Two-Sided Matching with Indifferences

Aytek Erdil* Haluk Ergin†

September 2006

Abstract

In two-sided matching literature it has been a standard assumption that agents are not indifferent between any two members of the opposite side, despite the existence of such indifferences in various actual settings. A number of issues arise if such an assumption is abandoned and weak preferences are allowed. Most importantly, stability no longer implies Pareto efficiency, and the deferred acceptance algorithm can not be applied to produce a Pareto efficient or a worker/firm optimal stable matching.

In this paper, we allow ties in preference rankings and explore the Pareto domination relation on stable matchings, as well as the two relations defined via workers' welfare and firms' welfare. Our structural results lead to *fast* algorithms to compute a Pareto efficient and stable matching, and a worker [or firm] optimal stable matching.

1 Introduction

1.1 Background

Several entry-level labor markets appear to suffer from coordination failures. The somewhat chaotic nature of their decentralized structure leads to congestion or costly unraveling occurring over time due to strategic behavior of the participants. (Kagel and Roth

*Harvard Business School. Email: aerdil@hbs.edu
†MIT, Department of Economics. Email: hergin@mit.edu

2000; Roth 2002). It has been understood that in many cases it is possible to design a centralized clearinghouse that performs better and operates more easily than the decentralized bilateral contracting. In the US, examples include matching programs regarding Medical Residencies, Postdoctoral Dental Residencies, Pharmacy Practice Residencies, Clinical Psychology Internships, Reform Rabbis; in the UK, medical residencies; and in Canada, Clerks with Law Firms, and Medical Residencies.

In the context of matching, *stability* is a notion that, in some sense, captures the competitive nature of a decentralized market working well. It requires that once a matching is announced no two agents would rather be matched with each other instead of whoever their matches are. Indeed as reported in Roth (2002), there is strong correlation between a clearinghouse being successful and its delivering stable matchings. The various regional markets for new physicians and surgeons in the UK provide field data on this, and the lab experiments by Kagel and Roth (2000) confirm this prediction in a controlled environment. In the context of public resource allocation on the basis of priorities, the very same notion captures the idea of *respecting priorities.* Accordingly stability has been a property expected to be satisfied by most centralized matching schemes.

Gale and Shapley (1962) might be the first to study a two-sided matching environment where both sides are assumed to have preferences over the opposite side that can be represented by linear orders. They proved the existence of stable matchings by demonstrating an algorithm which also turns out to be not only of polynomial time, but very fast in practice. Their algorithm, called the *Deferred Acceptance Algorithm*, has been central to the design of matching programs across several institutions and markets.

1.2 Motivation

Gale and Shapley's stylized model captured some economic environments correctly, but certainly not all of them fit perfectly. The design of economic institutions may be an abstract endeavor in its purest form, yet the actual tasks require the theorist to respond to the details and complications of the real world problems, which their models are supposed to address. The motivation of this paper is studying one of those "details" which has been assumed away in economic modelling, yet turns out to be an important one. This is the issue of *indifference* in matching markets. As opposed to another detail, the issue of *couples*, which has been recognized better, the indifferences do not bring in extra constraints, and therefore their presence do not lead to a problem of existence of stable matchings. In fact, typically their appearance suggests less constraints, hence a larger number of stable matchings.

The usual practice in dealing with indifferences is to break the ties arbitrarily and artificially "make the problem fit into the model," as opposed to enrich the model so that it fits to the problem. This could very well lead to an efficiency loss, even though there is always a way of avoiding such loss. When efficiency concerns are based only one side's preferences, one obvious example is the school choice programs that are in practice in several cities in the U.S. Schools rank students according to priorities as defined in the legislation, whereas students have preferences over the schools which constitute the welfare criteria. Usually the priority rankings involve big classes of indifference for which the common practice is employing random tie-breaking rules.

But indifference might occur even in preferences due to a variety of reasons. Lack of detailed information on the set of alternatives is one example.[1] This is reasonable when it is costly to acquire such information, especially when the set of alternatives is big. When faced with a large number of candidates, an employer might not want to invest into figuring out a strict ranking, depending on the nature of the work. For instance in programs matching students to professors (such as in the Freshman Seminar Program at Harvard), or students to alumni for mentors (such as Student-Alumni Mentor Program at Harvard Business School), it is often the case that the professors or the alumni do not have a strict preference ranking over the set of students, though they may have weak preferences with large indifference classes. In most cases, they would not invest the effort of learning in detail the qualifications of the potential matches.

1.3 Our approach

Since Gale and Shapley (1962), the two-sided matching literature has been built on the assumption that agents are not indifferent between any two potential matches. A lot

[1] Another example is *kidney exchange*: Roth, Sönmez & Ünver (2005) address such preferences, i.e., where patients are assumed or required to have 0-1 preferences in the sense that they find a kidney either acceptable or unacceptable.

of structure in the set of stable matchings is lost if agents may be indifferent between members of the opposite side and the classical results do not carry over to the more general framework. Most importantly, stability no longer implies Pareto efficiency, and the deferred acceptance algorithm can not be applied to produce a Pareto efficient or a worker/firm optimal stable matching.

In this paper, we allow agents to have indifferences and address the following questions: Does there exist stable and Pareto efficient matchings? Are there algorithms that produce Pareto efficient and worker/firm-optimal stable matchings in polynomial time, generalizing the deferred acceptance algorithm to weak preferences? What is the relation between Pareto efficiency and uniform tie-breaking?

Our model consists of a finite set of workers W and a finite set of of firms F. Each worker w can work for at most one firm and each firm f can hire at most a quota q_f of workers. Workers have weak preferences over firms and being unemployed. Firms have weak preferences over workers and maintaining an empty position. A *matching* determines an assignment of the workers to the firms such that each worker works for at most one firm and no firm hires more workers than its quota.

A matching is *individually rational* if no worker would like to quit a position to which she is hired and no firm would like to fire a hired worker. A worker firm pair (w, f) is a *blocking pair* for a given matching if (i) the worker w strictly prefers f to her current match and (ii) the firm f strictly prefers w to a currently hired worker, or f has an empty position and would like to hire w. A matching is *stable* if it is individually rational and there are no blocking pairs.

In the context of strict preferences it is well-known that stability is a sufficient condition for Pareto efficiency. The next example illustrates that there could be significant efficiency loss in a stable matching when agents have weak preferences.

Example 1 Consider a market consisting of $n \geq 2$ workers and an equal number of firms each having one position to fill. Every agent finds those on the other side acceptable. Firms are indifferent between any two workers. Each worker w_i has a strict preference top ranking firm f_i and bottom ranking firm $f_{i-1} \pmod{n}$ for $i = 0, 1, \ldots, n-1$. Both the matching μ which assigns w_i to f_i and the matching ν which assigns each w_i to f_{i-1} $\pmod{n}$ are stable. The size of the Pareto inefficiency in the stable matching μ is n in terms of the number of affected agents and $n(n-1)$ in terms of the total steps up the preference lists of the agents. ◇

We show in Lemma 1 that if a stable matching is Pareto dominated by another matching, then the latter is necessarily stable. Since there always exists a stable matching and the model is finite, this observation proves the existence of a Pareto efficient and stable matching. We introduce notions of Pareto Improvement (PI)-cycles and PI-chains. We show in Theorem 1 that a matching is Pareto efficient if and only if it does not admit PI-cycles and PI-chains. Combining these two results, we introduce the *Efficient and Stable Efficient Algorithm (ESMA)*, which produces a stable and Pareto efficient matching in polynomial time.

We next give procedures that compute worker/firm optimal stable matchings. Our current results generalize the main finding[2] from an earlier paper (Erdil and Ergin, 2005), where we allowed for indifferences in only the firms' preferences, to weak preferences on both sides. This requires the introduction of *stable improvement chains* in addition to the generalization of *stable improvement cycles* with potentially indifferent agents. We show in Theorem 2 that a stable matching is considered to be worse than another matching from the point of view of the workers if and only if it admits a stable worker improvement cycle or chain. In addition to classifying through what moves one can improve upon a stable matching, this naturally leads to an algorithm with two different versions, namely the *Worker Optimal Stable Matching Algorithm (WOSMA)* and the *Firm Optimal Stable Matching Algorithm (FOSMA)*.

