Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 36

Procedurally Fair and Stable Matching

Bettina Klaus and Flip Klijn

November 2003

Procedurally Fair and Stable Matching*

Bettina Klaus† Flip Klijn‡

Barcelona Economics WP n° 36

This Version: November 2003

Abstract

We study procedurally fair matching mechanisms that produce stable matchings for the so-called marriage model of one-to-one, two-sided matching. Our main focus is on two such mechanisms: *employment by lotto* introduced by Aldershof *et al.* (1999) and the *random order mechanism* due to Roth and Vande Vate (1990) and Ma (1996). For both mechanisms we give various examples of probability distributions on the set of stable matchings and discuss properties that differentiate employment by lotto and the random order mechanism. Furthermore, we correct some misconceptions by Aldershof *et al.* (1999) and Ma (1996) that exist on the probability distribution induced by both mechanisms. Finally, we consider an adjustment of the random order mechanism, *the equitable random order mechanism.*

Keywords: procedural fairness, random mechanism, stability, two-sided matching.

JEL classification: C78, D63

1 Introduction

The marriage model describes a two-sided matching market without money where the two sides of the market for instance are workers and firms (job matching) or medical students and hospitals (matching of students to internships). We use the common terminology in the literature and refer to one side of the market as "men" and to the other as "women." An outcome for a marriage market is called a matching, which can simply be described by a collection of single agents and "married" pairs (consisting of one man and one woman). Loosely speaking, a matching is stable if all agents have acceptable spouses and there is no couple whose members both like each other better than their current spouses. Gale and Shapley (1962) formalized this notion of stability

*The work of the authors is partially supported by Research Grant BEC2002-02130 from the Spanish *Ministerio de Ciencia y Tecnología* and the Barcelona Economics Program of CREA. B. Klaus's research is supported by a Ramón y Cajal contract of the Spanish *Ministerio de Ciencia y Tecnología*. F. Klijn's research is supported by a Marie Curie Fellowship of the European Community programme "Improving Human Research Potential and the Socio-economic Knowledge Base" under contract number HPMF-CT-2001-01232.

†e-mail: `bettina.klaus@uab.es`.

‡Corresponding author. CODE and Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, Edifici B, 08193 Bellaterra, Spain. Tel. (34) 93 581 1720; Fax. (34) 93 581 2012; `fklijn@pareto.uab.es`

for marriage markets and provided an algorithm to calculate stable matchings. These classical results (Gale and Shapley, 1962) inspired many researchers to study stability not only for the marriage model, but for more general models as well. We refer to Roth and Sotomayor (1990) for a comprehensive account on stability for two-sided matching models.

In this paper we study a combination of fairness and stability in the marriage model. Masarani and Gokturk (1989) showed several impossibilities to obtain a fair deterministic matching mechanism within the context of Rawlsian justice. In contrast to this cardinal approach we focus on the ordinal aspects of the model and opt for an approach of procedural fairness. Since for any deterministic matching mechanism we can detect an inherit favoritism either for one side of the market or for some agents over others, in order to at least recover *ex ante* fairness, we consider probabilistic stable matching mechanisms that assign to each marriage market a probability distribution over the set of stable matchings. We do not intend to judge the fairness of a probabilistic stable matching mechanism by judging the assigned probability distributions, but by considering procedurally fair matching algorithms in which the sequence of moves for the agents is drawn from a uniform distribution. Hence, whenever an agent has the same probability to move at a certain point in the procedure that determines the final probability distribution, we consider the random stable matching mechanism to be *procedurally fair.* In other words, here we focus on "procedural justice" rather than on "endstate justice" (see Moulin, 1997,2003).

First, we analyze a random matching mechanism proposed by Aldershof *et al.* (1999) called *employment by lotto.* Loosely speaking, employment by lotto can be considered to be a random serial dictatorship on the set of stable matchings. A first agent is drawn randomly and can discard all stable matchings in which he/she is not matched to his/her best partner in a stable matching. Exclude the first agent and his/her partner from the set of agents and randomly choose the next agent who can discard all stable matchings in which he/she is not matched to his/her best partner in the reduced set of stable matchings. Continue with this sequential reduction of the set of stable matchings until it is reduced to a singleton. Using all possible sequences of agents, this mechanism induces a probability distribution on the set of stable matchings. The associated probabilistic matching mechanism of this probabilistic sequential dictatorship equals employment by lotto. We give various examples of probability distributions on the set of stable matchings induced by employment by lotto, show certain limitations of this mechanism (*e.g.*, complete information of all agents' preferences is needed), and disprove several conjectures about the distribution of probabilities made in Aldershof *et al.* (1999).

Next, we consider a random matching mechanism based on Roth and Vande Vate's (1990) results. We follow Ma (1996) and refer to this rule as the *random order mechanism.* The basic idea is as follows. Imagine an empty room with one entrance. At the beginning, all agents are waiting outside. At each step of the algorithm one agent is chosen randomly and invited to enter. Before an agent enters the matching in the room is stable. However, once an agent enters the room, the existing matching in the room may become unstable, meaning that the new agent can form a blocking pair with another agent that already is present in the room. By satisfying this (and possible subsequent) blocking pair(s) in a certain way a new stable matching including the entering agent is obtained for the marriage market in the room. After a finite number of steps a stable matching for the original marriage market is obtained. Using all possible sequences of agents, this mechanism induces a probability distribution on the set of stable matching. The associated probabilistic matching mechanism equals the random order mechanism. We give various examples of probability distributions on the set of stable matchings induced by the random order mechanism. Furthermore, we show that the probability distribution Ma (1996) presented is not correct. The mistake in the calculations by Ma (1996) is due to the fact that even though the example looks very symmetric, some of the calculations are not as "symmetric" since the random order mechanism does not satisfy what we call *independence of dummy agents*; that is, the final probability distribution on the set of stable matchings may crucially depend on preferences of agents who are matched to the same partner in all stable matchings. Moreover, we answer in the negative a question posed by Cechlárová (2002) on whether certain matchings can always be reached.

Finally, following a suggestion by Romero-Medina (2002), we briefly discuss an adjustment of the random order mechanism, *the equitable random order mechanism*. This mechanism limits the set of options available for each agent, trying to avoid the inherent favoritism of optimal matchings. We show that even for small markets the three mechanisms may give completely different and somewhat surprising outcomes.

In all our examples, we implement the mechanisms discussed so far in Matlab ©. In some examples the resulting probabilities are rounded.

The article is organized as follows. In Section 2 we introduce the marriage model and the concepts of stability and procedural fairness. In Section 3 we recall and study employment by lotto. In Section 4 we recall and study the random order mechanism and its adjustment, the equitable random order mechanism.

2 Matching Markets, Stability, and Procedural Fairness

2.1 Matching Markets

First we introduce the model of a two-sided matching market without money where each agent may be matched to (at most) one agent of the opposite side. For convenience we apply Gale and Shapley's (1962) interpretation of a "marriage market." For further details on the interpretation and standard results we refer to Roth and Sotomayor's (1990) comprehensive book on two-sided matching.

There are two finite and disjoint sets of agents: a set $M = \{m_1, \ldots, m_a\}$ of "men" and a set $W = \{w_1, \ldots, w_b\}$ of "women," where possibly $a \neq b$. The set of agents equals $N = M \cup W$. Let $n = |N|$. We denote a generic agent by i, a generic man by m, and a generic woman by w.

Each agent has a complete, transitive, and strict preference relation over the agents on the other side of the market and the prospect of being alone. Hence, man m's preferences $\succeq_m$ can be represented as a strict ordering $P(m)$ of the elements in $W \cup \{m\}$, for instance: $P(m) = w_3, w_2, m, w_1, \ldots, w_4$, which indicates that m prefers w_3 to w_2 and he prefers remaining single to any other woman. Similarly, woman w's preferences $\succeq_w$ can be represented as a strict ordering $P(w)$ of elements in $M \cup \{w\}$. Let P be the profile of all agents' preferences: $P = (P(i))_{i \in N}$.

We write $w \succ_m w'$ if m strictly prefers w to w' $(w \neq w')$, and $w \succeq_m w'$ if m likes w at least as well as w' ($w \succ_m w'$ or $w = w'$). Similarly we write $m \succ_w m'$ and $m \succeq_w m'$. A woman w is *acceptable* to a man m if $w \succ_m m$. Analogously, m is acceptable to w if $m \succ_w w$.

A *marriage market* is a triple (M, W, P). An outcome for a marriage market (M, W, P) is a *matching*, a one-to-one function μ from N to itself, such that for each $m \in M$ and for each $w \in W$ we have $\mu(m) = w$ if and only if $\mu(w) = m$, $\mu(m) \notin W$ implies $\mu(m) = m$, and similarly $\mu(w) \notin M$ implies $\mu(w) = w$. If $\mu(m) = w$, then man m and woman w are matched to one another (they are mates). If $\mu(i) = i$, then agent i is *single* or *unmatched*.

2.2 Stability

A key property of matchings is *stability*. First, since agents can always choose to be single, we require a voluntary participation condition. A matching μ is *individually rational* if each agent is acceptable to his/her mate, *i.e.*, $\mu(i) \succeq_i i$ for all $i \in N$. Second, if an agent can improve upon his present matching by switching to another agent such that this agent is better off as well, then we would expect this mutually beneficial "trade" to be carried out, rendering the given matching instable. For a given matching μ, a pair (m, w) is a *blocking pair* if they are not matched to one another but prefer one another to their mates at μ, *i.e.*, $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching is *stable* if it is individually rational and if there are no blocking pairs. With a slight abuse of notation, we denote the set of stable matchings for marriage market (M, W, P) by $S(P)$. Gale and Shapley (1962) proved that $S(P) \neq \emptyset$. Furthermore, any set of stable matchings has the structure of a (distributive) lattice, which we explain next.

For any two matchings μ and μ' we define a function $\lambda \equiv \mu \vee_M \mu'$ on N that assigns to each man his more preferred mate from μ and μ' and to each woman her less preferred mate. Formally, let $\lambda = \mu \vee_M \mu'$ be defined for all $m \in M$ by $\lambda(m) := \mu(m)$ if $\mu(m) \succ_m \mu'(m)$ and $\lambda(m) := \mu'(m)$ otherwise, and for all $w \in W$ by $\lambda(w) := \mu(w)$ if $\mu'(w) \succ_w \mu(w)$ and $\lambda(w) := \mu'(w)$ otherwise. In a similar way we define the function $\mu \wedge_M \mu'$, which gives each man his less preferred mate and each woman her more preferred mate. Knuth (1976) published the following theorem, but it is attributed to John Conway.

