**Decentralized Job Matching**

Guillaume Haeinger

And

Myrna Wooders

**No 688**

# WARWICK ECONOMIC RESEARCH PAPERS

**DEPARTMENT OF ECONOMICS**

# Decentralized Job matching*

Guillaume Haeringer†
*Universitat Autònoma de Barcelona*

Myrna Wooders‡
*Warwick University*

October 6, 2003

**Abstract**

This paper studies a decentralized job market model where firms (academic departments) propose sequentially a (unique) position to some workers (Ph.D. candidates). Successful candidates then decide whether to accept the offers, and departments whose positions remain unfilled propose to other candidates. We distinguish between several cases, depending on whether agents' actions are simultaneous and/or irreversible (if a worker accepts an offer he is immediately matched, and both the worker and the firm to which she is matched go out of the market). For all these cases, we provide a complete characterization of the Nash equilibrium outcomes and the Subgame Perfect equilibria. While the set of Nash equilibria outcomes contain all individually rational matchings, it turns out that in most cases considered all subgame perfect equilibria yield a unique outcome, the worker-optimal matching.

KEYWORDS: Two-sided matching, job market, subgame perfect equilibrium, irreversibilities.

*Journal of Economic Literature Classification*: C78, C62, J41.

*We thank Jordi Massò, Tayfun Sönmez, Alejandro Neme and Francis Bloch for helpful comments. The first author acknowledges a financial support from the Spanish Ministry of Science and Technology through grant BEC2002- 02130.

†Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. Guillaume.Haeringer@uab.es

‡Warwick University, Department of Economics, Coventry, CV4 7AL, UK. M.Wooders@warwick.ac.uk

# 1 Introduction

Matching markets concern those environments where agents have incentives to form groups. The *College admission problem* where students choose which college to attend and colleges have to choose which students to accept, and the American market for intern physicians are two of the most well known examples. In spite of the very large literature on matching markets, however, one sort of matching market has received very little attention in the literature: job markets with commitment, for example, the academic job market for junior faculty members.

This paper is aimed at studying the strategic issues of markets that are similar to the academic job market, that is, markets that are *decentralized* with the possibility that agents' decisions are *irreversible*. To this end, we propose a model where firms make offers to workers, and then workers each decide which, if any, offer to accept. A firm that has had some of its offers rejected then may make further offers, and so on. To some extent, the model we propose can be seen as a decentralized version of the Gale-Shapley deferred acceptance algorithm.

By 'decentralization' we mean that each agent makes her own decisions, independently of the decisions made by others at the same stage of the game. For instance, a candidate is allowed to refuse an offer even if this is the only one she has and even if she prefers the offer to being unmatched at the end of the game. This contrasts with centralized markets employing the Gale-Shapley deferred acceptance algorithm to match agents. Also, in a centralized context, matching markets are typically modeled as normal form games (where a players' strategy consists of a preference list), while in our decentralized context the market can be better described as a sequential game.

Decentralized markets involve different strategic issues from those of centralized markets. To see this, consider the case of a candidate $i$, who receives an offer that is acceptable to her from a firm $f$ that is ranked low in her preferences.[1] Under the Gale-Shapley algorithm, unless she already has a better offer, $i$ would be assigned $f$, at least temporarily. By contrast, in a decentralized market, $i$ could decide to reject $f$'s offer. If she does so, then $f$ may propose to another candidate, say $k$, who might accept. The candidate $k$ may then reject further offers (or might not receive some offers) that are subsequently made to $i$ and ranked higher in $i$'s preference ordering, allowing $i$ to be better off than she would have been had she accepted the initial offer. In decentralized markets, there are also strategic considerations for the side of the market making offers. For example, as perhaps some academic departments have already intuited, for a firm, if it fears rejection by

[1] By an acceptable offer, we mean one that the candidate prefers to being unmatched at the end of the game.

its most preferred set of candidates, it can be advantageous to make its first offer to less preferred candidates. By initially making offers that are likely to be rejected, the firm will loose some time and may eventually end up with only candidates from among those that are minimally acceptable.

A second crucial aspect of the market studied in this paper is that candidates' decisions are not "*deferrable*". In a centralized market, holding an offer while waiting for a better one may "block" the market and leaves workers with their less preferred outcome. To see this, consider a candidate, say $i$, having received an offer from some department $x$, but holding this offer, expecting to have a better offer, say from $y$. She may well never receive this offer, for the other candidate who received the offer from $y$ (but preferring that from $x$) will accept it. As this second candidate does not receive any offer from $x$ (because it is "held" by $i$) she will accept the one from $y$, eliminating the possibility for $i$ to receive the offer from $y$. This sort of phenomenon is peculiar to the deferred acceptance algorithm, especially when all firms have a distinct most preferred worker and all firms are acceptable to their most preferred workers. Since a firm has to be unmatched to make an offer, no worker can expect to receive a better offer than the one she had in the first round of offers.

That candidates' decisions are not deferrable in our model means that as soon as a candidate, say $i$, has accepted an offer from a firm, say $f$, then $i$ and $f$ are immediately matched and the position offered to $i$ cannot, at any further stage in the game, be offered to anyone else. Although we are aware that in actual hiring procedures, this assumption may sometimes be violated we believe that the non-deferrability of decisions is a relevant feature of the academic job market for junior economists.[2] A close look at the hiring procedures may help to grasp the impact of this assumption. When a candidate receives an offer, this is usually (first) by e-mail or by phone. It is not until a couple of days or weeks later that she receives an official letter confirming the offer. Meanwhile, the candidate has notified the department whether she accepts the offer, also by e-mail or by phone. If a candidate accepts an offer, although she is typically not *officially* contracted until several weeks (if not months) later, it is usually the case that this candidate is regarded as being indeed off the market.[3] In this case, the candidate rejects all subsequent offers, explaining that she is no longer seeking a job. Consider a candidate, say Wendy, who receives an offer from University B. At the time she receives this offer, it is the best she has, and therefore she decides to accept. Suppose now that a couple of weeks later she receives an offer from University A, which she prefers to the offer from B. If her con-

[2]During the course of the research carried out in this paper we have heard that sometimes, and quite unfortunately, some candidates do reject an offer they have previously accepted.

[3]Although often the contract is not signed before the newly hired faculty member arrives at the university, when her contract begins.

tract with A is not yet signed, that is, if her notification of acceptance has only been informal, she could accept the offer from A and send a message to B saying that she has now decided to decline its offer. All this would be perfectly legal. Considering Wendy's case, reversing her acceptance of B would be very likely to jeopardize her reputation. Moreover, if A knows that she has accepted B, A is likely to refrain from making her an offer. Thus, although Wendy could legally reject, *ex-post*, the position offered by A, she cannot easily do it in practice. That is, candidates' decisions on the job market are *irreversible* or, said differently, workers and firms must *commit*. It is important to notice that this assumption is different from the fact that candidates are usually given several days to give their answer.[4]

We begin our study with a characterization of the Nash equilibrium. We first show that for any individually rational matching $\mu$ we can find a strategy profile that is a Nash equilibrium and whose outcome yields this matching $\mu$. Subgame perfect equilibrium yields more interesting results. The most delicate case, a benchmark upon which we build most of our results, is when agents' decisions are irreversible and *not* simultaneous. The main technical difficulty in this benchmark case is that agents' strategies cannot in general be identified with their preferences. For instance, firm $f$'s strategy may consist of proposing to worker $w$ if firm $f'$ proposed to worker $w'$ at a previous stage and to propose to worker $w''$ otherwise. Similarly, a worker who received offers from firms $f$ and $f'$ could decide to accept $f$'s offer if worker $w'$ has received an offer from $f''$ and to reject both offers otherwise. We are able, however, to identify a natural class of firms' strategies, which we call *sanitization strategies*, that can be, to some extent, identified with an order of workers. Such a strategy for a firm, say $f$, consist of making a plan of offers respecting some ordering of workers (which may not follows $f$' preferences), but taking into account that some subgame may be reached because $f$ did not follow its order of offers. Put differently, a firm's sanitization strategy is characterized by an ordered list of offers that does not depend on the strategies of other firms.[5]

Though useful to prove the following result, sanitization strategies turn out not to be the only optimal strategy for firms. We find that there exists a unique subgame perfect equilibrium *outcome* (i.e., a unique matching), the candidate optimal matching, which may be supported by a large number of different strategy profiles. Recall that, in centralized markets using

[4]In the academic job market, candidates are always given some time to make up their mind before deciding about an offer, and a department waiting for the answer of a candidate cannot make a proposal to another candidate (for the same position). In this paper, we will consider the perfect information case between each period (albeit not necessarily within periods. See section 3.1 for further details). Therefore, allowing for a delay in responding to an offer has no effect on the outcome.

[5]See Appendix A for a formal description of these strategies.

the deferred acceptance algorithm with firms proposing, the outcome is the firm optimal matching. Thus, our result demonstrates a situation in which decentralization gives more power to the workers.

Our result crucially depends on the assumption that workers' decisions are *not* simultaneous. When workers' decisions are simultaneous, the set of subgame perfect equilibria outcome is enlarged and includes (but may not coincide with) the set of stable matching. Yet, whether firms' decisions are simultaneous or not does not change our findings. More precisely, if we suppose that there is an "offer stage" during which firms make their offer and then an "acceptance-rejection" stage, or if we suppose that firms do not follow the same agenda,[6] the outcome remains unchanged.

It turns out that the irreversibility assumption has lesser effects than assuming that the market is decentralized. If we allow workers not to commit, then we also obtain that there is a unique subgame perfect equilibrium outcome, the worker optimal matching, irrespective of the timing of workers' decision (i.e., simultaneous or not). The reason for why this holds is quite simple. Under perfect information each participant knows completely the strategy employed by her opponents. Hence, if a candidate "knows" that she will receive an offer from some department there is no need for her to hold offers from less preferred departments. That is, holding an offer makes sense only when some uncertainty prevails about the set of subsequent offers that may be received. Here, all the relevant information comes from the workers – which offers are accepted and which are not. Hence, if firms act simultaneously and workers do not the results remain unchanged. To sum up, the irreversibility assumption has an impact only when workers decisions can be simultaneous.

In all cases, except when actions are simultaneous and irreversible, subgame perfect equilibrium strategies turn out to have a very specific structure. Roughly, for a firm, a strategy consists of proposing to any worker preferred to the worker-optimal mate (including this mate) before proposing to any worker less preferred than the worker-optimal mate.[7] Since there might be many workers preferred (or less preferred) to the worker-optimal mate, it follows that optimal strategies are multiple. For a worker, an optimal strategy is to accept any offer that is preferred to the worker-optimal mate (including her), and reject all others.

The paper is organized as follows. In section 2 we define the main concepts and notations. In section 3 we analyze the job market game where players' actions are irreversible and not simultaneous. In section 4 we study the other cases, i.e., when actions may be simultaneous, with and without

[6]For instance, we could have the case of one firm making three different offers while another firm has time to make only one offer.

[7]Firms' subgame perfect equilibrium strategies may not be unique. However, in some case the position (in a firm's plan of offers) of the most preferred worker among the set of workers less preferred than the worker-optimal mate may matter. See Section 5.2.

irreversibility. In section 5 we provide several examples illustrating the main findings of the paper. In section 6 we discuss some papers in the matching literature dealing with decentralized markets. A conclusion is given in section 7. Appendix A is devoted to a formal description of sanitization strategies and Appendix B contains the proofs of the results.

