Int J Game Theory (2011) 40:1–28
DOI 10.1007/s00182-009-0218-x

Decentralized job matching

Guillaume Haeringer · Myrna Wooders

Accepted: 10 December 2009 / Published online: 13 January 2010

Abstract This paper studies a multi-stage decentralized matching model where firms sequentially propose their (unique) positions to workers. At each stage workers sequentially decide which offer to accept (if any). A firm whose offer has been declined may make an offer to another worker in the next stage. The game stops when all firms either have been matched to a worker or have already made unsuccessful offers to any worker remaining in the market. We show that there is a unique subgame-perfect equilibrium outcome, the worker-optimal matching. Firms in this game have a weakly dominant strategy, which consists of making offers in the same order as given by their preferences. When workers play simultaneously any stable matching can be obtained as an equilibrium outcome, but an unstable matching can obtain in equilibrium.

Keywords Two-sided matching · Job market · Subgame perfect equilibrium · Commitment

JEL Classification C78 · C62 · J41

G. Haeringer (✉)
Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona,
08193 Bellaterra, Spain
e-mail: Guillaume.Haeringer@uab.es

M. Wooders
Department of Economics, Vanderbilt University, 415 Calhoun Hall, Nashville, TN 37235-1819, USA
e-mail: myrna.wooders@vanderbilt.edu
URL: www.myrnawooders.com

1 Introduction

Matching markets concern those environments where players have incentives to form groups. The *College admissions problem,* where students choose which college to attend and colleges choose which students to accept, and the American market for intern physicians are two of the most well known examples.[1] 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 in that they are *decentralized*, with the possibility that players' decisions are *irreversible*. To this end, we propose a model that alternates offering stages, in which firms make offers to workers, and responding stages, in which each worker each decides which, if any, offer to accept. A firm that has had its offer declined may make further offers, and so on. To some extent, the model we propose can be seen as a decentralized version of Gale and Shapley's (1962) deferred acceptance (DA) algorithm or the McVitie and Wilson's (1970) algorithm, depending on whether players make their decisions simultaneously.

By 'decentralized' we mean that each player makes her decisions independently of the decisions made by others. Decentralized markets involve different strategic issues from those of centralized markets. To see this, consider the case of a candidate i, who receives an acceptable offer from a firm f that is ranked low in her preferences.[2] Under the DA algorithm, unless she already has a better offer, i would be assigned to f, at least temporarily. By contrast, in a decentralized market, i could decide to decline f's offer. If she does so, then f may propose to another candidate, say k, who might accept. The candidate k may then decline further offers (or might not receive some offers) that are subsequently made to i. These subsequent offers may rank higher in i's preference ordering, allowing i to be better off than she would have been had she accepted the initial offer. Also, in a centralized context, matching markets are typically modeled as strategic form games (where a player's strategy consists of a preference list), while in our decentralized context the market can be better described as a sequential game.

We assume that both firms and workers behave strategically. This creates difficulties because in this case players' strategies cannot in general be identified with a preference ordering. For instance, a worker w 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 decline both offers otherwise. Similarly, a 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. Addressing these difficulties is one of the main tasks of this paper.

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 offer may "block" the market and leave some workers with less preferred

[1] See Roth and Sotomayor (1990) for a survey up to 1990.

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

outcomes than they otherwise might have obtained. 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.[3] A close look at the hiring procedures may illuminate the impact of this assumption. When a candidate receives an offer, it is usually (first) by email 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 email 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.[4] In this case, the candidate declines all subsequent offers, explaining that she is no longer seeking a job.

Let us motivate the feature that decisions are non-deferrable with an example. Consider a candidate, say Wendy, who receives an offer from University f_1. At the time she receives this offer, it is the best she has, and therefore she accepts the offer. Suppose now that a couple of weeks later she receives a preferred offer from University f_2. If her contract with f_1 is not yet signed, that is, if her notification of acceptance has only been informal, she could accept the offer from f_2 and send a message to f_1 saying that she has now decided to decline f_1's offer. All this would be perfectly legal. Reversing her acceptance of f_1 may jeopardize her reputation, however. Moreover, if f_2 knows that Wendy has accepted f_1, f_2 is likely to refrain from making her an offer. Thus, although Wendy could legally decline, *ex-post*, the position offered by f_1, in practice she might not be able to do this costlessly. Thus, in the academic job market it may be appropriate to assume that candidates' decisions are *irreversible* or, in other words, workers and firms must *commit*. It is important to notice that this assumption is different from the fact that candidates may be given several days to give their answer.[5]

Our decentralized job matching game is built upon a one-to-one matching model. In this model each firm and each worker has a strict preference relation over the players from the other side of the market (including oneself, to allow for unacceptable partners). A *matching* is simply an assignment of workers to firms such that no two workers are assigned to the same firm and no two firms are assigned to the same worker. Possibly, some workers and firms can be matched to themselves. The central concept in this model is *stability*. A matching is stable if it is individually rational and there are no blocking pairs. A matching is individually rational if no individual would prefer to be matched to himself than being matched to his current partner. A pair worker-firm is a blocking pair for a matching if they are not matched together and both

[3] We have heard that exceptions, although infrequent, do occur.

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

[5] In the academic job market, a candidate is given some time to make up his 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). Therefore, allowing for a delay in responding to an offer would have no effect on the outcome.

prefer to be matched together than being matched to their current partner. Gale and Shapley (1962) proved that the set of stable matching is never empty. Among all stable matchings, one is the most preferred by all workers, the worker-optimal matching, and one is the most preferred by all firms, the firm-optimal matching. The worker-optimal (resp. firm-optimal) matching is obtained when running the DA algorithm with workers (resp. firms) proposing.

In the most delicate case, a benchmark upon which we build most of our results, players' decisions are irreversible and *not* simultaneous. In this context, we first consider the game in which firms are not strategic and make offers to workers in the order of their preferences, independently of the history of the game.[6] In this case we show that there is a unique subgame perfect equilibrium, which yields the worker-optimal matching (Theorem 2). That is, our decentralized market with firms proposing yields the same outcome as the centralized matching market that uses the DA algorithm with workers proposing. We use this first result to consider the case when both firms and workers are strategic and prove our main result: there is a unique subgame perfect equilibrium outcome, the worker-optimal matching (Theorem 3). While the outcome is unique, it may be supported by a large number of different strategy profiles. However, our proof shows that for each firm, it is a weakly dominant strategy to make offers in the same order as its preferences. The reason is quite simple. When firms are strategic, although the sequence of offers may depend on the history, along the equilibrium path workers face only one sequence. For the workers, these sequences of offers are perceived as if they were representing firms' preferences. Hence, the game for firms (when they are strategic) is to some extent equivalent to a strategic form game in which firms have to choose and submit a preference ordering over the set of workers and the outcome is obtained by applying the DA algorithm with workers proposing using the preferences submitted by the firms. It is well known that in this game firms may have an incentive to misrepresent their preferences. However, it is a weakly dominant strategy for firms to submit a preference ordering such that the most preferred worker according to this ordering is indeed truly most preferred. Transposed to our context of a sequential game this property implies that for each firm it is a dominant strategy to make its first offer to its most preferred worker. Repeating this observation for the subsequent stages, we then easily obtain that for each firm it is a weakly dominant strategy to make an offer, at any stage of the game, to its current most preferred worker. Using our characterization result when firms are not strategic, we conclude that any subgame perfect equilibrium outcome yields the worker-optimal matching.

The fact that all subgame perfect equilibria give the worker-optimal matching crucially depends on the assumption that workers' decisions are *not* simultaneous. When they are, the set of subgame perfect equilibria outcomes is enlarged and includes (but may not coincide with) the set of stable matchings. Yet, whether or not firms' decisions are simultaneous does not change our findings.

[6] If a firm wants to make an offer to some worker at some stage this worker needs to be unmatched. So offers depend to some extent on the history of the game.

The irreversibility assumption turns out to have fewer effects than assuming that the market is decentralized. If we allow workers not to commit, then we again obtain a unique subgame perfect equilibrium outcome, the worker-optimal matching, irrespective of the timing of workers' decisions (i.e., whether simultaneous or not). The reason is quite simple. Under perfect information each participant knows the complete sequence of actions 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.

The paper is organized as follows. In Sect. 2 we define the main concepts and notation and the job matching game is presented in Sect. 3. In Sect. 4 we characterize the equilibrium outcomes in decentralized job matching games when firms are not strategic. The equilibrium when firms are strategic is analyzed in Sect. 5. We discuss in Sect. 6 the main hypothesis we make in our model, that matches are irreversible in the course of the game and show how the equilibrium outcomes may vary when firms and/or workers play simultaneously. We then discuss in Sect. 7 some papers in the matching literature dealing with decentralized markets. A conclusion is given in Sect. 8.

2 Preliminaries

2.1 Notations and definitions

We consider a (finite) set of workers and a (finite) set of firms, $W = \{w_1, \ldots, w_m\}$ and $F = \{f_1, \ldots, f_n\}$ respectively. Throughout the following, 'it' refers to a firm and 'he' refers to a worker. We use the symbols $\subset$ and $\subseteq$ to denote the strict and weak inclusions, respectively. The cardinality of a (finite) set Z is denoted $|Z|$.

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 $W \cup \{f\}$. For each player $v \in F \cup W$, let R_v denote the weak preference relation associated to P_v. For a worker $w \in W$ and any subset $S \subset F$ let $\max_{P_w} S$ be w's most preferred player in $S \cup \{w\}$ according to the preference relation P_w,

$$\max_{P_w} S \overset{\text{def}}{=} \{v \in S \cup \{w\} \mid v \, R_w \, v', \ \forall \, v' \in S \cup \{w\}\}.$$

Note that, since the maximum is taken over a nonempty finite set, $\max_{P_w} S \neq \emptyset$. Now, redefining S as a subset of W, we similarly define $\max_{P_f} S$ as f's most preferred player in $S \cup \{f\}$. The set of all possible preferences for a player $v \in W \cup F$ is denoted by $\mathcal{P}_v$, and an element P of

$$\mathcal{P} \overset{\text{def}}{=} \Pi_{w \in W} \mathcal{P}_w \times \Pi_{f \in F} \mathcal{P}_f$$

is called a *preference profile*. A *job market* is described by the triple

$$(W, F, P).$$

A *matching* μ is a function from $W \cup F$ into 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 $v \in W \cup F$, $\mu(\mu(v)) = v$.

Given a matching μ, we will refer to $\mu(v)$ as the *mate* of v under μ. 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 over the set of matchings $\mathcal{M}$ in the following way (abusing notation); $\mu P_w \mu'$ if and only if $\mu(w) P_w \mu'(w)$ and $\mu R_w \mu'$ if and only if $\mu(w) R_w \mu'(w)$.[7] Similarly, we can also extend firms' preferences to preferences over $\mathcal{M}$.

