable matchings. Clearly, the support of the probability distribution induced by $\tilde{\varphi}[Q]$ includes both $\mu^F[Q] = \{(f_1, w_1), (f_2, w_2)\}$ and

$\mu^W[Q] = \{(f_1, w_2), (f_2, w_1)\}$. By Proposition 1, Q is not an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. In fact, every worker can successfully deviate. For example, by using his true preferences, w_1 obtains his preferred firm f_1 with probability one.

In the context of deterministic mechanisms, Roth (1985) shows that by suitably falsifying their preferences, agents can induce any individually rational matching with respect to the true preferences. Unfortunately, this is not a very illuminating result: the set of individually rational matchings includes all the matchings that are remotely plausible. Moreover, the possibility of sustaining matchings where agents hold non-acceptable partners is not ruled out, although individual rationality appears to be a minimum requirement for an equilibrium outcome.

The results that follow establish that μ can be supported as an ordinal equilibrium if and only if it is individually rational. Hence, we provide a complete characterization of ordinal Nash equilibria outcomes in the game induced by random stable mechanisms. Furthermore, it can easily be shown that Proposition 3 can be extended to the deterministic case, providing a necessary condition for Nash equilibria in games induced by deterministic stable matching mechanisms.

Proposition 2 *Let μ be any individually rational matching for (F, W, P) and let $\tilde{\varphi}$ be a random stable matching rule. Then, there exists an ordinal Nash equilibrium Q that supports μ in the game $(\mathcal{P}, \tilde{\varphi}, P)$.*

Proof Let Q_w be such that $A(Q_w) = \{\mu(w)\}$, for every $w \in W$, and let Q_f be such that $A(Q_f) = \mu(f)$, for every $f \in F$. Clearly, $S(Q) = \{\mu\}$ and μ is reached with probability one. Moreover, no agent can profitably deviate. To see this, take an arbitrary worker w. If $\mu(w) \in F$, the only agent that accepts w is $\mu(w)$. Hence, w faces the choice of holding $\mu(w)$ or being unmatched. Since $\mu(w) P_w w$ by individual rationality of μ, w has no profitable deviation. If $\mu(w) = w$, no firm is willing to hire w, so that w has no profitable deviation: his only alternative is to remain unmatched. Now consider $f \in F$. If $\mu(f) \neq \emptyset$, only those workers in $\mu(f)$ are willing to accept filling a position in f. Moreover, by individual rationality of μ, $\mu(f) \overline{R}_f S$, for every $S \subseteq \mu(f)$. If $\mu(f) = \emptyset$, no worker accepts filling a position if f. In neither case can f improve upon $\mu(f)$ by deviating. Hence, Q is an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. □

Proposition 3 *Let Q be an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. Then, the unique equilibrium outcome $\tilde{\varphi}[Q]$ is individually rational for the true preferences P.*

Proof By Proposition 1, a degenerate probability distribution is achieved in any equilibrium play of $(\mathcal{P}, \tilde{\varphi}, P)$. Let us say $\tilde{\varphi}[Q] = \mu$. We will prove that μ is individually rational.

First, by contradiction, assume there exists a worker w such that $w P_w \mu(w)$. Suppose that, instead of acting according to Q_w, w uses the strategy Q'_w such that $A(Q'_w) = \emptyset$ and define $Q' = (Q'_w, Q_{-w})$. By considering every firm unacceptable, w is alone under every matching in $S(Q')$. Hence, $1 = \tilde{\varphi}_w[Q'](U_{P_w}(w)) > \tilde{\varphi}_w[Q](U_{P_w}(w))$ and Q_w does not stochastically P_w-dominate Q'_w. It follows that Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$.

Now suppose that there is a firm f and a set of workers $S^G \subsetneq \mu(f)$ such that $S^G \bar{P}_f \mu(f)$. Let S^G be, among all the subsets of $\mu(f)$, the one that is preferred by f. Consider Q'_f, an alternative strategy for f, where only the elements of S^G are considered acceptable. We will show that Q_f does not stochastically P_f-dominate Q'_f.