Given a weak preference profile, a *tie-breaking* is a strict preference profile that is obtained from the original profile by ordering members within each indifference class. In real-life matching markets, indifferences are typically treated by arbitrarily breaking the ties and then using the deferred acceptance procedure. The ties in indifferences can be broken directly by using a lottery (e.g., in school choice), or indirectly through forcing participants to submit strict preference listings which do not allow them to indicate indifferences. Conventional wisdom has been that it should be best, from a welfare perspective, to use the same tie-breaking rule uniformly across the members of each side

[2] Abdulkadiroğlu, Pathak and Roth (2006) report that had the *Stable Improvement Cycles Algorithm* proposed in Erdil and Ergin (2005) been used in the 2003-04 NYC High School Match, 8150 students (9.5 % of 86,049 students) would have been placed at schools they prefer more without violating anyone's priority.

of the market. Although this turns out not to be true in general (Example 2), we show in Theorem 3 that the conventional wisdom is verified in a special case of our model. More precisely, when workers have strict preferences, a stable matching is Pareto efficient if it is stable with respect to a uniform tie-breaking.

We argue that where ever centralized matching mechanisms are in use, it is possible to “take advantage” of existing indifference classes. We note that forcing agents to deliver strict rankings over alternatives they are indifferent about can lead to efficiency loss even when the welfare criteria take both sides' preferences into account. In fact, it can even be encouraged or required that preferences be expressed involving less than a certain number of indifference classes.

2 The Model

Let W and F denote disjoint finite sets of workers and firms, respectively. Let $A = W \cup F$ stand for the set of all agents. Let $q = (q_f)_{f \in F}$ where $q_f \geq 1$ denotes the number of positions that firm f would like to fill, i.e., the maximum number of workers it can hire. A **preference profile** is a vector of weak orders (complete and transitive relations) $R = (R_a)_{a \in A}$ where R_w denotes the preference of worker w over $F \cup \{\emptyset\}$ and R_f denotes the preference of firm f over $W \cup \{\emptyset\}$. Being matched to the empty set is interpreted as not being employed (for a worker) or keeping an empty position (for a firm). Let P_a and I_a denote the antisymmetric and symmetric parts of R_a, respectively. Throughout, we will assume that there is no worker w and firm f such that $w I_f \emptyset$ or $f I_w \emptyset$. We will call this the **no indifference to the empty set (NI$\emptyset$)** assumption. A worker w is said to be **acceptable** to firm f if $w P_f \emptyset$; similarly a firm f is acceptable to worker w if $f P_w \emptyset$. A preference profile $R = (R_a)_{a \in A}$ is **strict** if R_a is anti-symmetric for each $a \in A$.

A **matching** is a function $\mu \colon W \to F \cup \{\emptyset\}$ such that $|\mu^{-1}(f)| \leq q_f$ for each $f \in F$. A matching μ is **individually rational** if $\mu(w) R_w \emptyset$ for each worker w; and $v R_f \emptyset$ for each $v \in \mu^{-1}(f)$ and firm f. Given a matching μ, a worker firm pair (w, f) is said to form a **blocking pair** if (i) $f P_w \mu(w)$, and (ii) $w P_f v$ for some $v \in \mu^{-1}(f)$, or $|\mu^{-1}(f)| < q_f$ and $w P_f \emptyset$. A matching μ is **stable** if it is individually rational and if there is no blocking pair. Our definition of a blocking pair requires both sides to strictly prefer each other to their current matches. It is easy to see that if we relax this condition to allow one side to be indifferent, stable matchings may fail to exist.

Note that our model specifies firms' preferences only over workers. This preference information is enough to check for stability of a given matching. However in order to conduct welfare analysis, we also need to specify how firms rank *sets* of workers. Given a firm f, a preference $\tilde{R}_f$ over 2^W is **responsive** (Roth 1985) if it is complete, transitive, and for any $I, J, K \subset W$ where $I \cap K = J \cap K = \emptyset$ and $|I|, |J| \leq 1$:

$$(I \cup K)\ \tilde{R}_f\ (J \cup K) \iff I\, \tilde{R}_f\, J.$$

Responsiveness can be thought of as relating preferences over sets of workers to preferences over individual workers in a natural way.[3] We will extend the preference R_f over $W \cup \{\emptyset\}$, to a reflexive and transitive (but typically incomplete) preference over 2^W by: $I R_f J$ if and only if $I \tilde{R}_f J$ for any responsive extension $\tilde{R}_f$ of R_f.[4] It is straightforward to verify that $I R_f J$ if and only if the sets I and J can be indexed as $I : i_1, \ldots, i_n$ and $J : j_1, \ldots, j_n$, where for each worker short of n a copy of $\emptyset$ is written and $i_t R_f j_t$ for each $t \in \{1, \ldots, n\}$.

We define the partial orders $\geq_W$, $\geq_F$ and $\geq_A$ on the set of matchings as follows. Let $\mu \geq_W \nu$, if $\mu(w) R_w \nu(w)$ for each $w \in W$; let $\mu \geq_F \nu$, if $\mu^{-1}(f) R_f \nu^{-1}(f)$ for each $f \in F$; and let $\mu \geq_A \nu$ if $\mu \geq_W \nu$ and $\mu \geq_F \nu$. Let $\sim_W$, $\sim_F$, and $\sim_A$ denote the symmetric parts, whereas $>_W$, $>_F$, and $>_A$ denote the asymmetric parts of these relations. A matching μ **Pareto dominates** ν if $\mu >_A \nu$. This is equivalent to the requirement that all workers and firms weakly prefer μ to ν, and at least one worker or a firm strictly prefers μ to ν. A matching is **Pareto efficient** if it is not Pareto dominated by any other matching. A stable matching μ is called **W-optimal** if there is no stable matching ν such that $\nu >_W \mu$. Similarly a stable matching μ is called **F-optimal** if there is no stable matching ν such that $\nu >_F \mu$.

Gale and Shapley (1962) described an algorithm, which is polynomial-time in the number of workers and firms, that yields a stable matching for a strict preference profile R. This is known as the worker proposing **deferred acceptance (DA)** algorithm:

[3] It basically says that between two matchings that differ in only one position from the perspective of a firm, that firm [weakly] prefers the matching treating that position in its [weakly] preferred manner.

[4] An extension is naturally defined as: the preference $\tilde{R}_f$ is an **extension** of R_f if (i) for any $w, v \in W$: $\{w\}\tilde{R}_f\{v\}$ if and only if $wR_f v$ and (ii) for any $w \in W$, $\{w\}\tilde{R}_f\emptyset$ if and only if $wR_f\emptyset$.

At the first step, every worker applies to her favorite acceptable firm. For each firm f, q_f most preferred acceptable applicants (or all if there are fewer than q_f) are placed on the waiting list of f, and the others are rejected.

At the kth step, those applicants who were rejected at step $k-1$ apply to their next best acceptable firms. For each firm f, the most preferred acceptable q_f workers among the new applicants and those in the waiting list are placed on the new waiting list and the rest are rejected.

The algorithm terminates when every worker is either on a waiting list or has been rejected by every firm that is acceptable to her. After this procedure ends, firms admit workers on their waiting lists which yields the desired matching. The firm proposing deferred acceptance algorithm is defined analogously by changing the roles of workers and firms, with the modification that each firm f may simultaneously apply up to q_f workers at each step. When R is strict, $DA_W(R)$ and $DA_F(R)$ denote the outcome of the worker and firm proposing DA algorithms, respectively.

Theorem *(Gale and Shapley, 1962) When preferences are strict, the worker [firm] proposing deferred acceptance algorithm returns the unique worker [firm] optimal stable matching.*

Note that the DA algorithm is not well-defined when the preference profile R is not strict and the above theorem does not hold when weak preferences are allowed. However, in situations involving indifferences, the above algorithm is employed after the ties are exogenously broken. Since a matching that is stable with respect to a tie-breaking R' of R is also stable with respect to R, an immediate corollary of the above theorem is that there always exists a stable matching in our model.

Corollary 1 *There exists a stable matching.*

Even though Gale and Shapley's result guarantees the existence of a stable matching, it does not say much about how to find a worker optimal or a firm optimal stable matching. Again, in contrast with the case of strict preferences, a stable matching is not necessarily Pareto efficient in the presence of indifferences, as shown in Example 1.

3 Pareto Efficient and Stable Matchings

We start by noting that a matching must be stable if every agent weakly prefers it to some stable matching.

Lemma 1 *If $\nu \geq_A \mu$ for some stable matching μ, then ν is also stable.*

Proof. Since μ is individually rational and every agent is weakly better-off at ν, ν is also individually rational. Next consider any worker firm pair (w, f). There exist enumerations $\nu^{-1}(f) : i_1, \ldots, i_{q_f}$ and $\mu^{-1}(f) : j_1, \ldots, j_{q_f}$ such that for each worker short of q_f a copy of $\emptyset$ is inserted and $i_t R_f j_t$ for $t \in \{1, \ldots, q_f\}$. If (w, f) is a blocking pair for ν, then $f P_w \nu(w)$ and $w P_f i_t$ for some $t \in \{1, \ldots, q_f\}$. But then $f P_w \nu(w) R_w \mu(w)$ and $w P_f i_t R_f j_t$ implying that (w, f) is a blocking pair for μ, a contradiction. □

Lemma 1 implies that a stable matching that is not Pareto efficient is Pareto dominated by a stable matching. Therefore, starting from an arbitrary stable matching, it is possible to reach a Pareto efficient and stable matching through a finite sequence of Pareto improving stable matchings.

