Employment by Lotto Revisited*

Bettina Klaus† Flip Klijn‡

January 2006

Abstract

We study *employment by lotto* (Aldershof *et al.*, 1999), a procedurally fair matching algorithm for the so-called stable marriage problem. We complement Aldershof *et al.*'s (1999) analysis in two ways. First, we give an alternative and intuitive description of employment by lotto in terms of a probabilistic serial dictatorship on the set of stable matchings. Second, we show that Aldershof *et al.*'s conjectures are correct for small matching markets but not necessarily correct for large matching markets.

Keywords: employment by lotto, probabilistic mechanism, two-sided matching, stability.

1 Introduction

The so-called marriage model is concerned with (two-sided, one-to-one) matching markets where the two sides of the market, for instance, are firms and workers. A matching is then a partition of all firms and workers into pairs and unmatched firms and/or workers. Such a matching is "stable" if each firm and worker has an acceptable match, and no firm and worker prefer one another to their respective matches. Gale and Shapley (1962) were the first to formalize this notion of stability and provide an algorithm to calculate stable matchings. Their results inspired many researchers to study stability and its importance for real-life matching markets; see Roth and Sotomayor (1990) for an excellent survey.

An important property of the set of stable matchings is its lattice structure and, as a consequence, the polarization between stable matchings. In particular, there always exists a best stable matching for firms (workers) which is at the same time the worst stable matching for the workers (firms). Using Gale and Shapley's (1962) deferred acceptance algorithm both optimal matchings can be easily calculated. However, by choosing optimal stable matchings one side of the market is clearly favored over the other side. Masarani and Gokturk (1989) showed several impossibilities to obtain a fair deterministic matching mechanism within the context of Rawlsian justice based on cardinal preference information. One way to recover fairness is to use probabilistic (stable) matching mechanisms that are *ex ante* fair and/or 'procedurally fair;' see for instance Aldershof *et al.* (1999), Klaus and Klijn (2006), and Ma (1996).

*B. Klaus's and F. Klijn's research was supported by Ramón y Cajal contracts of the Spanish *Ministerio de Ciencia y Tecnología*. F. Klijn's work was also partially supported by the Barcelona Economics Program of CREA. This paper is part of the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme. This article reflects only the authors' views and the Community is not liable for any use that may be made of the information contained therein.

†Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; e-mail: B.Klaus@algec.unimaas.nl

‡Corresponding author. Institut d'Anàlisi Econòmica (CSIC), Campus UAB, 08193 Bellaterra (Barcelona), Spain; Tel. (34) 93 580 6612; Fax. (34) 93 580 1452; e-mail: Flip.Klijn@uab.es

In this paper we focus on employment by lotto, a probabilistic stable matching algorithm introduced by Aldershof *et al.* (1999) to avoid the inherent favoritism explained above. Aldershof *et al.* (1999, p. 281) introduced employment by lotto as a new procedure that combines a proposal algorithm and a refining process of the set of linear inequalities that describe the set of stable matchings. Our first contribution is to propose an alternative and easy description/interpretation of 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 (a firm or a worker)[1] is drawn randomly and can discard all stable matchings in which it is not matched to its best partner (possibly itself) in a stable matching. Exclude the first agent and its partner from the set of agents and randomly choose the next agent who can discard all stable matchings in which it is not matched to its 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 the same probability distribution on the set of stable matchings as employment by lotto.

Our second contribution is to address Aldershof *et al.*'s (1999, p. 288) conjectures on a symmetry and a monotonicity property of the probability distribution induced by employment by lotto. The first conjecture says that if the two sides of the market have the same cardinality then the two extreme stable matchings are equiprobable. We prove that this is indeed true if there are no more than three agents on each side of the market. We also show that for markets with more than three agents on each side the conjecture is not necessarily true. The second conjecture states that the probability distribution on the lattice of stable matchings is in some sense unimodal. We show that there is a problem with Aldershof *et al.*'s (1999) implicit assumption that the lattice can be partitioned in certain "natural" levels. Moreover, even for markets where this problem does not arise, the probability distribution may not be unimodal. In all our examples, we implement employment by lotto using Matlab ©.

The article is organized as follows. In Section 2 we introduce the matching model. In Section 3 we first recall Aldershof *et al.*'s (1999) definition of employment by lotto. Then, we describe our alternative random "serial-dictatorship" mechanism and prove its equivalence with employment by lotto. Finally, in Section 4 we present our results concerning Aldershof *et al.*'s (1999) conjectures.

2 The Matching Model

There are two finite and disjoint sets of agents of equal size:[2] a set $F = \{f_1, \ldots, f_n\}$ of firms and a set $W = \{w_1, \ldots, w_n\}$ of workers. Thus, $|F| = |W| = n$ and $|F \cup W| = 2n$. We denote a generic agent by i, a generic firm by f, and a generic worker 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, firm f's preferences $\succeq_f$ can be represented as a strict ordering $P(f)$ of the elements in $W \cup \{f\}$, for instance: $P(f) = w_3\, w_2\, f\, w_1 \ldots w_4$, which indicates that f prefers w_3 to w_2 and it prefers remaining single to any other worker. Similarly, worker w's preferences $\succeq_w$ can be represented as a strict ordering $P(w)$ of the elements in $F \cup \{w\}$. Let P be the profile of all agents' preferences: $P = (P(i))_{i \in F \cup W}$.