## 2 Preliminaries

We consider a (finite) set of workers and a (finite) set of firms, denoted $W$ and $F$ respectively. Throughout the paper, feminine pronouns refer to firms, whereas masculine pronouns refer to workers. Each worker $w \in W$ has a strict, complete, transitive, and asymmetric preference relation $P_w$ over $F \cup \{w\}$. Similarly, each firm $f \in F$ has a strict, complete, transitive, and asymmetric preference relation $P_f$ over $W \cup \{f\}$. The cardinality of a (finite) set $A$ is denoted $\sharp A$. Let $\max_{P_v} S$ be the maximal element of $S$ with respect to $P_v$;

$$\max_{P_v} S \stackrel{\text{def}}{=} \{v' \in S \cup \{v\} \mid v' R(v) v'' \ \forall \ v'' \in S \cup \{v\}\}$$

where $R(v)$ denotes 'preferred or indifferent.' The set of all possible preferences for an individual $v \in W \cup F$ is denoted by $\mathcal{P}_v$, and any element $P$ of $\mathcal{P} \stackrel{\text{def}}{=} \Pi_{w \in W} \mathcal{P}_w \times \Pi_{f \in F} \mathcal{P}_f$ is called a *preference profile.* A *job market* is described by the triple $(W, F, P)$.

A matching $\mu$ is a function from $W \cup F$ onto itself such that

*(i)* For all $w \in W$, $\mu(w) \in F \cup \{w\}$ ;

*(ii)* For all $f \in F$, $\mu(f) \in W \cup \{f\}$ ;

*(iii)* For all $v \in W \cup F$, $\mu(\mu(v)) = v$ .

Let $\mathcal{M}$ denote the set of all matchings. Given a preference relation $P_w$ of a worker $w$ over $F \cup \{w\}$, we can extend $P_w$ to a complete, transitive and reflexive preference relation $R(w)$ (with $P_w$ and $I(w)$ denoting respectively the strict preferences and the indifference relations) over the set of matchings $\mathcal{M}$ in the following way: $\mu P_w \mu'$ if and only if $\mu(w) P_w \mu'(w)$, and $\mu I(w) \mu'$ if and only if $\mu(w) = \mu'(w)$. Similarly, we can also extend firms' preferences to preferences over $\mathcal{M}$.

A matching $\mu$ is *blocked by agent* $i$ if $i P_i \mu(i)$. A matching $\mu$ that is not blocked by any agent is called *individually rational.* A matching $\mu$ is *blocked by a pair* $(w, f)$ if $f P_w \mu(w)$ and $w P_f \mu(f)$. A matching $\mu$ that is individually rational and is not blocked by any pair under $P$ is called *stable.* For a job market $(W, F, P)$, we denote by $S(P)$ the set of stable matching. The firm-optimal matching (resp. the worker-optimal matching), denoted

$\mu_f$ (resp. $\mu_W$), is the matching preferred by all firms (resp. workers) among all stable matchings.

Gale and Shapley [6] and Knuth [7] respectively proved the following properties,[8]

- The deferred acceptance algorithm with firms (workers) proposing produces the firm-optimal (worker-optimal) matching;
- The firm-optimal (worker-optimal) matching is the workers' (firms') least preferred matching among all stable matchings.

# 3 Job matching with commitment

We consider now the situation where firms can make sequential offers to workers, and each worker who receives an offer from a firm has to make an irreversible decision, to either *reject* or *accept* the offer. The irreversibility assumption means that if a worker rejects an offer, he cannot decide later to accept it. Moreover, if a worker accepts an offer he cannot decide at a subsequent stage to reject the offer. This implies that once a worker, say $w$, has accepted an offer from a firm, say $f$, both $w$ and $f$ go out of the market: worker $w$ cannot receive other offers *and* firm $f$ cannot make offers to other workers.

## 3.1 The sequential game

Let $\mathcal{G}(P)$ denote the sequential game where firms make the offers and workers accept or reject the offer and all agent's decisions are irreversible. Since the game analyzed in this section is a game of perfect information, the order of play in our context has little importance. For the sake of clarity we shall nevertheless specify an order of play.

The game is divided into several *stages*. Each stage is divided into two sub-stages: First and offering stage (during which firms act), and second a responding stage (during which workers act). Workers and firms are assigned index numbers, so the $k^{th}$ firm (worker) is denoted by $f_k$ ($w_k$). At each step, it is assumed that players act in the order of their index numbers, i.e., firm $f_1$ plays first, then $f_2$, ..., and then worker $w_1$ plays first, then $w_2$, ...

**Stage 1.1**: Each firm sequentially offers its position to at most one worker. If a firm does not offer her position to any worker, then she exits the market.[9]

[8] See Roth and Sotomayor [10] for further references.

[9] This assumption makes thus the game finite, and avoids, for instance, the case where at each stage some firms decide not to make any offer.

**Stage 1.2**: If a worker did not receive any offer at stage 1, then he does nothing and waits for the next stage. If a worker received one or more offers at stage 1, then he can either reject all or accept one of them. If a worker, say $i$, accepts an offer from a firm, say $j$, then $i$ and $j$ are matched and both go out of the market. If a firm does not offer her position to any worker, then she exits the market.

...

**Stage t.1**: Each remaining firm decides either not to make any offer or to make an offer to at most one agent remaining in the market to whom she did not make an offer yet. If such an agent does not exists at stage $s$, then the firm cannot make any offer and goes out of the market.

**Stage t.2**: If a worker did not receive any offer at stage $s$, then he does nothing and waits for the next stage. If a worker received one or more offers at stage $s$, then he can either reject all or accept one of them. If a worker, say $i$, accepts an offer from a firm, say $j$, then $i$ and $j$ are matched and both go out of the market.

**Final stage**: The game stops when all firms are out of the market (whether matched or not), or when for each remaining firm there is no worker left in the market to whom she has not already made an offer.

Dividing the timing of the game into stages is for convenience only. Our results would still hold if, for example, one firm were allowed to make more offers at each stage than another firm.[10]

**Remark 1** The sequence described here makes the game actually closer to a decentralized version of the McVitie-Wilson algorithm [8]. A decentralized version of the deferred acceptance algorithm would instead more closely resemble a game where, within each sub-stage, firms and workers act simultaneously (see section 4).

Players' actions are taken at *decision nodes* (also called information sets), which are typically denoted by $\alpha$. We write $s(\alpha)$ for the stage in wich node $\alpha$ occurs. At each node, each active player $v \in W \cup F$ chooses an action $a \in A_v$. If $f \in F$, $A_f = W \cup \{\varnothing\}$, and $a = w$ means that at node $\alpha$ firm $f$ makes an offer to $w$. For an active worker at node $\alpha$, denote by $X_w^\alpha$ the set of offers $w$ receives at that node. Then, his action set is $A_w = X_w^\alpha \cup \{w\}$ where $a = f$ means that $w$ accepted the offer of $f$ and $a = w$ means that $w$ rejected all his offers. For each node $\alpha$, we write $\gamma(\alpha)$ the subgame starting at node $\alpha$.

A *strategy profile* $\sigma$ specifies a strategy for each player in the market, i.e., a plan of action for each node of the game where this player may have to act. We write $\sigma_i$ for the strategy of player $i$ and $\sigma_{-i}$ the strategy vector $\{\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_k\}$, where $k = \sharp(W \cup F)$. For $S \subseteq W \cup F$, let

[10] See footnote 6.

$\Sigma_S = \Pi_{v \in S} \Sigma_v$, where $\Sigma_v$ is $v$'s strategy set. For a node $\alpha$ and a strategy $\sigma_i$, let $\sigma_i|_\alpha$ be the strategy played by $i$ in the subgames that comes *after* node $\alpha$ .

Finally, we denote by $\mu[\sigma]$ the outcome obtained when the strategy profile $\sigma$ is played.

### 3.2 Nash equilibrium

Nash equilibria outcomes appear to be easily characterized. Individual rationality turns out to be a necessary and sufficient condition for a matching to be supported by a Nash equilibrium. Necessity is obvious. To see that it is also sufficient, take any individually rational matching $\mu$ (possibly unstable), and construct the following strategy profile $\sigma$,

*(i)* For all $f \in F$, make an offer to $\mu(f)$ only;

*(ii)* For all $w \in W$, accept only the offer from $\mu(w)$ .

Clearly, any player's strategy is a best response given the strategy of the others. We have just proved the following result.

**Proposition 1** *A strategy profile $\sigma$ is a Nash equilibrium of $\mathcal{G}(P)$ if and only if $\mu[\sigma]$ is individually rational.*

### 3.3 Subgame perfect equilibrium

The previous section showed that the Nash equilibrium concept does not require that we work with the full definition of a player's strategy, i.e. a plan of actions for any decision node where this player has to act. Subgame perfect equilibrium (SPE) strategies turn out not to be as simple Nash equilibrium strategies and, as expected, actions off the equilibrium path play a significant role.

With the use of 'sanitized strategies' in the proof, we are able to obtain the following result characterizing the set of subgame perfect equilibria for our matching game. It turns out that there a unique equilibrium outcome, the worker-optimal matching.

**Theorem 1** *Let $\sigma$ be a subgame perfect equilibrium. Then $\mu[\sigma] = \mu_W$ .*

Although Theorem 1 states that the outcomes of any subgame perfect equilibrium are identical, nothing is said about the number of subgame perfect equilibria. Typically, subgame perfect equilibria are not unique (see the examples in Section 5).

### 3.3.1 Sanitized strategies

A natural way to look at firms' strategy would be to define a strategy as merely a sequence of offers, which would imply that firms' strategy spaces are just the set of all possible orderings of workers (union themselves), i.e., equivalent to the set of all preferences.[11] Since a player's strategy must specify a list of actions, one for each of her information sets, it is not possible to describe a firm's strategy as a mere ordering. To see this, suppose that a firm, say $f$, decided to make her offers based on some ordering of workers (which may not necessarily coincides with her preferences), at any subgame. Quite obviously, there are subgames that can be reached only if this very firm $f$ did not follow its ordering. If such subgames are attained, the firm's sequence of offers will not necessarily coincide with the original ordering. More generally, it could well be the case that a firm follows some order of offers for some subgames and another order of offers for other subgames.(Guillaume, could this be much shorter?)

Of course, if we look at the execution path of a strategy profile we can identify for each firm and worker a partial order. For firms, this partial order consists of the sequence of offers made. For a worker, if he rejects first firms $f_1$ and $f_2$ and finally accept that of $f_3$ the order consists of having $f_3$ before $f_1$ and $f_2$ (but without any indication about the relative order between $f_1$ and $f_2$). Blum, Roth and Rothblum [4] identify the partial order thus obtained as "*revealed preferences*". While such preferences may serve as a proxy of agents' strategies (their game has a high level of imperfect information) this serve little in our case.

In spite of the above remarks, it is nevertheless possible to define a class of strategies where a firm's strategy can be identified as a simple sequence of offers, or more precisely, as an ordering of workers. A firm's sanitization strategy is characterized by a consistency requirement which dictates that, given her ordering of workers at her initial information set, the firm's offers at every decision node must be consistent with that ordering, even at decision nodes that would not be reached had that ordering been used.