Corollary 2 *There exists a stable and Pareto efficient matching.*

The argument behind Corollary 2 suggests a constructive method to find a Pareto efficient and stable matching. However the argument does not explicitly specify (1) how to check whether a given stable matching μ is Pareto efficient, and (2) if not, how to find a matching that Pareto dominates it. Since the model is finite, one can imagine answering these questions by comparing μ exhaustively to every other matching. However such an approach is computationally infeasible, since the number of matchings grows exponentially in $\min\{|W|, |F|\}$. In order to produce a stable and Pareto efficient matching for a centralized matching market, it is therefore necessary to provide polynomial time methods to answer these questions.

Given a preference profile R and a matching μ, we will next introduce and discuss two tests: the existence of *Pareto improvement cycles* and the existence of *Pareto improvement chains*. The existence of these cycles or chains will immediately imply that μ is not Pareto efficient. Conversely we will prove in Theorem 1, that if such cycles or chains do not exist for a stable matching μ, then μ is Pareto efficient. We will use these findings to describe a polynomial time method for producing a stable and efficient matching.

Definition 1 A **Pareto improvement (PI) cycle** consists of distinct workers $w_1, \ldots, w_n \equiv w_0$ ($n \geq 2$) such that:

(i) Each w_t is matched to some firm,

(ii) $\mu(w_{t+1}) R_{w_t} \mu(w_t)$ and $w_t R_{\mu(w_{t+1})} w_{t+1}$ for $t \in \{0, 1, \ldots, n-1\}$,

(iii) At least one of the preferences in (ii) is strict for some $t \in \{0, 1, \ldots, n-1\}$.

Each worker w_t in a PI-cycle *weakly desires* the position of the following worker w_{t+1}, and the employer $\mu(w_{t+1})$ of the latter would not mind replacing w_{t+1} with w_t. Moreover, at least one worker strictly *envies* the following worker or at least one firm $\mu(w_{t+1})$ prefers w_t to w_{t+1}. If there is a PI-cycle, then the matching μ can be Pareto improved, where the Pareto dominating matching μ' is obtained by letting each worker move into the firm of the next worker:

$$\mu'(w) = \begin{cases} \mu(w_{t+1}) & \text{if } w = w_t \text{ for some } t \in \{0, \ldots, n-1\}, \\ \mu(w) & \text{otherwise.} \end{cases}$$

Definition 2 A **Pareto improvement (PI) chain** consists of distinct workers $w_1, \ldots, w_n$ ($n \geq 2$) and a firm f with an empty position such that:

(i) a. w_1 is unmatched,

b. w_t is matched with some firm for $t \in \{2, \ldots n\}$,

(ii) a. $\mu(w_{t+1}) R_{w_t} \mu(w_t)$ and $w_t R_{\mu(w_{t+1})} w_{t+1}$ for $t \in \{1, \ldots, n-1\}$.

b. $f R_{w_n} \mu(w_n)$ and $w_n R_f \emptyset$.

Each worker w_t in a PI-chain except w_n, weakly envies the following worker w_{t+1}, and as in a PI-cycle, the employer $\mu(w_{t+1})$ of the latter would not mind replacing w_{t+1} with w_t. The last worker w_n weakly desires the empty position of f and is acceptable to f. Note that if there is a PI-chain, then the matching μ can be Pareto improved, where the Pareto dominating matching μ' is obtained by letting each worker other than w_n move into the firm of the next worker and letting w_n move to f:

$$\mu'(w) = \begin{cases} \mu(w_{t+1}) & \text{if } w = w_t \text{ for some } t \in \{1, \ldots, n-1\}, \\ f & \text{if } w = w_n, \\ \mu(w) & \text{otherwise.} \end{cases}$$

By **carrying out a PI-cycle or a PI-chain**, we mean constructing the new matching μ' which Pareto dominates μ as in above.[5] Our next theorem proves a converse to the above observations: if μ is stable and there are no PI-cycles nor PI-chains, then we can conclude that μ is Pareto efficient.

A directed graph $G = (V, E)$ consists of a set V of vertices and a set E of directed edges, where a directed edge is an ordered pair of vertices, i.e., an element of the cartesian product $V \times V$. The word 'directed' will be omitted throughout the text. We will write an edge (x, y) as $x \to y$ as we will visualize the vertices as nodes, and the edges as arrows between these nodes. A **directed cycle** in G consists of distinct vertices $x_0, \ldots, x_{n-1}$ ($n \geq 2$) such that $x_0 \to x_1 \to \cdots \to x_{n-1} \to x_n \equiv x_0$.[6] We will simply refer to these as 'cycles' for the rest of the text unless we prefer to emphasize the directed structure. Note that if each vertex has exactly one arriving and one leaving edge, then each edge is part of a cycle.

Theorem 1 *A stable matching is Pareto efficient if and only if it does not admit PI-cycles nor PI-chains.*

Proof. It only remains to prove the "if" part. Assume that μ is stable but not Pareto efficient and let ν be a matching that Pareto dominates μ. Then by NI$\emptyset$, every worker matched at μ is matched at ν, and each firm is matched with at least as many

[5]Note that in the definition of a PI-chain, we do not need to require that at least some of the preferences in (ii) is strict, since the NI$\emptyset$ assumption guarantees that $\mu(w_2) P_{w_1} w_1 = \mu(w_1)$ and $w_n P_f \emptyset$. Moreover the requirement that w_1 is not matched is crucial for μ' to Pareto dominate μ, because otherwise w_1's employer could be worse-off at μ'. Note also that if the matching μ is stable, then in part (ii) of the definition of a PI-cycle and part (ii.a) of the definition of a PI-chain, at least one of the preferences should be an indifference for each t. Similarly in part (ii.b) of the definition of a PI-chain, we must have $f I_{w_n} \mu(w_n)$.

[6]Note that according to our definition of digraph, we allow for self pointing edges $x \to x$, but do not call them cycles.

workers at ν as it is matched at μ. Let $W' = \{w \in W \mid \mu(w) \neq \nu(w)\}$ and note that by NI$\emptyset$, each worker in W' is matched to a firm at ν. For each firm f fix enumerations $\nu^{-1}(f) : i_1^f, \ldots, i_{q_f}^f$ and $\mu^{-1}(f) : j_1^f, \ldots, j_{q_f}^f$ such that (1) for each worker short of q_f a copy of $\emptyset$ is inserted, (2) $\nu(j_t^f) = f \Rightarrow i_t^f = j_t^f$, and (3) $i_t^f R_f j_t^f$, for $t \in \{1, \ldots, q_f\}$. Construct a digraph G with the vertex set W' as follows. For any $w \in W'$, consider the unique t such that $w = i_t^{\nu(w)}$, and let $w \to j_t^{\nu(w)}$ if $j_t^{\nu(w)} \neq \emptyset$. Note that if $w \to v$ then $\mu(v) R_w \mu(w)$, and $w R_{\mu(v)} v$. Call an edge of G *strict* if one of these preferences is strict and denote a strict edge by $w \twoheadrightarrow v$.

If there is no extra worker matched at ν, each firm must be matched with the same number of workers in μ and ν. In particular each vertex in G has exactly one leaving edge and one arriving edge. Therefore each edge in this digraph must be part of a cycle. Since ν Pareto dominates μ, G must have a strict edge. In particular each strict edge is part of a cycle, leading to a PI-cycle.

If there is a worker w_1 who is matched at ν but not at μ, then by NI$\emptyset$ and stability of μ, $\nu(w_1)$ can not have an empty position at μ. Therefore there exists a worker w_2 such that $w_1 \twoheadrightarrow w_2$. Then either w_2 moved to a firm with an empty position at μ or there is a worker w_3 such that $w_2 \to w_3$. In the first case, w_1, w_2, and $\nu(w_2)$ form a PI-chain. In the second case, w_3 must have moved to a firm which had an empty position at μ, or there is a worker w_4 such that $w_3 \to w_4$. In the first case, w_1, w_2, w_3, and $\nu(w_3)$ form a PI-chain. Proceeding analogously, we find a PI-chain in at most $|W'|$ steps. □

The above theorem naturally suggests an algorithm which returns a stable and Pareto efficient matching: First obtain a stable matching by applying the DA algorithm to a tie-breaking. So long as the matching is not Pareto efficient, by Theorem 1, there will be a PI-cycle or a PI-chain. If so, find one and carry it out to obtain a Pareto improving matching. Since the original matching is stable, the new matching continues to be stable by Lemma 1. Repeat this as long as the obtained matching has a PI-cycle or a PI-chain. A more precise description can be found in Appendix A.2.