To start, consider the matching μ' such that $\mu'(f) = S^G$ and $\mu'(\hat{f}) = \mu(\hat{f})$, for every $\hat{f} \neq f$. Let $S^B = \mu(f) \backslash S^G$ (note that $S^B \neq \emptyset$) and $Q' = (Q'_f, Q_{-f})$. Now consider the matching market $(F, W \backslash S^B, Q'^R)$, where Q'^R is the same profile as Q', but restricted to $W \backslash S^B$. We will prove that μ' is stable for Q'^R in this reduced market. Note that, when Q is considered, $\mu(w)$ is acceptable for every worker w, all elements in $\mu(\bar{f})$ are acceptable for every firm $\bar{f} \neq f$, and S^G is the preferred subset of $\mu(f)$ for f. It follows that μ' is individually rational for Q'^R. Now suppose that $(\hat{f}, w)$ blocks μ', i.e., $w \notin \mu'(\hat{f})$, but $\hat{f} Q'^R_w \mu'(w)$, $w Q'^R_{\hat{f}} \hat{f}$, and either (i) $|\mu'(\hat{f})| < q_{\hat{f}}$ or (ii) if $|\mu'(\hat{f})| = q_{\hat{f}}$ then there exists $w' \in \mu'(\hat{f})$ such that $w Q'^R_{\hat{f}} w'$. Since only the elements of $\mu'(f)$ are considered acceptable in $Q'^R_{\hat{f}}$, we must have $\hat{f} \neq f$. Hence, $Q'^R_{\hat{f}} = Q^R_{\hat{f}}$, where $Q^R_{\hat{f}}$ is the same strategy as $Q_{\hat{f}}$, but restricted to $W \backslash S^B$. By definition of μ', we have $\mu'(\hat{f}) = \mu(\hat{f})$, for every $\hat{f} \neq f$, and $\mu'(w) = \mu(w)$, for every $w \in W \backslash S^B$. The above expression thus becomes $\hat{f} Q^R_w \mu(w)$, $w Q^R_{\hat{f}} \hat{f}$, and either (i) $|\mu(\hat{f})| < q_{\hat{f}}$ or (ii) if $|\mu(\hat{f})| = q_{\hat{f}}$ then there exists $w' \in \mu(\hat{f})$ such that $w Q^R_{\hat{f}} w'$. Hence, in the unrestricted market, $\hat{f} Q_w \mu(w)$, $w Q_{\hat{f}} \hat{f}$, and either (i) or (ii) holds with $w Q_{\hat{f}} w'$, for some $w' \in \mu(\hat{f})$. This means that $(\hat{f}, w)$ blocks μ under Q, contradicting $\mu \in S(Q)$. Thus, μ' is stable in $(F, W \backslash S^B, Q'^R)$. Note that, since f is matched to S^G under a stable matching, it must hold exactly S^G under the firm-optimal stable matching for $(F, W \backslash S^B, Q'^R)$, by definition of Q'^R_f and of the firm-optimal stable matching.

Suppose S^B join in. By Theorem 5.35 in Roth and Sotomayor (1990), every firm must be at least as well off in the new firm-optimal stable matching. Since only S^G are considered acceptable by f in the strategy Q'_f, f cannot improve upon S^G. Thus, it must be matched to S^G under the firm-optimal stable matching of the market (F, W, Q').

Finally, notice that since $|\mu(f)| \leq q_f$ and $S^B \neq \emptyset$, we have $|S^G| < q_f$. Hence, Theorem 5.13 in Roth and Sotomayor (1990) guarantees that f must hold the same workers under every stable matching in (F, W, Q'). Therefore, by deviating and acting according to Q'_f, f will get S^G with probability one instead of $\mu(f)$. Concluding, $1 = \tilde{\varphi}_f[Q'](U_{\bar{P}_f}(S^G)) > \tilde{\varphi}_f[Q](U_{\bar{P}_f}(S^G))$ and Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. $\square$

The above result is as uninformative as large the set of individually rational matchings may be. Ma (2002) shows that one way to make a sharper prediction of equilibrium outcomes and guarantee stability is to go as far as refining the notion of Nash equilibrium to strong Nash and require the use of a particular kind of strategies: truncations at the match point (i.e., deleting the $(m+1)$th and less preferred partners when matched to the mth choice). We provide a different sufficient condition for stability in the game induced by a random stable mechanism: every ordinal Nash equilibrium where firms behave straightforwardly is stable for the true preferences. Truth telling by firms is natural in markets where firms obey some kind of objective criterion to fill their positions (e.g., universities admit students on the basis of examination scores, student placement mechanisms assign students to public schools according to the area of residence, firms hire workers according to scores given by recruiting agencies). Moreover, in situations involving uncertainty agents may have no clue about the form that effective strategies might have and straightforward behavior is always an easy resort.

Proposition 4 *Let $Q = (P_F, Q_W)$ be an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. Then, the unique equilibrium outcome $\tilde{\varphi}[Q]$ is stable for the true preferences P.*

Proof By Proposition 1, a unique matching is achieved as the outcome of an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. Let us say that $\tilde{\varphi}[Q] = \mu$. By Proposition 3, $\mu \in IR(P)$. We will prove that $\mu \in S(P)$ by contradiction. Suppose that (f, w) blocks μ when the true preferences are considered, i.e., $w \notin \mu(f)$ but $f P_w \mu(w)$, $w P_f f$, and either (i) $|\mu(f)| < q_f$ or (ii) if $|\mu(f)| = q_f$ then there exists $w' \in \mu(f)$ such that $w P_f w'$.