Consider for example a firm having chosen a strategy that is representable by some ordering. To define formally such a strategy, we need to specify how the firm will act in those subgames that are reached even if she did not follow the original ordering of offers. Our consistency criterion amounts to suppose that this deviation was a "temporary mistake", and that the firm will make further offers following the original sequence. For instance, if the original sequence of offers was $\{w_1, w_2, w_3, w_4\}$, and instead of making an offer to $w_2$ the firm made an offer to $w_4$, then the order of offers after having proposed to $w_4$ will be $\{w_2, w_3\}$. After having proposed (by mistake) to $w_4$, the workers to whom she did not propose yet are $w_2$ and $w_3$.

[11] Blum, Roth and Rothblum [4] call such strategies "preference strategies".

In the original order, $w_2$ comes before $w_3$. Therefore, in the subgame that starts after her "deviation", she must propose to $w_2$ before $w_3$.[12] Strategies that are consistent with this (informal) criterion will be called **sanitization strategies**, and the ordering $\pi = (w_1, w_2, w_3, w_4)$ will be called a **representation** of the firm's strategy (see Appendix A for a formal definition). If a firm uses a sanitization strategy that represents the preference of that firm we shall say that the firm's strategy is a **strong sanitization strategy**.

A brief explanation why any SPE yields the worker-optimal matching is the following (we can also refer the reader to see the examples in section 5). When firms make their offers it is always profitable for a worker to reject those offers that are less preferred than his worker-optimal mate. Indeed, these offers are the worker-optimal mates of some other workers. In other words, as long as workers receive such offers it is in their common interest to reject them. To some extent, the game has some flavor of strategic complementarity between the workers. However, when workers receive offers from their worker-optimal mates this strategic complementarity disappears, and workers start vying with each other for better outcomes. If a worker rejects his worker-optimal mate in order to get a better offer (that he will indeed receive) it must be the case that there is another worker who accepts being matched to a firm less preferred than his worker-optimal mate. This property is mainly derived from the fact that the worker-optimal matching is a stable matching (see lemma 3 in Appendix B).

### 3.3.2 A sketch of the proof of Theorem 1

To prove Theorem 1 we proceed in several steps. First, we consider the case where firms do not act strategically. That is, firms are required to act as robots and make offers according to their preferences; that is, they are required to use strong sanitization strategies. With this restriction, we show that any subgame perfect equilibrium yields the worker-optimal matching (Lemma 6). From this result, we deduce that if at some SPE some firms are worse off than at the worker-optimal matching, it must be the case that one of those firms did not use a strong sanitization strategy (Lemma 7). It follows that at any SPE firms cannot be strictly worse off than in the worker-optimal matching (Lemma 10). The penultimate step of the proof consists of deducing that workers are matched to their worker-optimal mate.

A key concept used in throughout the proofs is that of "sub-markets,"i.e., the job market corresponding to a particular subgame. Observe that for any subgame the history of play can be summarized by

- A list of matched workers and firms;

[12] In the above example, it may be the case that after deviating (i.e., proposing to $w_4$ instead of $w_2$), worker $w_2$ may be matched to another firm and thus not available anymore in the market. In that case, the firm will propose to $w_3$ (if he is available).

- For each firm, the set of workers that have rejected her offers.

Let $\alpha$ be any node of the game, and let $\mathbf{O}_f^\alpha$ be the set of workers to whom $f$ has already made a proposal (and who have rejected $f$'s offer) and workers already matched at the nodes *preceding* $\alpha$. That is, $W \backslash \mathbf{O}_f^\alpha$ is the set of workers that are still in the market at node $\alpha$ and to whom $f$ has not yet made an offer.

For a subgame starting at node $\alpha$, let $F(\alpha)$ and $W(\alpha)$ respectively denote the set of firms and workers **not** yet matched. For a firm $f \in F(\alpha)$, let $P_f^\alpha$ denote firm $f$'s preferences over the where all the workers in $\mathbf{O}_f^\alpha$ are taken as unacceptable (since they have already rejected $f$), and leaving unchanged the relative rankings of those reminaing workers:

$$\begin{aligned} \forall\, v, v' \in W(\alpha) \cup \{f\} \backslash \mathbf{O}_f^\alpha \qquad & v P_f^\alpha v' \quad \Leftrightarrow \quad v P_f v' \;, \\ \forall\, w \in \mathbf{O}_f^\alpha \qquad & f P_f^\alpha w \;. \end{aligned}$$

For a worker $w \in W(\alpha)$, let $\mathbf{O}_w^\alpha$ be the set of firms from which $w$ has rejected an offer (and have rejected $f$'s offer) and firms already matched at the nodes *preceding* $\alpha$. For a worker $w \in W(\alpha)$, let $P_w^\alpha$ denote workers $w$'s preferences where all firms in $\mathbf{O}_w^\alpha$ are taken as unacceptable, and leaving unchanged the relative ranking of all other firms:

$$\begin{aligned} \forall\, v, v' \in F(\alpha) \cup \{w\} \backslash \mathbf{O}_w^\alpha \qquad & v P_w^\alpha v' \quad \Leftrightarrow \quad v P_w v' \;, \\ \forall\, f \in \mathbf{O}_w^\alpha \qquad & w P_w^\alpha f \;. \end{aligned}$$

Let $P^\alpha$ be the preference profile of agents in $F(\alpha) \cup W(\alpha)$ that satisfy the above conditions. It is obvious to see that, in terms of subgame perfection (or even with Nash equilibria), the job market $(F, W, P)$ at the subgame starting at node $\alpha$ is equivalent to the job market $(F(\alpha), W(\alpha), P^\alpha)$. Indeed, in Proposition 1 we saw that only individually rational matchings are possible candidates for subgame perfection. Therefore, if for a firm in $F(\alpha)$ we put all workers in $\mathbf{O}_f^\alpha$ as being unacceptable, we are sure that $f$'s subgame perfect equilibrium strategy will never consider making an offer to workers in $\mathbf{O}_f^\alpha$. For a subgame starting at a node $\alpha$, let $\mu_W^\alpha$ denote the worker-optimal matching corresponding to the market $(F(\alpha), W(\alpha), P^\alpha)$.

A direct corollary of Theorem 1 is that for a job market $(F, W, P)$, at any node $\alpha$ the subgame perfect equilibrium is such that workers and firms that are not matched yet will be matched to the individual assigned under the worker-optimal matching $\mu_W^\alpha$.[13]

# 4 Alternative market organizations

In the case studied in section 3 it was assumed that decisions are irreversible and that agents' actions are not simultaneous. In this case, it is obvious

[13]Note that for some node $\alpha$, there may exist a firm $f \in F(\alpha)$ such that $\mu_W^\alpha(f) \neq \mu_W(f)$.

that the irreversibility assumption has no impact. Hence, we can deduce that if actions are simultaneous, but not irreversible, then we obviously get the same result: there is a unique subgame perfect equilibrium outcome, the worker-optimal matching, and the equilibrium strategies are identical to those of the case where irreversibility holds. If a firm orders workers acccording to its true preferences, but then it gets rejected for its first $k$ offers then it could make those offers in any order.)

Two other possibles cases are when firms' moves are simultaneous (resp. not simultaneous) and workers' moves are not simultaneous (resp. simultaneous). If firms' moves are simultaneous but workers' moves are not, it is easy to see (the proofs remain unchanged) that we are in the same situation as in our benchmark case, where we have perfect information.

Things differ, however, when we consider the situation where firms' moves are not simultaneous but workers' moves are simultaneous. Let $\mathcal{G}^*$ be the game where firms actions are not simultaneous but workers' actions are. More precisely, any proposing stage is with perfect information, any responding stage is a simultaneous game, and between each sub-stage there is perfect information. That is, each firm can observe the offers made by the firms who decided before her, and all workers observe the offers made by all firms. However, at each stage active workers do not observe the action taken by the other workers *during* the stage, although they do observe the workers' decisions taken at previous stages. In this case, we find that any stable matching can be supported by a subgame perfect equilibrium.

**Proposition 2** *Let $\mu \in S(P)$. Then there exists $\sigma$ in $\mathcal{G}^*$ such that $\mu[\sigma] = \mu$ and $\sigma$ is an SPE of $\mathcal{G}^*$.*

It turns out, however, that an unstable matching can be the outcome of an SPE. To see this, consider the following job market where $F = \{f_1, f_2, f_3\}$, $W = \{w_1, w_2, w_3\}$ and the preferences are

$$\begin{aligned} P(w_1) &= f_3, f_2, f_1, w_1, & P(f_1) &= w_1, w_3, w_2, f_1, \\ P(w_2) &= f_2, f_1, f_3, w_2, & P(f_2) &= w_3, w_2, w_1, f_2, \\ P(w_3) &= f_1, f_3, f_2, w_3, & P(f_3) &= w_3, w_2, w_1, f_3, \end{aligned}$$

which can be read as "firm $f_1$'s first choice is $w_1$, then $w_3$ and then $w_2$, and they are all acceptable for her (i.e., preferred to herself)", and similarly for the other players.

The worker-optimal matching for this market is

$$\mu_W = \{(f_1, w_3), (f_2, w_2), (f_3, w_1)\}$$

and the firm-optimal matching is

$$\mu_F = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\}.$$

Suppose the game is such that $f_1$ decides first, then $f_2$ and finally $f_3$, and consider the following strategy profile. Let $f_1$ propose to $w_1$, $f_2$ to $w_3$ and $f_3$ to $w_2$. After this chain of proposal, let each worker accept the offer he received. Clearly, for each worker, given that the two others accept their offer it is a best response to accept (otherwise he remains unmatched as there is no other available firm in the market). Hence, after $f_3$'s proposal, the actions chosen by each worker is a Nash equilibrium, and therefore an SPE of the subgame considered. (Note that this is not the only equilibrium of this subgame.) The matching obtained is then

$$\hat{\mu} = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\}.$$

Clearly, $\hat{\mu}$ is not stable as it is blocked by the pair $(f_3, w_3)$. If one (or more) firm does not follow the above mentioned sequence of offers, consider the following workers' strategy profile: accept only the offer from the worker-optimal mate and reject all other offers. For the subgames that are reached after the deviation of one or several firms, we have an SPE. For any order of proposal we attain $\mu_W$. Since any worker is acceptable, firms are therefore indifferent between all the possible order of offers they may follow. For the workers, they are all matched to their best mate, so any deviation is strategically unconceivable. Hence, it remains to check that the initial order made during the first stage is a Nash equilibrium when considering the outcomes obtained at all other subgames.

Consider now $f_3$'s optimal choice. If $f_3$ follows the plan of offers, she is matched to $w_2$. If she deviates, she obtains $\mu_W(f_3) = w_1$. Since $w_2 P_{f_3} w_1$, it is a best response for $f_3$ not to deviate. Likewise, it is in their best interest for $f_2$ and $f_1$ not to deviate, for they can obtain their most preferred worker. By deviating, they both would be strictly worse off. Hence, no firm wants to deviate from the original plan of offers. We therefore have $\hat{\mu}$, an unstable matching, that can be supported as an SPE.

This result contrasts to those obtained by Peleg [9], Alcalde *et al.* [2], and Alcalde and Romero-Medina [3].[14] These authors proposed a model similar to ours, but with the difference that there is only one stage. Like us, they show that any stable matching can be supported by an SPE. However, they show that the converse also holds ([9] for the marriage model, and [2] and [3] for the College Admission problem). The reasons for this are twofold. First, there is no continuation stage if at the end of the first stage there are still available workers and firms unmatched and acceptable to each other.