[1]To avoid the distinction between genders "he, she, and it," we simply refer to any agent as "it."

[2]Employment by lotto (and our alternative description in Section 3) also applies to unequal sizes. We make the assumption of equal size to simplify notation and because it is crucial for Aldershof *et al.*'s (1999) conjectures.

We write $w \succ_f w'$ if f strictly prefers w to w' ($w \neq w'$), and $w \succeq_f w'$ if f likes w at least as well as w' ($w \succ_f w'$ or $w = w'$). Similarly we write $f \succ_w f'$ and $f \succeq_w f'$. A worker w is *acceptable to a firm f* if $w \succ_f m$. Analogously, a firm f is *acceptable to a worker w* if $f \succ_w w$. Also, any agent is acceptable to itself.

A *matching market* is a triple (F, W, P). A matching for (F, W, P) is a function $\mu : F \cup W \to F \cup W$ such that for all $f \in F$ and $w \in W$ it holds that $\mu(f) = w \Leftrightarrow \mu(w) = f$, $\mu(f) \notin W \Rightarrow \mu(f) = f$, and $\mu(w) \notin F \Rightarrow \mu(w) = w$. If $\mu(f) = w$, then firm f and worker w are matched to one another. If $\mu(i) = i$, then agent i is *single*. We call $\mu(i)$ the *match of agent i* at μ. When denoting a matching μ we list the workers that are matched to firms $f_1, f_2, \ldots$; *e.g.*, $\mu = w_3, w_4, f_3, w_1$ denotes a matching where f_1 is matched to w_3, f_2 to w_4, f_3 to itself, and f_4 to w_1. Alternatively, a matching is denoted as a collection of matched agents and single agents; *e.g.*, $\{(f_1, w_3), (f_2, w_4), (f_4, w_1), f_3, w_2\}$ denotes matching $\mu = w_3, w_4, f_3, w_1$.

A key property of matchings is *stability*. First, since agents can always choose to be single, we require *individual rationality*: all matches are acceptable, *i.e.*, for all $i \in F \cup W$, $\mu(i) \succeq_i i$. Second, if an agent can improve upon its present match by switching to another agent such that this agent is better off as well, then this blocking clearly would cause instability. For a given matching μ, a firm-worker pair (f, w) is a *blocking pair* if they are not matched to one another but prefer one another to their current match at μ, *i.e.*, $w \succ_f \mu(f)$ and $f \succ_w \mu(w)$. A matching is *stable* if it is individually rational and if there are no blocking pairs. In view of stability, we henceforth denote a preference list, or list for short, by its list of acceptable agents. For instance, $P(f) = w_3\, w_2\, f\, w_1 \ldots w_4$ is denoted by $P(f) = w_3\, w_2\, f$. With a slight abuse of notation, we denote the set of stable matchings for matching market (F, 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 the function $\lambda := \mu \vee_F \mu'$ on $F \cup W$ that assigns to each firm its more preferred match from μ and μ' and to each worker its less preferred match: for all $f \in F$, $\lambda(f) := \mu(f)$ if $\mu(f) \succ_f \mu'(f)$ and $\lambda(f) := \mu'(f)$ otherwise, and for all $w \in W$, $\lambda(w) := \mu(w)$ if $\mu'(w) \succ_w \mu(w)$ and $\lambda(w) := \mu'(w)$ otherwise. Similarly, the function $\mu \wedge_F \mu'$ assigns to each firm its less preferred and to each worker its more preferred match. The following theorem (published by Knuth, 1976, but attributed to John Conway) establishes the lattice structure of the set of stable matchings.

Theorem 2.1 [Lattice Theorem] *If $\mu, \mu' \in S(P)$, then $\mu \vee_F \mu', \mu \wedge_F \mu' \in S(P)$.*

From Theorem 2.1 and the existence of a stable matching it follows easily that for any matching market (F, W, P) there is a stable matching μ_F^P that is optimal for all firms in the sense that no other stable matching μ gives to any firm f a match $\mu(f)$ that it prefers to $\mu_F^P(f)$. Similarly, there is a stable matching μ_W^P that is optimal for all workers. In fact, Gale and Shapley (1962) already proved the existence of μ_F^P and μ_W^P by providing an algorithm, the deferred acceptance algorithm, to calculate these matchings.

Since preferences are strict, the set of matched agents does not vary from one stable matching to another (Roth, 1982), *i.e.*, the set of single agents is the same for all stable matchings.

Theorem 2.2 *For all $i \in F \cup W$ and all $\mu, \mu' \in S(P)$, $\mu(i) = i \Rightarrow \mu'(i) = i$.*

3 Employment by Lotto

We first recall Aldershof *et al.*'s (1999) definition of employment by lotto. Next, we prove its equivalence with our alternative and intuitive definition which describes employment by lotto as a random serial dictatorship mechanism on the set of stable matchings.

Aldershof *et al.* (1999, p. 287) introduced employment by lotto as a new procedure that consists of repeated and alternate application of (full runs of) a so-called refinement algorithm and an adaptation of the set of linear inequalities that describe the set of stable matchings. Aldershof *et al.* (1999, pp. 283-284) provide a rather technical description of their refinement algorithm, but noticed (Aldershof *et al.*, 1999, p. 286) that a full run of the algorithm is equivalent to the construction of the profile of reduced preference lists described in Roth and Sotomayor (1990, pp. 61-62). Therefore, instead of using the original definition of employment by lotto, we take a short-cut by applying Roth and Sotomayor's reduced preference lists.