Theorem 2.1 [Lattice Theorem, Conway] *If $\mu, \mu' \in S(P)$, then also $\mu \vee_M \mu', \mu \wedge_M \mu' \in S(P)$.*

From Theorem 2.1 and the existence of a stable matching it follows easily that there is a stable matching μ_M that is optimal for all men in the sense that no other stable matching μ gives to any man m a mate $\mu(m)$ that he prefers to $\mu_M(m)$. Similarly, there is a stable matching μ_W that is optimal for all women. In fact, Gale and Shapley (1962) already proved the existence of μ_M and μ_W, and provided an algorithm, called the deferred acceptance procedure, to calculate these matchings.

Since preferences are strict, the set of matched agents does not vary from one stable matching to another. In other words, the set of unmatched agents is the same for all stable matchings.

Theorem 2.2 [McVitie and Wilson (1970), Roth (1982)] *For all $i \in N$ and all $\mu, \mu' \in S(P)$, $\mu(i) = i$ implies $\mu'(i) = i$.*

2.3 Procedural Fairness

We are interested in matching mechanisms that produce stable matchings and that can be considered "fair." Before explaining the concept of *procedural fairness* that we apply here, we define stable matching mechanisms. A *stable matching mechanism* μ is a function that for any marriage market (M, W, P) assigns a stable matching $\mu(M, W, P)$.

Two well-known and widely applied stable matching mechanisms are the *man-optimal* and the *woman-optimal deferred acceptance (DA)* algorithm by Gale and Shapley (1962). As discussed in Section 2.2, for any marriage market (M, W, P), the man-optimal DA algorithm yields the (unique) stable matching preferred by all men and the woman-optimal DA algorithm yields the (unique) stable matching preferred by all women. However, although stable, for all marriage markets where the man-optimal matching differs from the woman-optimal matching, which is the rule rather than the exception, each of the matching mechanisms clearly favors one side of the market. If there is no obvious reason why one side of the market should be favored, this favoritism can be considered "unfair."

This inherit incompatibility between stability and fairness is not restricted to the man-optimal and the woman-optimal DA algorithm, but in fact extends to all deterministic matching rules. Given the lattice structure of the set of stable matchings, for some marriage markets any deterministic matching mechanism is bound to favor one side of the market; for instance whenever the set of stable matchings consist of a man-optimal and a woman-optimal matching. Even if the matching mechanism does not choose a man-optimal or woman-optimal matching whenever possible, depending on the lattice structure of stable matchings, some agents may have to be favored relative to other agents on both sides of the market. Therefore, in order to formulate fairness without sacrificing stability, we consider probabilistic stable matching mechanisms, that is, for each marriage problem (M, W, P) a *probabilistic stable matching mechanism* assigns a probability distribution $\mathcal{P}(M, W, P)$ over the set of stable matchings $S(P)$.

We do not intend to judge the fairness of a probabilistic stable matching mechanism by judging the assigned probability distributions, but by considering procedurally fair matching algorithms in which the sequence of moves for the agents is drawn from a uniform distribution. Loosely speaking, whenever each agent has the same probability to move at a certain point in the procedure that determines the final probability distribution, we consider the respective probabilistic stable matching mechanism to be *procedurally fair*.

3 Procedural Fairness: Employment by Lotto

Aldershof *et al.* (1999) proposed a probabilistic stable matching mechanism, called employment by lotto, that is intended to avoid the inherent favoritism of optimal matchings by using randomization. Loosely speaking employment by lotto can be considered to be a random serial dictatorship on the set of stable matchings. A first agent is drawn randomly and can discard all stable matchings in which he/she is not matched to his/her best partner in a stable matching. Note that now the first agent is matched to the same partner in all remaining stable matchings. Exclude the first agent and his/her partner from the set of agents and randomly choose the next agent who can discard all stable matchings in which he/she is not matched to his/her best partner in the reduced set of stable matchings. Continue with this sequential reduction of the set of stable matchings until it is reduced to a singleton. Using all possible sequences of agents, this mechanism induces a probability distribution on the set of stable matchings. The associated probabilistic matching mechanism of this probabilistic sequential dictatorship mechanism equals employment by lotto. An alternative definition of employment by lotto is given by Aldershof *et al.* (1999).

3.1 The Employment by Lotto Algorithm

As mentioned before, we opt for a different description of the procedure than Aldershof *et al.* (1999). The description of employment by lotto as a probabilistic sequential dictatorship mechanism on the set of stable matchings enables us to avoid introducing further notation and technicalities.

Employment by Lotto (EL) Algorithm

Input
A marriage market (M, W, P).
Set $N_1 := N$, $S_1 := S(P)$, and $t := 1$.

Step t
Choose an agent i_t from N_t at random.
Match agent i_t to his most preferred mate $ch(i_t)$ among $\{j : j = \mu(i_t) \text{ for some } \mu \in S_t\}$.

If $N_t \backslash \{i_t, ch(i_t)\} = \emptyset$, then stop and define $\{EL(P)\} := S_t$.
Otherwise set $N_{t+1} := N_t \backslash \{i_t, ch(i_t)\}$, $S_{t+1} := S_t \backslash \{\mu \in S_t : \mu(i_t) \neq ch(i_t)\}$, and go to Step $t := t + 1$.

Recall that $|M| = a$ and $|W| = b$. It is easy to see that the algorithm ends in a finite number r ($\max\{a, b\} \leq r \leq a + b$) of steps that only depends on the preferences (this follows from Theorem 2.2). The outcome is a random stable matching $EL(P) \in S(P)$, generated by a sequence of agents $(i_1, \ldots, i_r)$. Let Q be the set of such sequences and let $q = |Q|$. Moreover, for any $\mu \in S(P)$, let $Q_\mu \subseteq Q$ be the (possibly empty) set of sequences that lead to μ. Denote $q_\mu = |Q_\mu|$. Note that if $a = b$ and if all men and women are mutually acceptable, then $r = a$ and $q = 2a \cdot (2a - 2) \cdot \ldots \cdot 2$.

The employment by lotto algorithm induces in a natural way a probability distribution $\mathcal{P} = \{p_\mu\}_{\mu \in S(P)}$ over the set of stable matchings: for any $\mu \in S(P)$, the probability that $EL(P) = \mu$ equals $p_\mu = \frac{q_\mu}{q}$. Aldershof *et al.* (1999) observe that if a stable matching μ does not match any agent to his/her man/woman optimal mate, then $p_\mu = 0$. More precisely, if for all $i \in N$ it holds that $\mu_M(i) \neq \mu(i) \neq \mu_W(i)$, then $p_\mu = 0$. We demonstrate this characteristic of the EL algorithm in the following example. In addition, we show how the example can be adjusted to prove that the converse is not true, *i.e.*, $p_\mu = 0$ does not necessarily imply that for all $i \in N$, $\mu_M(i) \neq \mu(i) \neq \mu_W(i)$. More importantly, the example proves that a stable matching that constitutes a "perfect compromise" between contrary preferences on both sides of the market may never result from employment by lotto.

Example 3.1 Employment by lotto may never find the perfect compromise
Let (M, W, P) with $a = b = 3$ and P such that

$$\begin{array}{rcllll}
P(m_1) & = & \underline{w_1}, & \widetilde{w_2}, & \overline{w_3}, & m_1 \\
P(m_2) & = & \underline{w_3}, & \widetilde{w_1}, & \overline{w_2}, & m_2 \\
P(m_3) & = & \underline{w_2}, & \widetilde{w_3}, & \overline{w_1}, & m_3 \\
P(w_1) & = & \overline{m_3}, & \widetilde{m_2}, & \underline{m_1}, & w_1 \\
P(w_2) & = & \overline{m_2}, & \widetilde{m_1}, & \underline{m_3}, & w_2 \\
P(w_3) & = & \overline{m_1}, & \widetilde{m_3}, & \underline{m_2}, & w_3
\end{array}$$

There exist three stable matchings, $\underline{\mu}$, $\overline{\mu}$, and $\widetilde{\mu}$, where men m_1, m_2, m_3 are matched to

$$w_1,\ w_3,\ w_2, \tag{$\underline{\mu}$}$$

$$w_2,\ w_1,\ w_3, \text{ and} \tag{$\widetilde{\mu}$}$$

$$w_3,\ w_2,\ w_1. \qquad (\overline{\mu})$$

Note that in matching $\underline{\mu}$ all men are matched to their most preferred mate and all women are matched to their least preferred mate ($\underline{\mu} = \mu_M$ is underlined at preference profile P)[1]. Matching $\overline{\mu}$ establishes the other extreme: all women are matched to their most preferred mate and all man are matched to their least preferred mate ($\overline{\mu} = \mu_W$). At matching $\widetilde{\mu}$ agents are matched neither to their most, nor to their least preferred mate. In fact, at $\widetilde{\mu}$ all agents are matched to their second choice, which is why we consider $\widetilde{\mu}$ to be a perfect compromise in this situation. We depict the corresponding lattice in Figure 1. The nodes denote the stable matchings and the first number in each series is the corresponding probability resulting from employment by lotto (the other two numbers are probabilities from other random matching mechanisms that we discuss later). The solid arcs denote comparability or unanimity on each side of the market. For instance $\underline{\mu} \to \widetilde{\mu}$ in Figure 1 means that all men at least weakly prefer their mates at $\widetilde{\mu}$ to their mates at $\underline{\mu}$ and all women at least weakly prefer their mates at $\underline{\mu}$ to their mates at $\widetilde{\mu}$ (*i.e.*, $\underline{\mu} \vee_M \widetilde{\mu} = \widetilde{\mu}$).

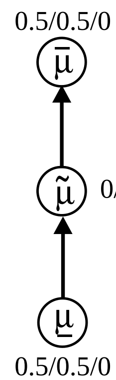

Figure 1: Lattice of Example 3.1

It is easy to check that whenever agent i_1 in the EL algorithm is a man, then $EL(P) = \mu_M$, and whenever agent i_1 in the EL algorithm is a woman, then $EL(P) = \mu_W$ ($p_{\mu_M} = \frac{1}{2} = p_{\mu_W}$). Hence, for the perfect compromise matching $\widetilde{\mu}$, $p_{\widetilde{\mu}} = 0$.