2.2 Stability and deferred acceptance algorithm

For a firm f (resp., worker w), a worker w (resp., firm f) is *acceptable* if f (resp. w) prefers to be matched to w (resp. f) rather than remaining unmatched, i.e., $w P_f f$ (resp. $f P_w w$). For a player $v \in F \cup W$, $A(P_v)$ denotes the set of acceptable players with respect to the preference relation P_v.[8]

A matching μ is *individually rational* at a profile P if for each player $v \in F \cup W$ it holds that $\mu(v) R_v v$, i.e., $\mu(v) \in A(P_v) \cup \{v\}$. The set of individually rational matchings at P is denoted $IR(P)$. A matching μ is *blocked by a pair at* P (w, f) if $f P_w \mu(w)$ and $w P_f \mu(f)$. A matching μ that is individually rational and is not blocked by any pair is called *stable*. We denote by $S(P)$ the set of stable matchings at P.

A stable matching can be obtained using Gale and Shapley's (1962) *Deferred Acceptance* (DA) algorithm. Their algorithm, with workers making proposals to firms, works as follows:

Step 1: Each worker w proposes to his most preferred firm among the ones that are acceptable for him. If there is no such firm then the worker is matched to himself.
Each firm declines all but its most preferred worker among the workers who proposed to this firm and are acceptable to it (if any).

Step k, $k \geq 2$: Each worker who has been declined in the previous step proposes to his most preferred firm among the firms that have not yet declined him and are acceptable for him. If there is no such firm then the worker is matched to himself.
Each firm declines all but its most preferred worker among the workers who proposed to this firm and are acceptable to it and the worker it did not decline in the previous step (if any).

[7] Preferences over mates are assumed to be strict. However, since at two different matchings an agent can be matched to the same individual, and thus be indifferent between these two matchings (i.e., there are no externalities), strict preferences over mates naturally extend to weak preferences over matchings.

[8] Note that by convention we set $v \notin A(P_v)$, for all $v \in W \cup F$.

The algorithm stops when every worker is either matched to a firm or to himself. Given a preference profile P, we denote by $\mu_W(P)$ the matching obtained by the DA algorithm we just described and by $\mu_F(P)$ the matching that is obtained when we invert the roles of workers and firms in the algorithm, i.e., firms propose to workers and workers reject or accept workers' proposals.

Theorem 1 (Knuth 1976) *The matching $\mu_F(P)$ (resp. $\mu_W(P)$) is the workers' (firms') least preferred matching among all stable matchings and $\mu_W(P)$ (resp. $\mu_F(P)$) is the workers' (firms') most preferred matching among all stable matchings.*

Henceforth, we shall refer to the matching $\mu_W(P)$ as the *worker-optimal matching* (W-optimal matching), and $\mu_F(P)$ as the *firm-optimal matching* (F-optimal matching).

3 Sequential job matching

Throughout the paper we consider situations where firms make offers to workers and each worker who receives an offer from a firm has to make an irreversible decision, to either *decline* or *accept* the offer. The situation we consider is that of a multi-stage game. Each stage is divided into two sub-stages: first, an offering stage during which firms act and, second, a responding stage during which workers act. We assume that a firm can make at most one offer to each worker, that is, if at some stage a worker w declined an offer from a firm f, then firm f cannot make an offer to w at a later stage. Once a worker, say w, has accepted an offer from a firm, say f, both w and f exit the market: worker w cannot receive further offers *and* firm f cannot make further offers to other workers.[9] A firm f whose offer has been declined by a worker can make an offer to any worker who has not already received and declined an offer from it. Unless otherwise specified, we assume that the game is with perfect information. The game ends when all firms have exited (whether matched or not). If there remain any worker in the market when all firms have exited, then these workers are matched to themselves.

Offering stage: According to the ordering given by their index numbers, each remaining firm $f_k, k = 1, \ldots, n$ offers its position to an acceptable worker among the remaining workers who have not previously declined that firm, or exits the market if none of these workers are acceptable.

Acceptance stage: According to the ordering given by their index numbers, each worker $w_k, k = 1, \ldots, m$ either accepts one of his offers (if he has any) or declines all his offers. A worker cannot "hold" an offer and accept or decline it at a later stage.[10] A worker who has not received any offer waits for the next stage. If a worker, say w_i, accepts an offer from a firm, say f_j, then w_i and f_j are matched and exit.

9 In a recent paper Diamantoudi et al. (2006) study the role of commitment in a repeated matching game and show how the outcome may change depending on whether firms can dismiss employees.

10 That is, firms make "exploding" offers. The other possibility would be that firms make "open" offers, i.e., offers that workers can hold on to in order to have the offer still available at a later stage. See Sect. 6.2 for a discussion.

Remark 1 We fixed the order of firms' play in the offering stage as well as that of workers' play in the responding stage. This is for pure convenience only. As we shall see, the equilibria do not depend on a specific order of play within each sub-stage.

Remark 2 When firms and workers act simultaneously during the offering and acceptance stages respectively, the sequential job market game can be seen as a decentralized version of the deferred acceptance algorithm. When the sequential job matching game is with perfect information then it can be seen as a decentralized version of the McVitie–Wilson algorithm (McVitie and Wilson 1970), with the exception that we assume matched workers exit.[11] While the DA and the McVitie–Wilson algorithms are equivalent in a strategic form game we shall see that their equilibrium outcomes may differ when considering their decentralized versions.

3.1 Histories

Formally, a decentralized job matching game consists of a finite collection of finite histories (see Osborne and Rubinstein 1994), with generic element denoted by h, and a mapping $p : \mathcal{H} \to W \cup F$, which associates to each non-terminal history a player $v \in W \cup F$ whose turn it is so act, where $\mathcal{H}$ is the set of all non-terminal histories.

A history is an ordered collection of actions taken by the players in $W \cup F$. For a firm, an action consists of either making an offer to some worker or exiting. For a worker, an action consists of accepting an offer and declining all other offers (if any), or declining all offers. So, for instance, at the beginning of the play, i.e., at the empty history, we have $p(\emptyset) = f_1$, and $p(h) = f_2$ for any action h that firm f_1 took at the empty history. Similarly, if h denotes the history "firm f_1 made an offer to w in the first stage and firm f_2 made an offer to w' in the first stage," then $p(h) = f_3$.

Given a non-terminal history h, an *immediate history after* h is a history h' such that $h \subset h'$ and there is no history h'' such that $h \subset h'' \subset h'$. Also, given a non-empty history h, a history h' is a sub-history of h if $h' \subseteq h$.

We use the convention that if there is a history h such that $p(h) \in W$, say, $p(h) = w$, then worker w received in the corresponding stage at least one offer. In other words, if at some stage a worker did not receive any offer, then in *this* stage this worker does not have to play.

For each terminal history there is a corresponding matching $\mu \in \mathcal{M}$. The *length* of an history h is the number of distinct histories h' such that $h' \subseteq h$. For convenience, we set the length of the empty history, h_0, to 1.

A strategy for a player $v \in W \cup F$ is a mapping σ_v that associates to each non-terminal history h such that $p(h) = v$ an action $\sigma_v(h)$. We use the usual game theoretic notation to write the profile of a particular set of players, i.e., given a strategy profile σ and a set of players in $S \subseteq W \cup F$, let $\sigma_S = (\sigma_v)_{v \in S}$ and $\sigma_{-S} = (\sigma_v)_{v \notin S}$. Thus, for each firm $f \in F$ and each history h satisfying $p(h) = f$, an action at h consists

[11] The McVitie–Wilson algorithm differs from the DA algorithm in that at each step there is only one firm (randomly chosen) making an offer. It stops when each firm is either matched to a worker or has proposed to all its acceptable workers. Both the DA and the McVitie–Wilson algorithms yield the same matching, the F-optimal matching.

of making an offer to some worker $v \in (W \cup \{f\}) \backslash \{Z_f^h\}$, where making an offer to itself is interpreted as exiting and Z_f^h is the set of workers to whom firm f has already made an offer at some history $h' \subset h$. For each worker $w \in W$ and each history h satisfying $p(h) = w$, an action consists of accepting at most one offer that w holds at history h or declining all such offers.

Given a strategy profile σ, we denote by $\mu[\sigma]$ the outcome (the matching) that obtains when players play according to σ. Given a history h, we denote by $\mu[\sigma|h]$ the outcome that is obtained if at history h players use the strategy profile σ. Note that it may be that $\mu[\sigma] \equiv \mu[\sigma|h_0] \neq \mu[\sigma|h]$, i.e., playing according to σ from the empty history h_0 may not lead to history h.

Let $\mathcal{G}(P)$ denote the sequential game we just described, in which workers and firms evaluate the outcomes according to their preferences P_W and P_F, respectively. Note that in $\mathcal{G}(P)$ firms are strategic.

Let $S \subset W \cup F$ be a subset of players, and let σ_S be a strategy profile for players in S. We denote by $\mathcal{G}(P|\sigma_S)$ the decentralized matching game in which the strategy of each player $v \in S$ is fixed to σ_v.

3.2 Equilibrium and sub-markets

The equilibrium concept we consider in this paper is the subgame perfect equilibrium (SPE).[12] Formally, a strategy profile σ is a SPE if for each player $v \in F \cup W$, and for each history h satisfying $p(h) = v$, there is no strategy σ'_v such that $\mu[\sigma_{-v}, \sigma'_v|h] P_v \mu[\sigma_{-v}, \sigma_v|h]$.

Let $SPE(\mathcal{G}(P))$ be the set of subgame perfect equilibria when firms and workers evaluate outcomes according to the preference profile P. Correspondingly, let $SPE(\mathcal{G}(P|\sigma_F))$ be the set of subgame perfect equilibria when firms play according to the profile σ_F and workers evaluate outcomes according to the preference profile P_W. Note that an equilibrium of $\mathcal{G}(P|\sigma_F)$ is a strategy profile for workers (i.e., not including firms' strategies).

Notice that for a firm, for each history at which that firm is not matched yet, there are two sets of workers (among the workers who are not matched yet) to whom the firm will never make an offer in a SPE of $SPE(\mathcal{G}(P))$. The first set is made of workers that are unacceptable to it and the second set contains all the workers whom it made an offer to at an earlier history but was rejected by. It follows that, in a SPE, for a firm not matched yet at some history, we can consider the set of workers who have rejected its offer at a previous history as unacceptable to it. We capture this property by defining, for each non-terminal history h, an associated *sub-market* (W^h, F^h, P^h). The central idea in the definition of a sub-market at a history h is that, in a subgame perfect

[12] We shall not consider Nash equilibrium. It is not difficult to see that the set of Nash equilibrium outcomes coincides with the set of individually rational matchings. To this end, consider any individually rational matching μ. For each firm $f \in F$, let σ_f be such that firm f only proposes to $\mu(f)$ and, if declined, exits the market. As for workers, each worker $w \in W$ only accepts the offer of $\mu(w)$. This strategy profile is a Nash equilibrium (with dominated strategies, though). (See Alcalde (1996) for a similar argument in a strategic form game.)

equilibrium, matchings that are not individually rational matchings of (W^h, F^h, P^h) cannot be equilibrium outcomes.