Consider Q'_w, an alternative strategy for w, such that $f Q'_w v'$ and $v Q'_w v'$ if and only if $v Q_w v'$, for every $v, v' \in F\backslash\{f\} \cup \{w\}$. Let $Q' = (Q'_w, Q_{-w})$. By stability under Q', if w is not matched to f with positive probability under $\tilde{\varphi}[Q']$—so that we cannot show Q'_w is a deviation for w—then, under every matching in $\tilde{\varphi}[Q']$, each position of f is filled with a worker f finds better than w. We will prove that, in this case, f has a profitable deviation when the other agents use Q_{-f}, so that Q is not an ON equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$.

Let μ' be such that $\mu'(f) \bar{R}_f \hat{\mu}(f)$, for every $\hat{\mu} \in \operatorname{supp} \tilde{\varphi}[Q']$ and let Q_f be such that $A(Q_f) = \mu'(f)$. Note that $|A(Q_f)| = q_f$. Let $\hat{Q} = (Q_f, Q_{-f})$. We will show that f will get $\mu'(f)$ under every matching in $\tilde{\varphi}[\hat{Q}]$, so that Q_f is a profitable deviation for f.

By contradiction, assume μ' is not stable for $\hat{Q}$. Since the definitions of $\hat{Q}$ and of Q'_w ensure $\mu' \in IR(\hat{Q})$, this implies that there exists a blocking pair $(\tilde{f}, \tilde{w})$ for μ' when $\hat{Q}$ is considered. As, by definition of Q_f, we have $\tilde{f} \neq f$, this means $\tilde{f} \hat{Q}_{\tilde{w}} \mu'(\tilde{w})$, $\tilde{w} P_{\tilde{f}} \tilde{f}$, and either (i) $|\mu'(\tilde{f})| < q_{\tilde{f}}$ or (ii) if $|\mu'(\tilde{f})| = q_{\tilde{f}}$ then there exists $\bar{w} \in \mu'(\tilde{f})$ such that $\tilde{w} P_{\tilde{f}} \bar{w}$. As a consequence, $(\tilde{f}, \tilde{w})$ blocks μ' for Q' unless $\tilde{w} = w$ and $\mu'(\tilde{w}) = f$. But this contradicts the assumption that, under μ', each position of f is filled with a worker f finds better than w. Hence, $\mu' \in S(\hat{Q})$. It follows from $|\mu'(f)| = q_f$ and from Theorem 5.12 in Roth and Sotomayor (1990) that f has all positions filled under every matching in $S(\hat{Q})$; as the set of agents f finds acceptable in Q_f is exactly $\mu'(f)$, f is matched to $\mu'(f)$ under every matching in $S(\hat{Q})$ and, in particular, under every matching in $\tilde{\varphi}[\hat{Q}]$. □

Two remarks are in order. First, this result can easily be applied to games arising from deterministic stable mechanisms: stability for the true preferences is obtained in any Nash equilibrium where firms are truthful. Second, in accordance with the claims in Roth and Sotomayor (1990) concerning deterministic mechanisms, the analogous result with workers telling the truth and firms acting strategically does not hold, although it would hold when all quotas equal one.[8] The college admissions problem, unlike the marriage problem, is not symmetric between the two sides of the market and there are substantial differences between the two when strategic issues are contemplated. Any firm with a quota greater than one resembles something like a coalition rather than an individual. Hence, allowing for manipulation on the firms' side is similar to giving such powers to sets of agents in a marriage market and, in equilibria where workers tell the truth, stability is lost.

The converse result is given in Proposition 5, asserting that every stable matching for the true preferences can be supported as the outcome of an ordinal Nash equilibrium where firms act according to the true preferences. In fact, workers can compel any jointly achievable outcome in the game induced by a random stable mechanism, while firms behave straightforwardly.

Proposition 5 *Let μ be any stable matching for (F, W, P) and let $\tilde{\varphi}$ be a random stable matching rule. Then, there exists an ordinal Nash equilibrium $Q = (P_F, Q_W)$ that supports μ in the game $(\mathcal{P}, \tilde{\varphi}, P)$.*

Proof Define Q_w such that $A(Q_w) = \{\mu(w)\}$ for every $w \in W$. Clearly, $S(Q) = \{\mu\}$ and μ is reached with probability one.

