Multilateral Matching Under Dichotomous Preferences

Sophie Bade*

February 21, 2015

Abstract

For any multilateral matching problem with dichotomous preferences, in the sense that agents are indifferent between all acceptable teams, there exist some strategy proof, individually rational and efficient mechanisms. Moreover some random matching mechanisms simultaneously satisfy the three crucial desiderata of mechanism design: efficiency, fairness and strategyproofness, which typically conflict even in simple matching problems. The results hold for a novel criterion of fairness. I prove a non-existence result for the weakest fairness criterion that is standardly used: no ex-post Pareto optimal and individually rational mechanism for multilateral matching problem treats equals equally.

Keywords: . *JEL Classification Numbers*: C78.

*Royal Holloway College, University of London and Max Planck Institute for Research on Collective Goods, Bonn

1 Introduction

Are there any good mechanisms for multilateral matching problems that allow for teams of different sizes and compositions? Think of a generalized roommate problem: a university needs to match students to dorm rooms. The available rooms have different sizes, some have four beds others only two. There might be some additional restrictions. Girls might only be allowed to room with girls. Students in wheelchairs need rooms that are accessible to them. Minors are not allowed to share rooms with other minors. The prospects to find good mechanisms for such problems look pretty bleak when we allow for the full domain of preferences. Even if we restrict attention to classic roommate problems, where all rooms have two beds and where there are no restrictions on who may room with whom, there are no strategy proof, individually rational and efficient matching mechanisms. The same holds for marriage problems which only allow for teams one man and one woman each (Alcalde and Barbera [1]).

This non-existence problem disappears on the domain of dichotomous preferences which classify any two matchings under which the agent belongs an acceptable team as indifferent. Theorem 1 shows that there exist efficient, strategy proof and individually rational mechanisms. Such a mechanism can be defined to respect any given priorities of the designer, in the sense that it maps any profile of preferences to a matching that has highest priority among the set of all Pareto-optima. In the context of the above housing example, the designer might prioritise housing the maximum number of students or the maximum number of students in wheelchairs. Even fairness can be reconciled with strategy proofness and efficiency: Theorem 2 shows that there exists a strategy proof, individually rational, fair, and efficient random matching mechanisms for any multilateral matching problem with dichotomous preferences.

My definitions of strategy proofness, individual rationality and efficiency are standard: truthfully revealing one's preference is a dominant strategy in a strategy proof mechanism. Efficient and individually rational mechanisms map profiles of preferences to Pareto undominated matchings, in which no matched agent would rather be single. A lottery over individually rational matchings is efficient (ex-ante Pareto optimal) if there exists no other lottery over individually rational matchings that is weakly preferred by all and strictly some. A random matching mechanism is efficient if it matches every profile of preferences to an efficient lottery.

With respect to fairness I show that even in bilateral matching problems the standard requirement of equal treatment of equals conflicts with efficiency and individual rationality. Consider a marriage problem with one man, Alfred, and two women, Berta and Charlotte. Suppose that the two women would both like to marry Alfred, who only finds Berta acceptable. Equal treatment of equals requires that the two women, who have the same preference, would face the same probability of a marriage to Alfred. Given that Alfred would not voluntarily marry Charlotte, individual rationality requires that this probability is zero. Since Alfred and Berta would rather be a married, efficiency is violated. The notion of fairness used in Theorem 2 is not vulnerable to the above criticism. This notion requires that agents who are "relatively equal" in the sense that they are equal in terms of the environment as well as in the eyes of all other agents are treated equally and do not envy each other.

The present paper echoes some results by Bogomolnaia and Moulin [2] as well as Roth, Sönmez and Ünver [7] who have shown that there are some strategy proof, individually rational, efficient and fair mechanisms for marriage and classic roommate problems when preferences are dichotomous. My results apply to a much larger set of matching problems. There is no upper limit on the size of teams. Complementarities between teams, in the sense that the formation of some team is necessary for the formation of another team, are allowed. A team of children may, for example, only be allowed if there is also a team of adults. Moreover there may be any number of agent-types, not just one as in classic roommate problems or two as in marriage problems. We could for example think of matching patients to doctors and nurses and assume that teams have to contain at least one agent of each type. The case of matchings between agents and some objects who have no preferences of their own can also be accommodated. In terms of the medical care example we could require that each team consist of a doctor, a patient, a nurse and an operating theatre. The designer on the other hand may hold any priorities - as long as they do not conflict with Pareto ranking. A hospital operator might use medical urgency or profit maximization as his priority when matching doctors to nurses and patients.

The second major difference with respect to the work of Bogomolnaia and Moulin [2] and Roth, Sönmez and Ünver [7] is that I only use elementary mathematics to derive my results. My proof of Theorem 1 is reminiscent of Hatfield's [4] proof that there exist efficient, strategy proof and individually rational mechanisms for generalized kidney matching problems under dichotomous preferences. To prove the existence of a fair, efficient, individually rational and strategy proof mechanism (Theorem 2), I fix the set of all strategy proof, Pareto optimal and individually rational mechanism that maximize the number of matched agents. Theorem 1 ensures that this set is non-empty. To generate a fair mechanism I uniformly randomize over all mechanisms in this set.