By finiteness of our model one can not keep Pareto improving indefinitely, hence the procedure will stop after finitely many steps and yield a Pareto efficient matching. The fact that we started with a stable matching guarantees that each matching along the procedure, and in particular the final matching, is stable. We call this procedure the **Efficient and Stable Matching Algorithm (ESMA)**. We show in Proposition 1 in

Appendix A.2, that the ESMA is polynomial in the number of workers and the total number of positions.

4 Worker-Optimal Stable Matchings

We next turn to the question of how to compute W-optimal stable matchings. Let μ be a stable matching for some fixed R. We will say that a worker w **weakly [strictly] desires** firm f if $\mu(w) \neq f$ and she weakly [strictly] prefers f to her match at μ, that is $fR_w\mu(w)$ [$fP_w\mu(w)$]. Let D_f^μ denote the set of workers who weakly desire f and are acceptable to f, such that there is no other worker who strictly desires f and ranks strictly higher in R_f. Clearly D_f^μ depends on R, too, but for notational simplicity we suppress the dependence of D_f^μ on the preference profile.

Definition 3 A **stable worker improvement (SWI) cycle** consists of distinct workers $w_1, \ldots, w_n \equiv w_0$ ($n \geq 2$) such that:

(i) Each w_t is matched to some firm,

(ii) $w_t \in D^\mu_{\mu(w_{t+1})}$ for each $t \in \{0, \ldots, n-1\}$,

(iii) $\mu(w_{t+1})P_{w_t}\mu(w_t)$ for some $t \in \{0, 1, \ldots, n-1\}$.

Each worker w_t in an SWI-cycle weakly desires the firm of the following worker w_{t+1} and the employer of the latter $\mu(w_{t+1})$ finds w_{t+1} acceptable.[7] There is also no other worker who strictly desires $\mu(w_{t+1})$ and is ranked strictly higher than w_t by $\mu(w_{t+1})$.

If μ is a stable matching which admits an SWI-cycle, then it can be improved from the workers' perspective, to another stable matching μ', obtained by letting each worker move into the firm of the next worker:

$$\mu'(w) = \begin{cases} \mu(w_{t+1}) & \text{if } w = w_t \text{ for some } t \in \{0, \ldots, n-1\}, \\ \mu(w) & \text{otherwise.} \end{cases}$$

Note that although workers improve from μ to μ', firms may become worse-off in the transition, since unlike in a PI-cycle, in an SWI-cycle we do not require that $w_tR_{\mu(w_{t+1})}w_{t+1}$.

[7] Note that the definition does not rule out the possibility that some workers in a cycle are matched with the same firm, but no two consecutive workers are.

Hence we can not make use of Lemma 1 to conclude that μ' is stable. Instead condition (ii) is key in guaranteeing that the new matching μ' continues to be stable. Suppose for instance that (w, f) were a blocking pair for μ'. Then $f = \mu(w_t)$ for some t. Because otherwise the set of workers matched with f would be the same at μ and μ', and since each worker is weakly better off at the latter matching, (w, f) would form a blocking pair at μ, too. On the other hand, for (w, f) to be a blocking pair at μ', one must have w desiring f at μ' and hence at μ. Stability of μ and condition (ii) imply that whoever moved to f via the SWI-cycle generating μ' from μ is weakly preferred to w by f, contradicting with (w, f) being a blocking pair.

Definition 4 A **stable worker improvement (SWI) chain** consists of distinct workers $w_1, \ldots, w_n$ $(n \geq 2)$ and a firm f with an empty position such that:

(i) a. If w_1 is matched to a firm, then there is no worker who strictly desires and is acceptable to $\mu(w_1)$,

b. w_t is matched to some firm for each $t \in \{2, \ldots, n\}$,

(ii) $w_t \in D^{\mu}_{\mu(w_{t+1})}$ for each $t \in \{1, \ldots, n-1\}$, and $w_n \in D^{\mu}_{f}$,

(iii) $\mu(w_{t+1}) P_{w_t} \mu(w_t)$ for some $t \in \{1, \ldots, n-1\}$.

Each worker w_t in an SWI-chain except w_n, weakly desires the firm of the following worker w_{t+1}, and as in an SWI-cycle, the employer $\mu(w_{t+1})$ of the latter finds w_t acceptable. The last worker w_n weakly desires and is acceptable to f. Also, there is no other worker who strictly desires $\mu(w_{t+1})$ and is ranked strictly higher than w_t by $\mu(w_{t+1})$.

If there is an SWI-chain, then the matching μ can be improved from the workers' perspective, to a new stable matching μ' obtained by letting each worker other than w_n move into the firm of the next worker and letting w_n move to f:

$$\mu'(w) = \begin{cases} \mu(w_{t+1}) & \text{if } w = w_t \text{ for some } t \in \{1, \ldots, n-1\}, \\ f & \text{if } w = w_n, \\ \mu(w) & \text{otherwise.} \end{cases}$$

As in an SWI-cycle, although workers improve from μ to μ', firms may become worse-off in the transition and we can again not make use of Lemma 1 to conclude that μ' is stable. Here again, condition (ii) plays the analogous key role in guaranteeing that μ' is stable.

In the same vein as in the previous section, by **carrying out an SWI-cycle or an SWI-chain**, we mean constructing the new stable matching μ' which improves μ from the workers' perspective, as done above. The next theorem proves a converse to the above observations: if μ is stable and admits no SWI-cycles nor SWI-chains, then we can conclude that μ is W-optimal. The proof can be found in Appendix A.1.

Theorem 2 *A stable matching μ is W-optimal if and only if there are no SWI-cycles nor SWI-chains.*

The above theorem naturally leads to an algorithm, which returns a W-optimal stable matching: First obtain a stable matching by applying the DA algorithm to a tie-breaking. So long as the stable matching is not W-optimal, by Theorem 2, there will be an SWI-cycle or an SWI-chain. If that is the case, find an SWI-cycle or an SWI-chain and carry it out to obtain a new stable matching that improves the original one from the workers' perspective. Repeat this as long as the the obtained stable matching has an SWI-cycle or SWI-chain. A precise description of this procedure is in Appendix A.2.

By finiteness of our model, the procedure will stop after finitely many steps and yield a W-optimal stable matching. We call this procedure the **Worker Optimal Stable Matching Algorithm (WOSMA)**. We show in Proposition 2 in Appendix A.2, that the WOSMA is polynomial in the number of workers and the number of firms.

Suppose that a stable matching μ admits an SWI-chain $w_1, \ldots, w_n$ and f. Then at μ, f has an empty position and w_n weakly desires f (in particular w_n is not matched to f). The worker w_n must be indifferent between $\mu(w_n)$ and f for otherwise (w_n, f) would form a blocking pair for μ. In particular if μ is stable and if the workers have strict preferences, then μ does not admit any SWI-chains. Moreover if workers have strict preferences and μ admits an SWI-cycle, then each worker strictly envies the following worker in that SWI-cycle.

Definition 5 A **SWI*-cycle** is an SWI-cycle $w_1, \ldots, w_n \equiv w_0$ where $\mu(w_{t+1}) P_{w_t} \mu(w_t)$ for each $t \in \{0, 1, \ldots, n-1\}$.

As a result, in the special case where workers have strict preferences, Theorem 2 has the following corollary.

Corollary 3 *Suppose that workers have strict preferences. Then a stable matching μ is W-optimal if and only if there are no SWI*-cycles.*

Theorem 1 in Erdil and Ergin (2005) is slightly stronger than Corollary 3. The former says that, if workers have strict preferences, and μ and ν are stable matchings such that $\nu >_W \mu$, then there exist stable matchings $\mu_1, \ldots, \mu_n$ such that $\mu = \mu_1$, $\nu = \mu_n$, and μ_{t+1} is obtained by carrying out an SWI* cycle at μ_t, for $t \in \{1, \ldots, n-1\}$. That is, if the stable matching ν does better than another stable matching μ from the point of view of the workers, then ν can be reached from μ by a sequence of SWI*-cycles.[8] An analogue of this result is not true in our framework where both sides have weak preferences, i.e., if μ and ν are stable matchings such that $\nu >_W \mu$, then ν may not be reached from μ by a sequence of SWI-cycles and SWI-chains. An intuition behind this is that there may be moves from one matching to the other that do not effect any agent's welfare and it may be impossible to recover such moves by improving chains or cycles.