RED(P) - Reduction of preference lists (Roth and Sotomayor, 1990, pp. 61-62):
Input: A matching market (F, W, P).
Step 1: For all f, remove from f's list all w that are more preferred than $\mu_F(f)$.
Similarly, for all w, remove from w's list all f that are more preferred than $\mu_W(w)$.
Step 2: For all f, remove from f's list all w that are less preferred than $\mu_W(f)$.
Similarly, for all w, remove from w's list all f that are less preferred than $\mu_F(w)$.
Step 3: If worker w is not acceptable to firm f, then remove f from w's list.
Similarly, if firm f is not acceptable to worker w, then remove w from f's list.
Output: A profile of reduced preference lists $RED(P)$.

A reduction of the preference lists does not alter the set of stable matchings.

Lemma 3.1 *For any matching market* (F, W, P)*, if* $Q = RED(P)$*, then* $S(Q) = S(P)$.

Proof: Follows from Roth and Sotomayor's (1990) Proposition 3.10 with $\mu = \mu_W^P$ or $\mu = \mu_F^P$.□

Remark 3.2 Let (F, W, P) be a matching market and let $Q = RED(P)$ be the profile of reduced preferences. Note that for any $f \in F$, $\mu_F^P(f)$ is the most preferred agent for Q_f. Similarly, for any $w \in W$, $\mu_W^P(w)$ is the most preferred agent for Q_w. By Lemma 3.1, μ_F^P and μ_W^P are also stable for profile Q. Hence, $\mu_F^Q = \mu_F^P$ and $\mu_W^Q = \mu_W^P$.

Using the reduced preference lists, Aldershof *et al.*'s (1999) definition of employment by lotto boils down to the following procedure.

Algorithm I: Aldershof *et al.*'s (1999, p. 287) employment by lotto → EL*(P)
Input: A matching market (F, W, P). Set $\mu := \emptyset$, $N_1^* := F \cup W$, $Q^1 := P$, and $t := 1$.
Step *t*:
- Compute $P^t := RED(Q^t)$.
- Choose an agent i_t^* from N_t^* at random.
- Match agent i_t^* to its most preferred match $ch^*(i_t^*)$ in $P_{i_t^*}^t$ by setting $\mu := \mu \cup \{(i_t^*, ch^*(i_t^*))\}$.
- If $N_t^* \backslash \{i_t^*, ch^*(i_t^*)\} = \emptyset$, then set $EL^*(P) := \mu$ and stop.

Otherwise set $N^*_{t+1} := N^*_t \backslash \{i^*_t, ch^*(i^*_t)\}$ and define Q^{t+1} by
a) setting $Q^{t+1}(i^*_t) := ch^*(i^*_t) i^*_t$, *i.e.*, $ch^*(i^*_t)$ is the only acceptable agent for i^*_t, and
b) for all $i \in F \cup W \backslash \{i^*_t\}$, $Q^{t+1}(i) := P^t(i)$.
Go to Step $t := t + 1$.

Since the set of agents is finite, Algorithm I ends in a finite number r^* ($n \leq r^* \leq 2n$) of steps. As Aldershof *et al.* (1999) note, 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 all firms and workers are mutually acceptable, then $r^* = n$ and $q^* = 2n \cdot (2n-2) \cdot \ldots \cdot 2$. Algorithm I 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 $p^*_\mu = \frac{q^*_\mu}{q^*}$.

Below we give an alternative definition, which describes the employment by lotto algorithm as a random serial dictatorship mechanism on the set of stable matchings.

Algorithm II: employment by lotto as a serial dictatorship → EL(P)

Input: A matching market (F, W, P). Set $N_1 := F \cup W$, $S_1 := S(P)$, and $t := 1$.

Step *t*:
- Choose an agent i_t from N_t at random.
- Match agent i_t to its most preferred match $ch(i_t)$ in $\{j : j = \mu(i_t)$ for some $\mu \in S_t\}$.
- If $N_t \backslash \{i_t, ch(i_t)\} = \emptyset$, then set $\{EL(P)\} := S_t$ and stop.

Otherwise set $N_{t+1} := N_t \backslash \{i_t, ch(i_t)\}$, $S_{t+1} := \{\mu \in S_t : \mu(i_t) = ch(i_t)\}$, and go to Step $t := t+1$.

Since the set of agents is finite, Algorithm II ends in a finite number r ($n \leq r \leq 2n$) of steps. In fact, by Theorem 2.2, r only depends on the preferences (and hence all executions of the algorithm end in r steps). It is clear that 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 all firms and workers are mutually acceptable, then $r = n$ and $q = 2n \cdot (2n-2) \cdot \ldots \cdot 2$. Algorithm II 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 $p_\mu = \frac{q_\mu}{q}$.

Theorem 3.3 states that our alternative algorithm to define employment by lotto (Algorithm II) is indeed equivalent to Aldershof *et al.*'s (1999) original algorithm (Algorithm I).