Let h be a non-terminal history. The sub-market (W^h, F^h, P^h) is such that:

- W^h is the set of workers that have not exited (whether matched to a firm or to themselves) at history h;[13]
- F^h is the set of firms that are not matched (to a worker or to itself) at history h;
- P^h is a preference profile of players in $W^h \cup F^h$ that satisfies:[14]
  - For each worker $w \in W^h$, P^h_w is a preference relation over $F^h \cup \{w\}$. Similarly, for each firm $f \in F^h$, P^h_f is a preference relation over $W^h \cup \{f\}$;
  - For each worker $w \in W^h$, and for each $f \in F^h$, $f P^h_w w$ if, and only if, $f P_w w$. For each pair $f, f' \in F^h$, $f P^h_w f'$ if, and only if, $f P_w f'$;
  - For each firm $f \in F^h$ and $w \in W^h$, if f has not made any offer to w at any history $h' \subset h$, then $f P^h_f w$ if, and only if, $f P_f w$. If f has made an offer to w at some history $h'' \subset h$ but w declined that offer at some history h' such that $h'' \subset h' \subset h$, then $f P^h_f w$. If f has made an offer to w at some history $h'' \subset h$ but there is no history h' such that $h'' \subset h' \subset h$ and w declined that offer at history h', then $f P^h_f w$ if, and only if, $f P_f w$;
  - For each firm $f \in F^h$ and for any w, w' such that f has not made any offer to w and w' at any sub-history h' of h, it holds that $w P^h_f w'$ if, and only if, $w P_f w'$. If f has not made an offer to w' at any history $h' \subset h$ and if f has made an offer to w at some history $h'' \subset h$ but there is no history h', $h'' \subset h' \subset h$ at which w declined that offer then $w P^h_f w'$ if, and only if, $w P_f w'$.

3.3 Firms' strategies

We analyze two distinct games of decentralized matching markets. In the first game, firms are not strategic; they make offers to workers in the order of their preferences over workers. In the second game we allow firms to strategically choose the worker to whom they make an offer.

When firms are not strategic, at any history of the game at which a firm f has to act it makes an offer to its most preferred worker among those workers who are not matched yet and who have not received an offer from firm f at an earlier stage (and declined that offer). Strategies that involve making offers in a specific order (i.e., the sequence of offers does not depend on the history of the play) are known as "preference

[13] Since workers do not make offers a worker cannot be matched to himself during the game. That is, a worker that finds all remaining firms in the market unacceptable must wait the end of the game to be assigned to himself. This is not the case for a firm since it can opt to make an offer to itself and exit the market.

[14] Notice that we do not specify the relative ranking of workers that are unacceptable. Thus in general there are many preference profiles that satisfy our below requirements. Since, in equilibrium, at any history firms never make an offer to an unacceptable worker this loose description of the profile P^h is without loss of generality. Notice also that it is sufficient to "update" firms' preferences. That is, if a firm considers a worker as unacceptable whether that worker considers the firm acceptable has no impact.

strategies" (see Blum et al. 1997; Pais 2008). We assume that the ordering of offers of a firm is identical to the ordering given by the firm's preference.

Definition 1 For each firm $f \in F$, the strategy σ_f^* consists of making offers to workers in the order given by f's preferences. Formally, for each history h such that $p(h) = f, \sigma_f^*(h)$ is an offer to $\max_{P_f^h}(W^h)$). If $\max_{P_f^h}(W) = \{f\}$ then f exits the market.

When firms are not strategic the decentralized job matching game is denoted by $\mathcal{G}\left(P|\sigma_F^*\right)$.

4 Equilibria when firms are not strategic

We start by characterizing the equilibria when firms are not strategic. The main result of this section is the following.

Theorem 2 *Let σ_W be a subgame perfect equilibrium of $\mathcal{G}(P|\sigma_F^*)$. Then $\mu[\sigma_W, \sigma_F^*] = \mu_W$. Furthermore, the equilibrium is unique.*

The proof of Theorem 2 is in Appendix A. A brief explanation why any SPE yields the W-optimal matching is the following. In a SPE it is always profitable for a worker to decline those offers that are less preferred than his W-optimal mate. Anticipating the behavior of the other workers, a worker w can safely decline an offer from a firm less preferred than his W-optimal mate: This rejection will trigger a chain of rejections by the other workers with the consequence that worker w will eventually receive an offer from his W-optimal mate. As long as some workers receive offers from firms less preferred than their W-optimal mates and reject these offers, they help each other. However, when workers receive offers from their W-optimal mates, this strategic complementarity aspect disappears, and workers vie with each other for better outcomes. In this case, if the strategies of the other workers allow a worker to obtain an offer preferred to his W-optimal mate, it must be the case that there is another worker whose strategy dictates accepting a firm less preferred than his W-optimal mate.

Notice that for each worker $w \in W$, at each history h where w has to act, and for any strategy profile σ, worker w is not indifferent between the various actions he may take. If a worker has an offer from a firm f and declines it, he cannot obtain that offer again at a later stage. That is, by declining the offer it is impossible to obtain the same payoff he would receive had he accepted f's offer. So if each worker is matched to a firm at the W-optimal matching, then the equilibrium in the game $\mathcal{G}(P|\sigma_F^*)$ is unique.

Remark 3 If we allow workers to exit, then there is still a unique subgame perfect equilibrium outcome, namely the W-optimal matching, but the equilibrium may not be unique. To see this, consider the case of 3 workers, w_1, w_2 and w_3, and two firms f_1 and f_2, and the preferences (over acceptable mates) are $P_f = w_1, w_2, w_3$, for $f = f_1, f_2$, and $P_{w_1} = f_1, f_2$ and $P_{w_2} = P_{w_3} = f_2, f_1$. It is easy to check that there are two equilibria, one in which worker w_3 exits in the first stage and one in which he exits in the second stage.

5 Equilibria when firms are strategic

We now consider the decentralized job matching game when firms are strategic. The principal difficulty in this case is that now firms' sequence of offers may depend on the history of play. In spite of this difficulty we can show that the set of equilibrium outcomes does not change.

Theorem 3 *Let* $\sigma \in SPE(\mathcal{G}(P))$. *Then* $\mu[\sigma] = \mu_W$.

The proof of Theorem 3 relies on Theorem 2. The intuition is the following. When facing the sequence of offers made by firms, workers roughly behave as if firms have preferences over workers that agree with the sequence of offers. That is, if a firm f makes first an offer to a worker, say, w_1 and then to another worker, say, w_2, from the viewpoint of a worker it is the same as if f actually prefers worker w_1 to worker w_2. Hence, we can always construct a preference profile $\widetilde{P}_F$ for the firms such that for each firm f the order of offers it made is the same as that dictated by $\widetilde{P}_f$. Theorem 2 then implies that when workers play optimally the outcome will be $\mu_W(P_W, \widetilde{P}_F)$. However, since firms' offers can depend on the history of play we may well have different such alternative preference profiles $\widetilde{P}_F$, one for each terminal history. This problem is partly circumvented using an induction argument on the length of the game, which will reduce the problem of finding the "correct" alternative profile $\widetilde{P}_F$ to finding the most preferred worker according to $\widetilde{P}_f$, for each firm $f \in F$. For each firm $f \in F$, for the workers that are less preferred than the most preferred worker in $\widetilde{P}_f$, their relative ranking in $\widetilde{P}_f$ and in P_f are the same.

The proof of Theorem 3 illuminates one aspect of the strategic issue faced by firms. The main part of the proof consists of showing that for each firm making its offers in the same order as in its preferences is a weakly dominant strategy. However, although the equilibrium outcome is unique, an equilibrium may not be unique. A simple example with two firms and two workers illustrates this point.

Example 1 Let $W = \{w_1, w_2\}$ and $F = \{f_1, f_2\}$ and consider the preference profile P given below.

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

which is to be read as "firm f_1's first choice is w_1 and his second choice is w_2 (both are acceptable)," and similarly for the other players.

There are two stable matchings, μ_F and μ_W where $\mu_F(w_i) = f_i$ and $\mu_W(w_i) = f_{3-i}$ for all $i = 1, 2$. One can easily show that it suffices that at least one worker declines the offer from his F-optimal mate to yield as a SPE the W-optimal matching. This game has several equilibria. In one of these each firm makes an offer in the first stage to its F-optimal mate, and each worker declines the offer he receives in the first stage. In another equilibrium each firm makes an offer to its W-optimal mate in the first stage and each worker accepts the offer he received in the first stage. □

6 Discussion

6.1 Simultaneous play

So far we have assumed that firms and workers' moves are not simultaneous. If firms' moves are simultaneous but workers' moves are not, we are in the same situation as in our benchmark case with perfect information (the proofs remain unchanged or become simpler).

Things differ, however, when we consider situations where workers' moves are simultaneous. Let $\mathcal{G}^{sim}(P)$ be a game with these features. For simplicity, assume that firms' moves are not simultaneous. 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. Thus, each firm can observe the offers made by the firms who act before it, and all workers observe the offers made by all firms. However, although they do observe the workers' decisions taken at previous stages, at each stage active workers do not observe the action taken by the other workers *during* the stage. In this case, we find that any stable matching can be supported by a subgame perfect equilibrium. Note that as for the game $\mathcal{G}(P)$ we can build many different games satisfying the conditions of $\mathcal{G}^{sim}$ for the same set of workers, firms and preference profile, each game depending on the order in which firms act during the offering stages.

Proposition 1 *Let $\mu \in S(P)$. Then there exists a subgame perfect equilibrium strategy profile σ for the game $\mathcal{G}^{sim}(P)$ satisfying $\mu[\sigma] = \mu$.*

Proof Let $\mu \in S(P)$ and let σ be a strategy profile that satisfies the following:

- (i) In the first stage each firm f makes an offer to $\mu(f)$ if $\mu(f) \in W$, and exits the market otherwise.
- (ii) If all firms played according to *(i)* then each worker who receives an offer in the first stage accepts it. If one or several firms do not play according to *(i)* then each worker $w \in W$, at each subgame, declines all offers that are strictly less preferred to $\mu_W(w)$ and accepts any offer from $\mu_W(w)$ or from firms preferred to $\mu_W(w)$.[15]

By construction, $\mu[\sigma] = \mu$. Any worker w who declines his first stage offer will be unmatched at the end of the game or matched to a less preferred firm than $\mu(w)$ (by the stability of μ, such firms are unmatched at this matching). Hence, no worker declines his first stage offer. Also, it is not difficult to see that, for any subgame that may be attained if a firm deviates in the first stage, σ restricted to that subgame is a SPE for that subgame and yields the outcome μ_W.

Hence, to check that $\sigma \in SPE(\mathcal{G}^{sim}(P))$ it suffices to check that no firm has a profitable deviation in the first stage. Since $\mu \in S(P)$, Theorem 1 implies that for each firm $f \in F$, $\mu R_f \mu_W$. So no firm can be better off deviating in the first stage and thus $\sigma \in \mathcal{G}^{sim}(P)$. □

[15] Notice that it may be that a worker w received an offer in the first stage from $\mu_W(f)$. In this case he accepts this offer.

The following example demonstrates that when workers' moves are simultaneous, there are subgame perfect equilibria with unstable outcomes.