5 Uniform Tie-breaking and Pareto Efficiency

The conventional wisdom was that the efficiency loss arising from tie-breaking in matching markets with indifferences has to do with tie-breaking rules across different lists being different. When only one side of the market's preferences constitute the welfare criteria (e.g., students in school choice), it has been noted that even a *uniform* tie-breaking would not deliver optimal outcomes. In Erdil and Ergin (2005) we addressed that issue and offered a possible remedy. And in this paper we noted that even when both sides' preferences are taken into account for efficiency considerations, stability at a uniform tie-breaking is still not sufficient to ensure Pareto efficiency.

Example 2 Let $W = \{w, v\}, F = \{f, g\}$, and $q_f = q_g = 1$. Assume that each side finds those on the other side acceptable. R is given by:

[8] We use this in proving Theorem 3 below.

R_w	R_v
f	f, g
g	

R_f	R_g
w, v	w, v

The stable matching (wg, vf) is stable with respect to the uniform tie-breaking that favors f over g and v over w, yet it is not Pareto efficient at R. ◇

In the above example, note that had the ties been broken favoring w over v, still uniformly that is, a Pareto efficient matching would be stable. On the other hand, it may not be possible to reach all stable and Pareto efficient matchings by focusing on uniform tie-breaking rules as the following example demonstrates.

Example 3 Let $W = \{w, v\}, F = \{f, g\}$, and $q_f = q_g = 2$. Assume that each side finds those on the other side acceptable. R is given as:

R_w	R_v
f, g	f, g

R_f	R_g
w	w
v	v

Then the stable matching $\mu = (wf, vg)$ is not stable at any uniform tie-breaking. This also shows that there may be Pareto efficient (also W-optimal) stable matchings which can not be reached by using the (worker proposing) DA algorithm after all possible ways of uniform tie-breaking. ◇

A profile $R' \in \mathcal{T}(R)$ is called a **uniform** tie-breaking of R, if members of the same side resolve their indifferences in the same way, i.e., if there exist bijections $\phi^W : W \to \{1, \ldots, |W|\}$ and $\phi^F : F \to \{1, \ldots, |F|\}$ such that:

$$f \, I_w \, g \implies [f \, R'_w \, g \iff \phi^F(f) \leq \phi^F(g)],$$

and

$$w \, I_f \, v \implies [w \, R'_f \, v \iff \phi^W(w) \leq \phi^W(v)],$$

for all $w, v \in W$ and $f, g \in F$.

Theorem 3 *Assume that workers have strict preferences at R and that there exists a uniform tie-breaking R' such that μ is stable with respect to R'. Then μ is Pareto efficient at R.*

Proof. Let $\phi^W : W \to \{1, \ldots, |W|\}$ be a bijection that induces the tie-breaking R'_F. Suppose for a contradiction that μ is Pareto dominated by a matching ν at R. Since the workers have strict preferences, their being indifferent between μ and ν would imply that $\mu = \nu$, therefore some worker(s) must strictly prefer ν to μ at R. We also know by Lemma 1 that ν is stable, therefore μ is not W-optimal at R.

By Theorem 1 in Erdil and Ergin (2005), there exist stable matchings $\mu_1, \ldots, \mu_n$ such that $\mu = \mu_1$, $\nu = \mu_n$, and μ_{t+1} is obtained by carrying out an SWI*-cycle at μ_t, for $t \in \{1, \ldots, n-1\}$. Note that $\mu_t \geq_F \mu_{t+1}$, for otherwise if the tth SWI*-cycle rematches a firm f to a worker w such that $wP_f w'$ for some $w' \in \mu_t^{-1}(f)$, then (w, f) would block μ_t, a contradiction. Hence $\mu = \mu_1 \geq_F \mu_2 \geq_F \cdots \geq_F \mu_n = \nu$. We also have that $\nu \geq_F \mu$ since ν Pareto dominates μ, therefore $\mu = \mu_1 \sim_F \mu_2 \sim_F \cdots \sim_F \mu_n = \nu$.

Let $w_1, \ldots, w_n$ be the SWI-cycle at $\mu = \mu_1$ above. Then $\mu(w_{t+1}) P_{w_t} \mu(w_t)$ and $w_t I_{\mu(w_t)} w_{t+1}$ for $t \in \{0, 1, \ldots, n-1\}$, which, by the definition of the uniform tie-breaking R'_F, implies that $\phi^W(w_0) < \phi^W(w_1) < \cdots < \phi^W(w_{n-1}) < \phi^W(w_n) = \phi^W(w_0)$, a contradiction. □

A directed graph G is **acyclic** if it has no cycles. A **topological ordering** of a directed graph is a bijection $\phi : X \to \{1, \ldots, |X|\}$ such that $x \to y$ implies that $\phi(x) \geq \phi(y)$. It is not hard to see that a digraph is acyclic if and only if it is topologically ordered.

Theorem 4 *If each firm has one position and μ is stable and Pareto efficient at R, then there exists a uniform tie-breaking R' such that μ is stable with respect to R'.*

Proof. Assume that μ is Pareto efficient and stable at R. We will construct the tie breaking R' in two steps, by first breaking the ties in firms' preferences and then those in workers' preferences.

Consider a directed graph G with vertex set W, where $w \to v$ if v is matched to a firm, $\mu(v) P_w \mu(w)$, and $w I_{\mu(v)} v$. Such an edge means that w strictly envies v, and the firm $\mu(v)$ would not mind replacing v with w. Pareto efficiency of μ implies that this graph is acyclic: If the graph has a cycle $w_0 \to w_1 \to \cdots \to w_{n-1} \to w_n \equiv w_0$, then the new matching obtained by rematching each worker w_t in the cycle to $\mu(w_{t+1})$ for $t \in \{0, \ldots, n-1\}$, would Pareto dominate μ. Let $\phi^W : W \to \{1, \ldots, |W|\}$ be a bijection inducing a topological ordering of G. Let R'_F denote the uniform tie-breaking of R_F

induced by ϕ^W.

By NI$\emptyset$ and individual rationality of μ before the tie-breaking, μ continues to be individually rational after the tie-breaking. Suppose that (w, f) blocks μ at (R_W, R'_F), i.e. $fP_w\mu(w)$ and $wP'_f\mu^{-1}(f)$. Stability of μ at R implies that $wI_f\mu^{-1}(f)$, in particular by NI$\emptyset$, f is matched to a worker v. Note that $w \to v$ since at R, w strictly envies v, and the firm $f = \mu(v)$ would not mind replacing v with w. Hence $\phi^W(w) \geq \phi^W(v)$, a contradiction to wI_fv and wP'_fv. We conclude that μ is stable at (R_W, R'_F).

Next consider an analogous directed graph G' with vertex set F, where $f \to g$ if g is matched to a worker, $\mu^{-1}(g)P'_f\mu^{-1}(f)$, and $fI_{\mu^{-1}(g)}g$. Suppose that there is a cycle $f_0 \to f_1 \to \cdots \to f_{n-1} \to f_n \equiv f_0$. Consider the new matching μ' obtained by rematching each firm f_t in the cycle to $\mu^{-1}(f_{t+1})$ for $t \in \{0, \ldots, n-1\}$. At (R_W, R'_F), all workers, as well as the firms not involved in the cycle are indifferent between μ and μ', whereas all the firms involved in the cycle strictly prefer μ' to μ. No firm strictly prefers μ' to μ at R, since otherwise μ' would Pareto dominate μ at R. Since $\mu^{-1}(f_{t+1})P'_{f_t}\mu^{-1}(w_t)$ and $\mu^{-1}(f_{t+1})I_{f_t}\mu^{-1}(w_t)$ for $t \in \{0, \ldots, n-1\}$, by the definition of the uniform tie-breaking R'_F, we have $\phi^W(\mu^{-1}(f_0)) = \phi^W(\mu^{-1}(f_n)) < \phi^W(\mu^{-1}(f_{n-1})) < \cdots < \phi^W(\mu^{-1}(f_1)) < \phi^W(\mu^{-1}(f_0))$, a contradiction. Therefore G' is acyclic, let $\phi^F : F \to \{1, \ldots, |F|\}$ be a bijection inducing a topological ordering of G'. Let R'_W be the uniform tie-breaking of R_W induced by ϕ^F. By the same argument as in the above paragraph switching the roles of firms and workers, we conclude that μ is stable with respect to (R'_W, R'_F). $\square$

Corollary 4 *Assume that each firm has one position and one side has strict preferences at R. Then μ is stable and Pareto efficient at R if and only if there exists a uniform tie-breaking R' such that μ is stable at R'.*

A Appendix

A.1 Proof of Theorem 2

Apart from SWI-chains and SWI-cycles, it will be useful for the purposes of the proof to consider chains and cycles that do not change any worker's welfare.