Theorem 3.3 *For any matching market (F, W, P) and any stable matching $\mu \in S(P)$, $p^*_\mu = p_\mu$.*

Proof: Let (F, W, P) be a matching market. Note that any execution of Algorithm I generates a sequence of quadruples $(N^*_t, i^*_t, S(P^t), ch^*(i^*_t))$. Similarly, any execution of Algorithm II generates a sequence of quadruples $(N_t, i_t, S_t, ch(i_t))$.

We prove the theorem by showing that the algorithms generate the same sequences and that each sequence results in the same stable matching. In fact, we will show that for any sequence of quadruples $(N_t^*, i_t^*, S(P^t), ch^*(i_t^*))$ generated by an execution of Algorithm I there is an execution of Algorithm II that generates a sequence of quadruples $(N_t, i_t, S_t, ch(i_t))$ such that for each Step $t \geq 1$ of Algorithm I:

(i) $N_t = N_t^*$, and therefore we can indeed pick $i_t = i_t^*$,

(ii) $S_t = S(P^t)$, and

(iii) $ch(i_t) = ch^*(i_t^*)$.

It can be proven in a very similar way that any sequence of Algorithm II corresponds to a sequence of Algorithm I, and hence this is left to the reader.

Let $(N_t^*, i_t^*, S(P^t), ch^*(i_t^*))$ be a sequence of quadruples generated by an execution of Algorithm I. We will show that there is exactly one execution of Algorithm II that generates a sequence of quadruples $(N_t, i_t, S_t, ch(i_t))$ such that (i), (ii), and (iii) hold for each Step $t \geq 1$ of Algorithm I.

Step 1:

(1.i) Follows from the fact that both algorithms have as input $F \cup W = N_1 = N_1^*$. Let $i_1 = i_1^*$.

(1.ii) Next, since in Algorithm I, $P^1 = RED(Q^1) = RED(P)$, by Lemma 3.1, $S(P^1) = S(P)$. In Algorithm II, $S_1 := S(P)$. Hence, $S_1 = S(P^1)$.

(1.iii) In Algorithm I, $ch^*(i_1^*) = ch^*(i_1)$ equals agent i_1's most preferred match in $P^1_{i_1^*}$. Since $P^1 = RED(Q^1)$ and $Q^1 = P$, by Remark 3.2, $ch^*(i_1^*)$ is the most preferred match of agent i_1 at any matching in $S(P)$.

In Algorithm II, $ch(i_1)$ equals agent i_1's most preferred match in $\{j : j = \mu(i_t) \text{ for some } \mu \in S_1\}$. Since $S_1 = S(P)$, $ch(i_1)$ is the most preferred match of agent i_1 at any matching in $S(P)$.

Thus, $ch(i_1) = ch^*(i_1^*)$.

Induction Hypothesis: We assume that (k.i), (k.ii), and (k.iii) hold for $k = t - 1$.

Step *t*:

(t.i) Since in Step $t-1$, $N_{t-1} = N_{t-1}^*$, $i_{t-1} = i_{t-1}^*$, and $ch(i_{t-1}) = ch^*(i_{t-1}^*)$, we have $N_t = N_{t-1} \backslash \{i_{t-1}, ch(i_{t-1})\} = N_{t-1}^* \backslash \{i_{t-1}^*, ch^*(i_{t-1}^*)\} = N_t^*$. Hence, $N_t = N_t^*$.

(t.ii) Since in Step $t-1$, $S_{t-1} = S(P^{t-1})$, $i_{t-1} = i_{t-1}^*$, and $ch(i_{t-1}) = ch^*(i_{t-1}^*)$, we have $S_t = \{\mu \in S_{t-1} : \mu(i_{t-1}) = ch(i_{t-1})\} = \{\mu \in S(P^{t-1}) : \mu(i_{t-1}^*) = ch^*(i_{t-1}^*)\}$. Hence, it suffices to prove that $S(P^t) = S(Q^t) = \{\mu \in S(P^{t-1}) : \mu(i_{t-1}^*) = ch^*(i_{t-1}^*)\}$.

The first equality follows directly from Lemma 3.1 with $P = Q^t$ and $Q = P^t$.

We prove the second equality in two steps. First let $\mu \in S(P^{t-1})$ be such that $\mu(i_{t-1}^*) = ch^*(i_{t-1}^*)$. By construction of Q^t, μ is individually rational for Q^t. Suppose $\mu \notin S(Q^t)$. Then there is a blocking pair (f, w) for μ and Q^t. By construction of Q^t, (f, w) is a blocking pair for μ and P^{t-1}, a contradiction. Hence, $\mu \in S(Q^t)$.

Finally, let $\mu \in S(Q^t) = S(P^t)$. We have to prove that $\mu(i_{t-1}^*) = ch^*(i_{t-1}^*)$ and $\mu \in S(P^{t-1})$.