[14] Suh and Wen [15] have an example of a sequential marriage market with perfect information where an unstable matching can be the outcome of a subgame perfect equilibrium. However, they assume that there is only one stage, i.e., each agent can only propose at most once. In their model, both men and women (firms and workers in our case) can propose, but their example of an SPE yielding an unstable matching fits to our model in the sense that first all men propose and then all women either accept one of their proposal (if any) or reject all of them.

Second, firms do act simultaneously. When there is only one stage, the proof to show that any SPE must yield a stable matching is quite simple. Suppose that there is a blocking pair, $(f, w)$. Since we are at an SPE, $w$ did not receive any offer from $f$. Hence, $f$ can deviate and make an offer to $w$. Since firms play simultaneously, firms playing "after" $f$ will not change their action. After this deviation, worker $w$ has a better offer than the one he would have received had $f$ not deviated. Subgame perfection implies then that $w$ will accept $f$'s offer, making therefore $f$'s deviation a profitable one.

# 5 Examples

This section is devoted to the study of two examples, one with two firms and two workers, and one with three firms and three workers. The first example mainly serves to show how crucial is the assumption that workers decisions are not simultaneous. We shall not provide the complete game tree with the second case (a readable tree would not fit in one page). This second example serves as an illustration of some possible equilibrium strategies.

## 5.1 Two firms - two workers

Let $F = \{f_1, f_2\}$ and $W = \{w_1, w_2\}$, with the following preferences:

$$\begin{aligned} P_{f_1} &= w_1, w_2, f_1 & \qquad P_{w_1} &= f_2, f_1, w_1 \\ P_{f_2} &= w_2, w_1, f_2 & \qquad P_{w_2} &= f_1, f_2, w_2 \end{aligned}$$

This job market has two stable matchings:

$$\begin{aligned} \mu_F &= \{\{f_1, w_1\}, \{f_2, w_2\}\} \ , \\ \mu_W &= \{\{f_1, w_2\}, \{f_2, w_1\}\} \ . \end{aligned}$$

The five other matchings are:

$$\begin{aligned} \mu_1 &= \{\{f_1, w_1\}, \{f_2\}, \{w_2\}\} \ , \\ \mu_1' &= \{\{f_2, w_1\}, \{f_1\}, \{w_2\}\} \ , \\ \mu_2 &= \{\{f_2, w_2\}, \{f_1\}, \{w_1\}\} \ , \\ \mu_2' &= \{\{f_1, w_2\}, \{f_2\}, \{w_1\}\} \ , \\ \mu_\varnothing &= \{\{f_1\}, \{f_2\}, \{w_1\}, \{w_2\}\} \ . \end{aligned}$$

Consider subgame where $f_1$ has already asked to $w_1$ but he refused the offer, and $f_2$ proposed to $w_2$ who also refused. In this subgame, $f_1$ has the choice between proposing to $w_2$ or to drop out of the market. Since $w_2$ is acceptable for $f_1$, in a subgame perfect equilibrium it is easy to see that it is never optimal for $f_1$ not to propose a job to $w_2$. The same applies for $f_2$ and $w_1$. Therefore, we can easily deduce that the SPE outcome for this subgame is $\mu_W$. Similarly, if $f_1$ proposed first to $w_2$ and $f_2$ to $w_1$ and both refused their offer, we obtain another subgame where the only subgame perfect equilibrium is $\mu_F$. Now, consider the original root of the game. For the same reason as before (i.e., the game tree would be to large), we shall not consider the case where a firm decides to drop out of the market without having proposed to all available workers, or when a worker decides not to be matched at the end of the game.

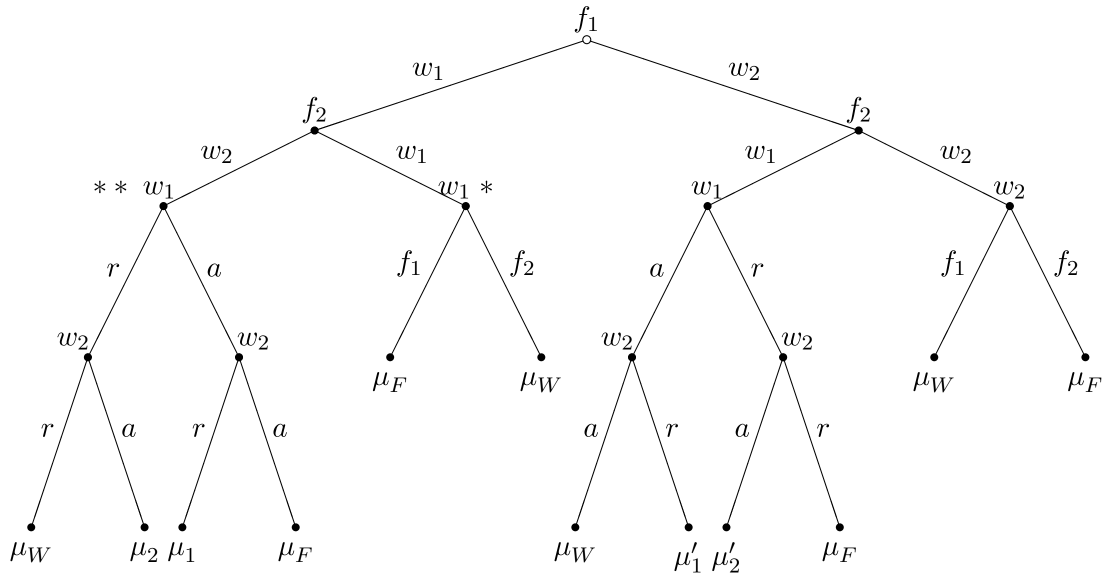


Figure 1: Two firms and two workers

The game tree reads as follows. At node $*$ where $w_1$ has a choice between $f_1$ and $f_2$, $w_1$ received two offers, and $w_2$ none. Then, $w_1$ can either accept $f_1$ (action $f_1$) or accept $f_2$ (action $f_2$). If he accepts $f_1$, then $f_2$ will make an offer in the next stage (not drawn) to $w_2$ whose best response is to accept. We then have the matching $\mu_F$. On the contrary, if $w_1$ chooses $f_2$, then $f_1$ will make an offer to $w_2$, which yields $\mu_W$. The same applies in the subgame where $w_2$ received two offers and $w_1$ none. Hence, the outcome in these two subgames is $\mu_W$. In the other subgames, it is obvious to see that $\mu_W$ is also the outcome obtained with backward induction.

Obviously, whatever the firms' decision during the first stage is, we obtain the worker-optimal matching. Hence, even in this simple game firms have at some subgame multiple best response, which entails in the existence of several SPE's.

It is also easy to see that making firms' decision simultaneously does not affect the outcome. The only part of the game where this simultaneity may affect the game is during the first stage. However, since the workers' sub-stage is with complete information, workers have still the possibility to "swap" between each their offers and attain the worker-optimal matching. Suppose now that workers' decisions are simultaneous (but firms' decision are not), and consider for instance the subgame $**$ of the game depicted in Figure 1 where $w_1$ received an offer from $f_1$ and $w_2$ from $f_2$. The tree at this stage is given in Figure 2.

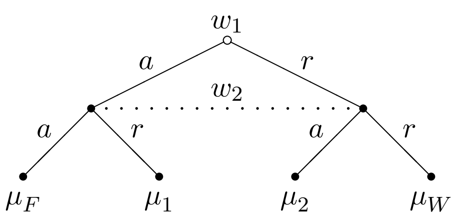


Figure 2: Workers' decisions are simultaneous.

In this subgame, there are two Nash equilibria: $(a, a)$ and $(r, r)$. If both workers choose to reject their offers the game continue and we reach the matching $\mu_W$. However, if they both accept, we obtain the matching $\mu_F$. Observe that any SPE yields to a stable matching, which would contradict the example proposed after Proposition 2. In fact, since both firms (resp. workers) are acceptable for both workers (resp. firms) it is easy to deduce that any SPE must yield to a matching in which all agents are matched. Since the only such matchings are $\mu_W$ and $\mu_F$ the observation easily follows. Thus, when workers' actions are simultaneous, the scope for cooperation is reduced. In this case, if a worker does reject his current offer he may not be able to receive a better one in exchange (at a later stage). For this to be the case, it must be the case that the other workers also do so. This may happen when we assume that workers' acceptance are deferrable (i.e., dropping the irreversibility assumption) and workers wait for a better offer before rejecting their current (most preferred) offer.

### 5.2 Three firms - three workers

Let $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, w_3\}$, with the following preferences:

$$P_{f_1} = w_2, w_3, w_1, f_1 \qquad P_{w_1} = f_1, f_2, f_3, w_1$$
$$P_{f_2} = w_1, w_2, w_3, f_2 \qquad P_{w_2} = f_2, f_1, f_3 w_2$$
$$P_{f_3} = w_3, w_1, w_2, f_3 \qquad P_{w_3} = f_1, f_3, f_2, w_3$$

Suppose that $f_1$ proposes first, then $f_2$ and then $f_3$, and assume that $f_1$ uses the following representation for her strategy: $\pi_{f_1} = w_2, w_1, w_3$, and $f_2$ and $f_3$ use for representation their own preferences. When the game starts, $w_1$ has an offer from $f_2$, $w_2$ has an offer from $f_1$ and $w_3$ from $f_3$.

Observe that if all workers accept their offer we obtain the worker-optimal matching. However, given the strategy used by $f_1$, it is optimal for $w_1$ to reject $f_2$'s offer. By doing so, $f_2$ will propose to her next best choice, $w_2$, who will accept (this is his most preferred firm). Then, $w_2$ rejects $f_1$. She then proposes to $w_1$, who accepts. Obviously, $f_1$ did not choose an optimal strategy. With the representation $\pi_{f_1} = P_{f_1}$ she would have been better off. Indeed, if $w_1$ rejects $w_2$, $w_2$ will still accept $f_2$ and reject $f_1$, but $f_1$ will now propose to $f_3$, who will accept. Then, $w_1$ is left with $f_3$, which makes him worse off.

Here, the key element in $f_1$'s optimal strategies is about the positions of the workers less preferred than the worker-optimal mate, $w_3$ and $w_2$. Observe that both $w_2$ and $w_3$ prefer $f_1$ to their worker-optimal mate. For a job market $(W, F, P)$, let

$$Q_f = \{w \in W \mid f P_w \mu_W(w) \text{ and } w P_f f\} ,$$

be the set of workers who prefer $f$ to their worker-optimal mate and are acceptable for $f$. By the stability of $\mu_W$, for all workers $w \in Q_f$ we have $\mu_W(f) P_f w$ . Define $\hat{w}_f$ as the worker such that

$$\hat{w}_f = \max_{P_f} Q_f .$$

That is, $\hat{w}_f$ is $f$'s most preferred acceptable worker among those worker who are less preferred than the worker-optimal mate and who prefer $f$ to their worker-optimal mate. In the job market we are considering in this example, $\hat{w}_{f_1} = w_3$. The fact that for any representation $\pi$ of $f$'s strategy we have,[15]

$$\pi(\hat{w}_f) = \pi(\mu_W) + 1 , \tag{1}$$

may be a key element for $f$'s to be matched to her worker-optimal mate (or any worker preferred to him). If Eq. (1) is satisfied, then $f$ will always propose to $\hat{w}_f$ before the other workers who are less preferred than $\mu_W(f)$. If not, then some workers may have an incentive to swap their mates ($w_1$ and $w_2$ in the example).