Definition 6 Given two matchings μ and ν, **a reversible cycle from μ to ν** consists of distinct workers $w_1, \ldots, w_n \equiv w_0$ $(n \geq 2)$ such that:

(i) Each w_t is matched to some firm both at μ and ν,

(ii) $\nu(w_t) = \mu(w_{t+1}) \neq \mu(w_t)$ for $t \in \{0, 1, \ldots, n-1\}$,

(iii) $\mu(w_t) I_{w_t} \nu(w_t)$ for $t \in \{1, \ldots, n\}$.

Definition 7 Given two matchings μ and ν, **a reversible chain from** μ **to** ν consists of distinct workers $w_1, \ldots, w_n$ $(n \geq 1)$ and a firm f with an empty position at μ such that:

(i) a. Each w_t is matched to some firm both at μ and ν,

b. $\mu(w_1)$ has an empty position at ν,

(ii) $\nu(w_n) = f$ and $\nu(w_t) = \mu(w_{t+1}) \neq \mu(w_t)$ for $t \in \{1, \ldots, n-1\}$,

(iii) $\mu(w_t) I_{w_t} \nu(w_t)$ for $t \in \{1, \ldots, n\}$.

If there is a reversible cycle [chain] from μ to ν, to **reverse** such a cycle [chain] will mean replacing ν with ν' by simply reassigning the workers who are involved in the cycle [chain] back to their firms at μ, i.e.,

$$\nu'(w) = \begin{cases} \mu(w) & \text{if } w \text{ is involved in the reversible cycle [chain]} \\ \nu(w) & \text{otherwise.} \end{cases}$$

Clearly, the reversing process does not effect the welfare of the workers.

Lemma 2 *Assume that μ and ν are stable matchings such that $\nu \geq_W \mu$. If ν' is obtained by reversing a reversible cycle or chain from μ to ν, then ν' is also stable.*

Proof. Let μ, ν, and ν' be as in above. Take any firm f and worker w such that $f = \nu'(w)$. Then by the definition of ν', $f = \nu(w)$ or $f = \mu(w)$. Since both μ and ν are individually rational, i.e., $f R_w \emptyset$ and $w R_f \emptyset$, ν' is individually rational, too.

Suppose for a contradiction that (w, f) is a blocking pair for ν'. Then (i) $f P_w \nu'(w)$ and (ii.a) $w P_f v$ for some $v \in \nu'^{-1}(f)$, or (ii.b) $w P_f \emptyset$ and f has an empty position at ν'. Since $\nu' \sim_W \nu \geq_W \mu$, we have (i.$\nu$) $f P_w \nu(w)$ and (i.μ) $f P_w \mu(w)$. In case (ii.a), f is matched to v at ν or μ, which along with (i.ν) and (i.μ) imply that ν or μ is unstable, a contradiction. In case (ii.b), f has an empty position at ν or μ, which along with (i.ν) and (i.μ) imply that ν or μ is unstable, again a contradiction. □

If μ is a stable matching that is not W-optimal, then there exists a stable matching ν^0 such that $\nu^0 >_W \mu$. If there are any reversible cycles or chains from μ to ν^0, by Lemma 2, we can arbitrarily select one and reverse it to obtain a new stable matching ν^1 such that $\nu^1 \sim_W \nu^0 >_W \mu$. If there exist any reversible chains or cycles from μ to ν^1, by Lemma 2, we can again arbitrarily select one and reverse it to obtain a yet another stable matching ν^2 such that $\nu^2 \sim_W \nu^1 \sim_W \nu^0 >_W \mu$. Proceeding analogously, we will eventually obtain a stable matching ν such that $\nu >_W \mu$ and there are no reversible cycles or chains from μ to ν. We summarize this observation in the following Lemma.

Lemma 3 *If μ is a stable matching that is not W-optimal, then there exists a stable matching ν such that $\nu >_W \mu$ and there are no reversible cycles nor chains from μ to ν.*

Lemma 4 *Let μ be a stable matching and ν be an individually rational matching such that $\nu >_W \mu$. Assume that μ does not admit an SWI-cycle nor an SWI-chain, and that there are no reversible cycles nor chains from μ to ν. Then each firm f is matched to at least as many workers at μ as at ν.*

Proof. Let $W' = \{w \in W : \mu(w) \neq \nu(w)\}$. For each firm f, if there exists a worker $u \in W$ who strictly desires f at μ and is acceptable to f, then fix u_f to be a highest ranked such u with respect to R_f. Otherwise we will say that "u_f does not exist." If u_f does not exist and there exists $v \in W'$ such that $\nu(v) = f$, then fix v_f to be any such v. Otherwise, i.e., if u_f exists *or* if there is no $v \in W'$ such that $\nu(v) = f$, we will say that "v_f does not exist." By definition u_f and v_f can not co-exist. If u_f exists then $f P_{u_f} \mu(u_f)$ and $u_f \in D^\mu_f$. If v_f exists, then $f = \nu(v_f) \neq \mu(v_f)$, $f = \nu(v_f) I_{v_f} \mu(v_f)$, and $v_f \in D^\mu_f$. [9]

A finite sequence $(w_1, \ldots, w_n)$ of $n \geq 1$ workers is of **Type I** if (i) they are all distinct, (ii) each one is matched to some firm both at μ and ν, (iii) $\nu(w_1)$ has an empty

[9]This last inclusion uses individual rationality of ν. v_f is matched to f at ν, so she must be acceptable to f. But since there are no workers that strictly desire f and are acceptable to f (since u_f does not exist), v_f does not strictly desire f (and since $\nu >_W \mu$, v_f must be indifferent between $f = \nu(v_f)$ and $\mu(f)$).

position at μ, (iv.a) $w_1 = v_{\nu(w_1)}$, and (iv.b) $w_{t+1} = v_{\mu(w_t)}$ for $t \in \{1, \ldots, n-1\}$. Note that in a Type I sequence, $\nu(w_t) I_{w_t} \mu(w_t)$ and $w_t \in D^{\mu}_{\nu(w_t)}$ for each $t \in \{1, \ldots, n\}$.

A finite sequence $(w_1, \ldots, w_n)$ of $n \geq 2$ workers is of **Type II** if (i) they are all distinct, (ii) there exists a $k \leq n-1$ such that: (ii.a) $(w_1, \ldots, w_k)$ is of Type I, (ii.b) each one of $w_{k+1}, \ldots, w_n$ is matched to some firm at μ, and (ii.c) $w_{t+1} = u_{\mu(w_t)}$ for $t \in \{k, \ldots, n-1\}$. Note that in a Type II sequence, $\nu(w_t) I_{w_t} \mu(w_t)$ and $w_t \in D^{\mu}_{\nu(w_t)}$ for each $t \in \{1, \ldots, k\}$; and $\mu(w_{t-1}) P_{w_t} \mu(w_t)$ and $w_t \in D^{\mu}_{\mu(w_{t-1})}$ for each $t \in \{k+1, \ldots, n\}$.

We will show in step 1 below that, if there exists a Type I sequence of length $n \geq 1$, then there exists a Type I or Type II sequence of length $n+1$. We will prove in step 2 that, if there exists a Type II sequence of length $n \geq 2$, then there exists a Type II sequence of length $n+1$. The two steps imply that there can not be any Type I sequence of length one, otherwise it is possible to generate an arbitrarily large sequence of distinct workers, contradicting finiteness of W. To see that this is enough to prove the lemma, suppose that there exists a firm f who is matched to less workers at μ than at ν. Then f must have an empty position at μ. By stability of μ and NI$\emptyset$, u_f does not exist. Since f is matched to more workers at ν, v_f exists. Since $f = \nu(v_f) I_{\nu_f} \mu(\nu_f)$, by NI$\emptyset$, v_f is matched to a firm at μ. Hence (v_f) constitutes a Type I sequence of length one, a contradiction. It remains to prove steps 1 and 2.

Step 1: Let $(w_1, \ldots, w_n)$ be a Type I sequence. Then $\mu(w_n)$ does not have an empty position at ν, since otherwise $w_n, \ldots, w_1$ (yes, in the reverse order) and $\nu(w_1)$ would constitute a reversible chain from μ to ν. Since $\nu(w_n) \neq \mu(w_n)$ and the positions of $\mu(w_n)$ are full at ν, there exists a worker in W' matched to $\mu(w_n)$ at ν. Hence either $u_{\mu(w_n)}$ or $v_{\mu(w_n)}$ exists.

If $u_{\mu(w_n)}$ exists, let $w_{n+1} = u_{\mu(w_n)}$. Since $\mu(w_n) \neq \mu(u_{\mu(w_n)})$, $w_{n+1} \neq w_n$. Also w_{n+1} is distinct from $w_1, \ldots, w_{n-1}$, because otherwise if $w_{n+1} = w_k$ for some $k \leq n-1$, then $w_{n+1}, w_n, w_{n-1}, \ldots, w_{k+1}$ (yes, in this order) would constitute an SWI-cycle. Moreover, w_{n+1} must be matched to a firm at μ, since otherwise $w_{n+1}, w_n, w_{n-1}, \ldots, w_1$ and $\nu(w_1)$ would constitute an SWI-chain. Hence in this case $(w_1, \ldots, w_n, w_{n+1})$ is a Type II sequence of length $n+1$.