Example 2 Consider a job market where $F = \{f_1, f_2, f_3\}$, $W = \{w_1, w_2, w_3\}$ and preferences are given by:

$$P(w_1) = f_3, f_2, f_1, w_1, \qquad P(f_1) = w_1, w_3, w_2, f_1,$$
$$P(w_2) = f_2, f_1, f_3, w_2, \qquad P(f_2) = w_3, w_2, w_1, f_2,$$
$$P(w_3) = f_1, f_3, f_2, w_3, \qquad P(f_3) = w_3, w_2, w_1, f_3,$$

The W-optimal and F-optimal matchings for this market are

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

Suppose the game is such that f_1 decides first, then f_2 and finally f_3, and consider a strategy profile, say, σ, that satisfies the following. 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's strategy consists of accepting the offer he received at this stage of the game. In this case, the game stops and we obtain the following matching,

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

Clearly, $\hat{\mu} \notin S(P)$ as it is blocked by the pair (f_3, w_3).

We first need to show that for each worker, after the above chain of proposals by the firms, his strategy is subgame perfect optimal. Note that the game continues to a second stage only if one or more worker declines his offer. For all the subgames reached if one or more worker declined his offer, assume that firms and workers' strategies constitute an SPE. It follows that we only need to check that each worker's action is a best response against the other players' actions. If a worker w_i accepts his offer, then he is matched to $\hat{\mu}(w_i)$, who is acceptable. Otherwise, the game enters the subgame where w_i and $\hat{\mu}(w_i)$ are the only players that are not matched. Since $\hat{\mu}(w_i)$ has already made an offer to w_i and there is no other worker available in the market the game stops and w_i remains un-matched (and $\hat{\mu}(w_i)$ as well). That is, w_i is better off accepting the offer from $\hat{\mu}(w_i)$, given that the other workers have accepted their offers. To sum up, *after* the above mentioned offers by the firms, there is a subgame perfect equilibrium with the property that each worker accepts his offer.

If one or more firms does not follow the sequence of offers described above, assume that σ also satisfies the following requirement: each worker $w_i, i = 1, 2, 3$ only accepts an offer from its most preferred firm (i.e., his worker-optimal mate in this example), and declines all other offers. For the subgames that are reached after the deviation of one or more firms, we have a SPE. For any order of proposals we attain μ_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 each matched to their best mate, so any deviation is strategically inconceivable. 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, it is matched to w_2. If it deviates, it 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 a SPE. □

Obtaining an unstable matching as an outcome of a SPE is in clear contrast with the results obtained by Peleg (1997), Alcalde et al. (1998), and Alcalde and Romero-Medina (2000).[16,17] 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 a SPE. However, they also show that only stable matchings can arise in equilibrium.

Niederle and Yariv (2009, Example 2) also find that an unstable matching can occur in equilibrium. Like us, they do need that workers' moves are simultaneous to sustain an unstable matching. However, they also need that firms can decide not to make an offer at some stage and resume offering their positions at a later stage.[18]

6.2 On the commitment assumption

When workers play sequentially, if we modify the model to allow workers to decline an offer they previously accepted, it is obvious that the equilibrium outcome remains unchanged. Indeed, $\mathcal{G}(P)$ is a game of perfect information and so workers do not have any incentive to hold an offer when they know that, if they decline it, they will eventually receive a better offer.[19]

The change in equilibrium outcomes when workers play simultaneously is more substantial if we drop the commitment assumption. To glean some intuition about the impact of the commitment assumption consider the proof of Theorem 2. When workers play sequentially, we saw that it was a dominant strategy for a worker to decline any offer less preferred than his worker-optimal mate. Indeed, declining such "bad" offers trigger declination chains which eventually lead workers to receive more preferred offers. Such declination chains may not be possible when workers play simultaneously since some workers' strategies may consist of accepting their current offer—see Example 2. However, if workers can resign from previous matches then

16 Peleg (1997) for the marriage model, and Alcalde et al. (1998), and Alcalde and Romero-Medina (2000) for the College Admission problem.

17 Suh and Wen (2008) have an example of a sequential marriage market with perfect information where an unstable matching can be the outcome of a subgame perfect equilibrium. They assume, however, 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 a 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 proposals (if any) or reject all of them.

18 In our context not making an offer is equivalent to exiting the market.

19 In a SPE, along the equilibrium path a worker will only accept the best offer he will receive and reject all other offers. The fact that decisions are irreversible may slow down the market, as will become clear, but otherwise is inconsequential.

these declination chains will occur (perhaps with some delay), thus leading to the same equilibrium outcome as in the game in which workers' move are sequential. The next Proposition simply summarizes these observations.

Proposition 2 *Suppose that workers acceptance decisions are not irreversible (but not necessarily simultaneous) and let* σ *be a subgame perfect equilibrium of* $\mathcal{G}(P)$*. Then* $\mu[\sigma] = \mu_W$.

6.3 Many-to-one matching

In this paper we have analyzed the case of a one-to-one matching model. A natural extension would be to consider the many-to-one matching model, i.e., when firms have more than one position to fill. In this case, firms would have preferences overs sets of workers instead of just preferences over workers. It is well known that in this case many strategic results that hold for the one-to-one case do not carry over (Roth 1985). In our case this is problematic because some results from the one-to-one matching model that are necessary in our proofs (e.g., the Blocking Lemma) may not be valid in many-to-one matching model.

One way out is to consider the *College Admissions* matching model. In this model, firms' preferences are *responsive*. For each firm $f \in F$, let q_f denote its *quota*, i.e., the maximum number of positions it has to fill. We say that a firm's preferences over sets of workers is responsive to its preferences over workers if for any set of workers S, $|S| \leq q_f$ and any two workers w, w' such that $w \in S$, $w' \notin S$, and $w' P_f w$ then $((S \backslash \{w\})) \cup \{w'\} P_f S$. When firms' preferences are responsive all the results obtained for the one-to-one matching model that we need are still valid.[20] So our results easily extend to the many-to-one model when firms have responsive preferences.[21]

7 The literature

Several authors have considered decentralized matching markets. The main papers are Roth and Xing (1997), Blum et al. (1997), Alcalde et al. (1998) and more recently Pais (2008), Diamantoudi et al. (2006) and Niederle and Yariv (2009).[22]

---

[20] The three results we need are the following. First, we need that the F-optimal matching is the workers' least preferred matching among the stable matchings (and firms' most preferred), and conversely that W-optimal matching is the firms' least preferred matching among the stable matchings (and workers' most preferred). Second, we need the Blocking Lemma. Third, we need the feature that for each firm it is a (weakly) dominant strategy to state a preference ordering over sets of workers such that the most preferred worker belongs to the declared most preferred set of workers. All these results are valid when firms' preferences are responsive.

[21] Another extension of firms' preferences over workers to preferences over sets of workers is the class of *substitutable preferences* (Kelso and Crawford 1982). Many properties of the set of many-to-one stable matchings that are valid when preferences are responsive are no longer valid when firms' preferences are substitutable. Martínez et al. (2000) propose a condition, *q-separability* that together with substitutability allows recovery of the usual results that hold for the case when firms' preferences are responsive.

[22] Roth and Xing (1997) study the market for clinical psychologist which is, like the market studied in this paper, decentralized. Roth and Xing focus their attention to the timing aspect of the market, an aspect not captured by our game.

The closest model to ours is that of Alcalde et al. (1998) who consider a game in which there is only one stage.[23] Another difference between their model and ours is that they consider a matching model with money. That is, an offer from a firm includes a wage proposal for the worker. There are obviously several clear contrasts between Alcalde et al.'s model and ours. Their results also differ substantially from us. They show that for a wide class of preferences, the set of subgame perfect equilibrium outcomes coincides with the set of stable matchings. For the particular case of additive preferences, they show that any undominated subgame perfect equilibrium yields the F-optimal matching, the opposite result to the main result of our paper. Alcalde's et al. crucial assumption that enables them to prove this result is that there is only one stage. In their setting, a firm whose offer has been rejected cannot make other offers. It follows that workers cannot "coordinate" by declining all offers less preferred than their W-optimal mate. Example 1 illustrates this point. Consider the first worker to respond in the first stage, say w_1, and consider the case when each worker received an offer from the other worker's W-optimal mate, i.e., w_1 and w_2 received an offer from f_1 and f_2 respectively. If w_1 declines f_1's offer then w_2 is strictly better off declining f_2's offer, for he will obtain an offer from f_1 in the second stage. Hence, if w_1 declines f_1's offer he obtains an offer from f_2 in the second stage.

Another paper considering decentralized matching markets is Diamantoudi et al. (2006). They consider a multi-stage game in which at each stage firms make offers to workers who then decide which offer to accept (if any). Their main concern is whether allowing firms to dismiss their employees can affect the equilibrium outcomes in the repeated matching game. In particular, they show that when firms cannot fire their employees, unstable matchings can be obtained as equilibrium outcomes. Pais (2008) considers a decentralized matching market when firms are randomly given the opportunity to make offers. She also considers the case when at the beginning of the game some workers and firms are already matched. She also assumes that the game is with imperfect information, i.e., workers only know the offers they receive (and do not know the offers other workers received) and firms only know which workers declined its offer. She shows that the ordinal subgame perfect equilibrium outcomes are stable matchings with respect to a certain preference profile. Niederle and Yariv (2009) also propose a model of decentralized job matching. Contrary to us, they assume that for each job market there is a unique stable matching and the equilibrium concept they use is the weakly undominated Nash equilibrium. When there is complete information, they show that any equilibrium yields the (unique) stable matching.

8 Conclusion

We found that whenever workers play sequentially, the W-optimal matching is the unique subgame perfect equilibrium outcome. When workers play simultaneously, we show that any stable matching can be obtained as an equilibrium outcome and that unstable matchings may also arise in equilibrium. These results are in marked

[23] See also Peleg (1997) and Alcalde and Romero-Medina (2000) for models similar to that of Alcalde et al. (1998) for the marriage model and the College Admission problem, respectively.

contrast with those obtained in centralized markets. Indeed, if a market is centralized and firms make the proposals, the deferred acceptance algorithm yields the F-optimal matching. Such a reversal of outcomes is not new in the matching literature, however. Alcalde (1996) studies the case of a centralized market in which the deferred acceptance algorithm is used to match players, and players act strategically. He shows that when the algorithm in use is the deferred acceptance with firms proposing the game is dominance solvable, yielding the W-optimal matching.[24] When matching is decentralized, it is in the interest of each worker who is matched to a firm less preferred than his W-optimal mates to decline such a matching. When they do so, they can help each other to attain the W-optimal matching; as long as we have not reached the W-optimal matching, workers' interests coincide. However, for matchings that are preferred to the W-optimal matching, implicit cooperation is no longer optimal; workers compete for firms preferred to their W-optimal mates. For a worker to be better off than he would be with his W-optimal mate, it must be the case that another worker is worse off, and indeed accepts being worse off (i.e., he declines his W-optimal mate).

Acknowledgments We thank Francis Bloch, Jordi Massó, Alejandro Neme, Joana Pais, Pablo Revilla and Flip Klijn. We are also extremely grateful to William Thomson and two anonymous referees for their thorough reading of the paper and their comments. The first author acknowledges financial support from the Spanish Ministry of Science and Innovation through grant "Consolidated Group-C" ECO2008-04756, and by the Generalitat de Catalunya, Departament d'Universitats, Recerca i Societat de la Informació through grant SGR2009-0419, and through the Barcelona Graduate School of Economics. This research was initiated when both authors were members of the Department of Economics of the University of Warwick.