We first prove that $\mu(i_{t-1}^*) = ch^*(i_{t-1}^*)$. Suppose without loss of generality that $i_{t-1}^* \in W$. By construction, $\mu_W^{P^{t-1}}(i_{t-1}^*) = ch^*(i_{t-1}^*)$. Furthermore, $\mu_W^{P^{t-1}} \in S(P^{t-1})$ by definition and $\mu_W^{P^{t-1}} \in S(Q^t)$ by the construction of preference profile Q^t out of P^{t-1} (recall that only agent i_{t-1}^*'s preferences are changed to make its most preferred match $ch^*(i_{t-1}^*)$ the only acceptable agent – a preference change that does not affect the stability of $\mu_W^{P^{t-1}}$). Thus, there exists one stable matching for Q^t at which agent i_{t-1}^* is matched to agent $ch^*(i_{t-1}^*)$. Hence, since $ch^*(i_{t-1}^*)$ is the only acceptable agent for i^*_{t-1} by Theorem 2.2 (with $P = Q^t$), agents i^*_{t-1} and agent $ch^*(i^*_{t-1})$ are matched at all stable matchings in $S(Q^t) = S(P^t)$. In particular, $\mu(i^*_{t-1}) = ch^*(i^*_{t-1})$.

It remains to show that $\mu \in S(P^{t-1})$. Suppose to the contrary that $\mu \notin S(P^{t-1})$. By construction of Q^t, μ is individually rational for P^{t-1}. So there is a blocking pair (f, w) for μ and P^{t-1}. Hence, $w \succ_f^{P^{t-1}} \mu(f)$ and $f \succ_w^{P^{t-1}} \mu(w)$.[3] We distinguish between the following two cases:

CASE (A): $f, w \neq i^*_{t-1}$. Thus, $Q^t_f = P^{t-1}_f$ and $Q^t_w = P^{t-1}_w$. Hence, $w \succ_f^{Q^t} \mu(f)$, $f \succ_w^{Q^t} \mu(w)$, and (f, w) is also a blocking pair for μ and Q^t, a contradiction.

CASE (B): $f = i^*_{t-1}$ or $w = i^*_{t-1}$. Suppose without loss of generality that $f = i^*_{t-1}$. Since (f, w) is a blocking pair for μ and P^{t-1}, $w \succ_f^{P^{t-1}} \mu(f) = \mu(i^*_{t-1}) = ch^*(i^*_{t-1})$. However, by definition, $ch^*(i^*_{t-1})$ is f's most preferred match in P^{t-1}_f, a contradiction.

(t.iii) Note that $i_t = i^*_t$.

In Algorithm I, $ch^*(i^*_t) = ch^*(i_t)$ equals agent i_t's most preferred match in $P^t_{i_t}$. Since $P^t = RED(Q^t)$, by the definition of $RED(Q^t)$, $ch^*(i_t)$ is the most preferred match of agent i_t at any matching in $S(Q^t)$. Since $S(Q^t) = S(P^t)$, $ch^*(i_t)$ is the most preferred match of agent i_t at any matching in $S(P^t)$.

In Algorithm II, $ch(i_t)$ equals agent i_t's most preferred match in $\{j : j = \mu(i_t)$ for some $\mu \in S_t\}$. So, $ch(i_t)$ is the most preferred match of agent i_t at any matching in $S_t = S(P^t)$.

Thus, $ch(i_t) = ch^*(i_t) = ch^*(i^*_t)$. □

Aldershof *et al.* (1999) observe that if a stable matching μ does not match any agent to its firm/worker optimal match, then $p_\mu = 0$. More precisely, if for all $i \in F \cup W$ it holds that $\mu_F(i) \neq \mu(i) \neq \mu_W(i)$, 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 $i \in F \cup W$, $\mu_F(i) \neq \mu(i) \neq \mu_W(i)$.

Example 3.4 Let (F, W, P) with $n = 3$ and P listed below.[4] The three stable matchings for this market are listed below as well. We depict the corresponding lattice in Figure 1.

Preferences				Stable Matchings			
$P(f_1) =$	w_1	w_2	w_3	$\mu_W =$	w_3	w_2	w_1
$P(f_2) =$	w_3	w_1	w_2	$\widetilde{\mu} =$	w_2	w_1	w_3
$P(f_3) =$	w_2	w_3	w_1	$\mu_F =$	w_1	w_3	w_2
$P(w_1) =$	f_3	f_2	f_1				
$P(w_2) =$	f_2	f_1	f_3				
$P(w_3) =$	f_1	f_3	f_2				

The nodes denote the stable matchings and the corresponding numbers are the probabilities resulting from employment by lotto. The solid arcs denote comparability or unanimity on each side of the market. For instance, $\mu_W \rightarrow \widetilde{\mu}$ in Figure 1 means that all workers weakly prefer their matches at μ_W to their matches at $\widetilde{\mu}$ and all firms weakly prefer their matches at $\widetilde{\mu}$ to their matches at μ_W.

[3]We add the index P^{t-1} to avoid any confusion.

[4]Note that in all our examples of matching markets all workers (firms) are acceptable to all firms (workers). We do not have to rely on more general matching markets to prove our results. Furthermore, in order to save space we do not depict the agent itself in an agent's preference list.

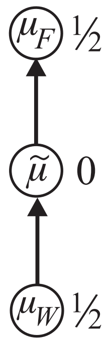

Figure 1: Lattice of Example 3.4

It is easy to check that whenever agent i_1 in the EL algorithm is a firm, then $EL(P) = \mu_F$, and whenever agent i_1 in the EL algorithm is a worker, then $EL(P) = \mu_W$. Hence, $p_{\mu_F} = \frac{1}{2} = p_{\mu_W}$ and $p_{\tilde{\mu}} = 0$.