If $v_{\mu(w_n)}$ exists, let $w_{n+1} = v_{\mu(w_n)}$. Since $\mu(w_n) \neq \mu(v_{\mu(w_n)})$, $w_{n+1} \neq w_n$. Also w_{n+1} is distinct from $w_1, \ldots, w_{n-1}$, for otherwise if $w_{n+1} = w_k$ for some $k \leq n-1$, then $w_{n+1}, w_n, w_{n-1}, \ldots, w_{k+1}$ would constitute a reversible cycle from μ to ν. Moreover,w_{n+1}

must be matched to a firm at μ, because of the NI$\emptyset$ assumptions, her indifference between μ and ν, and her being matched with $\mu(w_n)$ at ν. Thus, in this case $(w_1, \ldots, w_n, w_{n+1})$ is a Type I sequence of length $n+1$.

Step 2: Let $w_1, \ldots, w_n$ $(n \geq 2)$ be a Type II sequence where k is as in part (ii) of the definition of a Type II sequence. There exists a worker who strictly desires $\mu(w_n)$ at μ and is acceptable to $\mu(w_n)$, because otherwise $w_n, \ldots, w_1$ and $\nu(w_1)$ would constitute an SWI-chain. Hence $u_{\mu(w_n)}$ exists.

Let $w_{n+1} = u_{\mu(w_n)}$. Since $\mu(w_n) \neq \mu(u_{\mu(w_n)})$, $w_{n+1} \neq w_n$. Also w_{n+1} is distinct from $w_1, \ldots, w_{n-1}$, because otherwise if $w_{n+1} = w_k$ for some $k \leq n-1$, then $w_{n+1}, w_n, w_{n-1}, \ldots, w_{k+1}$ would constitute an SWI-cycle. Moreover, w_{n+1} must be matched to a firm at μ, for otherwise $w_{n+1}, w_n, w_{n-1}, \ldots, w_1$ and $\nu(w_1)$ would constitute an SWI-chain. Hence in this case $(w_1, \ldots, w_n, w_{n+1})$ is a Type II sequence of length $n+1$. $\square$

Lemma 5 *Let μ be a stable matching and ν be an individually rational matching such that $\nu >_W \mu$. Assume that μ does not admit an SWI-cycle nor an SWI-chain and that there are no reversible cycles nor chains from μ to ν. Let $W' = \{w \in W : \mu(w) \neq \nu(w)\}$ and $F' = \mu(W')$. Then:*

(i) For each firm f, the number of workers in W' who are matched to firm f is the same at μ and ν. In particular, $F' = \nu(W')$.

(ii) Each worker in W' is matched to a firm in both μ and ν.

Proof. By $\nu >_W \mu$, individual rationality of μ, and NI$\emptyset$, each worker in W' is matched to a firm at ν. To see part (i), note that Lemma 4 implies that $|W' \cap \mu^{-1}(f)| \geq |W' \cap \nu^{-1}(f)|$ for any firm f. Suppose that the inequality $|W' \cap \mu^{-1}(f)| \geq |W' \cap \nu^{-1}(f)|$ holds strictly for some firm f^*. Summing across all firms we have:

$$\sum_{f \in F} |W' \cap \mu^{-1}(f)| > \sum_{f \in F} |W' \cap \nu^{-1}(f)|.$$

That is, the number of workers in W' matched to some firm atμ is more than the number of workers in W' matched to some firm at ν. This implies that there exists a worker in

W' who is unmatched at ν, a contradiction. Part (ii) follows from part (i) and the fact that each worker in W' is matched to a firm at ν. □

Proof of Theorem 2. It only remains to prove the "if" part. Assume that μ is stable but not W-optimal. By Lemma 3, there exists a stable matching ν such that $\nu >_W \mu$ and there are no reversible cycles nor chains from μ to ν. Suppose for a contradiction that μ admits no SWI-cycle nor SWI-chain. Let W' and F' be as in Lemma 5.

For any $f \in F'$, let W'_f denote the set of workers in W' who weakly desire f and are acceptable to f, such that there is no other worker in W' who strictly desires f and ranks strictly higher in R_f. By Lemma 5, $f \in F' = \nu(W')$, hence there exist workers in W' who are matched to f at ν. Those workers weakly desire f and are acceptable to f, which shows that W'_f is nonempty.

If there is any worker in W'_f who is matched to f at ν, fix w_f to be such a worker who is ranked highest with respect to R_f. If not, those who weakly desire f and are in W' can not be in a single indifference class with respect to R_f. Therefore there exists a worker in W', and hence in W'_f, who strictly desires f, and fix w_f to be any such worker. Note that if $u \in W'$ is matched to f at ν, then $w_f R_f u$. Also note that $\mu(w_f) \in F'$ and $\mu(w_f) \neq f$.

We next show that $w_f \in D^{\mu}_f$. Suppose not, then there is a worker $v \notin W'$ who strictly desires f and is strictly higher in R_f than w_f. Since $v \notin W'$, $\nu(v) = \mu(v)$, therefore $f P_v \nu(v)$. Let u be a worker in W' who is matched to f at μ, then $v P_f w_f R_f u$, a contradiction to the stability of ν.

Now, consider a directed graph G with vertex set F', where for each firm f there is a unique incoming edge given by $\mu(w_f) \to f$. Since each firm in F' is pointed to by a different firm in F', there exists a cycle $f_1, \ldots, f_n = f_0$ in F'.[10] Let $w_t = w_{f_{t+1}}$ for $t \in \{0, \ldots, n-1\}$. Since $f_t \to f_{t+1}$ and $w_t = w_{f_{t+1}}$, we have $\mu(w_t) = f_t$. In particular $w_1, \ldots, w_n$ are distinct and each one is matched to some firm at μ. By construction $w_t \in D^{\mu}_{\mu(w_{t+1})}$, and if $\mu(w_{t+1}) I_{w_t} \mu(w_t)$, then $\nu(w_t) = \mu(w_{t+1})$, for $t \in \{0, 1, \ldots, n-1\}$. Hence $w_1, \ldots, w_n$ constitute either an SWI-cycle or a reversible cycle from μ to ν, a contradiction. □

[10] Remember that our definition of a cycle in a graph requires that the vertices are distinct and $n \geq 2$.

A.2 The Algorithms and Their Time Complexity Analysis

In this section we give precise descriptions of the algorithms announced earlier. In doing so we introduce the notion of a 2-labeled graph and a strict cycle, and then establish an upper bound on the time complexity of strict cycle search on a 2-labeled graph in Lemma 6. In Propositions 1 and 2, we use this result to establish that the algorithms introduced are polynomial time.

An algorithm being of polynomial time means that the time required in order for it to return its outcome or halt is a polynomial in the size of its input. This property becomes especially important as the size of the input grows. In the problems studied in this paper, even with 100 agents, there are 100! (more than 10^{145}) different ways of uniform tie-breaking, and many more arbitrary tie breaking rules. Therefore methods of exhaustion are not computationally feasible. On the other hand, *polynomial time* is a theoretical benchmark for 'algorithmically efficient' computation. Time complexity is expressed via the big-Oh notation, as the theory is concerned with the asymptotic behavior. However, such notation can hide arbitrarily large constants, and may not always give a realistic sense of what the actual running times in practice could be. Partly to attend this issue, we conducted simulations for our earlier paper, Erdil and Ergin (2005), where the indifference classes had several hundred agents. We confirmed that on an average desktop computer, with such data set as the input, the actual running time was always at most a few minutes.

Given a directed graph $G = (V, E)$, a **path** from a vertex x to a vertex y is a sequence of distinct vertices $x_1, \ldots, x_n$ such that $x = x_1 \rightarrow x_2 \rightarrow \cdots \rightarrow x_n = y$. A directed graph is called **strongly connected** if for every pair of vertices x and y there is a path from x to y and a path from y to x. The strongly connected components of a directed graph are its maximal strongly connected subgraphs. These form a partition of the graph.

A **2-labeled graph**[11] is a graph $G = (V, E)$ and a function $\ell : E \rightarrow \{0, 1\}$. That is, each edge is assigned one of the two labels.

We will denote the edges labeled 0 with $x \rightarrow y$, and those labeled 1 with $x \twoheadrightarrow y$. The edges labeled 1 will be called **strict edges of** G. A cycle of G with at least one strict edge on is called a **strict cycle**.