Let us now prove that Q is an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. Take an arbitrary worker w and suppose that there exists a firm f such that $f P_w \mu(w)$. We claim that w cannot deviate to get matched to f. In fact, the stability of μ with respect to P implies that either $f P_f w$—in which case f declares w unacceptable—or, if $w P_f f$, then $|\mu(f)| = q_f$ and $w' P_f w$, for every $w' \in \mu(f)$. In the latter case, since $\mu(w') = f$ for every $w' \in \mu(f)$, then $Q_{w'}$ satisfies $A(Q_{w'}) = \{f\}$ and f ends up matched to $\mu(f)$. Now consider firm f. The only workers willing to accept f are those in $\mu(f)$. Furthermore, individual rationality of μ implies that $\mu(f) \bar{R}_f S$, for every $S \subseteq \mu(f)$. It follows that f cannot improve upon $\mu(f)$ by deviating. In conclusion, Q is an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. □

Our next results establish a strong link between equilibria in games induced by random and by deterministic stable mechanisms. We start by pointing out that every ordinal Nash equilibrium of the random process must be a simple Nash equilibrium of a game induced by some mechanism where chance plays no role.

Proposition 6 *Let Q be an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$. Then, there exists a deterministic stable matching rule φ such that Q is a Nash equilibrium in the game $(\mathcal{P}, \varphi, P)$.*

Proof Assume that Q is an ON equilibrium that yields μ in $(\mathcal{P}, \tilde{\varphi}, P)$. Proposition 1 guarantees that μ is the only element in supp $\tilde{\varphi}[Q]$ and, by Proposition 3, $\mu \in IR(P)$. Now suppose, by contradiction, that there exists no game induced by a deterministic stable matching rule φ where Q is a Nash equilibrium. In particular, consider any φ such that $\varphi[Q] = \mu$—such a rule exists since $\mu \in S(Q)$—and assume that some agent has a profitable deviation.

Let such agent be a worker, w. Then, there exists a strategy Q'_w such that $\varphi_w[Q'] \; P_w \mu(w)$, with $Q' = (Q'_w, Q_{-w})$. This implies that $\varphi_w[Q'] \in F$ since

[8] See Roth (1985).

$\mu \in IR(P)$. Let $f = \varphi_w[Q']$ and define Q''_w such that $A(Q''_w) = \{f\}$. Observe that under any matching in $S(Q''_w, Q_{-w})$, w is matched to f—$\varphi[Q'] \in S(Q''_w, Q_{-w})$ since it remains individually rational and no blocking pairs emerge once w uses Q''_w. Therefore, under every matching in supp $\tilde{\varphi}[(Q''_w, Q_{-w})]$, w holds f and Q_w does not stochastically P_w -dominate Q''_w. We get a contradiction: Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$.

Now assume that $f \in F$ can profit by deviating from Q_f in $(\mathcal{P}, \varphi, P)$. This means that there exists Q'_f such that $\varphi_f[Q']\bar{P}_f\mu(f)$, with $Q' = (Q'_f, Q_{-f})$. Since $\mu \in IR(P)$, $\varphi_f[Q'] \neq \emptyset$. Define Q''_f such that only the workers in $\varphi_f[Q']$ are considered acceptable. Since $\varphi[Q'] \in S(Q')$, once only the workers in $\varphi_f[Q']$ are considered acceptable by f, we can guarantee that $\varphi[Q'] \in S(Q'')$. The definition of Q''_f and the fact that under every stable matching firms have the same number of positions filled (Theorem 5.12 in Roth and Sotomayor 1990) imply that f holds $\varphi_f[Q']$ in every element of $S(Q'')$. Therefore, $1 = \tilde{\varphi}_f[Q''](U_{\bar{P}_f}(\varphi_f[Q'])) > \tilde{\varphi}_f[Q](U_{\bar{P}_f}(\varphi_f[Q'])) = 0$ and Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$. □

In Proposition 7, we establish a partially converse statement: the set of ordinal Nash equilibria in the game induced by a random stable mechanism includes all the strategy profiles that are simultaneously equilibria in the games induced by the rules that yield the firm-optimal and the worker-optimal stable matchings.

Proposition 7 *Let Q be a Nash equilibrium in both $(\mathcal{P}, \varphi^F, P)$ and $(\mathcal{P}, \varphi^W, P)$. Then, Q is an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$ for any random sable matching rule $\tilde{\varphi}$.*

The following Lemma is useful in proving Proposition 7.