In order to show that $p_\mu = 0$ does not necessarily imply that for all $i \in F \cup W$, $\mu_F(i) \neq \mu(i) \neq \mu_W(i)$, we add two agents f_4, w_4 to the market above such that for all $i = 1, 2, 3, 4$, $f_4 \succ_{f_4} w_i$, $w_4 \succ_{w_4} f_i$, and f_4, w_4 are placed anywhere in the preferences of the other agents. Then, for stable matching $\mu = w_2, w_1, w_3, f_4$, we have $p_\mu = 0$ and $\mu(f_4) = \mu_F(f_4)$ and $\mu(w_4) = \mu_W(w_4)$. $\diamond$

Note that in order to compute the probability distribution that is induced by employment by lotto, in general it is necessary to calculate the complete set of stable matchings. Gusfield and Irving (1989, Section 3.5, pp. 121) provided a time- and space-optimal algorithm for enumerating all stable matchings. Their algorithm needs $O(n^2 + n|\sum(P)|)$ total time and $O(n^2)$ space.

4 Aldershof *et al.*'s Conjectures

Aldershof *et al.* (1999, p. 288) made the following conjectures about the probability distribution $\mathcal{P}$ generated by employment by lotto over the set of stable matchings.

Conjecture 4.1 *For any matching market,* $p_{\mu_F} = p_{\mu_W}$.

Conjecture 4.2 *"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."*

The following two results show for which matching markets Conjecture 4.1 is correct and for which matching markets it is not necessarily true.

Theorem 4.3 *If $n > 3$, then not necessarily* $p_{\mu_F} = p_{\mu_W}$.

Proof: Let (F, W, P) with $n = 4$ and P such that[5]

$$\begin{array}{lcllll} P(f_1) & = & w_1 & w_2 & w_3 & w_4 \\ P(f_2) & = & w_2 & w_1 & w_4 & w_3 \\ P(f_3) & = & w_3 & w_4 & w_1 & w_2 \\ P(f_4) & = & w_4 & w_3 & w_2 & w_1 \\ P(w_1) & = & f_4 & f_2 & f_3 & f_1 \\ P(w_2) & = & f_3 & f_4 & f_1 & f_2 \\ P(w_3) & = & f_2 & f_1 & f_4 & f_3 \\ P(w_4) & = & f_1 & f_2 & f_3 & f_4. \end{array}$$

There are seven stable matchings and for the firm and worker optimal matchings we find $p_{\mu_F} = \frac{2}{8} \neq \frac{3}{8} = p_{\mu_W}$. (There are three other stable matchings with positive probability.) For $n > 4$ one can simply add agents that find any other agent on the other side of the market unacceptable.□

Theorem 4.4 *If $n \leq 3$, then $p_{\mu_F} = p_{\mu_W}$.*

Proof: From Theorem 2.2 it follows that the probability distribution on 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 $N_1 := F \cup W \backslash \{i \in F \cup W : \mu(i) = i \text{ for some } \mu \in S(P)\}$. *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 $F \cup W$.

Case $n = 1$: Since $\mu_F(f_1) = \mu_W(f_1) = w_1$, it follows immediately that $p_{\mu_F} = p_{\mu_W} = 1$.

Case $n = 2$: Clearly, $S(P) \subseteq \{\mu_1(f_1, f_2) = (w_1, w_2), \mu_2(f_1, f_2) = (w_2, w_1)\}$. So, $|S(P)| \leq 2$. If $|S(P)| = 1$, then $\mu_F = \mu_W$, and hence, $p_{\mu_F} = p_{\mu_W} = 1$. If $|S(P)| = 2$, then the first agent i_1 in the EL algorithm being a firm or a worker determines the resulting matching, and hence, $p_{\mu_F} = p_{\mu_W} = \frac{1}{2}$.

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

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

$$\begin{array}{lcl} Q^1_{\mu_F} & = & \{(i_1, i_2, i_3) \in Q : i_1 \in \overline{N} \cap F\}, \\ Q^1_{\mu_W} & = & \{(i_1, i_2, i_3) \in Q : i_1 \in \overline{N} \cap W\}, \\ Q^2_{\mu_F} & = & \{(i_1, i_2, i_3) \in Q : i_1 \notin \overline{N}, i_2 \in \overline{N} \cap F\}, \text{ and} \\ Q^2_{\mu_W} & = & \{(i_1, i_2, i_3) \in Q : i_1 \notin \overline{N}, i_2 \in \overline{N} \cap W\}.^6 \end{array}$$

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

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

(a) $i_1 \in F$ and $\mu(i_1) = \mu_F(i_1)$ or

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

---

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

[6]Note that $|Q^2_{\mu_F}| = |Q^2_{\mu_W}| = 0$ if and only if $|\overline{N} \cap F| = |\overline{N} \cap W| = 3$.

We consider Case (a) (Case (b) is proven similarly). Without loss of generality let $i_1 = f_1$.