In order to show that $p_\mu = 0$ does not necessarily imply that for all $i \in N$, $\mu_M(i) \neq \mu(i) \neq \mu_W(i)$, we add two agents m_4, w_4 to the market above such that for all $i = 1, 2, 3, 4$, $m_4 \succ_{m_4} w_i$, $w_4 \succ_{w_4} m_i$, and m_4, w_4 are placed anywhere in the preferences of the other agents. Then, for stable matching μ, where man m_1, m_2, m_3, and m_4 are matched to w_2, w_1, w_3, m_4, $p_\mu = 0$ and $\mu(m_4) = \mu_M(m_4)$ and $\mu(w_4) = \mu_W(w_4)$. $\diamond$

Finally, one might think that the employment by lotto algorithm is equivalent to the following procedure: first pick an agent i_1 at random, match i_1 to $ch(i_1)$, and remove i_1 and $ch(i_1)$ from the marriage market and the preference lists of the remaining agents. Repeat this procedure with the reduced marriage market, *etc.*. Unfortunately, this procedure may not find a stable matching since, for instance, $ch(i_1)$ and $ch(i_2)$ thus obtained may form a blocking pair for the resulting matching. We demonstrate this using the marriage market introduced in Example 3.1. Suppose that m_1 first chooses w_1. In the reduced market there are two stable matchings: at μ_1

[1] Matchings $\overline{\mu}$ and $\widetilde{\mu}$ are marked in a similar way.

men m_2, m_3 are matched to w_3, w_2 and at μ_2 men m_2, m_3 are matched to w_2, w_3. Next, assume that w_2 can choose in the reduced market. Since w_2 prefers her mate at μ_2 over her mate at μ_1, the resulting matching for the original market matches men m_1, m_2, m_3 to w_1, w_2, w_3. However, this matching is not stable ((m_2, w_1) is a blocking pair).

Hence, in general it is necessary to calculate the complete set of stable matchings. For an algorithm polynomially bounded in n we refer to Irving and Leather (1986) and Roth and Sotomayor (1990).

3.2 Properties and Misconceptions of Employment by Lotto

We discuss two properties that set employment by lotto apart from the second procedurally fair matching mechanism that we consider in Section 4. First, we explain that employment by lotto is based on a strong information requirement. Next, we point out that the probability distributions obtained by employment by lotto do not depend on agents that are matched to the same partner in all stable matchings.

Complete Information needed: As mentioned in Section 3.1, in order to apply employment by lotto it is necessary to calculate the set of stable matchings. From an informational point of view that means that either a central planner or all agents need complete information of the preference profile.

We call an agent that is matched to the same partner (including being single) at all stable matchings a *dummy agent*. We call a stable matching mechanism *independent of dummy agents* if dummy agents do not have an influence on the final probability distribution in the following sense. Delete all dummy agents from the original set of agents and apply the matching mechanism to the obtained reduced marriage market. Then, the probabilities for the remaining agents do not change. In order to formalize this property, we need some notation. Let (M, W, P) be a marriage market and let $D \subseteq N$ be the set of all dummy agents. Then $M \backslash D$ denotes all men that are not dummy agents, $W \backslash D$ denotes all women that are not dummy agents, and $P_{N \backslash D} = (P(i)_{N \backslash D})_{i \in N \backslash D}$ denotes the profile of reduced preferences induced by $(P(i))_{i \in N \backslash D}$. Formally, for all $i \in M \backslash D$ and all $j, k \in \{i\} \cup W \backslash D$, if $j \succeq_i k$ at $P(i)$, then $j \succeq_i k$ at $P(i)_{N \backslash D}$. (Similarly for $i \in W \backslash D$.) Then, after eliminating all dummy agents, we obtain the reduced marriage market $(M \backslash D, W \backslash D, P_{N \backslash D})$. Note that there exists a one-to-one mapping between matchings in $S(P)$ and $S(P_{N \backslash D})$: by eliminating dummy agents from a matching $\mu \in S(P)$ we obtain a matching $\mu_{N \backslash D} \in S(P_{N \backslash D})$, and *vice versa*, by adding dummy agents with their respective matches to a matching $\mu_{N \backslash D} \in S(P_{N \backslash D})$ we obtain a matching $\mu \in S(P)$.

Independence of Dummy Agents: Let (M, W, P) be a marriage market and $\tilde{\mathcal{P}}$ the probability distribution on the corresponding set of stable matchings induced by a stable matching mechanism, that is, for all matchings $\mu \in S(P)$, $\tilde{\mathcal{P}}(\mu)$ denotes the probability that matching μ is chosen. Similarly, for the reduced marriage market $(M \backslash D, W \backslash D, P_{N \backslash D})$, $\tilde{\mathcal{P}}(\mu_{N \backslash D})$ denotes the probability that the reduced matching $\mu_{N \backslash D}$ is chosen.

Then, the matching mechanism satisfies *independence of dummy agents for* (M, W, P) if and only if for all matchings $\mu \in S(P)$, $\tilde{\mathcal{P}}(\mu) = \tilde{\mathcal{P}}(\mu_{N \backslash D})$. A matching mechanism satisfies *independence of dummy agents* if it satisfies independence of dummy agents for all marriage markets (M, W, P).

Since in the employment by lotto algorithm a dummy agent will never reduce the set of remaining stable matchings, it is easy to see that employment by lotto satisfies independence of dummy agents.

In the remainder of the section we correct some misconceptions by Aldershof *et al.* (1999) on the probability distributions induced by employment by lotto. Since in Aldershof *et al.* (1999) the assumption that $|M| = |W|$ is crucial, for the remainder of this section we assume that $a = |M| = |W| = b$.

Aldershof *et al.* (1999) made the following conjectures about the probability distribution $\mathcal{P}$ generated by employment by lotto over the set of stable matchings.

Conjecture 3.2 [Aldershof *et al.* (1999), p. 288] *For any marriage market,* $p_{\mu_M} = p_{\mu_W}$*.*

Conjecture 3.3 [Aldershof *et al.* (1999), p. 288] *"Consider a lattice of stable matchings for an instance of the stable matching problem. All matchings with rank* i *have the same probability* p_i *of resulting from employment by lotto. Also the function* $f(i) = 1 - p_i$ *is unimodal."*

We start with Conjecture 3.2. The following two results give a complete answer. The proof of Theorem 3.4 can be found in the Appendix.

Theorem 3.4 *If* $a \leq 3$*, then* $p_{\mu_M} = p_{\mu_W}$*.*

Theorem 3.5 *If* $a > 3$*, then not necessarily* $p_{\mu_M} = p_{\mu_W}$*.*

Proof. Let (M, W, P) with $a = 4$ and P such that[2]

$$\begin{array}{lcccccc}
P(m_1) & = & w_1, & w_2, & w_3, & w_4, & m_1 \\
P(m_2) & = & w_2, & w_1, & w_4, & w_3, & m_2 \\
P(m_3) & = & w_3, & w_4, & w_1, & w_2, & m_3 \\
P(m_4) & = & w_4, & w_3, & w_2, & w_1, & m_4 \\
P(w_1) & = & m_4, & m_2, & m_3, & m_1, & w_1 \\
P(w_2) & = & m_3, & m_4, & m_1, & m_2, & w_2 \\
P(w_3) & = & m_2, & m_1, & m_4, & m_3, & w_3 \\
P(w_4) & = & m_1, & m_2, & m_3, & m_4, & w_4.
\end{array}$$

There are seven stable matchings and for the man and woman optimal matchings we find $p_{\mu_M} = \frac{2}{8} \neq \frac{3}{8} = p_{\mu_W}$. There are three other matchings with positive probability. For $a > 4$ one can simply add agents that find any other agent on the other side of the market unacceptable. $\square$

Next, we consider Conjecture 3.3. Our first remark is that "rank" was not formally defined by Aldershof *et al.* (1999). It suggests that the matchings in any lattice can be partitioned in certain "natural" levels, which is true for many examples of lattices that are used throughout the literature on stable matching. The following example demonstrates that this notion of natural level/rank is not obvious at all. Given that Blair (1984) showed that every lattice can

[2]We switch the 2nd and 3rd position of agent w_1's preference in a marriage market taken from Knuth (1976).

be obtained as the set of stable matchings of some marriage market, this result is not surprising. In addition, the example also shows that even if two stable matchings are incomparable (*i.e.*, the men are not unanimous on which of the two is better) they may still have different probabilities of resulting from employment by lotto.

Example 3.6 Let (M, W, P) with $a = 5$ and P such that[3]

$$\begin{array}{lcllllll}
P(m_1) & = & w_1, & w_3, & w_2, & w_4, & w_5, & m_1 \\
P(m_2) & = & w_2, & w_3, & w_1, & w_4, & w_5, & m_2 \\
P(m_3) & = & w_3, & w_2, & w_1, & w_4, & w_5, & m_3 \\
P(m_4) & = & w_4, & w_5, & w_1, & w_2, & w_3, & m_4 \\
P(m_5) & = & w_5, & w_4, & w_1, & w_2, & w_3, & m_5 \\
P(w_1) & = & m_2, & m_1, & m_3, & m_4, & m_5, & w_1 \\
P(w_2) & = & m_3, & m_2, & m_1, & m_4, & m_5, & w_2 \\
P(w_3) & = & m_1, & m_2, & m_3, & m_4, & m_5, & w_3 \\
P(w_4) & = & m_5, & m_4, & m_1, & m_2, & m_3, & w_4 \\
P(w_5) & = & m_4, & m_5, & m_1, & m_2, & m_3, & w_5.
\end{array}$$

There are six stable matchings, where men m_1, m_2, m_3, m_4, m_5 are matched to

$$w_3,\ w_1,\ w_2,\ w_5,\ w_4, \qquad (\mu_1)$$

$$w_3,\ w_1,\ w_2,\ w_4,\ w_5, \qquad (\mu_2)$$

$$w_1,\ w_3,\ w_2,\ w_5,\ w_4, \qquad (\mu_3)$$

$$w_1,\ w_3,\ w_2,\ w_4,\ w_5, \qquad (\mu_4)$$

$$w_1,\ w_2,\ w_3,\ w_5,\ w_4, \text{ and} \qquad (\mu_5)$$

$$w_1,\ w_2,\ w_3,\ w_4,\ w_5. \qquad (\mu_6)$$

Note that $\mu_M = \mu_6$ and $\mu_1 = \mu_W$. We depict the corresponding lattice in Figure 2. The nodes denote the stable matchings and the numbers the corresponding probabilities resulting from employment by lotto. The solid arcs denote again comparability or unanimity on each side of the market. Dotted edges denote incomparability or disagreement on each side of the market. For instance $\mu_4 \cdots \mu_5$ in Figure 2 means that there is disagreement among the men (women) about which matching is better ($\mu_5(m_2) \succ_{m_2} \mu_4(m_2)$, but $\mu_4(m_4) \succ_{m_4} \mu_5(m_4)$).