[11] We restrict our attention to edge labeled graphs and assume that vertices are not labeled. It is worth noting that the notion of a *labeled graph* is different from that of *graph labeling.*

Lemma 6 *Strict cycle search on a 2-labeled graph $G = (V, E)$ is $\mathrm{O}\big(|V| + |E|\big)$.*

Proof. Note that G has a strict cycle if and only if a strongly connected component includes a strict edge. Identifying strongly connected components of G is $\mathrm{O}\big(|V| + |E|\big)$ by Tarjan (1972), checking for strict edges is $\mathrm{O}\big(|E|\big)$, and finding a cycle that includes a specific strict edge is $\mathrm{O}\big(|V|\big)$.[12] Hence strict cycle search is $\mathrm{O}\big(|V| + |E|\big)$. □

Efficient and Stable Matching Algorithm (ESMA)

Given a preference profile R, to a stable matching μ we will associate a 2-labeled graph Γ^μ with the vertex set $W \cup \{\emptyset\}$, and the edges and their labels specified as follows:

(i) $w \to v$ if $\mu(v)$ is a firm such that $\mu(v) R_w \mu(w)$ and $w R_{\mu(v)} v$.

(ii) $w \to \emptyset$ if there is a firm f with an empty position such that $f R_w \mu(w)$ and $w R_f \emptyset$.

(iii) $\emptyset \to w$ if $\mu(w) = \emptyset$.

Label the strict edges as follows:

(iv) $w \twoheadrightarrow v$ if $w \to v$ and one of the preferences in (i) is strict.

(v) $w \twoheadrightarrow \emptyset$ for each $w \to \emptyset$. [13]

Note that in Γ^μ, a strict cycle with [without] $\emptyset$ as one of its vertices, corresponds to a PI-chain [PI-cycle] at μ. Conversely any PI-chain or PI-cycle at μ corresponds to a strict cycle of Γ^μ.

In what follows, let us write Γ^k instead of Γ^{μ^k} for notational simplicity. Then the ESMA is described as:

Step 0:

Select a strict preference profile R' from $\mathcal{T}(R)$. Run the DA algorithm and obtain a temporary matching μ^0.

Step t $\geq$ 1:

$(t.a)$ Given μ^{t-1}, construct the associated 2-labeled graph Γ^{t-1}.

[12] If $x \twoheadrightarrow y$ is such an edge, we need to explore each vertex only once as we employ a depth-first search starting from y only checking whether there is an edge from the explored vertex back to x.

[13] Since the second preference in (ii) is always strict by NI$\emptyset$.

$(t.b)$ Find a strict cycle in Γ^{t-1} if there exists any, let the corresponding PI-cycle or the PI-chain take place to obtain μ^t, and go to step $(t+1.a)$. If there is no strict cycle, then return μ^{t-1} as the output of the algorithm.

Proposition 1 *The ESMA terminates in* $\mathrm{O}\big(|W|^3 \cdot Q\big)$ *time where* $Q = \sum_{f \in F} q_f$.

Proof. Each step t of the ESMA involves a strict cycle search in Γ^t which is $\mathrm{O}\big(|W \cup \{\emptyset\} + |E|\big)$, where E is the set of edges, by Lemma 6.

The DA algorithm which is conducted initially is $\mathrm{O}\big(|W| \cdot |F|\big)$, hence also $\mathrm{O}\big(|W|^3 \cdot Q\big)$ since $|F| \leq Q$. From the above paragraph, each subsequent step of the ESMA is $\mathrm{O}\big(|W|^2\big)$ since $|E| \leq (|W|+1)^2$. At each step, at least a worker or a firm improves, so these steps can be repeated at most $|W| \cdot |F|$ times in workers' favor and $|W| \cdot Q$ times in firms' favor. Hence the algorithm terminates in $\mathrm{O}\big(|W|^3 \cdot Q\big)$ time. □

Worker Optimal Stable Matching Algorithm (WOSMA)

Given a preference profile R, to a stable matching μ, let us associate a 2-labeled graph G^μ with the vertex set $W \cup \{\emptyset\}$, and the edges and their labels specified as follows:

(i) $w \rightarrow v$ if $\mu(v)$ is a firm such that $w \in D^\mu_{\mu(v)}$.

(ii) $w \rightarrow \emptyset$ if there is a firm f with an empty position such that $w \in D^\mu_f$.

(iii) $\emptyset \rightarrow w$ if $\mu(w) = \emptyset$ or there is no worker who strictly desires and is acceptable to $\mu(w)$.

Label the strict edges as follows:

(iv) $w \twoheadrightarrow v$ if $w \rightarrow v$ and $\mu(v) P_w \mu(w)$.

In G^μ a strict cycle with [without] $\emptyset$ as one of its vertices, corresponds to an SWI-chain [SWI-cycle]. Conversely any SWI-chain or SWI-cycle corresponds to a strict cycle of G^μ.

Let us write G^k instead of G^{μ^k} in what follows, for notational simplicity.

Step 0:

Select a strict preference profile R' from $\mathcal{T}(R)$. Run the DA algorithm and obtain a temporary matching μ^0.

Step t $\geq$ 1:

$(t.a)$ Given μ^{t-1}, let G^{t-1} be the associated 2-labeled graph as constructed above.

$(t.b)$ Find a strict cycle in G^{t-1}, if there exists any, let the corresponding SWI-cycle or the SWI-chain take place to obtain μ^t, and go to step $(t+1.a)$. If there is no strict cycle, then return μ^{t-1} as the output of the algorithm.

Proposition 2 *The WOSMA terminates in* $\mathrm{O}\big(|W|^3 \cdot |F|\big)$ *time.*

Proof. Each step t of the WOSMA involves a strict cycle search in G^t which is $\mathrm{O}\big(|E| + |W \cup \{\emptyset\}|\big)$ by Lemma 6.

The DA algorithm which is conducted initially is $\mathrm{O}\big(|W| \cdot |F|\big)$, hence also $\mathrm{O}\big(|W|^3 \cdot |F|\big)$. From the above paragraph, each subsequent step of the WOSMA is $\mathrm{O}\big(|W|^2\big)$ since $|E| \leq (|W|+1)^2$. At each step, at least a worker improves, so these steps can be repeated at most $|W| \cdot |F|$ times. Hence the algorithm terminates in $\mathrm{O}\big(|W|^3 \cdot |F|\big)$ time. □

References

[1] Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 2, May, 364–367.

[2] Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 2, May, 368–371.

[3] Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2006): "Stretegyproofness versus Efficiency: Redesigning the NYC High School Match," mimeo.

[4] Abdulkadiroğlu, A. and T. Sönmez (2003): "School choice: A mechanism design approach," *American Economic Review*, 93, 729–747.

[5] Bogomolnaia, A. and H. Moulin (2004): "Random Matching Under Dichotomous Preferences," *Econometrica*, 72 (1), 257–279.

[6] Cormen, T. H., C. E. Leiserson, R. L. Rivest (1997): *Introduction to Algorithms*, MIT Press, Cambridge.

[7] Erdil, A. and H. Ergin (2005): "What's the matter with tie-breaking? Improving efficiency in school choice," mimeo.

[8] Ergin, H. (2002): "Efficient resource allocation on the basis of priorities," *Econometrica*, 70, 2489–2497.

[9] Ergin, H. and T. Sönmez (2006): "Games of school choice under the Boston mechanism," *Journal of Public Economics*, 90, 215-237.

[10] Gale, D. and L. Shapley (1962): "College admissions and the stability of marriage," *American Mathematical Monthly*, 69, 9–15.

[11] Kesten, O. (2004): "Student placement to public schools in the US: Two new solutions," University of Rochester, mimeo.

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

[13] Roth, A. E. (1984): "The evolution of the labor market for medical interns and residents: A case study in game theory," *Journal of Political Economy*, 92 (6), 991–1016.

[14] Roth, A. E. (2002): "The economist as engineer: Game theory, experimentation, and computation as tools for design economics," *Econometrica*, 70, 1341–1378.

[15] Roth, A. E. (2003): "The origins, history, and design of the resident match," *Journal of the American Medical Association*, 289 (7), 909-912.

[16] Roth, A. E. and E. Peranson (1999): "The redesign of the matching market for American physicians: Some engineering aspects of economic design," *American Economic Review*, 89, 748–780.

[17] Roth, A. E. and U.G. Rothblum (1999): "Truncation strategies in matching markets –In search for advice for participants," *Econometrica*, 67, 21–43.

[18] Roth, A. E. and M. Sotomayor (1990): *Two-sided matching*, New York: Cambridge University Press.

[19] Shapley, L. and H. Scarf (1974): "On cores and indivisibility," *Journal of Mathematical Economics*, 1, 23–28.

[20] Tarjan, Robert E. (1972): "Depth-first search and linear graph algorithms," *SIAM Journal on Computing*, 1(2), 146–160.