A Appendix: Proof of Theorem 2

Given a strategy profile σ, define by $\mathcal{O}_w^\sigma$ the set of all possible offers w can receive when the other players play according to σ_{-w};

$$\mathcal{O}_w^\sigma = \{f \in F \mid \exists\, \widehat{\sigma}_w \text{ such that } \mu[\widehat{\sigma}_w, \sigma_{-w}](w) = f\}.$$

where $\mu[\widehat{\sigma}_w, \sigma_{-w}](w)$ denotes the mate assigned to w at $\mu[\sigma]$.

Remark 4 For a worker w and a strategy profile σ_{-w}, observe that the strategy of accepting only the offer from $\max_{P_w} \mathcal{O}_w^\sigma$ and declining all other offers is a best response.[25] Hence, since a subgame perfect equilibrium is necessarily a Nash equilibrium, if σ is a SPE of $\mathcal{G}(P)$ then we have, for all $w \in W$, $\mu[\sigma](w) = \max_{P_w} \mathcal{O}_w^\sigma$. That is, SPE's match each worker to the best mate he can obtain given the strategy of the other players.

Remark 5 Notice that if $\sigma_W \in SPE(\mathcal{G}(P|\sigma_F^*))$ then $\mu[\sigma_W, \sigma_F^*] \in IR(P)$. This is so because by construction, a firm never makes an offer to an unacceptable worker and

24 In Alcalde's game, a player's strategy set is the set of all possible preferences she can have (one of which is her 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.

25 Since we assumed that preferences are strict and F is finite, there is a unique maximal element.

a worker is always better off declining any offer rather than accepting the offer of an unacceptable firm.

We now prove Theorem 2. To begin with, we start by recalling two classic results of the literature.

Lemma 1 (Weak Pareto Efficiency, Roth (1982)) *Let (F, W, P) be a matching market and let μ_W be its W-optimal matching. There is no individually rational matching μ such that for all $w \in W$ we have $\mu P_w \mu_W$.*

Lemma 2 (Blocking Lemma, Gale and Sotomayor (1985)) *Let μ be any individually rational matching and let W° be the set of all workers who prefer μ to μ_W. If W° is nonempty, there is a pair (f, w) satisfying $w \in W \backslash W^\circ$ and $f \in \mu(W^\circ)$ that blocks μ.*

The next Lemma says that at a subgame perfect equilibrium, no worker can be matched to a firm preferred to his W-optimal firm. Interestingly, this Lemma holds independently of whether workers act simultaneously or sequentially at each stage.

Lemma 3 *Let $\sigma_W \in SPE(\mathcal{G}(P|\sigma_F^*))$. Then for all workers $w \in W$ it holds that $\mu_W R_w \mu[\sigma_W, \sigma_F^*]$.*

Proof Let $\sigma_W \in SPE(\mathcal{G}(P|\sigma_F^*))$ and suppose that for $w_1 \in W$, $\mu[\sigma_W, \sigma_F^*] P_{w_1} \mu_W$. Observe that since $\mu_W \in IR(P)$, $\mu[\sigma_W, \sigma_F^*](w_1) \in F$. Let $f_2 = \mu[\sigma_W, \sigma_F^*](w_1)$. Since $\mu_W \in S(P)$

$$\mu_W P_{f_2} \mu[\sigma_W, \sigma_F^*] \tag{1}$$

(for otherwise f_2 and w_1 could block μ_W) and thus $\mu_W(f_2) \in W$. Let $w_2 = \mu_W(f_2)$. We claim that $\mu[\sigma_W, \sigma_F^*] P_{w_2} \mu_W$. To see this, suppose on the contrary that $\mu_W P_{w_2} \mu[\sigma_W, \sigma_F^*]$. It follows that $f_2 \neq \max_{P_{w_2}} \mathcal{O}_{w_2}^{\sigma_W, \sigma_F^*}$, for otherwise w_2 could get an offer from f_2 or a more preferred firm. Thus, f_2 is matched before having the opportunity to make an offer to w_2. Since firms make offers in the order of in their preferences, we have $\mu[\sigma_W, \sigma_F^*] P_{f_2} \mu_W$. This contradicts Eq. 1, so $\mu[\sigma_W, \sigma_F^*] P_{w_2} \mu_W$.

Again $\mu_W \in IR(P)$ implies $\mu[\sigma_W, \sigma_F^*](w_2) \in F$. Let $f_3 = \mu[\sigma_W, \sigma_F^*](w_2)$. Since $\mu_W \in S(P)$,

$$\mu_W P_{f_3} \mu[\sigma_W, \sigma_F^*]. \tag{2}$$

Since $\mu[\sigma_W, \sigma_F^*] \in IR(P)$, $\mu_W(f_3) \in W$. Let $w_3 = \mu_W(f_3)$. Arguing as above, it can be shown that $\mu[\sigma_W, \sigma_F^*] P_{w_3} \mu_W$.

We can continue the process, and restart if necessary, to determine the largest set of workers $W^1 = \{w_1, w_2, w_3, w_4, \ldots\ldots, w_k\}$ where for each worker $w \in W^1$, $\mu[\sigma_W, \sigma_F^*] P_w \mu_W$ and $\mu[\sigma_W, \sigma_F^*](W^1) = \mu_W(W^1)$. We then have two cases, depending on whether $W^1 = W$ or $W^1 \subsetneq W$. If $W^1 = W$ then from Lemma 1 we deduce that $\mu[\sigma_W, \sigma_F^*] \notin IR(P)$, a contradiction since $\sigma_W \in SPE(\mathcal{G}(P|\sigma_F^*))$. So, $W^1 \subsetneq W$. We can then invoke the Blocking Lemma to deduce that there exists a pair (w, f) with $w \notin W^1$ and $f \in \mu(W^1)$ such that (w, f) blocks the matching $\mu[\sigma_W, \sigma_F^*]$. That is,

$$w P_f \mu[\sigma_W, \sigma_F^*](f) \quad \text{and} \quad f P_w \mu[\sigma_W, \sigma_F^*](w). \tag{3}$$

Since w prefers f to $\mu[\sigma_W, \sigma_F^*](w)$, $f \neq \max_{P_w} \mathcal{O}_w^{(\sigma_W, \sigma_F^*)}$, for otherwise w could get an offer from f or a more preferred firm. Thus, f is matched to $\mu[\sigma_W, \sigma_F^*](f)$ before having the opportunity to make an offer to w. Since firms make offers in the order of their preferences, we have $\mu[\sigma_W, \sigma_F^*] P_f w$, which contradicts Eq. 3. □

Consider a worker who is not matched to his W-optimal mate at some equilibrium of the game. According to the previous lemma, this worker is worse off than at the W-optimal matching. The next lemma states that this worker is matched to a firm, and the W-optimal worker of that firm is also matched to a firm less preferred than his W-optimal firm. The proof of this lemma exploits the argument that was used in the proof of Lemma 3.

Lemma 4 *Let $\sigma_W \in SPE(\mathcal{G}(P|\sigma_F^*))$. Assume that w satisfies $\mu_W P_w \mu[\sigma_W, \sigma_F^*]$. Then there exist $\widehat{W} = \{w_1, \ldots, w_k\}$ and $\widehat{F} = \{f_1, \ldots, f_k\}$ such that $w \in \widehat{W}$, $\mu[\sigma_W, \sigma_F^*](w_h) = f_{h+1}$ (modulo k), $\mu_W(w_h) = f_h$, and $\mu_W P_{w_h} \mu[\sigma_W, \sigma_F^*]$ for all $h = 1, \ldots, k$.*[26]

Proof Let σ be a SPE of $\mathcal{G}(P)$, and let w_1 be such that $\mu_W P_{w_1} \mu[\sigma]$. Since $\mu[\sigma]$ is individually rational (see Remark 5), $\mu_W(w_1) \in F$. Let $f_1 = \mu_W(w_1)$. Notice that $\mu_W P_{w_1} \mu[\sigma_W, \sigma_F^*]$ implies $f_1 \neq \max_{P_{w_1}} \mathcal{O}_{w_1}^{(\sigma_W, \sigma_F^*)}$, for otherwise w_1 could get an offer from f_1 or a more preferred firm. Thus, f_1 is matched before having the opportunity to make an offer to w_1. Since firms make offers in the order of their preferences dictate, this implies that $\mu[\sigma_W, \sigma_F^*] P_{f_1} \mu_W$. We then have $\mu[\sigma_W, \sigma_F^*](f_1) \in W$. Let $w_2 = \mu[\sigma_W, \sigma_F^*](f_1)$. Since μ_W is stable, $\mu_W P_{w_2} \mu[\sigma_W, \sigma_F^*]$. So, $\mu_W(w_2) \in F$. Let $f_2 = \mu_W(w_2)$. We then have either $f_2 = \mu[\sigma_W, \sigma_F^*](w_1)$ or $f_2 \neq \mu[\sigma_W, \sigma_F^*](w_1)$. In the first case we are done. In the second case, let $w_3 = \mu[\sigma_W, \sigma_F^*](f_2)$. Using the same reasoning as for w_2, we deduce that $\mu_W P_{w_3} \mu[\sigma_W, \sigma_F^*]$. Continuing this way, the finiteness of W implies that we end up with a firm f_k such that $\mu[\sigma_W, \sigma_F^*](f_k) = w_1$. We then obtain a set of firms and workers, $\widehat{F} = \{f_1, \ldots, f_k\}$ and $\widehat{W} = \{w_1, \ldots, w_k\}$ satisfying the properties that both μ_W and $\mu[\sigma_W, \sigma_F^*]$ map $\widehat{F}$ onto $\widehat{W}$, and for all $w \in \widehat{W}$, $\mu_W P_w \mu[\sigma_W, \sigma_F^*]$.[27] □

One of the main difficulties when proving Theorem 2 is the following. Consider a job market (W, F, P) and its W-optimal matching, μ_W. Let (w, f) satisfy $\mu_W(w) = f$. Consider now the sub-market $(W\backslash\{w\}, F\backslash\{f\}, P')$ where P' is the preference profile P restricted to the set of players $W\backslash\{w\}$, $F\backslash\{f\}$, and let μ'_W be the W-optimal

[26] Since our purpose is to prove that any subgame perfect equilibrium yields the worker-optimal matching Lemma 4 is stated considering a profile σ that is a SPE. The reader will notice that Lemma 4 is still valid if we only require σ to be a Nash equilibrium.

[27] There might be several pairs of such sets $\widehat{W}$ and $\widehat{F}$, i.e., sets $\widehat{W}_1, \ldots, \widehat{W}_l$ and $\widehat{F}_1, \ldots, \widehat{F}_l$ such that μ_F and μ_W map $\widehat{W}_h$ onto $\widehat{F}_h$, $h = 1, \ldots, l$. For the sake of simplicity we proceeded by considering only one such pair of sets.

matching of this sub-market. It is well known that for some $w' \in W \setminus \{w\}$, we may have $\mu'_W(w') \neq \mu_W(w')$. That is, the W-optimal matching operator is not consistent (see Ergin 2002). This lack of consistency creates a difficulty because, when pairs of workers and firms get matched as the game unfolds, their mate at the W-optimal matching of the sub-market at the history at which they are matched is not necessarily the same as that of the original job market.