The fact that there is no unanimity with respect to matchings μ_4 and μ_5 and also with respect to μ_2 and μ_5, but $\mu_2 \to \mu_4$, shows that a natural concept of "rank" is difficult to define. Moreover, for the two incomparable matchings μ_4 and μ_5 we have that $p_{\mu_4} = \frac{2}{24} \neq \frac{5}{24} = p_{\mu_5}$. $\diamond$

The following example shows that for $a > 3$ even if the matchings in a lattice can be partitioned in natural levels (*i.e.*, the notion of a "rank" can be defined), the function f in Conjecture 3.3 needs not be uni-modal (by the proof of Theorem 3.4 this part of the conjecture is true for $a \leq 3$).

[3]We complete the preferences of a marriage market taken from Blair (1984).

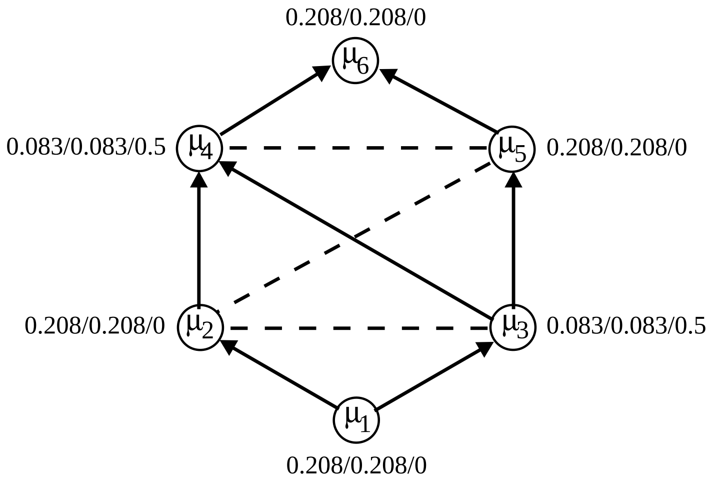

Figure 2: Lattice of Example 3.6

Example 3.7 Let (M, W, P) with $a = 4$ and P such that

$$\begin{array}{lcccccc}
P(m_1) & = & w_1, & w_2, & w_4, & w_3, & m_1 \\
P(m_2) & = & w_2, & w_1, & w_3, & w_4, & m_2 \\
P(m_3) & = & w_3, & w_4, & w_1, & w_2, & m_3 \\
P(m_4) & = & w_4, & w_3, & w_1, & w_2, & m_4 \\
P(w_1) & = & m_3, & m_2, & m_1, & m_4, & w_1 \\
P(w_2) & = & m_4, & m_1, & m_2, & m_3, & w_2 \\
P(w_3) & = & m_1, & m_2, & m_3, & m_4, & w_3 \\
P(w_4) & = & m_2, & m_1, & m_4, & m_3, & w_4.
\end{array}$$

There are six stable matchings, where men m_1, m_2, m_3, m_4 are matched to

$$w_3,\ w_4,\ w_1,\ w_2, \tag{ν_1}$$

$$w_4,\ w_3,\ w_1,\ w_2, \tag{ν_2}$$

$$w_4,\ w_1,\ w_3,\ w_2, \tag{ν_3}$$

$$w_2,\ w_3,\ w_1,\ w_4, \tag{ν_4}$$

$$w_2,\ w_1,\ w_3,\ w_4, \text{ and} \tag{ν_5}$$

$$w_1,\ w_2,\ w_3,\ w_4. \tag{ν_6}$$

We depict the corresponding lattice in Figure 3. Since $p_{\nu_2} = p_{\nu_5} = \frac{2}{48} < \frac{5}{48} = p_{\nu_3} = p_{\nu_4}$ it is clear that the function f as defined in Conjecture 3.3 is not uni-modal here. $\diamond$

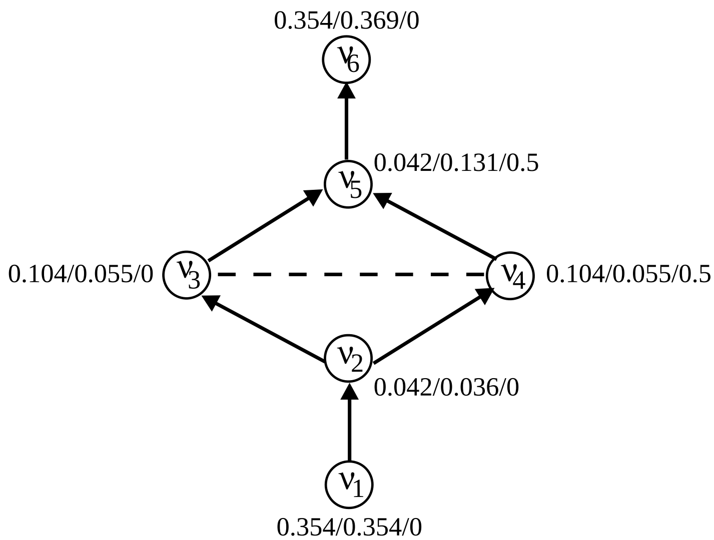

Figure 3: Lattice of Example 3.7

4 Procedural Fairness: the Random Order Mechanism

Ma (1996) described the random order mechanism, which is based on Roth and Vande Vate's (1990) random paths to stability. The basic idea is as follows. Imagine an empty room with one entrance. At the beginning, all agents are waiting outside. At each step of the algorithm, one agent is chosen randomly and invited to enter. Before an agent enters the matching in the room is stable. However, once an agent enters the room, the existing matching in the room may become unstable, meaning that the new agent can form a blocking pair with another agent that already is present in the room. By satisfying this (and possible subsequent) blocking pair(s) in a certain way (described below in full detail) a new stable matching including the entering agent is obtained for the marriage market in the room. Since at each step a new agent enters the room and no agent leaves the room, the final outcome is a stable matching for the original marriage market. Using all possible sequences of agents, this mechanism induces a probability distribution on the set of stable matchings. The associated probabilistic matching mechanism equals the random order mechanism.

4.1 The Random Order Mechanism

We first give a formal description of the random order mechanism.

Random Order (RO) Mechanism

Input
A marriage market (M, W, P).
Set $R_0 := \emptyset$, μ_0 such that for all $i \in N$, $\mu_0(i) = i$, and $t := 1$.

Step *t*
Choose an agent i_t from $N \backslash R_{t-1}$ at random. Set $R_t := R_{t-1} \cup \{i_t\}$.
Suppose $i_t = w \in W$. (Otherwise replace w by m in Step t.)

Stable Room Procedure

Case (i) There exists no blocking pair (m, w) for μ_{t-1} with $m \in R_t$:

Stop if $t = n$ and define $RO(P) := \mu_{t-1}$. Otherwise set $\mu_t = \mu_{t-1}$ and go to Step $t := t + 1$.

Case (ii) There exists a blocking pair (m, w) for μ_{t-1} with $m \in R_t$:

Choose the blocking pair (m^*, w) for μ_{t-1} with $m^* \in R_t$ that w prefers most.

If $\mu_{t-1}(m^*) = m^*$, then define μ_t such that $\mu_t(w) := m^*$, $\mu_t(m^*) := w$, and for all $i \in N \backslash \{w, m^*\}$, $\mu_t(i) := \mu_{t-1}(i)$. Stop if $t = n$ and define $RO(P) := \mu_t$. Otherwise go to Step $t := t + 1$.

If $\mu_{t-1}(m^*) = w' \in W$, then redefine $\mu_{t-1}(w) := m^*$, $\mu_{t-1}(m^*) := w$, $\mu_{t-1}(w') := w'$, and for all $i \in N \backslash \{w, m^*, w'\}$, $\mu_{t-1}(i) := \mu_{t-1}(i)$. Set $w := w'$, and repeat the Stable Room Procedure.

It is not difficult to see that the algorithm ends in exactly n steps. The outcome is a random stable matching $RO(P) \in S(P)$, generated by a sequence of agents $(i_1, \ldots, i_n)$. The set of possible sequences of agents equals the set of permutations of all agents denoted by Q^*. Hence, $|Q^*| = n!$. Moreover, for any $\mu \in S(P)$, let $Q^*_\mu \subseteq Q^*$ be the (possibly empty) set of sequences that lead to μ. Denote $q^*_\mu = |Q^*_\mu|$.

The random order mechanism induces in a natural way a probability distribution $\mathcal{P}^*$ over the set of stable matchings: for any $\mu \in S(P)$, the probability that $RO(P) = \mu$ equals $p^*_\mu = \frac{q^*_\mu}{n!}$. In all our examples, these probabilities are the second numbers in each series in the lattices.

Note that, similarly as employment by lotto, the random order mechanism never chooses the "perfect compromise" matching $\tilde{\mu}$ in Example 3.1.

4.2 Properties and Misconceptions of the Random Order Mechanism

We compare the random order mechanism with employment by lotto, using the same properties as in Section 3.2.

No Complete Information needed: An important advantage of the random order mechanism over employment by lotto is that it is not necessary to calculate the set of stable matchings beforehand. In order to be a part in the random order mechanism, each agent only needs to know his/her own preferences.

The following example shows however that the random order mechanism fails to satisfy independence of dummy players.