# 6 The literature

A startling feature of the literature on matching (whether applied or theoretical) is that decentralized market have received relatively little attention. In this respect, two comments are in order. First, some notable exceptions are Roth and Xiaolin [14], Blum, Roth and Rothblum [4], Peleg [9], and Alcalde, Pérez-Castrillo and Romero-Medina [2]. In [14], Roth and Xiaolin study the market for clinical psychologist which is, like the market studied

[15] See Appendix A for a definition of a representation.

in this paper, decentralized. A specific aspect of this market is that the proposing & acceptance-rejection game takes place in one single day. In this market candidates are allowed to hold an offer for a while, but their decisions are also irreversible. That an offer can be held for some period of time by a candidate is similar to the feature in the market for economists that the offer typically comes with a deadline. In the current paper, since the job market is modeled as a game of complete information, it is straightforward to see that giving the candidate some time to decide whether to accept the offer does not change the outcome of the game.[16] Indeed, all what matters in the complete information game is individuals' knowledge of the game. If, at some point a candidate happens to know that he will eventually receive an offer from a university $U$, which he plans to accept, there is no gain in delaying his responses to other, less preferred universities proposing to him earlier in the game than $U$. Roth and Xiaolin show that the market for clinical psychologist is actually somehow similar to a decentralized version of the well known *Deferred Acceptance algorithm* originally proposed by Gale and Shapley[6]. This calls for our second comment, which is about the very use of the Gale-Shapley algorithm in the literature. A quick glance at the plethora of papers about matching literature shows that a significant proportion these papers deal with Gale and Shapley's deferred acceptance algorithm or some of its refinements. All these algorithms, although being presented as *centralized* procedures where individuals have to submit their preferences, are actually described as purely *decentralized* procedures.As Gale and Shapley wrote:

> " *...let each boy propose to his favorite girl. Each girl who receives more than one proposal rejects all but her favorite among those who have proposed to her...*"

For a centralized market it would be more appropriate to write for instance (without altering the results):

> "*For each girl, take the set of boys who prefer her to any other girl, and match this girl to the one she prefers among these boys...*"

We are aware that this argument may be pedantic.[17] Nevertheless, it is our contention that, even if the same algorithm is employed, a centralized market and its decentralization version are not completely equivalent. In a centralized version, agents are matched by a "match-maker" (or the clearing

[16]Of course, as we said before, if we drop the irreversibility assumption holding offers may "block" the market and leave workers with outcomes less preferred than the one had they not held their offers during the course of the game.

[17]The distinction between centralized and decentralized presentation of the Deferred Acceptance algorithm has already pointed out by Roth and Rothblum [13].

house running the algorithm). This latter is usually assumed to be neutral, and thus there is no reason why it should un-match a firm and a worker previously matched when there is no second firm, not matched, that is preferred to the first one by the worker. Yet, in a decentralized version, a worker could well decide to refuse being matched to some (acceptable) firm even if there is no better alternative. This feature, characteristic of decentralized markets, is already pointed out by Blum, Roth and Rothblum [4]. Their work deals with markets for senior positions, which are usually decentralized. Although most of their paper is devoted to a centralized version of such markets,[18] their work also addresses the issue of decentralized markets. The game analyzed in [4] differs from ours. In contrast to our approach, they assume that there is a high level of uncertainty. Each firm only knows which of her offers have been rejected (by which worker), and each worker, at any stage, only knows which offer she has received so far. Furthermore, they restricted their study to the case where each agent do use "preference strategies", i.e., firms' strategies that follow a specific, fixed order of proposals, and workers' strategies that decide acceptance or rejection of offers following a unique ordering (which may not necessarily coincide with the workers' preferences). In this context, [4] analyze Nash equilibria and showed that they coincide with the equilibria of a centralized market.[19] Other papers related to decentralized matching markets are by Peleg [9], Alcalde, Pérez-Castrillo and Romero-Medina [2], and Alcalde and Romero-Medina [3]. Their models differs from ours, however, in an important respect. Like us, they study a sequential market but with the difference that firms can only propose once.[20] In contrast, we suppose here that there is a "continuation game" in the sense that firms whose offers have been rejected can propose their position to other workers (if any). This allows us then to construct an SPE where workers can credibly "threat" to reject all their offers if a firm deviates. When there is no continuation stages, the threat of rejecting all offers that differ from their worker-optimal mate is not credible.

# 7 Conclusion

In most of cases, we found that the worker-optimal matching is the unique subgame perfect equilibrium outcome. Firms' optimal strategies can shed some light on the reason why we obtain this outcome. For a firm, the order with which she proposes does not matter as long as it concerns those workers

[18] [4] studies how markets for senior positions can be re-stabilized after new positions have been created or after the retirement of some workers.

[19] Because in their game there is a random ordering under which firms make offers they use a different equilibrium concept than the Nash equilibrium (although similar in its spirit), which they called *Realization-Independent equilibrium*. This equilibria can be interpreted as an "Ex-post" Nash equilibrium.

[20] [2] only looks at the case where within each sub-stage agents' actions are simultaneous.

preferred to the worker-optimal mate. However, for workers less preferred to this latter, it must be the case that the sequence of offers follows, to some extent, her preferences. Roughly speaking, for any matching such that workers are matched to a firm less preferred to their worker-optimal mate, it is in the interest of all workers to reject such a matching. That way, they can help each other to attain the worker-optimal matching. In other words, as long as we have not reached the worker-optimal matching, workers' interests do coincide. For a firm, all workers that are preferred to her worker-optimal mate do consider this very firm as being less preferred than their own worker-optimal mate. Hence, whatever is the order of proposal between these workers, the firm will be rejected anyway. However, for matchings that are preferred to the worker-optimal matching, workers do enter into competition. For a worker to be better off, it must be the case that another worker is worse off, and does indeed accept to be worse off (i.e., he rejects his worker-optimal mate). Hence, to maintain this competition (and avoid being worse off the at the worker-optimal matching), firms do have to propose to workers less preferred than the worker-optimal mate with the same ordering that their preferences dictates (see section 5.2). Results differ when we workers' moves are simultaneous. In this case, we showed that the set of subgame perfect equilibrium outcomes includes the set of stable matchings. Results can be summarized in Table 1.

| | | Firms | |
|---|---|---|---|
| | Moves | simultaneous | non-simultaneous |
| Workers | simultaneous | includes $S(P)$ | includes $S(P)$ |
| | non-simultaneous | $\mu_W$ | $\mu_W$ |

Table 1: Outcomes of subgame perfect equilibria.

These results greatly contrast with those obtained in centralized market. Indeed, if the market is centralized and firms do make the proposals, the deferred acceptance algorithm yields to the firm optimal matching. Such an reversal of outcome is not new in the matching literature. Alcalde [1] studies the case of a centralized market where the deferred acceptance algorithm is used to match agents, and these latter act strategically. He shows that with the iterative deletion of dominated strategies we also obtain the worker optimal matching when firms do propose. In Alcalde's game, a player' strategy set is the set of all possible preferences she can have (one of which is the true preference list). We then have a normal form game where each player has to choose a ranking, which will be communicated to a central authority. This latter will match the agents using the deferred acceptance algorithm, and agents' "payoffs" are given by "how much" they like their mate with their true preferences. A natural question would be then to see if there is a link between our game and that of Alcalde, i.e., if the two games are equivalent. In fact, there is no relationship whatsoever. In our game, we have seen that the firms have several optimal strategies at the equilibrium, whereas Alcalde shows that there it is a dominant strategy for the firms to submit their true preferences.

## Appendix A: Sanitization strategies

For $f \in F$, let $\pi_f$ be any ordering of $W \cup \{f\}$ with $\pi_f(v)$ being the index of $v$ in $\pi_f$. Let $\pi^*_f$ be the ordering that follows the preferences of firm $f$, i.e.,

$$\pi^*_f(v) < \pi^*_f(v') \quad \Leftrightarrow \quad vP_f v', \, \forall\, v, v' \in W \cup \{f\}\,. \tag{2}$$

**Definition 1** *Given an ordering $\pi_f$, a firm $f$ and a node $\alpha$ where $f$ acts, $\pi_f$ is a **representation** of $\sigma_f$ at $\alpha$ if,*

$$a_f(\alpha) = w \quad \Leftrightarrow \quad \nexists\, w' \in W(\alpha)\backslash \mathbf{O}^\alpha_f \textit{ such that } \pi_f(w') < \pi_f(w)\,. \tag{3}$$

Note that given the action of a firm at some node, we may find several distinct representations. All that is required is that the worker to whom the firm is proposing is ranked first, according to the order $\pi$, among the workers who are still available to $f$. In order to complete the characterization of sanitization strategies we still need to specify firms' actions in those subgames reached when a firm deviated from its initial plan of offers.

We denote by $\sigma_f|_\alpha$ the strategy of firm $f$ in the subgame $\gamma(\alpha)$ where we have deleted $f$'s action at node $\alpha$. The reason for doing this is to make the definition of a sanitization strategy consistent with the fact that a firm may not follow her initial strategy at node $\alpha$. Consider for instance a firm $f$ whose plan of offers is $w_1, w_2, w_3, w_4, w_5$, but instead of making an offer to $w_2$ at the third stage (say node $\alpha$) of the game she makes an offer to $w_4$. It is easy to see that the consistency requirement would not make any sense had we considered $f$'s strategy in the subgame $\gamma(\alpha)$ *including* her action at node $\alpha$.

**Definition 2** *A strategy $\sigma_f$ is a **sanitization strategy** if, at any node $\alpha$ where $f$ has to play, $\sigma_f|_\alpha$ admits a representation $\pi$. The strategy $\sigma_f$ is a **strong sanitization strategy** if it admits for representation $\pi^*_f$.*

We denote by $\Sigma^*_f$ the set of strong sanitization strategies of firm $f$.

## Appendix B: Proofs

Without loss of generality we shall consider throughout the proofs a "truncated" version of the game where we delete those actions corresponding to a firm's offer to an unacceptable worker, offer to a worker by an unacceptable firm, and the actions of a worker if receiving such offers. In this truncated game, any strategy profile yields an individually rational matching for all workers and firms. Since the game is with perfect information, the original game and the truncated one are obviously strategically identical in terms of subgame perfection.

Workers' optimal strategies can be characterized quite easily. For a strategy profile $\sigma$, define by $\mathcal{O}_w^\sigma$ the set of possible offers $w$ can have if he opts for rejecting all his offers, given that all other players do not change their strategies.

**Lemma 1** *$\sigma \in SPE(\mathcal{G}(P))$. Then for all $w \in W$, $\mu[\sigma](w) = \max_{P_w} \mathcal{O}_w^\sigma$.*

**Proof** Since a subgame perfect equilibrium is necessarily a Nash equilibrium the claim easily follows from the fact that for any strategy profile $\sigma$ choosing $\max_{P_w} \mathcal{O}_w^\sigma$ is the best response of worker $w$. ■

**Lemma 2** (Roth [11]). *Let $(F, W, P)$ be a matching market and $\mu_W$ its worker-optimal matching. There is no individually rational matching $\mu$ such that for all $w \in W$ we have $\mu P_w \mu_W$.*

**Lemma 3** *Let $(F, W, P)$ be a matching market and $\mu_W$ its worker-optimal matching. Assume that there exists a set $\hat{W}$ and a matching $\bar{\mu}$ such that for all $w \in \hat{W}$, $\bar{\mu}(\mu(w)) \in \hat{W}$, $\bar{\mu}(w) P_w \mu_W(w)$. Then there exists $w_i, w_j \in \hat{W}$ and $w_k \notin \hat{W}$ with $w_j = \bar{\mu}(\mu_W(w_i))$ such that*

*(i)* $w_k P_{\mu_W(w_i)} w_j$ ,

*(ii)* $\mu_W(w_i) P_{w_k} \mu_W(w_k)$ .