Proof of Theorem 2 We prove the Theorem by induction on the highest number of acceptable workers across all firms.[28]

Step 1: For each firm $f \in F$, $|A(P_f)| \leq 2$.[29]

Let σ_W be a SPE of $\mathcal{G}(P|\sigma_F^*)$. Using Lemma 3, we can partition W in two sets, W^* and $\widehat{W}$, where for each $w \in W^*$, $\mu[\sigma_W, \sigma_F^*](w) = \mu_W(w)$ and for each $w \in \widehat{W}$, $\mu_W P_w \mu[\sigma_W, \sigma_F^*]$. To prove the theorem it then suffices to show that $\widehat{W} = \emptyset$. Suppose on the contrary that there exists $w \in \widehat{W}$. By Lemma 4, $\mu_W(w), \mu[\sigma_W, \sigma_F^*](w) \in F$. Let $f' = \mu[\sigma_W, \sigma_F^*](w)$. Again by Lemma 4, $\mu_W(f') \in \widehat{W}$. Let $w' = \mu_W(f')$. We claim that w is matched in the first stage under σ_W. Suppose not, i.e., w and f' are matched in the second stage. Thus, $f' \neq \max_{P_{w'}} \mathcal{O}_{w'}^{\sigma_W}$, which implies that w' obtains under σ_W a better offer than f'. So, $\mu[\sigma_W, \sigma_F^*] P_{w'} \mu_W$. This contradicts Lemma 3. Hence, $w \in \widehat{W}$ implies that w is matched in the first stage.

Let $\widehat{W} = \{w_1, \ldots, w_k\}$ and let $\mu[\sigma_W, \sigma_F^*](w_h) = f_{h-1}$ (modulo k) and $\mu_W(w_h) = f_h$, for all $h = 1, \ldots, k$. Let $\widehat{F} = \{f_1, \ldots, f_k\}$. Let w_1 be the first worker in $\widehat{W}$ to play in the first stage. Notice that since $|A(P_f)| \leq 2$ for all $f \in F$, we have $A(P_{f_h}) = \{w_h, w_{h+1}\}$, for all $h = 1, \ldots, k$ (modulo k). Hence, firm f_h makes an offer to w_{h+1} in the first stage and, if it is declined, it makes an offer to w_h in the second stage whenever this worker is available. Recall from the claim of the previous paragraph that w_1 accepts an offer from f_k in the first stage. Consider instead the subgame $\mathcal{G}'$ that starts just after w_1's hypothetical rejection of f_k's offer. In $\mathcal{G}'$, observe that it is a strictly dominant strategy for w_k to decline f_{k-1}'s offer, for he can get at least f_k in the second stage.[30] We can thus delete any strategy of w_k under which he accepts f_{k-1}'s offer in $\mathcal{G}'$. Thus in $\mathcal{G}'$ firm f_{k-1} makes an offer to w_{k-1} in the second stage, if w_{k-1} is unmatched in this second stage. So, in the first stage in $\mathcal{G}'$, worker w_{k-1} holds an offer from f_{k-2} and if he declines it, he will get an offer from f_{k-1} in the second stage. Since $f_{k-1} P_{w_{k-1}} f_{k-2}$, it is now a dominant strategy in $\mathcal{G}'$ to decline f_{k-2}'s offer. Repeating the deletion of strictly dominated strategy for workers $w_{k-2}, w_{k-3}, \ldots$, we eventually reach a worker w_2 whose acceptance of f_1's offer is strictly dominated. Since a subgame perfect equilibrium must survive the iterated deletion of strictly dominated strategies, any subgame perfect equilibrium of $\mathcal{G}'$ is such that w_1 obtains an offer from f_1. Thus, w_1 is strictly better off declining $\mu[\sigma_W, \sigma_F^*](w_1)$'s offer, which implies that σ_W is not a SPE, a contradiction. So, $\widehat{W} = \emptyset$. We then conclude that $|A(P_f)| \leq 2$ for all $f \in F$ and σ_W is a SPE of $\mathcal{G}(P|\sigma_F^*)$ implies $\mu[\sigma_W, \sigma_F^*] = \mu_W$.

[28] Note that this is tantamount to the highest number of stages in the game.

[29] Starting with $|A(P_f)| \leq 1$ would not help up much for our purpose because in this case there is a unique stable matching, which is both the W-optimal and the F-optimal matching.

[30] Note that it is a dominant strategy for w_k in this subgame, not necessarily in the whole game.

Step 2: $\exists\ f \in F$ such that $|A(P_f)| \geq 3$.

Let $k > 2$ and assume that Theorem 2 holds true whenever $\max_{f\in F} |A(P_f)| < k$. Suppose now that $\max_{f\in F} |A(P_f)| = k$. So, the game $\mathcal{G}(P|\sigma_F^*)$ has at most k stages. Let σ_W be a SPE of $\mathcal{G}(P|\sigma_F^*)$.

Note that from Lemma 3, it follows that if along the equilibrium path of σ_W a worker w receives an offer from $\mu_W(w)$ he must accept it. Hence, it suffices to show that along the equilibrium path, a worker declines any offer from a firm less preferred than his W-optimal mate.

Let h_k be the history that belongs to the equilibrium path such that h_k is the longest history that belongs to the first stage. Let $p(h_k) = w_k$.[31] Let $f_k = \mu_W^{h_k}(w_k)$ and let f be w_k's most preferred offer at h_k, where $\mu_W^{h_k}$ is the worker-optimal matching for the preferences P^{h_k} given by the sub-market at history h_k. Suppose that $f_k P_{w_k} f$. Let h' be the history attained if w_k declines f (and all other offers, if any). It is straightforward to check that $\mu_W^{h_k} = \mu_W^{h'}$. Notice that at h' there are at most $k-1$ stages left. So, the induction hypothesis applies and we deduce that $\mu[\sigma_W, \sigma_F^*|h'](w_k) = \mu_W^{h_k}(w_k)$. So at h_k worker w_k declines any offer less preferred than $\mu_W^{h_k}(w_k)$.

Consider now h_{k-1}, the history that comes just before history h_k (i.e., the penultimate history that belongs to stage 1 along the equilibrium path). Let $p(h_{k-1}) = w_{k-1}$. Let $f_{k-1} = \mu_W^{h_{k-1}}(w_{k-1})$ and let f be w_{k-1}'s most preferred offer at h_{k-1}. Abusing notation, let $f_k = \mu_W^{h_{k-1}}(w_k)$.[32] Suppose that $f_{k-1} P_{w_{k-1}} f$. We claim that w_{k-1}'s optimal action consists of declining all his offers. To see this, let h' be the history attained when w_{k-1} declines f's offer—and all other offers less preferred than that of f. (Note that if this is indeed w_{k-1}'s action at h_{k-1} under σ_W then $h' = h_k$.) Like in the previous paragraph, it is straightforward to check that

$$\mu_W^{h'} = \mu_W^{h_{k-1}}. \tag{4}$$

Let h'' be the immediate history attained after w_k's decision under σ_W. (So $p(h'') \in F$) Notice that at h'' there are at most $k-1$ stages left. So, the induction hypothesis applies and we deduce that $\mu[\sigma_W|h''](w_{k-1}) = \mu_W^{h''}(w_{k-1})$. We need to show that $\mu_W^{h''}(w_{k-1}) P_{w_{k-1}} f$. If $\mu_W^{h''}(w_{k-1}) = \mu_W^{h'}(w_{k-1})$, then we are done. Indeed, using Eq. 4 we obtain $\mu_W^{h''}(w_{k-1}) = f_{k-1}$, i.e., after declining f worker w_{k-1} ends up being matched to f_{k-1}.

Suppose then that $\mu_W^{h''}(w_{k-1}) \neq \mu_W^{h'}(w_{k-1})$. We claim that this implies that at h', worker w_k's strategy consists of accepting an offer. So at h'', w_k is matched, which implies $w_k \notin W^{h''}$. To see this, suppose on the contrary that worker w_k is unmatched at h''. This implies that w_k declined all his offers, and thus repeating the argument developed at history h_k we deduce that $\mu_W^{h''} = \mu_W^{h'}$, a contradiction. So, at history h,

[31] If we rule out the trivial case in which for each firm the most preferred mate is itself, it must be the case that at least one worker received an offer from a firm in stage 1. Hence, we can assume without loss of generality that $p(h_k) \in W$.

[32] Worker w_k's optimal mate at history h_k was already labelled f_k.

w_k is matched. Notice that if w_k is matched at h'' it is to $\mu_W^{h'}(w_k)$. Using Eq. 4 we deduce that w_k is matched to $\mu_W^{h_{k-1}}(w_k) = f_k$.

We now show that $\mu_W^{h''}(w_{k-1}) P_{w_{k-1}} f$. To this end, let $F^{h''} = F^{h'} \setminus \{f_k\}$ and $W^{h''} = W^{h'} \setminus \{w_k\}$ and consider the matching μ over $W^{h'} \cup F^{h''}$ defined as follows. For each $v \in F^{h''} \cup W^{h''}$, $\mu(v) = \mu_W^{h_{k-1}}(v)$. Notice that since $\mu_W^{h_{k-1}}(w_k) = f_k$, $\mu \cup (w_k, f_k) = \mu_W^{h_{k-1}}$, and $\mu(w_{k-1}) = f_{k-1}$. Also, for each worker $w \in W^{h''}$, and $v, v' \in W^{h''} \cup \{w\}$, $v P_w^{h''} v'$ if and only if $v P_w^{h'} v'$. It follows that $\mu \in IR(P^{h''})$ if, and only if, $\mu \in IR(P^{h'})$. Similarly, if $(f, w) \in F^{h''} \times W^{h''}$ is a blocking pair under μ for the profile $P^{h''}$ if, and only if, it is a blocking pair under μ for the profile $P^{h'}$. Since $\mu_W^{h_{k-1}} \in S(P^{h_{k-1}})$, and using Eq. 4, we have $\mu_W^{h_{k-1}} \in S(P^{h'})$, and thus $\mu \in S(P^{h''})$. Since for all workers the W-optimal matching is the most preferred among all stable matchings, we have $\mu_W^{h''} R_{w_{k-1}}^{h''} \mu$, and thus

$$\mu_W^{h''} R_{w_{k-1}} \mu. \tag{5}$$

Since $\mu(w_{k-1}) = f_{k-1}$ and (by assumption) $f_{k-1} P_{w_{k-1}} f$, Eq. 5 and the transitivity of $P_{w_{k-1}}$ imply $\mu_W^{h''}(w_{k-1}) P_{w_{k-1}} f$. We then conclude that at history h_{k-1} worker w_{k-1} is better off declining any offer less preferred than $\mu_W^{h_{k-1}}(w_{k-1})$.