Example 4.1 The random order mechanism does not satisfy independence of dummy agents. Let (M, W, P) with $a = b = 3$ and P such that

$$\begin{array}{rcllll} P(m_1) & = & w_1, & w_2, & w_3, & m_1 \\ P(m_2) & = & w_2, & w_1, & w_3, & m_2 \\ P(m_3) & = & w_3, & w_2, & w_1, & m_3 \\ P(w_1) & = & m_2, & m_1, & m_3, & w_1 \\ P(w_2) & = & m_1, & m_3, & m_2, & w_2 \\ P(w_3) & = & m_3, & m_2, & m_1, & w_3. \end{array}$$

There exist two stable matchings μ_M and μ_W, where men m_1, m_2, and m_3 are matched to

$$w_1,\ w_2,\ w_3, \tag{μ_M}$$

$$w_2,\ w_1,\ w_3. \tag{μ_W}$$

Some calculations give $(p^*_{\mu_M}, p^*_{\mu_W}) = (\frac{5}{12}, \frac{7}{12})$.
After elimination of the two dummy agents m_3 and w_3, we obtain the marriage market $(\hat{M}, \hat{W}, \hat{P})$ with $a = b = 2$ and $\hat{P}$ such that

$$\begin{array}{rcllll}
\hat{P}(m_1) & = & w_1, & w_2, & m_1 \\
\hat{P}(m_2) & = & w_2, & w_1, & m_2 \\
\hat{P}(w_1) & = & m_2, & m_1, & w_1 \\
\hat{P}(w_2) & = & m_1, & m_2, & w_2.
\end{array}$$

There exist two stable matchings $\hat{\mu}_M$ and $\hat{\mu}_W$, where men m_1 and m_2 are matched to

$$w_1,\ w_2, \tag{$\hat{\mu}_M$}$$

$$w_2,\ w_1. \tag{$\hat{\mu}_W$}$$

Some calculations give $(p^*_{\hat{\mu}_M}, p^*_{\hat{\mu}_W}) = (\frac{1}{2}, \frac{1}{2})$. $\diamond$

Ma (1996) showed that the random order mechanism may not reach all stable matchings. Although Ma's (1996) theorem is true, we show in Example 4.2 that one of the two claims on which the proof relies is not true.

Example 4.2 Let (M, W, P) with $a = b = 4$ and P such that[4]

$$\begin{array}{rcllllll}
P(m_1) & = & w_1, & w_2, & w_3, & w_4, & m_1 \\
P(m_2) & = & w_2, & w_1, & w_4, & w_3, & m_2 \\
P(m_3) & = & w_3, & w_4, & w_1, & w_2, & m_3 \\
P(m_4) & = & w_4, & w_3, & w_2, & w_1, & m_4 \\
P(w_1) & = & m_4, & m_3, & m_2, & m_1, & w_1 \\
P(w_2) & = & m_3, & m_4, & m_1, & m_2, & w_2 \\
P(w_3) & = & m_2, & m_1, & m_4, & m_3, & w_3 \\
P(w_4) & = & m_1, & m_2, & m_3, & m_4, & w_4.
\end{array}$$

There are ten stable matchings, where men m_1, m_2, m_3, m_4 are matched to

$$w_4,\ w_3,\ w_2,\ w_1, \tag{μ_1}$$

$$w_4,\ w_3,\ w_1,\ w_2, \tag{μ_2}$$

$$w_3,\ w_4,\ w_2,\ w_1, \tag{μ_3}$$

$$w_3,\ w_4,\ w_1,\ w_2, \tag{μ_4}$$

$$w_3,\ w_1,\ w_4,\ w_2, \tag{μ_5}$$

$$w_2,\ w_4,\ w_1,\ w_3, \tag{μ_6}$$

[4]This is a marriage market taken from Knuth (1976).

$$w_2,\ w_1,\ w_4,\ w_3, \tag{μ_7}$$

$$w_2,\ w_1,\ w_3,\ w_4, \tag{μ_8}$$

$$w_1,\ w_2,\ w_4,\ w_3, \text{ and} \tag{μ_9}$$

$$w_1,\ w_2,\ w_3,\ w_4. \tag{μ_{10}}$$

We depict the corresponding lattice in Figure 4. Ma (1996) claimed that $(p^*_{\mu_1}, p^*_{\mu_2}, p^*_{\mu_3}, p^*_{\mu_4}, p^*_{\mu_5}, p^*_{\mu_6}, p^*_{\mu_7}, p^*_{\mu_8}, p^*_{\mu_9}, p^*_{\mu_{10}}) = (\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, 0, 0, 0, 0, \frac{1}{8}, \frac{1}{8}, \frac{1}{4})$, but it is clear from Figure 4 that this is not true. Note however that EL does give these probabilities. A proof that Ma's (1996) claim on the probabilities in this example is wrong (that is, $\mathcal{P}^* \neq \mathcal{P}$) that does not rely on our computational results can be found in the Appendix. $\diamond$

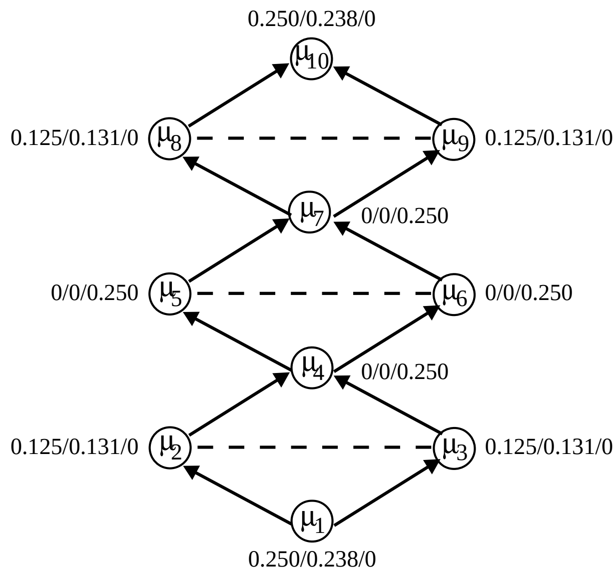

Figure 4: Lattice of Example 4.2

Cechlárová (2002) extended Ma's result showing that for any marriage market the only matchings that may be obtained are those that assign to at least one agent his/her best stable partner. One of the open problems Cechlárová (2002, p.4) mentioned is that "... it is not clear whether for each of those not excluded it is possible to find a suitable order of players [agents] to get it." The answer to this question can be found by repeating the procedure described in the last paragraph of Example 3.1. If we add two agents m_4, w_4 to the marriage market such that for all $i = 1, 2, 3, 4$, $m_4 \succ_{m_4} w_i$, $w_4 \succ_{w_4} m_i$, and m_4, w_4 are placed anywhere in the preferences of the other agents, then the extended matching $\widetilde{\mu}$, where men m_1, m_2, m_3, m_4 are matched to

$$w_2,\ w_1,\ w_3,\ w_4, \tag{$\widetilde{\mu}$}$$

respectively, is stable. It not difficult to see, however, that $p^*_{\widetilde{\mu}} = 0$.

4.3 The Equitable Random Order Mechanism

Romero-Medina (2002) adapted the random order mechanism in order to limit the set of options available for each agent, trying to avoid in this way the inherit favoritism of optimal matchings. Since the description of his algorithm would be a bit tedious and we only discuss briefly the differences between the three mechanisms in a few examples, we refer the reader to Romero-Medina (2002) for its definition. In fact, Romero-Medina (2002) defined the algorithm for a fixed order of the agents and only in his final remarks suggested an extension by randomizing the order of the agents. Henceforth, we will call this extension the equitable random order mechanism.

For any marriage market (M, W, P) and any $\mu \in S(P)$, let $\bar{p}_\mu$ be the probability that μ is the outcome of the equitable random order (ERO) mechanism. By $\bar{\mathcal{P}}$ we denote the associated probability distribution over the set of stable matchings. These probabilities are the third numbers in each series in the lattices.

In contrast to employment by lotto and the random order mechanism, the ERO mechanism chooses the "perfect compromise" matching $\tilde{\mu}$ in Example 3.1 not only with positive probability, but in fact with probability one. In the classical Example 4.2 the ERO mechanism demonstrates again nicely its avoidance of optimal matchings. The same occurs in Examples 3.6 and 3.7, although here probabilities seem to be split more arbitrarily.

In the example below we show that already for $a = b = 3$ the three mechanisms may give completely different and somewhat surprising outcomes. More specifically, it shows that the ERO mechanism may not always choose a probabilistic solution that compromises between both sides of the market: unlike the other two mechanisms, here the ERO mechanisms always chooses the woman optimal matching μ_W.

Example 4.3 Recall that for the matching market in Example 4.1 there are two stable matchings, where men m_1, m_2, m_3 are matched to

$$w_1,\ w_2,\ w_3, \text{ and} \tag{μ_M}$$

$$w_2,\ w_1,\ w_3. \tag{μ_W}$$

Some calculations give $(p_{\mu_M}, p_{\mu_W}) = (\frac{1}{2}, \frac{1}{2})$, $(p^*_{\mu_M}, p^*_{\mu_W}) = (\frac{5}{12}, \frac{7}{12})$, and $(\bar{p}_{\mu_M}, \bar{p}_{\mu_W}) = (0, 1)$. Note that the equitable random order mechanism fails to avoid the favoritism of one of the optimal matchings (μ_W). In contrast, the order two mechanisms, employment by lotto and the random order mechanism, spread probability over the two stable matchings, albeit in a slightly different way. $\diamond$

A Appendix

Proof of Theorem 3.4:

From Theorem 2.2 it follows that the probability distribution over the set of stable matchings does not change if we leave out all agents that are single in some (and hence all) stable matching(s). In other words, in the EL algorithm we can take $N_1 := N \backslash \{i \in N : \mu(i) = i$ for some $\mu \in S(P)\}$. *In order to simplify the proof, we assume that no agent is single in any stable matching.* Let P be any preference list for agents in N. Recall that $a = |M| = |W| = b$.

Case $a = 1$: Since $\mu_M(m_1) = \mu_W(m_1) = w_1$, it follows immediately that $p_{\mu_M} = p_{\mu_W} = 1$.

Case $a = 2$: Clearly, $S(P) \subseteq \{\mu_1(m_1, m_2) = (w_1, w_2), \mu_2(m_1, m_2) = (w_2, w_1)\}$. So, $|S(P)| \leq 2$. If $|S(P)| = 1$, then $\mu_M = \mu_W$, and hence, $p_{\mu_M} = p_{\mu_W} = 1$. If $|S(P)| = 2$, then the gender of the first agent i_1 in the EL algorithm determines the resulting matching, and hence, $p_{\mu_M} = p_{\mu_W} = \frac{1}{2}$.

Case $a = 3$: If $\mu_M = \mu_W$, then $p_{\mu_M} = p_{\mu_W} = 1$. Thus, let $\mu_M \neq \mu_W$.