Lemma 2 *Let Q be a Nash equilibrium in both $(\mathcal{P}, \varphi^F, P)$ and $(\mathcal{P}, \varphi^W, P)$. Then, the set $S(Q)$ is a singleton.*

Proof Assume that Q is a Nash equilibrium in both $(\mathcal{P}, \varphi^F, P)$ and $(\mathcal{P}, \varphi^W, P)$. Suppose, by contradiction, that $|S(Q)| \geq 2$. Clearly, this implies that $\varphi^F[Q] \neq \varphi^W[Q]$. Lemma 1 in Ma (2002) implies that, for any matching $\mu \in S(Q)$, we have $\varphi^F_w[Q]R_w\mu(w)$, for every $w \in W$. Since Q is an equilibrium in $(\mathcal{P}, \varphi^W, P)$, the same lemma guarantees that $\varphi^W_w[Q]R_w\mu(w)$, for every $w \in W$ and for any $\mu \in S(Q)$. It follows that $\varphi^F_w[Q] = \varphi^W_w[Q]$, for every $w \in W$ and we contradict the initial assumption that $\varphi^F[Q] \neq \varphi^W[Q]$. □

Proof of Proposition 7 Suppose that Q is a Nash equilibrium in both $(\mathcal{P}, \varphi^F, P)$ and $(\mathcal{P}, \varphi^W, P)$. By Lemma 2, $|S(Q)| = 1$. Let us say that $S(Q) = \{\hat{\mu}\}$ and assume, by contradiction, that there exists a random stable matching rule $\tilde{\varphi}$ such that Q is not an ON equilibrium in $(\mathcal{P}, \tilde{\varphi}, P)$.

Suppose then that there exists a worker $w \in W$ and an alternative strategy Q'_w such that Q_w does not stochastically P_w-dominate Q'_w. This implies that there exists $\mu \in$ supp $\tilde{\varphi}[Q'_w, Q_{-w}]$ such that $\mu(w)P_w\hat{\mu}(w)$. Note that, since Q is a Nash equilibrium in the game induced by a stable matching rule, $\hat{\mu} \in IR(P)$. Hence, $\hat{\mu}(w)R_w w$ and it must be the case that w is matched to a firm under every matching in $S(Q'_w, Q_{-w})$. Let $\mu'(w)$ be the best match for w in supp $\tilde{\varphi}[Q'_w, Q_{-w}]$ and define Q''_w such that $A(Q''_w) = \{\mu'(w)\}$. Since $\mu' \in S(Q''_w, Q_{-w})$ (it is still individually rational and no blocking pairs emerged), Theorem 5.12 in Roth and Sotomayor (1990) ensures that w is matched to $\mu'(w)$ under every matching in $S(Q''_w, Q_{-w})$. Then, in no game induced by a stable matching rule is Q a Nash equilibrium, since for every stable matching rule φ, $\varphi_w[Q''_w, Q_{-w}] = \mu'(w)$ and $\mu'(w) P_w \hat{\mu}(w)$. It follows that no worker can profitably deviate in the game induced by $\tilde{\varphi}$.

Then, there exists a firm f and a strategy Q'_f such that Q_f does not stochastically P_f-dominate Q'_f, i.e., there exists $\mu \in$ supp $\tilde{\varphi}[Q'_f, Q_{-f}]$ such that $\mu(f) \bar{P}_f \hat{\mu}(f)$. Since $\hat{\mu} \in IR(P)$, we have $\hat{\mu}(f) \bar{R}_f \emptyset$ and, under every matching in $S(Q'_f, Q_{-f})$, f has at least one position filled. Let μ' be such that $\mu'(f) \bar{P}_f \mu(f)$, for every $\mu \in$ supp $\tilde{\varphi}[Q'_f, Q_{-f}]$. Define Q''_f such that $A(Q''_f) = \mu'(f)$. Note that $\mu' \in IR(Q''_f, Q_{-f})$ and that no pair of agents blocks μ' under the preference profile (Q''_f, Q_{-f}). Therefore, $\mu' \in S(Q''_f, Q_{-f})$ and, since firms have the same positions filled under every stable matching (Theorem 5.12 in Roth and Sotomayor (1990)), the definition of Q''_f guarantees that f holds $\mu'(f)$ in every element of $S(Q''_f, Q_{-f})$. Finally, for every stable matching rule φ, $\varphi_f[Q''_f, Q_{-f}] = \mu'(f)$ and $\mu'(f) \bar{P}_f \hat{\mu}(f)$. It follows that there exists no stable matching rule φ such that Q is a Nash equilibrium in $(\mathcal{P}, \varphi, P)$, contradicting the initial assumption. □

The proof of the above result reveals that a sufficient condition for an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$ is in fact being a Nash equilibrium in every game $(\mathcal{P}, \varphi, P)$, i.e., in every game induced by a deterministic stable mechanism. This appears to be an extremely strong condition to fulfill. Nevertheless, we will now describe a class of random matching rules for which such condition becomes necessary for an ordinal Nash equilibrium.