Continuing this way with the last history before h_{k-1} (along the equilibrium path), until reaching the empty history h_0 of the game we have that for any history h along the equilibrium path, if $p(h)$ is a worker, say, w, then w accepts the offer from $\mu_W^h(w)$ if he holds such an offer and declines all other offers. Now, considering the first worker to act in the first stage, say w_1 and let h be the corresponding history. Observe that we obviously have $\mu_W^h(w_1) = \mu_W(w_1)$. Without loss of generality, suppose that w_1 receives and accepts this offer. Let $f_1 = \mu_W(w_1)$ and let h' be the history attained after this action and w_2 the worker who has to act.[33] From the previous argument, w_2 accepts an offer from $\mu_W^{h'}(w_2)$. We need to show that $\mu_W^{h'}(w_2) = \mu_W(w_2)$. To this end, let $F^{h'} = F \setminus \{\mu_W(f_1)\}$ and $W^{h'} = W \setminus \{w_1\}$ and consider the matching μ over $W^{h'} \cup F^{h'}$ defined as follows. For each $v \in F^{h'} \cup W^{h'}$, $\mu(v) = \mu_W^h(v)$. Notice that since $\mu_W^h(w_1) = f_1$, $\mu \cup (w_1, f_1) = \mu_W^h$. Like before, it can be checked that $\mu \in S(P^{h'})$, which implies that $\mu_W^{h'} R_{w_2} \mu$, i.e., $\mu_W^{h'} R_{w_2} \mu_W$. Since $\mu[\sigma_W, \sigma_F^*](w_2) = \mu_W^{h'}(w_2)$ and $\mu_W R_{w_2} \mu[\sigma_W, \sigma_F^*]$ (Lemma 3), we must have $\mu_W^{h'}(w_2) = \mu_W(w_2)$. Continuing this way until worker w_k we then obtain $\mu[\sigma_W, \sigma_F^*] = \mu_W$, the desired result. □

B Appendix: Proof of Theorem 3

Recall from the discussion following Theorem 3 that the main intuition to prove Theorem 3 consists of showing that if σ is a SPE of $\mathcal{G}(P)$, then we can identify for each firm $f \in F$ a preference ordering over workers $\widetilde{P}_f$ such that $\mu[\sigma] = \mu_W(P_W, \widetilde{P}_F)$. The next result, is key in identifying who should be the most preferred worker according

[33] Since w_1 received an acceptable offer in the first stage we have from the DA algorithm $\mu_W(w_1) \in F$.

to $\widetilde{P}_f$. Consider the strategic-form game in which firms and workers submit a preference ordering and a matching is computed using the DA algorithm in which workers propose to firms. So, if P is the submitted preference profile, then the outcome is $\mu_W(P)$. It is well known that in this game submitting one's true preference is not necessarily a dominant strategy for firms.

Proposition 3 (Roth 1982, Corollary 5.1) *In the strategic form game in which workers and firms submit a preference ordering and the matching that is obtained is the worker-optimal matching with respect to the submitted preferences, firms have no incentive to misrepresent their most preferred worker.*

We now consider a very specific form of manipulation by firms. Suppose that a firm, say f, submits a preference ordering P'_f such that one worker, who is not the most preferred worker in P_f, is put on the highest position in P'_f, and the relative ranking of all other workers is unchanged. The next lemma says that if, with such a manipulation, firm f is matched to the same worker as the one without the manipulation then the match of any other worker and firm is also unchanged.[34]

Lemma 5 *Let P be a preference profile, and let $f \in F$. Let P'_f be the preference such that (a)* $\max_{P_f} W \neq \max_{P'_f} W$*; (b) for any $w, w' \neq \max_{P'_f} W$, $w P_f w' \Leftrightarrow w P'_f w'$; and (c) for each $w \in W$, $w P_f f \Leftrightarrow w P'_f f$. If $\mu_W(P_{-f}, P_f)(f) = \mu_W(P_{-f}, P'_f)(f)$ then $\mu_W(P_{-f}, P_f) = \mu_W(P_{-f}, P'_f)$.*

Proof Let $f \in F$ and let P'_f satisfy the conditions of the lemma. Let $w = \max_{P'_f} W$. So, $\mu_W(P_{-f}, P_f)(f) = \mu_W(P_{-f}, P'_f)(f) \neq w$. Let $\mu_W = \mu_W(P_{-f}, P_f)$ We distinguish among three cases.
Case 1: $w P_f \mu_W(f)$. Notice that when running the Deferred Acceptance Algorithm with the profile P firm f again never receives an offer from w. So under the profile (P_{-f}, P'_f) firm f never receives an offer from w either and thus firm f has no way to alter the temporary matches that occur under the Deferred Acceptance Algorithm. Hence, $\mu_W(P_{-f}, P_f) = \mu_W(P_{-f}, P'_f)$.
Case 2: $w = \mu_W(f)$. It is easy to see that the sequences of acceptance and rejections in the DA algorithm with preferences P and (P_{-f}, P'_f) are identical, so $\mu_W(P_{-f}, P_f) = \mu_W(P_{-f}, P'_f)$.
Case 3: $\mu_W(f) P_f w$. Suppose that in the DA algorithm with preferences P firm f receives an offer from worker w. So in the DA algorithm with preferences (P'_f, P_{-f}) firm f also receives an offer from worker w, and thus $\mu_W(P_{-f}, P'_f)(f) = w$, a contradiction. So, in the DA algorithm with preferences P firm f never receives an offer from worker w. Like in *Case 2*, it is easy to see that the sequences of acceptance and rejections in the DA algorithm with preferences P and (P_{-f}, P'_f) are identical. (In particular, notice that under P' worker w also does not propose to f at any step.) So, we have $\mu_W(P_{-f}, P_f) = \mu_W(P_{-f}, P'_f)$. □

[34] The property established in Lemma 5 is known in Social Choice Theory as *non-bossiness* of social choice functions, with the exception that here we restrict the set of manipulations that firms can do.

The manipulation considered in Lemma 5 is exactly the type of manipulation that takes place in the game $\mathcal{G}(P)$. The induction argument we use assumes that from the second stage onwards firms make offers in the same order as in their preferences. Thus, the problem for a firm is the choice of its first offer. So choosing the first offer and then making subsequent offers (if needed) in the same order as in the preference is, roughly speaking, equivalent to choosing a preference ordering that satisfies the condition of Lemma 5.

The next lemma, treating the case where firms each have at most two acceptable workers, is a "small" version of the Theorem we want to prove. It will serve to prove Theorem 3 with an induction argument.

Lemma 6 *Suppose that for each firm $f \in F$, $|A(P_f)| \leq 2$. Let $\sigma \in SPE(\mathcal{G}(P))$. Then $\mu[\sigma] = \mu_W$. Furthermore, $(\sigma_W, \sigma_F^*) \in SPE(\mathcal{G}(P))$.*

Proof Notice that since for each firm $f \in F$ we have $|A(P_f)| \leq 2$, a firm whose offer has been declined in the first stage has no option other than making an offer to the remaining worker that is acceptable in its preferences (if any) as long as this worker is not matched in the first stage. That is, firms' actions in the second stage cannot be conditional on the sequence of offers made by other firms in the second stage.

Let h be any history of length n. So, $p(h) = f_n$. For each firm f_i, $i \leq n$, we define a preference ordering, denoted $P_{f_i}(h)$, as follows—and denoting by $P_F(h)$ the profile $(P_f(h))_{f \in F}$. Let h' be the sub-history of h at which f_i plays, i.e., $p(h') = f_i$. Let $A(P_{f_i}(h)) = A(P_{f_i})$. If $A(P_{f_i}) \neq \emptyset$, let $\max_{P_{f_i}(h)} W$ be the worker to whom f_i has made an offer at h' according to $\sigma_{f_i}(h')$. Observe that since in the second stage unmatched firms have no option other than making an offer to their unique acceptable worker, the game for workers is identical to the game $\mathcal{G}(P_W, P_F(h))$, and we have $\mu[\sigma|h] = \mu_W(P_W, P_F(h))$. Consider the strategy $\sigma_{f_n}^*$, i.e., firm f_n makes an offer to its most preferred worker in the first stage to its second most preferred acceptable worker in the second stage or to itself if there is no such worker. By Proposition 3, $\mu_W(P_W, P_{F \setminus \{f_n\}}(h), P_{f_n}) R_{f_n} \mu_W(P_W, P_F(h))$. By the same reasoning that leads to $\mu[\sigma|h] = \mu_W(P_W, P_F(h))$, $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*] = \mu_W(P_W, P_{F \setminus \{f_n\}}(h), P_{f_n})$. Thus, $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*] R_{f_n} \mu[\sigma|h]$. Either $\sigma_{f_n}(h) = \sigma_{f_n}^*(h)$ or $\sigma_{f_n}(h) \neq \sigma_{f_n}^*(h)$. In either case, since σ is a SPE and $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*] R_{f_n} \mu[\sigma|h]$, we have $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^* | h](f_n) = \mu[\sigma|h](f_n)$. By Lemma 5, we have $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n} | h] = \mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^* | h]$. Repeating this argument for all other histories of length n we obtain that $\sigma \in SPE(\mathcal{G}(P))$ implies $(\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*) \in SPE(\mathcal{G}(P))$, and $\mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}] = \mu[\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*]$.

Consider now the strategy profile $(\sigma_W, \sigma_{F \setminus \{f_n\}}, \sigma_{f_n}^*)$. Since f_n's action in the first stage is independent of the action played by the other firms in the first stage, we can repeat the above reasoning and use the strategy $\sigma_{f_{n-1}}^*$. Continuing this way with all other firms we obtain that if $\sigma \in SPE(\mathcal{G}(P))$ then $(\sigma_W, \sigma_F^*) \in SPE(\mathcal{G}(P))$, and $\mu[\sigma] = \mu[\sigma_W, \sigma_F^*] = \mu_W$, where the second equality comes from Theorem 2. □

We can now prove Theorem 3. The proof relies on an argument similar to that of Lemma 6.

Proof of Theorem 3 We prove the theorem by induction on the highest number of acceptable workers across all firms. Let $k > 2$ and assume that whenever $\max_{f \in F} |A(P_f)| < k$, if $\sigma \in SPE(\mathcal{G}(P))$ then $(\sigma_W, \sigma_F^*) \in SPE(\mathcal{G}(P))$ and $\mu[\sigma] = \mu[\sigma_W, \sigma_F^*] = \mu_W$. Suppose now that $\max_{f \in F} |A(P_f)| = k$, so the game $\mathcal{G}(P)$ has at most k stages. Let $\sigma \in SPE(\mathcal{G}(P))$.