Subcase 1: $p_{\mu_M} + p_{\mu_W} = 1$. Let $\overline{N} = \{i \in N : \mu_M(i) \neq \mu_W(i)\}$. Note that $i \in \overline{N} \cap M$ implies that there exist $j, k \in \overline{N} \cap W$ such that $j \neq k$. Similarly, $i \in \overline{N} \cap W$ implies that there exist $j, k \in \overline{N} \cap M$ such that $j \neq k$. Thus, $|\overline{N} \cap M| = |\overline{N} \cap W| \geq 2$. Hence, the set Q of EL sequences is the union of the following disjoint sets

$$\begin{aligned}
Q^1_{\mu_M} &= \{(i_1, i_2, i_3) \in Q : i_1 \in \overline{N} \cap M\}, \\
Q^1_{\mu_W} &= \{(i_1, i_2, i_3) \in Q : i_1 \in \overline{N} \cap W\}, \\
Q^2_{\mu_M} &= \{(i_1, i_2, i_3) \in Q : i_1 \notin \overline{N}, i_2 \in \overline{N} \cap M\}, \text{ and} \\
Q^2_{\mu_W} &= \{(i_1, i_2, i_3) \in Q : i_1 \notin \overline{N}, i_2 \in \overline{N} \cap W\}.^5
\end{aligned}$$

Note that $Q^1_{\mu_M} \cup Q^2_{\mu_M} \subseteq Q_{\mu_M}$ and $Q^1_{\mu_W} \cup Q^2_{\mu_W} \subseteq Q_{\mu_W}$. Since $|Q^1_{\mu_M}| = |Q^1_{\mu_W}|$ and $|Q^2_{\mu_M}| = |Q^2_{\mu_W}|$, it follows that $p_{\mu_M} = p_{\mu_W} = \frac{1}{2}$.

Subcase 2: $p_{\mu_M} + p_{\mu_W} < 1$. There exists a stable matching $\mu \notin \{\mu_M, \mu_W\}$ with $p_\mu > 0$. Thus, there exists a sequence $(i_1, i_2, i_3) \in Q_\mu$. Therefore, either

(a) $i_1 \in M$ and $\mu(i_1) = \mu_M(i_1)$ or

(b) $i_1 \in W$ and $\mu(i_1) = \mu_W(i_1)$.

We consider Case (a) (Case (b) is proven similarly). Without loss of generality let $i_1 = m_1$.

First we show that at matching μ at most one man can be matched to his man optimal match. Assume there exist $i, j \in M$, $i \neq j$ such that $\mu(i) = \mu_M(i)$ and $\mu(j) = \mu_M(j)$. Then, $\mu = \mu_M$, a contradiction. Hence, $\mu(m_1) = \mu_M(m_1)$, $\mu(m_2) \neq \mu_M(m_2)$, $\mu(m_3) \neq \mu_M(m_3)$, and $i_2 \in W$. After Step 1 of the EL algorithm only 4 agents are still to be matched. So, $|S_2| \leq 2$. But then $\mu \neq \mu_M$ yields $S_2 = \{\mu, \mu_M\}$. Similarly as above it follows that at matching μ at most one woman can be matched to her woman optimal match. In the remainder of the proof we will denote $\mu = (\mu(m_1), \mu(m_2), \mu(m_3))$. Without loss of generality assume that $\mu = (w_1, w_2, w_3)$. Then by $\mu_M(m_1) = \mu(m_1)$ and the assumption that no agent is single we have $\mu_M = (w_1, w_3, w_2)$.

Next, we consider the case $\mu_M(m_1) = \mu(m_1)$, $\mu_M(m_2) \neq \mu(m_2) \neq \mu_W(m_2)$, and $\mu_M(m_3) \neq \mu(m_3) \neq \mu_W(m_3)$.

Since, $\mu_W(m_2) \neq \mu(m_2)$, $\mu_W(m_2) \neq w_2$. Furthermore, $\mu_W(m_2) \neq \mu(m_2)$ implies $\mu_M(m_2) \succ_{m_2} \mu_W(m_2)$. Thus, $\mu_W(m_2) \neq w_3$. Hence, $\mu_W(m_2) = w_1$. However, applying the same arguments to agent m_3, we obtain $\mu_W(m_3) = w_1$ as well; a contradiction.

Now, the only case that remains is $\mu_M(m_1) = \mu(m_1) \neq \mu_W(m_1)$ and, without loss of generality,[6] $\mu_M(m_2) \neq \mu(m_2) = \mu_W(m_2)$, and $\mu_M(m_3) \neq \mu(m_3) \neq \mu_W(m_3)$.

Since $\mu_M(m_1) = \mu(m_1) = w_1$, $\mu_M(m_2) \neq \mu(m_2)$ and $\mu_M(m_3) \neq \mu(m_3)$ we have $\mu_M = (w_1, w_3, w_2)$. Since $\mu_W(m_2) = \mu(m_2) = w_2$, $\mu_W(m_1) \neq \mu(m_1)$ and $\mu_W(m_3) \neq \mu(m_3)$ we have $\mu_W = (w_3, w_2, w_1)$.

In fact, $S(P) \backslash \{\mu, \mu_M, \mu_W\} = \emptyset$. Suppose not. Let $\mu' \in S(P) \backslash \{\mu, \mu_M, \mu_W\}$. Then $\mu' \in \{(w_2, w_1, w_3), (w_2, w_3, w_1), (w_3, w_1, w_2)\}$. However, it can easily be checked that in all three cases $\mu \vee_M \mu'$ is not a well-defined matching, contradicting Theorem 2.1.

[6] The roles of m_2 and m_3 can be switched.

Finally, we calculate the probabilities p_{μ_M}, p_{μ_W}, and p_μ.[7] Note that after Step 2 of the EL algorithm only 2 agents remain, which hence will be matched to one another. Thus, it suffices to consider the sets $\hat{Q}_\mu, \hat{Q}_{\mu_M}, \hat{Q}_{\mu_W}$, where $\hat{Q}_\mu := \{(i_1, i_2) :$ there is an agent i_3 s.t. $(i_1, i_2, i_3) \in Q_\mu\}$ (the sets $\hat{Q}_{\mu_M}$ and $\hat{Q}_{\mu_W}$ are defined similarly).[8] One easily verifies that $\hat{Q}_{\mu_M} = \{(m_1, m_2),$ $(m_1, m_3),$ $(m_2, m_1),$ $(m_2, m_3),$ $(m_2, w_1),$ $(m_2, w_2),$ $(m_3, m_1),$ $(m_3, m_2),$ $(m_3, w_1),$ $(m_3, w_3)\}$, $\hat{Q}_{\mu_W} = \{(w_1, w_2),$ $(w_1, w_3),$ $(w_1, m_1),$ $(w_1, m_2),$ $(w_2, w_1),$ $(w_2, w_3),$ $(w_3, w_1),$ $(w_3, w_2),$ $(w_3, m_2),$ $(w_3, m_3)\}$, $\hat{Q}_\mu = \{(m_1, w_2),$ $(m_1, w_3),$ $(w_2, m_1),$ $(w_2, m_3)\}$. Thus, $|\hat{Q}_{\mu_M}| = 10 = |\hat{Q}_{\mu_W}|$ and $|\hat{Q}_\mu| = 4$. So, $p_{\mu_M} = p_{\mu_W} = \frac{10}{24} = \frac{5}{12}$ and $p_\mu = \frac{4}{24} = \frac{2}{12}$. □

Proof that $\mathcal{P}^* \neq \mathcal{P}$ in Example 4.2: We consider the marriage market (M, W, P) and the corresponding set of stable matchings $S(P) = \{\mu_1, \ldots, \mu_{10}\}$ as specified in Example 4.2. To prove that $\mathcal{P}^* \neq \mathcal{P}$ we calculate $\mathcal{P}^*$ by checking which stable matchings the random order mechanism induces for various sequences $(i_1, \ldots, i_8)$. Whenever we refer to a unique stable matching obtained for a marriage market not containing all agents, we calculated the man-optimal and the woman-optimal matching for the "submarket" using the deferred acceptance algorithm and detected that they coincide (this calculation is not included in the proof). Furthermore, whenever we "satisfy" a blocking pair, the (unique) proposing agent does not propose to agents that are better than his/her previous match (all these proposals would be rejected).

Case a: m_1 enters last; *i.e.*, the sequence of agents is $(i_1, \ldots, m_1)$. There are only two stable matchings μ' and μ'' when the set of agents consists of all women W and the remaining three men $\{m_2, m_3, m_4\}$:

$$\begin{array}{cccc} w_1 & m_2 & m_3 & m_4 \\ | & | & | & | \\ w_1 & w_2 & w_3 & w_4 \end{array} \qquad (\mu')$$

$$\begin{array}{cccc} w_1 & m_2 & m_3 & m_4 \\ | & | & | & | \\ w_1 & w_2 & w_4 & w_3 \end{array} \qquad (\mu'')$$

When m_1 enters last, he proposes to the single woman w_1, who accepts. So, matching μ' implies matching μ_{10} and matching μ'' implies μ_9.

[7]Recall that for all men $m_i \in M$, $\mu_M(m_i) \succeq_{m_i} \mu(m_i) \succeq_{m_i} \mu_W(m_i)$. Similarly, for all women $w_i \in W$, $\mu_M(m_i) \preceq_{w_i} \mu(w_i) \preceq_{w_i} \mu_W(w_i)$. Under the assumptions made in the proof without loss of generality, we can conclude that for μ to be reached with positive probability using the EL algorithm, the agents' preferences look as follows (by $*$ we indicate possible positions for the man/woman that is not specified in the preference lists of some agents such that the two agents that are underlined in the preference list of each agent are the best and worst partner to be matched to in a stable matching):

$$\begin{array}{rclllllll} P(m_1) & = & *, & \underline{w_1}, & *, & \underline{w_3}, & *, & m_1, & * \\ P(m_2) & = & *, & \underline{w_3}, & *, & \underline{w_2}, & *, & m_2, & * \\ P(m_3) & = & & \underline{w_2}, & w_3, & \underline{w_1}, & & m_3 & \\ P(w_1) & = & *, & \underline{m_3}, & *, & \underline{m_1}, & *, & w_1, & * \\ P(w_2) & = & *, & \underline{m_2}, & *, & \underline{m_3}, & *, & w_2, & * \\ P(w_3) & = & & \underline{m_1}, & m_3, & \underline{m_2}, & & w_3 & \end{array}$$

[8]Note that $|\hat{Q}_{\mu_M}| + |\hat{Q}_{\mu_W}| + |\hat{Q}_\mu| = 6 \cdot 4 = 24$.