**Proof** Let $(F, W, P)$ be a matching market and let $w_i, w_j \in \hat{W}$ such that $\bar{\mu}(w_i) = \mu_W(w_j)$. Consider the execution of the deferred acceptance algorithm with workers proposing. Worker $w_i$ will eventually ask to $\mu_W(w_j)$ before $\mu_W(w_i)$. Since he will be rejected by $\mu_W(w_j)$, there should be another worker, say $w_l$, such that $\mu_W(w_j)$ rejects $w_i$ in favor of $w_l$, i.e., $w_h P_{\mu_W(w_j)} w_i$ . We then have two cases.

*Case (i):* $w_l \notin \hat{W}$. Since $w_l$ is the player that $\mu_W(w_j)$ prefers to $w_i$, we have $w_l P_{\mu_W(w_j)} w_i$ . To obtain the lemma, it suffices now to show that $\mu_W(w_j) P_{w_l} \mu_W(w_l)$ . This is easily done, since we can observe that $w_l$ proposes to $\mu_W(w_j)$ before the algorithm stops for him,[21] which implies that $w_l$

[21] For each player on the proposing side, the deferred acceptance algorithm stops when this player is matched to his optimal mate.

prefers $\mu_W(w_j)$ to his optimal mate. We then have *(i)* $\mu_W(w_j)P_{w_l}\mu_W(w_l)$, and *(ii)* $w_l P_{\mu_W(w_j)} w_i$ . Relabel now $w_l$ into $w_k$, and $w_i := w_j$ , $w_j := w_i$ and we obtain the statement of the lemma.
*Case (ii):* $w_l \in \hat{W}$. Let $\hat{W} = \{w_1, \ldots, w_h, w_i, w_j\}$. Worker $w_j$ was rejected by $\mu_W(w_i)$ in favor of another worker, say $w_1$. Hence, $w_1$ was rejected by his previous mate by another worker, say $\bar{w}$. Either $\bar{w} \in \hat{W}$ or $\bar{w} \notin \hat{W}$. In the latter case, we can use the argument developed in *(i)*. If $\bar{w} \in \hat{W}$, then assume that $\bar{w} = w_2$. Again, for $w_2$ to propose $w_1$'s previous mate, it must be the case that $w_2$ was rejected by his own previous mate. Continuing this way, we arrive eventually to $w_{i-1}$ who has been rejected by his previous mate. If for one of these workers, his previous mate did rejected him in favor of a worker not in $\hat{W}$, then we are back to case *(i)*. Recall that a worker must be unmatched to make a proposal. Therefore, it must be the case that when workers in $\hat{W}$ "swap" their mate, one of them, say $w_j$, is unmatched before making his proposal. That is, $w_j$ is unmatched because he was rejected by $\mu_W(w_i)$. Since it cannot be in favor of $w_i$ (who is matched to $\mu_W(w_{i-1})$), it must be in favor of another player, say $w_l$, not in $\hat{W}$. This implies that *(i)* $\mu_W(w_i)P_{w_l}\mu_W(w_l)$, and *(ii)* $w_l P_{\mu_W(w_i)} w_j$ . Statement *(i)* is from the fact that $w_l \neq w_i$. and statement *(ii)* is from the fact that $w_j$ was rejected by $\mu_W(w_i)$ in favor of $w_l$ . Relabel now $w_l$ into $w_k$ and we obtain the statement of the lemma. ■

Let $\Phi(P, F, W)$ be the normal form game with workers and firms where a strategy of $v \in F \cup W$ consists of sending submitting a preference profile and the outcome are obtained using the Deferred-Acceptance algorithm with workers proposing with the submitted preferences, but agents payoffs are evaluated using their original "true" preferences.[22]

**Lemma 4** (Dubins and Freedman [5]). *In the game $\Phi(P, F, W)$ it is a dominant strategy for any agent in $W$ to declare his true preferences.*

**Lemma 5** *Let $(W, F, P)$ be a job market. Consider a worker $w$ such that $\mu_W(w) \neq w$. Let $f := \mu_W(w)$. Let $P'$ such that for all $v \neq w$, $P_v = P'_v$, and $P'_w$ such that for all $v, v' \neq \mu_W(w)$, $vP_w v'$ if and only if $vP'_w v'$ and $wP'_w f$. Then, with $\hat{v}$ being the mate of $w$ under the worker-optimal matching under the preference profile $P'$ we have $fP_w\hat{v}$ .*

**Proof** Immediate by Lemma 4. ■

The next result states that if all firms use a sanitization strategy with representation being their preferences, then a subgame perfect equilibrium must yields the candidate-optimal matching $\mu_W$. Note that this game differs from the original one since firms have less strategic options: they are

[22]See [1, 5, 12] for a description of the game.

forced to use a sanitization strategy. Let denote $SPE(\mathcal{G}(P))$, set of subgame perfect equilibria of the game $\mathcal{G}(P)$.

**Lemma 6** *Let $\mathcal{G}^F(P)$ be the game such that workers' strategy set are the same as in $\mathcal{G}(P)$, and for all firms $f \in F$, $\Sigma_f = \Sigma_f^*$ and for all nodes $\alpha$, $P_f$ is the representation of $\sigma_f|_\alpha$. Let $\sigma \in SPE(\mathcal{G}^F(P))$. Then $\mu[\sigma] = \mu_W$.*

**Proof** We prove the Lemma by induction on the number of stages needed to end the game. When there is only one stage left the result is obvious: Each unmatched firm (if any) has only one available worker left to propose to. Suppose then that there are two stages (i.e., two "offering" stages and two "responding" stages).[23] Let $\gamma(\alpha)$ be any subgame where there are only two stages left and $\sigma$ be an SPE. ($\alpha$ is the first node of this subgame.) Consider the preference profile $P^\alpha$. Since there are only two stages left, each firm $f \in F(\alpha)$ has at most to acceptable workers in her preferences $P_f^\alpha$.[24] Since $\sigma_F = \sigma_f^*$, and $\sigma$ is an SPE, $w$ such that $\mu[\sigma](w) \in F^\alpha$ $\Leftrightarrow \mu_F^\alpha(w) \in F^\alpha \Leftrightarrow \mu_W^\alpha(w) \in F^\alpha$. Indeed, if a worker receives an (acceptable) offer he is sure to be matched at the end of the game at any SPE. It follows, using the deferred acceptance algorithm that he's matched at $\mu_F$, and thus at $\mu_W$, too. Hence, $f$ such that $\mu[\sigma](f) \Leftrightarrow \mu_W^\alpha(f) \in W^\alpha$. Suppose that $\sigma$ is such that in the subgame $\gamma(\alpha)$ there is a worker, say $w_1$ such that $\mu_W^\alpha(w_1) P_{w_1}^\alpha \mu[\sigma](w_1)$. Let $f_1 = \mu_W^\alpha(w_1)$. Since $\mu[\sigma](w_1) \neq f_1$, $f_1 \notin \max_{P_w} \mathcal{O}_w^\sigma$. Therefore, $\mu[\sigma](f_1) \in W(\alpha)\backslash\{w_1\}$. Let $w_2 = \mu[\sigma](f_1)$. Since there are only two stages left, $f_1$ has only two acceptable workers in $P_{f_1}^\alpha$. Moreover, $\sigma_{f_1} = \sigma_{f_1}^*$ implies that $w_2 P_{f_1}^\alpha w_1$. Since $w_2$ is matched, it follows that $\mu_W^\alpha(w_2) \in F(\alpha)$. Let $f_2 = \mu_W^\alpha(w_2)$. By the stability of $\mu_W^\alpha$, we have $f_2 P_{w_2}^\alpha f_1$.

If $f_2 \neq \mu[\sigma](w_1)$ then we can deduce that there exists a worker, say $w_3$, such that $\mu[\sigma](w_3) = f_2$ and for some $f_3 \in F^\alpha$, $\mu_W^\alpha(w_3) = f_3$. Again, we have either $f_3 = \mu[\sigma](w_1)$ or $f_3 \neq \mu[\sigma](w_1)$. Since the number of workers and firms is finite we will eventually reach a firm $f_k$ such that $\mu[\sigma](f_k) = w_1$ and $\mu_W^\alpha(f_k) = w_k$. (The case where $f_2 = \mu[\sigma](w_1)$ corresponds to $k = 2$.) We then obtain a set of firms and workers, $\hat{F} = \{f_1, \ldots, f_k\}$ and $\hat{W} = \{w_1, \ldots, w_k\}$ such that both $\mu_W^\alpha$ and $\mu_F^\alpha$ map $\hat{F}$ onto $\hat{W}$, and for all $v \in \hat{W} \cup \hat{F}$, $\mu_W^\alpha(v) \neq \mu_F^\alpha(v)$. Observe that this implies under $\sigma$ each worker $w \in \hat{W}$ is matched at the penultimate stage.[25]

Consider first the worker in $\hat{W}$ who has to decide, say $w_1$.[26] If $w_1$ rejects all offers he has, then we enter a new subgame than the one had $w_1$ accepted

[23]This case is similar to the example depicted in Section 5.1.

[24]Recall we are using the truncated game as defined in the beginning of Appendix B.

[25]Clearly, there might be several "couples" of such sets $\hat{W}$ and $\hat{F}$, i.e., sets $\hat{W}_1, \ldots, \hat{W}_l$ and $\hat{F}_1, \ldots, \hat{F}_l$ such that $\mu_F^\alpha$ (resp. $\mu_W^\alpha$) maps $\hat{W}_h$ onto $\hat{F}_h$, $h = 1, \ldots, l$. For the sake of simplicity we shall continue the proof assuming that there is only one such couple of sets.