Let h be any history that corresponds to the offering stage of the second stage. That is, at history h a firm is making an offer to some worker. Notice that for each firm $f \in F^h$, $|A(P_f^h)| < k$, so the induction hypothesis applies and we can replace σ_F by the strategy profile $\widetilde{\sigma}_F$ such that for each firm $f \in F^h$, for any history $h' \supseteq h$ such that $p(h') = f$, the strategy $\widetilde{\sigma}_f(h')$ consists of firm f making an offer to $\max_{P_f^{h'}} W^{h'}$, i.e., to its most preferred worker at history h'. For any history $h' \subset h$, let $\widetilde{\sigma}_f(h') = \sigma_f(h')$. We claim that $(\sigma_W, \widetilde{\sigma}_F) \in SPE(\mathcal{G}(P))$. To see this, observe that by the induction hypothesis, for any sub-game that starts at the beginning of stage 2 or later, $(\sigma_W, \widetilde{\sigma}_F)$ is a SPE for that particular sub-game. Hence it remains to show that it is also a SPE for the whole game. By the one deviation property (see Osborne and Rubinstein 1994, Lemma 98.2), it suffices to show that no worker and no firm has a profitable deviation at any history h' that belongs to the first stage. By construction, for any such history $\mu[\sigma|h'] = \mu[\sigma_W, \widetilde{\sigma}_F|h']$. So if at h' player $p(h')$ has a profitable deviation when the strategy profile is $(\sigma_W, \widetilde{\sigma}_F)$ it also has a profitable deviation at h' when the strategy profile is σ, which contradicts $\sigma \in SPE(\mathcal{G}(P))$ and proves the claim. So we have $(\sigma_W, \widetilde{\sigma}_F) \in SPE(\mathcal{G}(P))$ and

$$\mu[\sigma] = \mu[\sigma_W, \widetilde{\sigma}_F]. \tag{6}$$

Let $\widetilde{P}_F$ be the following preference profile. For each firm f, $\max_{\widetilde{P}_f} W$ is the worker to which f is making an offer in the first stage *along the execution path* of σ. For all other workers, their relative ranking in $\widetilde{P}_f$ and P_f are unchanged. Formally, for any $w, w' \neq \max_{\widetilde{P}_f} W$, $w \widetilde{P}_f w' \Leftrightarrow w P_f w'$, and for each $w \neq \max_{\widetilde{P}_f} W$, $w \widetilde{P}_f f \Leftrightarrow w P_f f$. Let $\widetilde{P} = (P_W, \widetilde{P}_F)$.

Clearly, $(\sigma_W, \widetilde{\sigma}_F) \in SPE(\mathcal{G}(P))$ implies $\sigma_W \in SPE(\mathcal{G}(P|\widetilde{\sigma}_F))$. So by Theorem 2 we have $\mu[\sigma_W, \widetilde{\sigma}_F] = \mu_W(P_W, \widetilde{P}_F)$. Combining with Eq. 6 we obtain

$$\mu[\sigma] = \mu_W(P_W, \widetilde{P}_F). \tag{7}$$

To prove the theorem it suffices then to show that $\mu_W(P_W, \widetilde{P}_F) = \mu_W(P)$. To this end, consider any history h of $\mathcal{G}(P)$ of length n. So $p(h) = f_n$. At history h, all firms $f_1, \ldots, f_{n-1}$ have already made an offer to some worker, say, to $w_1^h, w_2^h, \ldots, w_{n-1}^h$, respectively. Let w_n^h be the worker to whom f_n makes an offer at h according to $\widetilde{\sigma}_{f_n}(h)$. For each firm f_i, $i \leq n$, let $P_{f_i}(h)$ be the following preference relation:

$$\max_{P_{f_i(h)}} W = w_i^h, \tag{8}$$

$$\text{for any } w, w' \in W \backslash \{w_i^h\}, \; w P_{f_i}(h) w' \quad \Leftrightarrow \quad w P_{f_i} w', \tag{9}$$

$$\text{for each } w \in W, \; w P_{f_i}(h) f_i \quad \Leftrightarrow \quad w P_{f_i} f_i. \tag{10}$$

After history h firms play according to $\widetilde{\sigma}$. So at the subgame that starts at the immediate history after history h, say $\mathcal{G}'$, the strategy profile $(\sigma_W, \widetilde{\sigma}_F)$ restricted to $\mathcal{G}'$ is a SPE of $\mathcal{G}'$. The same holds for all the subgames of $\mathcal{G}'$. Let $\widetilde{\widetilde{\sigma}}_F$ be the strategy profile for firms such that history h belongs to its execution path and for all other histories $h' \supseteq h$, and firm f such that $p(h') = f$, $\widetilde{\widetilde{\sigma}}_f(h') = \widetilde{\sigma}_f(h')$. Clearly, $\sigma_W \in SPE(\mathcal{G}(P|\widetilde{\widetilde{\sigma}}_F))$. This is so because the only difference between $\widetilde{\widetilde{\sigma}}_F$ and $\widetilde{\sigma}_F$ take place in the actions taken in the first sub-stage, before any worker has to play. So by Theorem 2, $\mu[\sigma_W, \widetilde{\widetilde{\sigma}}_F] = \mu_W(P_W, P_F(h))$. By construction, $\mu[\sigma_W, \widetilde{\widetilde{\sigma}}_F] = \mu[\sigma_W, \widetilde{\sigma}_F|h]$ and thus $\mu[\sigma_W, \widetilde{\sigma}_F|h] = \mu_W(P_W, P_F(h))$. Suppose now that at history h firm f_n makes an offer to its most preferred worker, $\max_{P_f} W$. Notice that from the second stage onwards (after f makes an offer to $\max_{P_f} W$), σ^*_f and $\widetilde{\sigma}_f$ agree. That is, firm f uses now the strategy σ^*_f. So, for the same reason as for $\widetilde{\widetilde{\sigma}}_F$, we have $\sigma_W \in SPE(\mathcal{G}(P|(\widetilde{\widetilde{\sigma}}_{F\setminus\{f_n\}}, \sigma^*_{f_n})))$. By Theorem 2, $\mu[\sigma_W, \widetilde{\widetilde{\sigma}}_{F\setminus\{f_n\}}, \sigma^*_{f_n}] = \mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n})$. Since $\mu[\sigma_W, \widetilde{\sigma}_{F\setminus\{f_n\}}, \sigma^*_{f_n}|h] = \mu[\sigma_W, \widetilde{\widetilde{\sigma}}_{F\setminus\{f_n\}}, \sigma^*_{f_n}]$ we then obtain $\mu[\sigma_W, \widetilde{\sigma}_{F\setminus\{f_n\}}, \sigma^*_{f_n}|h] = \mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n})$.

By Proposition 3, $\mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n}) R_{f_n} \mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n}(h))$. Since $(\sigma_W, \widetilde{\sigma}) \in SPE(\mathcal{G}(P))$, we have $\mu_W(P_W, P_F(h))(f_n) = \mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n})(f_n)$. That is, at history h firm f_n is no worse off making an offer to its most preferred worker (and then making offers in the same order as in P_f) than making an offer to $\max_{P_f(h)} W$.[35] By Lemma 5 this implies $\mu_W(P_W, P_F(h)) = \mu_W(P_W, P_{F\setminus\{f_n\}}(h), P_{f_n})$. Repeating the argument for all other histories h of length n, we obtain $(\sigma_W, \widetilde{\sigma}_F) \in SPE(\mathcal{G}(P))$ implies $(\sigma_W, \widetilde{\sigma}_{F\setminus\{f_n\}}, \sigma^*_{f_n}) \in SPE(\mathcal{G}(P))$ and

$$\mu[\sigma] = \mu_W(P_W, \widetilde{P}_{F\setminus\{f_n\}}, P_{f_n}). \tag{11}$$

Consider now any history h of length $n-1$ (so $p(h) = f_{n-1}$) and the strategy profile $(\sigma_W, \widetilde{\sigma}_{F\setminus\{f_n\}}, \sigma^*_{f_n})$. Note that firm f_n makes an offer in the first stage to $\max_{P_{f_n}} W$, independently of the offers made previously by the other firms. So we can repeat the reasoning that led to Eq. 11 and deduce that $(\sigma_W, \widetilde{\sigma}_{F\setminus\{f_n\}}, \sigma^*_{f_n}) \in SPE(\mathcal{G}(P))$ and implies $(\sigma_W, \widetilde{\sigma}_{F\setminus\{f_{n-1},f_n\}}, \sigma^*_{f_{n-1},f_n}) \in SPE(\mathcal{G}(P))$ and $\mu[\sigma] = \mu_W(P_W, \widetilde{P}_{F\setminus\{f_{n-1},f_n\}}, P_{f_{n-1},f_n})$.

Continuing this way until we reach the empty history h_0 we obtain $(\sigma_W, \widetilde{\sigma}_F) \in SPE(\mathcal{G}(P))$ implies $(\sigma_W, \sigma^*_F) \in SPE(\mathcal{G}(P))$, $\mu[\sigma] = \mu_W(P)$, the desired result. □

References

Alcalde J (1996) Implementation of stable solutions to marriage problems. J Econ Theory 69:240–254

Alcalde J, Romero-Medina A (2000) Simple mechanisms to implement the core of college admission problems. Games Econ Behav 31:294–302

Alcalde J, Pérez-Castrillo D, Romero-Medina A (1998) Hiring procedures to implement stable allocations. J Econ Theory 82:469–480

Blum Y, Roth AE, Rothblum UG (1997) Vacancy chains and equilibration in senior-level labor markets. J Econ Theory 76:362–411

[35] If $\max_{P_f} W = \max_{P_f(h) W}$ then this is obvious.

Diamantoudi E, Miyagawa E, Xue L (2006) Decentralized matching: the role of commitment. Mimeo
Ergin HI (2002) Efficient resource allocation on the basis of priorities. Econometrica 70:2489–2497
Gale D, Shapley L (1962) College admission and the stability of marriage. Am Math Mon 69:9–15
Gale D, Sotomayor M (1985) Some remarks on the stable matching problem. Discrete Appl Math 11:223–232
Kelso A, Crawford V (1982) Job matching, coalition formation, and gross substitutes. Econometrica 50:1483–1504
Knuth DE (1976) Mariages stables et leurs relations avec d'autres problèmes combinatoires. Les Presses de l'Université de Montréal, Montreal, Québec. Introduction à l'analyse mathématique des algorithmes, Collection de la Chaire Aisenstadt
Martínez R, Massó J, Neme A, Oviedo J (2000) Single agents and the set of many-to-one stable matchings. J Econ Theory 91:91–105
McVitie D, Wilson LB (1970) Stable marriage assignment for unequal sets. BIT 10:295–309
Niederle M, Yariv L (2009) Decentralized matching with aligned preferences. Mimeo
Osborne MJ, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge
Pais J (2008) Incentives in decentralized random matching markets. Games Econ Behav 64:632–649
Peleg B (1997) Implementation of the core of a marriage problem. Discussion paper ♯132. Hebrew University of Jerusalem
Roth AE (1982) The economics of matching: stability and incentives. Math Oper Res 7:617–628
Roth AE (1985) The college admissions problem is not equivalent to the marriage problem. J Econ Theory 36:277–288
Roth AE, Sotomayor M (1990) Two-sided matching: a study in game-theoretic modeling and analysis. Econometric Society Monographs, No. 18. Cambridge University Press, Cambridge
Roth AE, Xing X (1997) Turnaround time and bottlenecks in market clearing: decentralized matching in the market for clinical psychologists. J Polit Econ 105:284–329
Suh S-C, Wen Q (2008) Subgame perfect implementation of stable matchings in marriage problems. Soc Choice and Welfare 31:163–174