First we show that at matching μ at most one firm can be matched to its firm optimal match. Assume there exist $i, j \in F$, $i \neq j$ such that $\mu(i) = \mu_F(i)$ and $\mu(j) = \mu_F(j)$. Then, $\mu = \mu_F$, a contradiction. Hence, $\mu(f_1) = \mu_F(f_1)$, $\mu(f_2) \neq \mu_F(f_2)$, $\mu(f_3) \neq \mu_F(f_3)$, and $i_2 \in W$. In the remainder of the proof we will denote $\mu = (\mu(f_1), \mu(f_2), \mu(f_3))$. Without loss of generality assume that $\mu = (w_1, w_2, w_3)$. Then by $\mu_F(f_1) = \mu(f_1)$ and the assumption that no agent is single we have $\mu_F = (w_1, w_3, w_2)$.

Next, we consider the case $\mu_F(f_1) = \mu(f_1)$, $\mu_F(f_2) \neq \mu(f_2) \neq \mu_W(f_2)$, and $\mu_F(f_3) \neq \mu(f_3) \neq \mu_W(f_3)$.

Since $\mu_W(f_2) \neq \mu(f_2)$, $\mu_W(f_2) \neq w_2$. Furthermore, $\mu_W(f_2) \neq \mu(f_2)$ implies $\mu_F(f_2) \succ_{f_2} \mu_W(f_2)$. Thus, $\mu_W(f_2) \neq w_3$. Hence, $\mu_W(f_2) = w_1$. However, applying the same arguments to agent f_3, we obtain $\mu_W(f_3) = w_1$ as well; a contradiction.

Now, the only case[7] that remains is $\mu_F(f_1) = \mu(f_1) \neq \mu_W(f_1)$ and, without loss of generality,[8] $\mu_F(f_2) \neq \mu(f_2) = \mu_W(f_2)$, and $\mu_F(f_3) \neq \mu(f_3) \neq \mu_W(f_3)$.

Recall that $\mu = (w_1, w_2, w_3)$ and $\mu_F = (w_1, w_3, w_2)$. Since $\mu_W(f_2) = \mu(f_2) = w_2$ and $\mu_W(f_1) \neq \mu(f_1)$ we have $\mu_W = (w_3, w_2, w_1)$.

In fact, $S(P) \backslash \{\mu, \mu_F, \mu_W\} = \emptyset$. Suppose not. Let $\mu' \in S(P) \backslash \{\mu, \mu_F, \mu_W\}$. 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 $\mu \vee_F \mu'$ is not a well-defined matching, contradicting Theorem 2.1.

Finally, we calculate the probabilities p_{μ_F}, p_{μ_W}, and p_μ.[9] 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 $\hat{Q}_\mu, \hat{Q}_{\mu_F}, \hat{Q}_{\mu_W}$, 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 $\hat{Q}_{\mu_F}$ and $\hat{Q}_{\mu_W}$ are defined similarly).[10] One easily verifies that $\hat{Q}_{\mu_F} = \{(f_1, f_2), (f_1, f_3), (f_2, f_1), (f_2, f_3), (f_2, w_1), (f_2, w_2), (f_3, f_1), (f_3, f_2), (f_3, w_1), (f_3, w_3)\}$, $\hat{Q}_{\mu_W} = \{(w_1, w_2), (w_1, w_3), (w_1, f_1), (w_1, f_2), (w_2, w_1), (w_2, w_3), (w_3, w_1), (w_3, w_2), (w_3, f_2), (w_3, f_3)\}$, $\hat{Q}_\mu = \{(f_1, w_2), (f_1, w_3), (w_2, f_1), (w_2, f_3)\}$. Thus, $|\hat{Q}_{\mu_F}| = 10 = |\hat{Q}_{\mu_W}|$ and $|\hat{Q}_\mu| = 4$. So, $p_{\mu_F} = p_{\mu_W} = \frac{10}{24} = \frac{5}{12}$ and $p_\mu = \frac{4}{24} = \frac{2}{12}$. □

Next, we consider Conjecture 4.2. 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 the marriage model. The following example demonstrates that this notion of natural

[7]We already know that $\mu(f_1) = \mu_F(f_1)$, $\mu(f_2) \neq \mu_F(f_2)$, and $\mu(f_3) \neq \mu_F(f_3)$. Now, if $\mu(f_i) = \mu_W(f_i)$ for some $i = 2, 3$, then we can assume that $\mu(f_j) \neq \mu_W(f_j)$ for all $j \neq i$. (Otherwise $\mu = \mu_W$, a contradiction.)

[8]The roles of f_2 and f_3 can be switched.

[9]Recall that for all f, $\mu_F(f) \succeq_f \mu(f) \succeq_f \mu_W(f)$. Similarly, for all w, $\mu_F(w) \preceq_w \mu(w) \preceq_w \mu_W(w)$. Under the assumptions made in the proof without loss of generality, $\mu_F = (w_1, w_3, w_2)$, $\mu = (w_1, w_2, w_3)$, and $\mu_W = (w_3, w_2, w_1)$. This allows us to 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 firm/worker that is not specified in the preference lists of some agents):

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

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

level/rank is not obvious at all. Given Blair's (1984) result that every lattice can be obtained as the set of stable matchings of some matching market, this result is not very surprising. In addition, the example also shows that even if two stable matchings are incomparable (*i.e.*, the firms are not unanimous on which of the two is better) they may still have different probabilities of resulting from employment by lotto.

Example 4.5 Let (F, W, P) with $n = 5$ and P listed below.[11] The six stable matchings for this market are listed below as well. We depict the corresponding lattice in Figure 2.