In the particular case that μ^I is the empty matching, Roth and Vande Vate (1990) have shown that, in the marriage model, every element of the set of stable matchings for the revealed preferences can be achieved with positive probability when the random matching rule they define is applied. In fact, starting from a situation in which all agents are unmatched, by successively satisfying all the pairs of a stable matching, we can guarantee that this matching is reached with positive probability. This random process is an instance of what we will name as *really* random stable matching rule.

A *really random stable matching rule* $\tilde{\varphi}$ assigns positive probability to at least two different elements of the set of stable matchings, i.e., $|\text{supp}\, \tilde{\varphi}[Q]| \geq 2$ for every Q such that $|S(Q)| \geq 2$. In Example 1, the rule that assigns probability 0.5 to the firm-optimal and to the worker-optimal stable matchings is clearly a really random stable matching rule. The following result is an implication of Propositions 6 and 7 in the particular case that $\tilde{\varphi}$ is really random.

Corollary 1 *Let $\tilde{\varphi}$ be a really random stable matching rule. Then, the profile of strategies Q is an ordinal Nash equilibrium in the game $(\mathcal{P}, \tilde{\varphi}, P)$ if and only if the set of stable matchings $S(Q)$ is a singleton and there exists a deterministic stable matching rule φ such that Q is a Nash equilibrium in the game $(\mathcal{P}, \varphi, P)$.*

Proof Follows directly from Propositions 6 and 7, and the fact that Proposition 1 implies supp $\tilde{\varphi}[Q] = S(Q)$ for a really random stable matching rule $\tilde{\varphi}$. □

For illustration, consider once more Example 1 and note that the set of stable matchings for truth telling is a singleton; further, it can easily be shown that it is an equilibrium in the game induced by the matching rule that yields, say, the firm-optimal stable matching. Corollary 1 thus implies that straightforward behavior is an ordinal Nash equilibrium in the game induced by the random stable matching rule described in the example.

5 Non-preference strategies

We have explored the game induced by a random matching mechanism, claiming that one of the main motivations of this paper is the study of some decentralized markets. This may be objected on the grounds that up to this point we have restricted our analysis to a one-period game where strategies are preference lists, which perfectly mirrors the functioning of a centralized market, but falls short of an illustration of a decentralized market. In particular, in matching processes of the kind described by Roth and Vande Vate (1990), at each moment in time, a pair of randomly chosen agents meets and (temporarily) matches if this is consistent with both agents' strategies. This clearly fits the structure of a sequential game. In this context, restricting each agent to hold the potential partner that is higher on some fixed preference ordering sustains the validity of the results of the preceding section. However, in a sequential game, agents can be expected to use richer strategies, conditioning behavior on the history of the game, and not necessarily acting consistently with a unique preference ordering. The strategy of matching with the first partner one meets and rejecting every other agent is an example of such kind of strategies.

One of the difficulties that arises in attempting to capture such complex forms of behavior concerns the very essence of the matching rule that, following Roth and Vande Vate (1990), we assume to be stable with respect to the revealed preferences. In fact, such definition is compromised when, for some play of the game, no list of preferences is compatible with the strategy of a player. Hence, the set of feasible strategies of the sequential game is simply too large and precludes analysis in the theoretical framework we have been using. One potential course of action is therefore to impose that under any play of the sequential game the choices actually made are consistent with some preference ordering, even though they may correspond to incompatible preference orderings when several plays are considered. We can then speak of preference orderings that are "revealed" in the course of the play. A worker w that entertains the described strategy in the example above, would match the first firm to tender an offer to him under any play of the game, and reveal that this firm is preferred to every other firm that he eventually meets in the course of that play. Since meeting is random, this worker would reveal distinct preference lists under different plays of the game.

Hence, consider a sequential game where, starting from an arbitrary matching, at each moment in time, a pair of randomly chosen agents, composed of a firm and a worker, meets. Agents match upon meeting if this is consistent with their strategies. We assume that strategies are restricted to those strategies compatible with a preference ordering for each play of the game, the revealed preference ordering, even though the information gathered in the course of the play might allow for other forms of behavior.[9] According to Roth and Vande Vate (1990), once the probability

[9] The lack of precision in defining what each player knows along the game is deliberate. The result that follows is valid in a perfect information setting, as well as when agents are only partially aware of the history of the game.

that a given pair of agents meets is bounded away from zero, each play of the game yields a matching stable with respect to the revealed orderings in the course of that play. Hence, given a profile of strategies that meets the above requirement, every outcome obtained with positive probability is stable for some revealed profile of preferences. We let $\mathcal{G}(P)$ denote this sequential game.

In Proposition 8, we show that ordinal Nash equilibria in preference strategies, which correspond to those obtained for the one-period game, are robust to the enlarged strategy space. In fact, given a profile of preference strategies, if by means of a strategy that is not consistent with a unique preference ordering, an agent may improve his position, he is certainly capable of doing so using a simple preference strategy.