[26]Since there is perfect information there is only one such worker.

the offer from $\mu[\sigma](w_1) = f_k$. So now $w_k = \mu_W^\alpha(f_k)$ knows that he will receive an offer from $f_k$ if he rejects his current offers.[27]

We claim that if $w_k$ rejects $f_{k-1}$'s offer he cannot receive a better offer than $f_k$. To see this, suppose that there is $\tilde{f}_1$ such that $\tilde{f}_1$ makes an offer to $w_k$ if he rejects $f_{k-1}$'s offer, where $\tilde{f}_1 P_{w_k}^\alpha f_k$. Clearly, $\tilde{f}_1 \notin \hat{F}$, otherwise it would mean that $\tilde{f}_1$ (say $\tilde{f}_1 = f_h$, $h \neq 1, k, k-1$) has three acceptable workers left, namely $w_h, w_{h+1}$, $w_2$, a contradiction since there are only two stages left. Hence, $\tilde{f}_1 \notin \hat{F}$ and she has at most two acceptable workers left in her preferences, $P_{\tilde{f}_1}^\alpha$. The only way to sustain the stability of $\mu_W^\alpha$ is that $\exists\ \tilde{w}_1 \notin \hat{W}$ such that $\mu_W^\alpha(\tilde{f}_1) = \tilde{w}_1$ and $\tilde{w}_1 P_{\tilde{f}_1}^\alpha w_k$. If $\tilde{f}_1$ proposes to $w_k$ in the last stage, it means that $\tilde{w}_1$ rejected her in the penultimate stage, for a better offer, say by $\tilde{f}_2$, where $\tilde{f}_1 \notin \hat{F}$. Hence, $\tilde{f}_2 P_{\tilde{w}_1}^\alpha \tilde{f}_1$. For $\mu_W^\alpha$ to be stable, it must be the case that $\exists\ \tilde{w}_2 \notin \hat{W}$ such that $\mu_W^\alpha(\tilde{f}_2) = \tilde{w}_2$ and $\tilde{w}_2 P_{\tilde{f}_2}^\alpha \tilde{w}_1$. Thus $\tilde{w}_2$ rejects $\tilde{f}_2$'s offer. Continuing this way, we shall eventually obtain a set of workers, $\tilde{W} = \{\tilde{w}_1, \ldots, \tilde{w}_l\}$ and a set of firms, $\tilde{F} = \{\tilde{f}_1, \ldots, \tilde{f}_l\}$ such that $\mu_W^\alpha(\tilde{w}_h) = \tilde{f}_h$, and in the subgame obtained when $w_1$ rejects $f_2$'s offer, the SPE is such that $\tilde{w}_h$ rejects $\tilde{f}_h$'s offer to obtain in the last stage an offer from $\tilde{f}_{h+1}$, for all $h = 1, \ldots, l$ modulo $l$. In particular, we have $\tilde{w}_k$ rejects $\tilde{f}_k$'s offer to accept that of $\tilde{f}_1$. That is, both $w_k$ and $\tilde{w}_l$ are expecting an offer from $\tilde{f}_1$ in the last stage, a contradiction. It follows that $w_k$ will not receive in the last stage a better offer than that of $f_k$, if he rejects that of $f_{k-1}$ in the penultimate stage, which proves the claim.

Hence, the SPE must be such that if $w_1$ rejects $f_2$'s offer in the penultimate stage, $w_k$ rejects $f_{k-1}$'s offer in the penultimate stage. Continuing this way, we attain worker $w_2$ who rejects $f_1$'s offer in the penultimate state. Then, $w_1$ will receive in the last stage an offer from $f_1$. For the same reason than with $w_2$, it can be shown that $w_1$ cannot expect to receive a better offer than $f_1$. In any case, $w_1$ is better off rejecting $f_2$'s offer, which contradict the fact that $\sigma$ is an SPE. To complete the proof of the induction hypothesis, observe that if $\exists w$ such that $\mu[\sigma](w) P_w^\alpha \mu_W^\alpha(w)$ then either $\mu[\sigma]$ is not individually rational or $\exists w' \neq w$ such that $\mu_W^\alpha(w') P_{w'}^\alpha \mu[\sigma](w')$. Therefore, when there are two stages left, any SPE gives the worker-optimal matching.

Suppose now that for all subgames $\gamma(\alpha)$ such that there are $k$ stages left any SPE gives the matching $\mu_W^\alpha$. Consider now any subgame $\gamma(\alpha')$ with $k+1$ stages left. By the stability of $\mu_W^{\alpha'}$ and because $\sigma_F = \sigma_F^*$, no worker $w \in W(\alpha')$ receives at stage $k+1$ an offer from a firm $f$ such that $f P_w^\alpha \mu_W^{\alpha'}(w)$. Otherwise we have $w P_f \mu_W^\alpha(f)$, contradicting the stability of $\mu_W^\alpha$. If a worker $w \in W(\alpha')$ receives at stage $k+1$ an offer from $\mu_W^{\alpha'}(w)$ we claim that he must accept it. Suppose he rejects such an offer. Then, in the

[27]Note that since firms make their offers first and then workers respond, the set of offers made to $w_k$ during the penultimate stage does not change if $w_1$ deviates. The same applies for the other workers.

next stage, say starting at node $\alpha''$, we have $wP_w^{\alpha''}\mu_W^{\alpha'}(w)$, and according to the induction hypothesis $w$ will be matched to $\mu_W^{\alpha''}(w)$. By Lemma 5 it follows that $\mu_W^{\alpha''}(w)P_w\mu_W^{\alpha'}(w)$. That is, $w$ is worse off rejecting $\mu_W^{\alpha'}(w)$'s offer, which proves the claim. To sum up, when there are $k+1$ stages left, (active) workers can receive offers from their worker-optimal mate or from less preferred firms. The issue now is that if for worker $w$ all his offers at stage $k+1$ are less preferred than $\mu_W^{\alpha'}(w)$. In this case, another worker $w'$ must have received an offer from $\mu_W^{\alpha'}(w)$. It may happen that, if $w$ rejects all his offers that $w'$ accepts $\mu_W^{\alpha'}(w)$'s offer. If not, then $w$ is ensured by the induction hypothesis to be matched to $\mu_W^{\alpha'}(w)$ when the game ends. To complete the proof we then have to show that all such workers $w$ must reject all the offers they have. Consider the last worker, say $w_h$ who has to decide at the end of the stage beginning at $\alpha'$ among those workers not having received an offer from their worker-optimal mate (by the previous argument we can get rid of the other workers). Let $\alpha_h$ be any the node where $w_h$ acts in the subgame $\gamma(\alpha')$. Note that in the subgame $\gamma(\alpha')$ there are several such nodes $\alpha_h$, each one depending on the history of decisions by the previous workers acting during the stage beginning at $\alpha'$. For all such nodes $\alpha_h^1, \ldots, \alpha_h^l$, the job market $(F(\alpha_h^r), W(\alpha_h^r), P^{\alpha_h^r})$, $r = 1, \ldots, l$ uniquely determines a worker-optimal mate for $w_h$. Of course, we may have $\mu_W^{\alpha_h^r}(w_h) \neq \mu_W^{\alpha_h^{r'}}(w_h)$ for some $r \neq r'$. If $w_h$ rejects his offer(s), then in the next subgame there are only $k$ stages left and the induction hypothesis applies, i.e., $w$ will be matched to $\mu_W^{\alpha_h^r}$, for all $r = 1, \ldots, l$. Therefore, $w$ is better off rejecting the offer(s) he has. Consider now the worker who has to decide just before $w_h$, say $w_{h-1}$. As we said above, the only reason that would lead $w_{h-1}$ to rejects his offer(s) is because $w_h$ holds $f_{h-1}$'s offer (workers $w_1, \ldots, w_{h-2}$, have already taken their decision). Since $w_h$ will reject this offer, $w_{h-1}$ knows that by rejecting his he at node $\alpha_{h-1}^r$ he will be matched to $\mu_W^{\alpha_{h-1}^r}(w_{h-1})$ at the end of the game. Since at the node $\alpha_{h-1}^r$ he does not have a better offer, his best response is to reject his offer(s). Repeating the argument with workers $w_{h-2}$, ..., $w_1$ we can deduce that all workers will reject their offers, and thus the outcome of any SPE when there are $k+1$ stages left gives the worker-optimal matching. ■

**Lemma 7** *Let $\sigma \in SPE(\mathcal{G}(P))$ and $\mu[\sigma] \neq \mu_W$, such that for at least one firm $f$, $\mu_W(f)P_f\mu[\sigma](f)$ . Then there exists $\tilde{f} \in F$ such that*

- $\sigma_f \neq \sigma_f^*$ ;
- $\mu_W(f)P_f\mu[\sigma](f)$ .

**Proof** Denote $\hat{F} = \{f \in F : \sigma_f \in \Sigma_f^*\}$. Let $f_1$ such that $w_1 = \mu_W(f_1)P_{f_1}\mu[\sigma](f_1)$. Observe that because preferences are strict we have

$\mu_W(f_1) \in W$. If $f_1 \notin \hat{F}$ then we are done. Hence, suppose that $f_1 \in \hat{F}$. It follows that $f_1$ did propose to $w_1$ (or was planing to if $w_1$ did accepted another offer) before proposing to $\mu[\sigma](f_1)$. Since $w_1$ could ensure to have at least $f_1$, it follows that $\mu[\sigma](w_1) \in F$. Let $f_2 = \mu[\sigma](w_1)$. We then have $f_2 P_{w_1} f_1$, which implies that $\mu_W(f_2) \in W$ and $w_2 P_{f_2} w_1$, where $w_2 := \mu_W(f_2)$. If $\sigma_{f_2} \neq \sigma^*_{f_2}$ we are done. Suppose then that $\sigma_{f_2} = \sigma^*_{f_2}$. It follows that there is $f_3$ such that $\mu[\sigma](w_2) = f_3$ and $f_3 P_{w_2} f_2$. Again, we deduce that there is $w_3$ such that $w_3 := \mu_W(f_3)$ and $w_3 P_{f_3} w_2$. Then, either $\sigma_{f_3} \neq \sigma^*_{f_3}$ or $\sigma_{f_3} = \sigma^*_{f_3}$. In the first case, we are done. Suppose then that $\sigma_{f_3} = \sigma^*_{f_3}$. We then have two cases: *(i)* $f_3 = f_1$, or *(ii)* $f_3 \neq f_1$. Our claim is that in both cases, if $f_3 \in \hat{F}$ we can deduce that there is another firm which is worse off than at the worker-optimal matching.

*Case (i).* We then have the conditions of Lemma 3 where $\hat{W} = \{w_1, w_2\}$. Applying Lemma 3 we deduce that there exists another worker, say $w_3 \neq w_1, w_2$ such that for a firm, say $f_2$, we have $w_3 P_{f_2} w_1$, and $f_2 P_{w_3} \mu_W(w_3)$. Since $\sigma_{f_2} \in \Sigma^*_{f_2}$, it follows that $w_3$ is matched to another firm, say $f_3$. Indeed, given the strategies played by $f_2$ (i.e., proposing to $w_3$ before $w_1$, this latter being her match under $\sigma$), $w_3$ could eventually receive an offer from $f_2$ but finally took a better offer: the one from $f_3$. That is, we have $f_3 := \mu[\sigma](w_3) P_{w_3} \mu_W(w_3)$, which implies that, by the stability of $\mu_W$, there exists $w_4 \in W$ such that $w_4 = \mu_W(f_3)$ and $w_4 P_{f_3} w_3$. If $f_3 \notin \hat{F}$ then we are done. In the other case, observe that we have now identified 3 firms in $\hat{F}$: $f_1$, $f_2$ and $f_3$. It follows that $w_4$ is matched to another firm, say $f_4$, under $\sigma$ such that $f_4 P_{w_4} \mu_W(w_4)$ and $\mu_W(f_4) P_{f_4} w_4$. We can repeat the argument in *(i)* if $f_4 \in \{f_1, f_2, f_3\}$, or the argument in *(ii)* otherwise. In both cases, we will eventually find another firm, say $f_5$, such that $\mu_W(f_5) P_{f_5} \mu[\sigma](f_5)$.