Preferences						Stable Matchings					
$P(f_1) =$	w_1	w_3	w_2	w_4	w_5	$\mu_1 =$	w_3	w_1	w_2	w_5	w_4
$P(f_2) =$	w_2	w_3	w_1	w_4	w_5	$\mu_2 =$	w_3	w_1	w_2	w_4	w_5
$P(f_3) =$	w_3	w_2	w_1	w_4	w_5	$\mu_3 =$	w_1	w_3	w_2	w_5	w_4
$P(f_4) =$	w_4	w_5	w_1	w_2	w_3	$\mu_4 =$	w_1	w_3	w_2	w_4	w_5
$P(f_5) =$	w_5	w_4	w_1	w_2	w_3	$\mu_5 =$	w_1	w_2	w_3	w_5	w_4
$P(w_1) =$	f_2	f_1	f_3	f_4	f_5	$\mu_6 =$	w_1	w_2	w_3	w_4	w_5
$P(w_2) =$	f_3	f_2	f_1	f_4	f_5						
$P(w_3) =$	f_1	f_2	f_3	f_4	f_5						
$P(w_4) =$	f_5	f_4	f_1	f_2	f_3						
$P(w_5) =$	f_4	f_5	f_1	f_2	f_3						

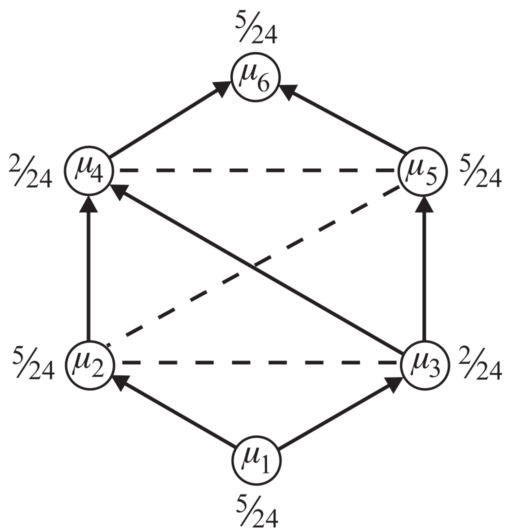

Figure 2: Lattice of Example 4.5

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 firms (workers) about which matching is better (for instance, $\mu_5(f_2) \succ_{f_2} \mu_4(f_2)$, but $\mu_4(f_4) \succ_{f_4} \mu_5(f_4)$).

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 $n > 3$ 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

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

Conjecture 4.2 needs not be unimodal (by the proof of Theorem 4.4 this part of the conjecture is true for $n \leq 3$).

Example 4.6 Let (F, W, P) with $n = 4$ and P listed below. The six stable matchings for this market are listed below as well. We depict the corresponding lattice in Figure 3.

Preferences					Stable Matchings				
$P(f_1) =$	w_1	w_2	w_4	w_3	$\nu_1 =$	w_3	w_4	w_1	w_2
$P(f_2) =$	w_2	w_1	w_3	w_4	$\nu_2 =$	w_4	w_3	w_1	w_2
$P(f_3) =$	w_3	w_4	w_1	w_2	$\nu_3 =$	w_4	w_1	w_3	w_2
$P(f_4) =$	w_4	w_3	w_1	w_2	$\nu_4 =$	w_2	w_3	w_1	w_4
$P(w_1) =$	f_3	f_2	f_1	f_4	$\nu_5 =$	w_2	w_1	w_3	w_4
$P(w_2) =$	f_4	f_1	f_2	f_3	$\nu_6 =$	w_1	w_2	w_3	w_4
$P(w_3) =$	f_1	f_2	f_3	f_4					
$P(w_4) =$	f_2	f_1	f_4	f_3					

Since $p_{\nu_2} = p_{\nu_5} = \frac{2}{48} < \frac{5}{48} = p_{\nu_3} = p_{\nu_4}$ the function f as defined in Conjecture 4.2 is not unimodal here. $\diamond$

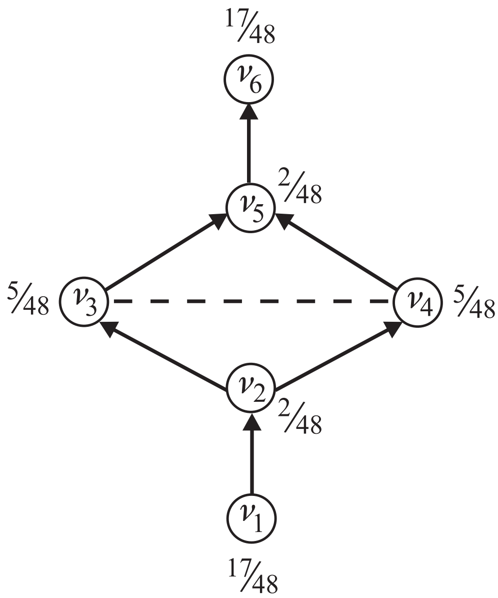

Figure 3: Lattice of Example 4.6

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.

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

Gusfield, D. and Irving, R.W. (1989) The Stable Marriage Problem: Structure and Algorithms. Cambridge: MIT Press.

Klaus, B. and Klijn, F. (2006) "Procedurally Fair and Stable Matching," *Economic Theory*, **27**, 431-447.

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

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.

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.