Proposition 8 *In the sequential game $\mathcal{G}(P)$, for any collection of stated preferences Q_{-v} for agents other than an arbitrary agent v, agent v always has a best response that is consistent with a unique preference ordering.*

Proof First, consider an arbitrary worker w and fix Q_{-w}. Let s_w denote an arbitrary strategy for w, revealing a preference ordering (not necessarily the same) under each play of the game. Denote by Q^i_w the preference ordering that is consistent with s_w under some play i. In general, we have supp $\tilde{\varphi}[s_w, Q_{-w}] = \{\mu_1, \ldots, \mu_k\}$, where $\mu_i \in S(Q^i_w, Q_{-w})$, for $i = 1, \ldots, k$. Now let Q_w be such that $A(Q_w) = \{\mu_j(w)\}$ where $\mu_j(w) R_w \mu_i(w)$, for all $\mu_i \in \{\mu_1, \ldots \mu_k\}$. Since $\mu_j \in S(Q^j_w, Q_{-w})$, we must have $\mu_j \in S(Q_w, Q_{-w})$ (it is still individually rational and there are fewer blocking pairs). Hence, given that the same agents are matched under any two elements of the set of stable matchings and the only firm w finds acceptable is $\mu_j(w)$, this worker is matched to $\mu_j(w)$ under every matching in $S(Q_w, Q_{-w})$. It follows that any lottery over $S(Q_w, Q_{-w})$ gives w a partner at least as good as any lottery over $S(s_w, Q_{-w})$. Since s_w and Q_{-w} are arbitrary, this completes the proof for a worker w.

Now take an arbitrary firm f. Let s_f denote a strategy for f with the same properties as the strategy for w above. Define Q^i_f as the preference ordering over individual workers that is consistent with s_f for some play i of the game. Let supp $\tilde{\varphi}[s_f, Q_{-f}] = \{\mu_1, \ldots, \mu_k\}$, where $\mu_i \in S(Q^i_f, Q_{-f})$, for $i = 1, \ldots, k$. Consider any alternative strategy Q_f for f such that $A(Q_f) = \mu_j(f)$ where $\mu_j(f) \bar{R}_f \mu_i(f)$, for all $\mu_i \in \{\mu_1, \ldots, \mu_k\}$ and for every responsive extension $\bar{R}_f$ of R_f. Then, $\mu_j \in IR(Q_f, Q_{-f})$ since $\mu_j \in IR(Q^j_f, Q_{-f})$. Moreover, $\mu_j \in S(Q_f, Q_{-f})$ since $\mu_j \in S(Q^j_f, Q_{-f})$ and no blocking pairs emerged. Given that the same positions of a firm are filled under any element of a set of stable matchings and by definition of Q_f, f is matched to $\mu_j(f)$ under every matching in $S(Q_f, Q_{-f})$. Since s_f and Q_{-f} are arbitrary, this completes the proof. □

Nevertheless, this is far from being a characterization of equilibria in this new setting. In fact, the set of ordinal Nash equilibria is larger here, as the following example demonstrates.

Example 2 (Example 1 revisited) Consider the matching market in Example 1. Let the strategy of each agent be defined as follows: s_{f_i} = "match only with

w_i if f_1 is the first firm to meet a worker; match only with w_j otherwise" and s_{w_i} = "match only with f_i if f_1 is the first firm to meet a worker; match only with f_j otherwise", for $i = 1, 2$. This strategy profile leads to a non-degenerate probability distribution over matchings. Namely, both $\mu = \{(f_1, w_1), (f_2, w_2)\}$ and $\hat{\mu} = \{(f_1, w_2), (f_2, w_1)\}$ are obtained with a 50% probability. Hence, Proposition 1 rules out the possibility that s can be reproduced by an equilibrium in preference strategies. Still, s is an ordinal Nash equilibrium, since any unilateral deviation of a firm or worker may either leave the probability distribution unchanged or leave the deviator unmatched with positive probability.