Case a.1: m_2 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, m_2, m_1)$.

Case a.1.1: m_3 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, m_2, m_1)$. The unique stable matching before m_3, m_2, and m_1 enter matches m_4 to w_4 and everybody else to themselves. Next, when m_3 enters he proposes to w_3, who accepts. Similarly, when m_2 enters he proposes to w_2, who accepts. Thus, w_1 is single and the resulting matching is μ'. Hence, all 5! sequences induce μ_{10}.

Case a.1.2: m_4 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, m_2, m_1)$. Similarly as in *Case a.1.1*, all 5! sequences induce μ_{10}.

Case a.1.3: w_1 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_1, m_2, m_1)$. There are only two stable matchings $\tilde{\mu}'$ and $\tilde{\mu}''$ before w_1, m_2, and m_1 enter:

$$\begin{array}{ccc} w_2 & m_3 & m_4 \\ | & | & | \\ w_2 & w_3 & w_4 \end{array} \qquad (\tilde{\mu}')$$

$$\begin{array}{ccc} w_2 & m_3 & m_4 \\ | & | & | \\ w_2 & w_4 & w_3 \end{array} \qquad (\tilde{\mu}'')$$

It is easy to check that half of the partial sequences $(i_1, \ldots, i_5)$ with $\{i_1, \ldots, i_5\} \cap \{w_1, m_2, m_1\} = \emptyset$ result in $\tilde{\mu}'$, the other half in $\tilde{\mu}''$: if $[i_5 \in \{m_3, m_4\}]$ then $(i_1, \ldots, i_5)$ results in $\tilde{\mu}'$, if $[i_5 \in \{w_3, w_4\}]$ then $(i_1, \ldots, i_5)$ results in $\tilde{\mu}''$, if $[i_4 \in \{m_3, m_4\}$ and $i_5 = w_2]$ then $(i_1, \ldots, i_5)$ results in $\tilde{\mu}'$, and if $[i_4 \in \{w_3, w_4\}$ and $i_5 = w_2]$ then $(i_1, \ldots, i_5)$ results in $\tilde{\mu}''$. After agents w_1, m_2, and m_1 enter, $\tilde{\mu}'$ induces μ'. Similarly, $\tilde{\mu}''$ induces μ''. Hence, $\frac{5!}{2}$ sequences induce μ_9 and $\frac{5!}{2}$ sequences induce μ_{10}.

Case a.1.4: w_2 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_2, m_2, m_1)$. Similarly as in *Case a.1.3*, $\frac{5!}{2}$ sequences induce μ_9 and $\frac{5!}{2}$ sequences induce μ_{10}.

Case a.1.5: w_3 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_3, m_2, m_1)$. The unique matching before agents w_3, m_2, and m_1 enter matches m_3 to w_4, m_4 to w_2, and w_1 to herself. When w_3 enters she proposes to m_4, who accepts. Now w_2 is single. Next, when m_2 enters he proposes to w_2, who accepts. Thus, w_1 is single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.1.6: w_4 enters before m_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_4, m_2, m_1)$. Similarly as in *Case a.1.5*, all 5! sequences induce μ_9.

Summary Case a.1: 360 sequences $(i_1, \ldots, m_2, m_1)$ induce μ_9 and 360 sequences $(i_1, \ldots, m_2, m_1)$ induce μ_{10}.

Case a.2: m_3 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, m_1)$. The unique stable matching before m_3 and m_1 enter matches m_2 to w_2, m_4 to w_4, and w_1 and w_3 to themselves. When m_3 enters he proposes to w_3, who accepts. Thus, w_1 is single and the resulting matching is μ'. Hence, all 6! sequences induce μ_{10}.

Summary Case a.2: all 720 sequences $(i_1, \ldots, m_3, m_1)$ induce μ_{10}.

Case a.3: m_4 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, m_1)$. Similarly as in *Case a.2*, all 6! sequences induce μ_{10}.

Summary Case a.3: all 720 sequences $(i_1, \ldots, m_4, m_1)$ induce μ_{10}.

Case a.4: w_1 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, w_1, m_1)$.

Case a.4.1: m_2 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_2, w_1, m_1)$. Note that the marriage market before agents m_2, w_1, and m_1 enter is the same as in *Case a.1.3*. Similarly as in *Case a.1.3*, $\frac{5!}{2}$ sequences induce μ_9 and $\frac{5!}{2}$ sequences induce μ_{10}.

Case a.4.2: m_3 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, w_1, m_1)$. The unique stable matching before m_3, w_1, and m_1 enter matches m_2 to w_2, m_4 to w_4, and w_3 to herself. When m_3 enters he proposes to w_3, who accepts. Next, when w_1 enters all her proposals are rejected. Thus, w_1 is single and the resulting matching is μ'. Hence, all 5! sequences induce μ_{10}.

Case a.4.3: m_4 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, w_1, m_1)$. Similarly as in *Case a.4.2* all 5! sequences induce μ_{10}.

Case a.4.4: w_2 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_2, w_1, m_1)$. The unique stable matching before agents w_2, w_1, and m_1 enter matches m_2 to w_4, m_4 to w_3, and m_3 to himself. When w_2 enters she proposes to m_3, who accepts. Next, when w_1 enters she proposes to m_4, who rejects, then to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Thus, w_1 is single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.4.5: w_3 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_3, w_1, m_1)$. The unique stable matching before agents w_3, w_1, and m_1 enter matches m_2 to w_4, m_3 to w_2, and m_4 to himself. When w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Next, when w_1 enters she proposes to m_4, who rejects, then to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.4.6: w_4 enters before w_1 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_4, w_1, m_1)$. Similarly as in *Case a.4.5* all 5! sequences induce μ_9.

Summary Case a.4: 420 sequences $(i_1, \ldots, w_1, m_1)$ induce μ_9 and 300 sequences $(i_1, \ldots, w_1, m_1)$ induce μ_{10}.

Case a.5: w_2 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, w_2, m_1)$.

Case a.5.1: m_2 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_2, w_2, m_1)$. Note that the marriage market before agents m_2, w_2, and m_1 enter is the same as in *Case a.1.4*. Similarly as in *Case a.1.4*, $\frac{5!}{2}$ sequences induce μ_9 and $\frac{5!}{2}$ sequences induce μ_{10}.

Case a.5.2: m_3 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, w_2, m_1)$. The unique stable matching before m_3, w_2, and m_1 enter matches m_2 to w_1, m_4 to w_4, and w_3 to herself. When m_3 enters he proposes to w_3, who accepts. Next, when w_2 enters she proposes to m_3, who rejects, then to m_4, who rejects, and then to m_2, who accepts. Now w_1 is single, but all her proposals are rejected. Thus, the resulting matching is μ'. Hence, all 5! sequences induce μ_{10}.

Case a.5.3: m_4 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, w_2, m_1)$. Similarly as in *Case a.5.2* all 5! sequences induce μ_{10}.

Case a.5.4: w_1 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_1, w_2, m_1)$. The unique stable matching before agents w_1, w_2, and m_1 enter matches m_2 to w_4, m_4 to w_3, and m_3 to himself. When w_1 enters she proposes to m_4, who rejects, then to m_3, who accepts. Next, when w_2 enters she proposes to m_3, who rejects, then to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. Thus, w_1 is single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.5.5: w_3 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_3, w_2, m_1)$. The unique stable matching before agents w_3, w_2, and m_1 enter matches m_2 to w_4, m_4 to w_1, and m_3 to himself. When w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Next, when w_2 enters she proposes to m_3, who rejects, then to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. Thus, w_1 is single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.5.6: w_4 enters before w_2 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_4, w_2, m_1)$. Similarly as in *Case a.5.5* all 5! sequences induce μ_9.

Summary Case a.5: 420 sequences $(i_1, \ldots, w_2, m_1)$ induce μ_9 and 300 sequences $(i_1, \ldots, w_2, m_1)$ induce μ_{10}.

Case a.6: w_3 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, w_3, m_1)$.

Case a.6.1: m_2 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_2, w_3, m_1)$. The unique stable matching before agents m_2, w_3, and m_1 enter matches m_3 to w_4, m_4 to w_2, and w_1 to herself. When m_2 enters he proposes to w_2, who rejects, then to w_1, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_2 is single and proposes to m_2, who accepts. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.6.2: m_3 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, w_3, m_1)$. The unique stable matching before agents m_3, w_3, and m_1 enter matches m_2 to w_2, m_4 to w_4, and w_1 to herself. Next, when m_3 enters he proposes to w_4, who accepts. Now m_4 is single and proposes to w_2, who accepts. Now m_2 is single and proposes to w_1, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_2 is single and proposes to m_2, who accepts. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.6.3: m_4 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, w_3, m_1)$. The unique stable matching before agents m_4, w_3, and m_1 enter matches m_2 to w_2, m_3 to w_4, and w_1 to herself. When m_4 enters he proposes to w_4, who rejects, then to w_2, who accepts. Now m_2 is single and proposes to w_1, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_2 is single and proposes to m_2, who accepts. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.6.4: w_1 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_1, w_3, m_1)$. The unique stable matching before agents w_1, w_3, and m_1 enter matches m_2 to w_4, m_3 to w_2, and m_4 to himself. When w_1 enters she proposes to m_4, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.6.5: w_2 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_2, w_3, m_1)$. The unique stable matching before agents w_2, w_3, and m_1 enter matches m_2 to w_4, m_4 to w_1, and m_3 to himself. When w_2 enters she proposes to m_3, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.6.6: w_4 enters before w_3 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_4, w_3, m_1)$. There are only two stable matchings $\bar{\mu}'$ and $\bar{\mu}''$ before w_4, w_3, and m_1 enter:

$$\begin{array}{ccc} m_2 & m_3 & m_4 \\ | & | & | \\ m_2 & w_1 & w_2 \end{array} \tag{$\bar{\mu}'$}$$

$$\begin{array}{ccc} m_2 & m_3 & m_4 \\ | & | & | \\ m_2 & w_2 & w_1 \end{array} \tag{$\bar{\mu}''$}$$

Suppose that the sequence of agents is such that it induces $\bar{\mu}'$ before w_4, w_3, and m_1 enter. When w_4 enters she proposes to m_2, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_2 is single and proposes to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''.

Suppose that the sequence of agents is such that it induces $\bar{\mu}''$ before w_4, w_3, and m_1 enter. When w_4 enters she proposes to m_2, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''.

Hence, all 5! sequences induce μ_9.