*Case (ii)* Then, $\exists$ $f_4$ such that $f_4 = \mu[\sigma](w_3) P_{w_3} f_3$. By stability, it follows that $w_4 := \mu_W(f_4) P_{f_4} w_3$. If $f_4 \notin \hat{F}$ then we are done. If $f_4 \in \hat{F}$ then we have $f_5 = \mu[\sigma](w_4)$ such that $f_5 P_{w_4} f_4$ (because $f_4$ did propose or was planning to propose to $w_4$). By stability of $\mu_W$, we have $w_5 := \mu_W(f_5) P_{f_5} w_4$. If $f_5 \in \{f_1, f_2, f_3, f_4\}$ we can use the argument given in *(i)*, or the argument in *(ii)* otherwise. In both cases, we will eventually find another firm, say $f_6$, such that $\mu_W(f_6) P_{f_6} \mu[\sigma](f_6)$. Continuing this way, we may eventually obtain a firm $f$ such that $\mu[\sigma](f) P_f \mu_W(f)$ and $f \notin \hat{F}$, otherwise we obtain $\hat{F} = F$, which would contradicts lemma 6. ■

The following lemma states that at any subgame perfect equilibrium all firms can ensure to be matched to their candidate optimal mate worker corresponding to the worker-optimal matching or to a worker they prefer to him. Such an equilibrium can be obtained when firms use sanitization strategies.

**Lemma 8** $\sigma \in SPE(\mathcal{G}(P)) \Rightarrow \forall\, f \in F,\ \mu[\sigma](f) R_f \mu_W(f)$ .

**Proof** Suppose now that the lemma is not true. Let $\sigma \in SPE(\mathcal{G}(P))$ such that for some $f \in F$, $\mu_W(f) P_f \mu[\sigma](f)$. By lemma 6, there exists an non-empty strict subset of $F$, $\hat{F}$, such that $\sigma_{\hat{F}} \notin \Sigma^*_{\hat{F}}$ . We proceed by induction on the size of $\hat{F}$ to show that $\sigma$ cannot be an SPE.

Suppose first that $\sharp\hat{F} = 1$ and let $\hat{F} = \{f_1\}$. By lemma 7, $\mu_W f_1 P_{f_1} \mu[\sigma](f_1)$ . If $f_1$ switch to a strategy in $\Sigma^*_{f_1}$, then all firms use a sanitization strategy, and thus, using Lemma 6, $f_1$ obtains $\mu_W(f_1)$ . Therefore, $\sigma$ cannot be an SPE, a contradiction. Suppose now that for all $h \leq k$ such that $\sharp\hat{F} \leq h$ it is true that $\sigma$ cannot be an SPE and that it is optimal for all firms to play $\sigma^*$. Suppose now that $\sharp\hat{F} = k + 1$. Again by lemma 7, there is at least one $f \in \hat{F}$ such that $\mu_W(f) P_f \mu[\sigma](f)$ . If $f$, switches to $\sigma^*_f$, then we have the strategy profile $\sigma' = (\sigma^*_f, \sigma_{-f})$, which necessarily implies that there are only $k$ firms not using a sanitization strategy at any subgame. Thus, the induction hypothesis applies and we deduce that at any subgame the other firms will play $\sigma^*$. In this case $f$ obtains $\mu_W(f)$, which shows that if $\sigma$ is such that $\sharp\hat{F} = k + 1$ cannot be an SPE either. ■

**Lemma 9** *If for some $w \in W$, $\mu_W(w) = w$ then for all $\sigma \in SPE(\mathcal{G}(P))$, $\mu[\sigma](w) = w$.*

**Proof** Let $w$ such that $\mu_W(w) = w$. Suppose that for some $\sigma \in SPE(\mathcal{G}(P))$, $\mu[\sigma](w) \neq w$. Denote $f := \mu[\sigma](w)$. Since $\sigma$ is an SPE, $\mu[\sigma]$ is individually rational, and therefore we have $f P_w w$. Moreover, $\mu_W(f) \neq \mu[\sigma](f)$. It follows by lemma 8 that $w P_f \mu_W(f)$, which contradicts the stability of $\mu_W$. ■

**Lemma 10** $\sigma \in SPE(\mathcal{G}(P)) \;\Rightarrow\; \forall\, w \in W,\ \mu_W(w) R_w \mu[\sigma](w)$ .

**Proof** Immediate by Lemma 2. ■

**Proof of Theorem 1** By Lemma 9 we can get rid of those workers $w$ for whom we have $\mu_W(w) = w$. We prove the lemma by induction on the number of steps. When there are only two steps left the result easily follows: the argument is almost identical to that in the proof of Lemma 6. The main difference is when there is at some $\sigma \in SPE(\mathcal{G}(P))$ a worker $w$ such that $\mu_W(w) P_w \mu[\sigma](w)$. In this case, $f = \mu_W(w)$ is matched to another worker $w'$ in the penultimate stage. In the proof of Lemma 6 we deduce that $w' P_f w$ because $\sigma_f = \sigma^*_f$. Since the game now is such that firms may not necessarily follow their strong sanitization strategy, the fact that $w' P_f w$ follows from Lemma 8.

Suppose now that for all subgames $\gamma(\alpha)$ such that there are $k$ stages left any SPE gives the matching $\mu^\alpha_W$. By Lemma 10 we can deduce that workers cannot obtain a better outcome than the worker-optimal matching at any SPE. Hence, it suffices to show that the optimal strategies for each worker consists of accepting the worker-optimal mate if receiving an offer from her and rejecting all offers that are less preferred than the worker-optimal mate. Consider now any subgame $\gamma(\alpha')$ with $k+1$ stages left. Let $\{w_1, \dots, w_h\} = W(\alpha')$, and let $w_h$ be the last worker deciding at the end of the sub-stage beginning at node $\alpha'$. Let $\alpha''$ be any node where $w_h$ acts in this stage. Since $w_h$ cannot obtain a better offer than $\mu_W^{\alpha''}$, we claim that his best response is to reject any firm $f$ such that $\mu_W^{\alpha''}(w_h) P_{w_h}^{\alpha''} f$, and accept the offer of $\mu_W^{\alpha''}(w_h)$. If he rejects $\mu_W^{\alpha''}(w_h)$'s offer, then by the induction hypothesis he will be matched in the following subgames to $\mu_W^{\alpha'''}(w_h)$, where $\alpha'''$ is the node following $\alpha''$. Since in node $\alpha'''$ we have $w_h P_{w_h}^{\alpha'''} \mu_W^{\alpha''}(w_h)$, it follows by Lemma 5 that $\mu_W^{\alpha''}(w_h) P_{w_h} \mu_W^{\alpha'''}(w_h)$. That is, $w_h$ is worse off if he rejects the offer from $\mu_W^{\alpha''}(w_h)$. It remains to show that $w_h$ is better off if he rejects any offer from firms less preferred than $\mu_W^{\alpha''}(w_h)$. To this end, it suffices to show that if he rejects all $f$ such that $\mu_W^{\alpha''}(w_h) P_{w_h}^{\alpha''} f$, then any node $\alpha'''$ that follows this decision is such that $\mu_W^{\alpha''}(w_h) = \mu_W^{\alpha'''}(w_h)$. This is easily done using the Deferred Acceptance algorithm. Observe that the only difference between $P_{w_h}^{\alpha''}$ and $P_{w_h}^{\alpha''}$ is that under $P_{w_h}^{\alpha'''}$ some firms less preferred than $\mu_W^{\alpha''}(w_h)$ are now unacceptable. Hence, running the Deferred Acceptance algorithm with workers proposing under the preference profile $(P_{w_h}^{\alpha''}, P_{-w_h}^{\alpha''})$ or $(P_{w_h}^{\alpha'''}, P_{-w_h}^{\alpha''})$ obviously gives the same output. Hence, $\mu_W^{\alpha''}(w_h) = \mu_W^{\alpha'''}(w_h)$. We can now repeat the same argument with worker $w_{h-1}$, then $w_{h-2}$, ... until worker $w_1$, which gives the desired conclusion. ■

**Proof of Proposition 2** Let $\mu \in S(P)$, and construct a strategy profile $\sigma$ such that all firm $f \in F$ propose to $\mu(f)$ at the first stage and each worker accepts the best offer he has at the end of the first stage (i.e., at all nodes that belong to stage 1.1). Clearly, $\mu[\sigma] = \mu$. If a worker $w$ decides to reject his offer at the first stage, then he will be unmatched at the end of the game or matched to a less preferred firm than $\mu(w)$ (by the stability of $\mu$, such firms are unmatched at this matching), i.e., worse off. If one or several firms do not propose to their partner under $\mu$ during the first stage let $\sigma$ be such that each worker $w$, at each subgame, does reject all offers that are strictly less preferred from $\mu_W(w)$ and accept any offer from $\mu_W(w)$ or firms preferred to $\mu_W(w)$. Clearly, the only way for a firm to be matched to a different worker than her worker-optimal mate is to be matched to a worker less preferred to her, and thus, no worker has an incentive to deviate from this profile. Hence, in the subgames attained if one or several firms did not propose to her partner under $\mu$, the profile constitutes am equilibrium and yields $\mu_W$. If a firm $f$ plays according to $\sigma$, then she is matched to $\mu(\sigma)(f)$, otherwise she is matched to $\mu_W(f)$. Since for all firms $f \in F$ and $\mu \in S(P)$, we have $\mu(f) P_f \mu_W(f)$, playing according to $\sigma$ is a best response for each firm, which gives the desired conclusion. ■

## References


[1] J. Alcalde, *Implementation of stable solutions to marriage problems*, Journal of Economic Theory, 69 (1996), pp. 240–254.

[2] J. Alcalde, D. Pérez-Castrillo, and A. Romero-Medina, *Hiring procedures to implement of stable allocations*, Journal of Economic Theory, 82 (1998), pp. 469–480.

[3] J. Alcalde and A. Romero-Medina, *Simple mechanisms to implement the core of college admission problems*, Games and Economic Behavior, 31 (2000), pp. 294–302.

[4] Y. Blum, A. E. Roth, and U. G. Rothblum, *Vacancy chains and equilibration in senior-level labor markets*, Journal of Economic Theory, 76 (1997), pp. 362–411.

[5] L. E. Dubins and D. A. Freedman, *Machiavelli and the Gale-Shapley algorithm*, The American Mathematical Monthly, 88 (1981), pp. 485–494.

[6] D. Gale and L. Shapley, *College admission and the stability of marriage*, American Mathematical Monthly, 69 (1962), pp. 9–15.

[7] D. E. Knuth, *Mariages stables et leurs relations avec d'autres problèmes combinatoires*, Les Presses de l'Université de Montréal, Montreal, Québec., 1976. Introduction à l'analyse mathématique des algorithmes, Collection de la Chaire Aisenstadt.

[8] D. McVitie and L. B. Wilson, *Stable marriage assignment for unequel sets*, BIT, 10 (1970), pp. 295–309.

[9] B. Peleg, *Implementation of the core of a marriage problem.* Discussion paper ♯132, Hebrew University of Jerusalem, 1997.

[10] A. Roth and M. Sotomayor, *Two-sided Matching: A study in Game-Theoretic Modeling and Analysis*, Econometric Society Monographs, No. 18, Cambridge University Press, Cambridge, England, 1990.

[11] A. E. Roth, *The economics of matching: stability and incentives*, Mathematics of Operations Research, 7 (1982), pp. 617–628.

[12] ——, *Misrepresentation and stability in the marriage problem*, Journal of Economic Theory, 34 (1984), pp. 383–387.

[13] A. E. Roth and U. G. Rothblum, *Truncation strategies in matching markets—in search of advice for participants*, Econometrica, 67 (1999), pp. 21–44.

[14] A. E. Roth and X. Xing, *Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists*, Journal of Political Economy, 105 (1997), pp. 284–329.

[15] S.-C. Suh and Q. Wen, *Stable matchings and subgame perfect equilibrium in an extensive form matching game.* mimeo., September 2002.