6 Concluding remarks

At the expense of using an ordinal equilibrium concept, we have provided a characterization of equilibria that arise in the game induced by a random stable mechanism. The analysis is set in the college admissions problem. First, we have proved that every ordinal Nash equilibrium yields a unique matching, while when agents act straightforwardly according to the true preferences several matchings may be obtained with positive probability. Hence, agents avoid uncertainty when behaving strategically. Furthermore, a matching can be reached at an ordinal Nash equilibrium if and only if it is individually rational for the true preferences. Ordinal equilibria where firms best reply by behaving straightforwardly always produce a matching stable for the true preferences. Conversely, every stable matching can be reached as the outcome of an equilibrium play of the game. In a different direction, we relate ordinal Nash equilibria in games induced by a random matching mechanism with Nash equilibria arising in the games induced by deterministic matching mechanisms. In particular, a preference profile is an ordinal equilibrium of the game induced by a matching rule that always assigns positive probability to two different matchings (if such matchings exists) if and only if the set of stable matchings is a singleton and it is a Nash equilibrium in the game induced by some deterministic stable rule. In the last section of the paper we have tried to extend the above results, derived for a one-period game where the set of available strategies coincides with the set of all possible lists of preferences, to the sequential game that may arise in a decentralized market. Here we assume agents may use strategies that correspond to different preference orderings when different plays of the game are considered. We have shown that ordinal Nash equilibria in preference strategies are robust to the enlarged strategy space.

In what the above results are concerned, a couple of remarks is in order. The first observation concerns fairness and random matching mechanisms. In opposition to deterministic mechanisms, which are bound to favor one side of the market over the other, we have claimed that random mechanisms promote procedural fairness.[10] Nevertheless, "endstate" justice is a different issue. Indeed, the results that relate equilibria in the games induced by random and deterministic mechanisms imply that every equilibrium outcome in the game induced by a random matching mechanism may be obtained by means of a deterministic mechanism. It follows that,

[10] For example, in the kind of process described in Roth and Vande Vate (1990), each pair of agents has the same probability of meeting at a certain point in the procedure, and this determines procedural fairness.

based on these results and in what "endstate" justice is concerned, we should not expect random matching rules to improve upon deterministic ones if equilibrium behavior is to be taken seriously.

Second, the aim of the last section is to shed some light on what happens once we move towards allowing for history-dependent strategies, preserving the stability of the mechanism. The purpose of this paper is to explore strategic behavior induced by random stable matching mechanisms, and not to provide a thorough analysis of the incentives agents face in decentralized markets.[11] Therefore, relaxing the restriction we impose over the strategy sets would compromise our main goal.

To conclude, equilibrium behavior in random mechanisms has barely been treated in the matching literature. One of the difficulties that arises in attempting to apply the common game theoretical tools stems from the need to compare the probability distributions over matchings generated by a random rules when preferences are ordinal. By means of the concept of ordinal Nash equilibrium we have taken a step towards filling the gap in the literature, providing a fairly complete characterization of equilibrium behavior.

References

Abdulkadiroglu, A., Sönmez, T.: House allocation with existing tenants. J Econ Theory **88**, 233–260 (1999)

Abdulkadiroglu, A., Sönmez, T.: School choice: a mechanism design approach. Am Econ Rev **93**(3), 729–747 (2003)

d'Aspremont, C., Peleg, B.: Ordinal bayesian incentive compatible representations of committees. Soc Choice Welfare **5**, 261–279 (1988)

Ehlers, L., Massó, J.: Incomplete information and small cores in matching markets. Mimeo. Universitat Autònoma de Barcelona (2003)

Gale, D., Shapley, L.S.: College admissions and the stability of marriage. Am Math Monthly **69**, 9–15 (1962)

Knuth, D.E.: Marriages Stables. Montréal: Les Presses de l'Université de Montréal (1976)

Ma, J.: Stable matchings and the small core in nash equilibrium in the college admissions problem. Rev Econ Des **7**, 117–134 (2002)

Majumdar, D.: Ordinally bayesian incentive compatible stable matchings. Mimeo (2003)

Majumdar, D., Sen, A.: Ordinally bayesian incentive compatible voting rules. Econometrica **72**, 523–540 (2004)

Moulin, H.: Procedural cum end state justice: an implementation viewpoint. Justice Polit Liberal Utilitarianism (forthcoming) (1997)

Moulin, H.: Fair Division and Collective Welfare. Cambridge: MIT Press (2003)

Pais, J.: On random matching markets: properties and equilibria. Mimeo. Universitat Autònoma de Barcelona (2004a)

Pais, J.: Incentives in decentralized random matching markets. Mimeo. Universitat Autònoma de Barcelona (2004b)

Roth, A.E.: The evolution of the labor market for medical interns and residents: a case study in game theory. J Polit Econ **92**, 991–1016 (1984)

Roth, A.E.: The college admissions problem is not equivalent to the marriage problem. J Econ Theory **36**, 277–288 (1985)

Roth, A.E., Sotomayor, M.A.O.: Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Cambridge: Cambridge University Press (1990)

Roth, A.E., Vande Vate, J.H.: Random paths to stability in two-sided matching. Econometrica **58**, 1475–1480 (1990)

Roth, A.E., Vande Vate, J.H.: Incentives in two-sided matching with random stable mechanisms. Econ Theory **1**, 31–44 (1991)

[11] In this connection, see Pais (2004a).