Summary Case a.6: all 720 sequences $(i_1, \ldots, w_3, m_1)$ induce μ_9.

Case a.7: w_4 enters before m_1; *i.e.*, the sequence is $(i_1, \ldots, w_4, m_1)$.

Case a.7.1: m_2 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_2, w_4, m_1)$. The unique stable matching before agents m_2, w_4, and m_1 enter matches m_3 to w_1, m_4 to w_3, and w_2 to herself. When m_2 enters he proposes to w_2, who accepts. Next, when w_4 enters she proposes to m_2, who rejects, then to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.7.2: m_3 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_3, w_4, m_1)$. The unique stable matching before agents m_3, w_4, and m_1 enter matches m_2 to w_2, m_4 to w_3, and w_1 to herself. Next, when m_3 enters he proposes to w_3, who rejects, then to w_1, who accepts. Next, when w_4 enters she proposes to m_2, who rejects, then to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.7.3: m_4 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, m_4, w_4, m_1)$. The unique stable matching before agents m_4, w_4, and m_1 enter matches m_2 to w_2, m_3 to w_3, and w_1 to herself. When m_4 enters he proposes to w_3, who accepts. Now m_3 is single and proposes to w_1, who accepts. Next, when w_4 enters she proposes to m_2, who rejects, then to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.7.4: w_1 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_1, w_3, m_1)$. The unique stable matching before agents w_1, w_3, and m_1 enter matches m_2 to w_4, m_3 to w_2, and m_4 to himself. When w_1 enters she proposes to m_4, who accepts. Next, when w_3 enters she proposes to m_2, who rejects, then to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. Now w_1 is single and proposes to m_2, who rejects. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.7.5: w_2 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_2, w_4, m_1)$. The unique stable matching before agents w_2, w_4, and m_1 enter matches m_2 to w_3, m_4 to w_1, and m_3 to himself. When w_2 enters she proposes to m_3, who accepts. Next, when w_4 enters she proposes to m_2, who accepts. Now w_3 is single and proposes to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''. Hence, all 5! sequences induce μ_9.

Case a.7.6: w_3 enters before w_4 and m_1; *i.e.*, the sequence is $(i_1, \ldots, w_3, w_4, m_1)$. Note that the marriage market before agents w_3, w_4, and m_1 enter is the same as in *Case a.6.6.* Thus the only two stable matchings $\bar{\mu}'$ and $\bar{\mu}''$ are:

$$\begin{array}{ccc} m_2 & m_3 & m_4 \\ | & | & | \\ m_2 & w_1 & w_2 \end{array} \tag{$\bar{\mu}'$}$$

$$\begin{array}{ccc} m_2 & m_3 & m_4 \\ | & | & | \\ m_2 & w_2 & w_1 \end{array} \tag{$\bar{\mu}''$}$$

Suppose that the sequence of agents is such that it induces $\bar{\mu}'$ before w_3, w_4, and m_1 enter. When w_3 enters she proposes to m_2, who accepts. Next, when w_4 enters she proposes to m_2, who accepts. Now w_3 is single and proposes to m_4, who accepts. Now w_2 is single and proposes to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''.

Suppose that the sequence of agents is such that it induces $\bar{\mu}''$ before w_3, w_4, and m_1 enter. When w_3 enters she proposes to m_2, who accepts. Next, when w_4 enters she proposes to m_2, who accepts. Now w_3 is single and proposes to m_4, who accepts. Now w_1 is single and proposes to m_3, who accepts. Now w_2 is single and proposes to m_4, who rejects, then to m_2, who accepts. Now w_4 is single and proposes to m_3, who accepts. This leaves w_1 single and the resulting matching is μ''.

Hence, all 5! sequences induce μ_9.

Summary Case a.7: all 720 sequences $(i_1, \ldots, w_4, m_1)$ induce μ_9.

Summary Case a: 2640 sequences $(i_1, \ldots, m_1)$ induce μ_9 and 2400 sequences $(i_1, \ldots, m_1)$ induce μ_{10}.

Case b: m_2 enters last; *i.e.*, the sequence is $(i_1, \ldots, m_2)$.

Because of the symmetry of the preferences, by changing the roles of agents [m_1 and m_2], [w_1 and w_2], [m_3 and m_4], and [w_3 and w_4] in the proof of *Case a* we can show that in *Case b* 2640 sequences $(i_1, \ldots, m_2)$ induce μ_9 and 2400 sequences $(i_1, \ldots, m_2)$ induce μ_{10}.

Case c: m_3 enters last; *i.e.*, the sequence is $(i_1, \ldots, m_3)$. There are only two stable matchings $\hat{\mu}'$ and $\hat{\mu}''$ when the set of agents consists of all women W and the remaining three men $\{m_1, m_2, m_4\}$:

$$\begin{array}{cccc} m_1 & m_2 & w_3 & m_4 \\ | & | & | & | \\ w_1 & w_2 & w_3 & w_4 \end{array} \qquad (\hat{\mu}')$$

$$\begin{array}{cccc} m_1 & m_2 & w_3 & m_4 \\ | & | & | & | \\ w_2 & w_1 & w_3 & w_4 \end{array} \qquad (\hat{\mu}'')$$

When m_3 enters last, he proposes to the single woman w_3, who accepts. So, matching $\hat{\mu}'$ implies matching μ_{10} and matching $\hat{\mu}''$ implies μ_8.

In order to determine which sequences induce matchings μ_{10} and μ_8, we change the roles of agents [m_1 and m_3], [w_1 and w_3], [m_2 and m_4], and [w_2 and w_4] in the proof of *Case a*. Note that after this change, matching $\hat{\mu}'$ corresponds to μ' in the proof of *Case a* and matching $\hat{\mu}''$ corresponds to μ'' in the proof of *Case a*. Similarly, matching μ_8 corresponds to μ_9 in the proof of *Case a* and μ_{10} corresponds to μ_{10} in the proof of *Case a*.

Thus, changing the roles of the agents as specified above in the proof of *Case a* implies that in *Case c* 2640 sequences $(i_1, \ldots, m_3)$ induce μ_8 and 2400 sequences $(i_1, \ldots, m_3)$ induce μ_{10}.

Case d: m_4 enters last; *i.e.*, the sequence is $(i_1, \ldots, m_4)$.

Because of the symmetry of the preferences, by changing the roles of agents [m_3 and m_4], [w_3 and w_4], [m_1 and m_2], and [w_1 and w_2] in the proof of *Case c* we can show that in *Case d* 2640 sequences $(i_1, \ldots, m_4)$ induce μ_8 and 2400 sequences $(i_1, \ldots, m_4)$ induce μ_{10}.

Summary Cases a to d: Let $m \in M$. Then, 5280 sequences $(i_1, \ldots, m)$ induce μ_8, 5280 sequences $(i_1, \ldots, m)$ induce μ_9, and 9600 sequences $(i_1, \ldots, m)$ induce μ_{10}.

Let $w \in W$. Similarly to *Cases a* to *d*, 5280 sequences $(i_1, \ldots, w)$ induce μ_2, 5280 sequences $(i_1, \ldots, w)$ induce μ_3, and 9600 sequences $(i_1, \ldots, w)$ induce μ_1.

Finally, the probability distribution induced by the random order mechanism equals:

$$\begin{array}{ccccccccccc}
(p^*_{\mu_1}, & p^*_{\mu_2}, & p^*_{\mu_3}, & p^*_{\mu_4}, & p^*_{\mu_5}, & p^*_{\mu_6}, & p^*_{\mu_7}, & p^*_{\mu_8}, & p^*_{\mu_9}, & p^*_{\mu_{10}}) & = \\
(\frac{9600}{40320}, & \frac{5280}{40320}, & \frac{5280}{40320}, & 0, & 0, & 0, & 0, & \frac{5280}{40320}, & \frac{5280}{40320}, & \frac{9600}{40320}) & = \\
(0.238, & 0.131, & 0.131, & 0, & 0, & 0, & 0, & 0.131, & 0.131, & 0.238) & \neq \\
(0.25, & 0.125, & 0.125, & 0, & 0, & 0, & 0, & 0.125, & 0.125, & 0.25) & = \\
(\frac{1}{4}, & \frac{1}{8}, & \frac{1}{8}, & 0, & 0, & 0, & 0, & \frac{1}{8}, & \frac{1}{8}, & \frac{1}{4}) & = \\
(p_{\mu_1}, & p_{\mu_2}, & p_{\mu_3}, & p_{\mu_4}, & p_{\mu_5}, & p_{\mu_6}, & p_{\mu_7}, & p_{\mu_8}, & p_{\mu_9}, & p_{\mu_{10}}). & \square
\end{array}$$

References

Aldershof, B., Carducci, O.M., and Lorenc, D.C. (1999) "Refined Inequalities for Stable Marriage," *Constraints* **4**, 281-292.

Blair, C. (1984) "Every Finite Distributive Lattice is a Set of Stable Matchings," *Journal of Combinatorial Theory* (Series A) **37**, 353-356.

Cechlárová, K. (2002) "Randomized Matching Mechanism Revisited," mimeo, Institute of Mathematics, P.J. Šafárik University, Slovakia.

Gale, D. and Shapley, L.S. (1962) "College Admissions and the Stability of Marriage," *American Mathematical Monthly* **69**, 9-15.

Irving, R.W. and Leather, P. (1986) "The Complexity of Counting Stable Marriages," *SIAM Journal of Computing* **15**, 655-667.

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

Ma, J. (1996) "On Randomized Matching Mechanisms," *Economic Theory* **8**, 377-381.

Masarani, F. and Gokturk, S.S. (1989) "On the Existence of Fair Matching Algorithms," *Theory and Decision* **26**, 305-322.

McVitie, D.G. and Wilson, L.B. (1970) "Stable Marriage Assignments for Unequal Sets," *BIT* **10**, 295-309.

Moulin, H. (1997) "Procedural cum End State Justice: An Implementation Viewpoint," Justice, Political Liberalism and Utilitarianism, editors, Maurice Salles and John Weymark. Proceedings of a conference in honor of John Harsanyi and John Rawls, Cambridge University Press, (forthcoming).

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

Romero-Medina, A. (2002) "Equitable Selection in Bilateral Matching Markets," mimeo, Universidad Carlos III de Madrid.

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

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

Roth, A.E. and Vande Vate, J.H. (1990) "Random Paths to Stability in Two-Sided Matching," *Econometrica* **58**, 1475-1480.