Not all of Bogomolnaia and Moulin's [2] and Roth, Sönmez and Ünver's [7] results for bilateral matching problems extend to multilateral ones. Bogomolnaia and Moulin [2] and Roth, Sönmez and Ünver [7] did show that any individually rational Pareto optimal matching in a marriage or roommate problem is matches equally many agents. Theorem 4 shows that this feature is specific to bilateral matching problems. The fact that multilateral matching problems generically have individually rational Pareto optima of different sizes undermines the appeal of the the random priority and the egalitarian mechanism as defined by Bogomolnaia and Moulin [2] and Roth, Sönmez and Ünver [7]. The random priority mechanism matches agents according to a priority ordering that is drawn from a uniform distribution over all such orderings. The egalitarian mechanism maximises the the minimal matching probability. These two mechanisms are strategy proof and efficient in marriage and roommate problems. However in multilateral matching problems the egalitarian mechanism is not strategy proof while the random priority mechanism is not ex-ante Pareto optimal.

This last result relates to a question by Yilmaz [10] who extended the methods developed in Bogomolnaia and Moulin [2] and Roth, Sönmez and Ünver [7] to construct an efficient and egalitarian random matching mechanisms for kidney exchange problems that allow for cadaveric donations in addition to pairwise exchanges. Leaving the question whether this egalitarian and efficient mechanism is strategy proof open, Yilmaz [10] suggests the study of mechanisms that are not only efficient and egalitarian but also strategy proof as a research subject. I show that even on the domain of dichotomous preferences these three properties cannot be reconciled in multilateral matching problems.

Matching is not the only area of mechanism design in which dichotomous preferences allow for the existence of mechanisms with desirable properties that are out of reach on more general domains of preferences. Collective choice problems with dichotomous preferences have been studied by Bogomolnaia and Moulin [3], Vorsatz [8] and [9], and Maniquet and Mongin [5]. The outstanding feature of such models is that approval voting, where the winner has been declared acceptable by the largest set of voters, is strategy proof and Pareto optimal. Models with utilities that are quasi-linear in money have also been studied under the assumption that agent's preferences over assignments are dichotomous Mishra and Roy [6].

2 Definitions

There is a set of agents $N := \{1, \cdots, n\}$. A **matching** $\mu := \{\mu(i)\}_{i \in N}$ consists of a set of mutually disjoint teams $\mu(i) \subset N$. We either have $\mu(i) = \emptyset$, then i is **single** under μ, or $\mid \mu(i) \mid \geq 2$, then i belongs to the **team** $\mu(i)$. If j is a member of i's team $\mu(i)$, then i is a member of $j's$ team, so $i \in \mu(j)$ implies that $\mu(i) = \mu(j)$. The set of **matched** agents under μ is $M(\mu) := \cup_{i=1}^{n} \mu(i)$. When μ matches exactly one team T, I denote μ by T. The **size** of a matching $\mid \mu \mid$ is the number of all matched agents under μ. There is a set of **feasible matchings** $\mathcal{F}$. The matching $\emptyset$ under which all agents are single is feasible.

Agent i's preferences $\succsim_i$ over the set $\mathcal{F}$ are dichotomous in the sense that i is indifferent between any two $\mu, \mu' \in \mathcal{F}$ if either $\mu(i) = \mu'(i)$ or $\mu, \mu' \succ_i \emptyset$. So agent i's preference over matchings only depends on whether and which team he belongs too; he is indifferent between all teams that he considers acceptable. Under dichotomous preferences agent i is indifferent between a matching μ and $\emptyset$ if and only if he is single under μ ($\mu(i) = \emptyset$). The set Ω

is the set of all possible profiles of preferences $\succsim = (\succsim_i)_{i\in N}$. Any profile $\succsim$ can be represented as $(\succsim_i, \succsim_{-i})$ with the understanding that $\succsim_{-i}$ denotes the preferences of all agents but i.

Fix a profile of preferences $\succsim$. A matching μ is **individually rational** if $\mu \succsim_i \emptyset$ holds for all $i \in N$, meaning that under μ no agent does worse than being single; μ is **efficient (Pareto optimal)** if there exists no $\mu' \in \mathcal{F}$ such that $\mu \succsim_i \mu'$ for all $i \in N$ and $\mu \succ_{i^*} \mu'$ for some $i^* \in N$. An individually rational matching μ is Pareto optimal if there exists no individually rational $\mu' \in \mathcal{F}$ such that $M(\mu) \subsetneq M(\mu')$. The set of all individually rational Pareto optima (at the fixed $\succsim$) is denoted $\mathcal{F}(\succsim)$. The set of teams that are acceptable to their members at $\succsim$ is $A(\succsim)\colon = \{T \mid \mu(i) = T \Rightarrow \mu \succ_i \emptyset \text{ for all } i \in T\}$. A matching μ is individually rational and Pareto optimal if and only if there exists no $\mu' \subset A(\succsim)$ with $M(\mu) \subsetneq M(\mu')$. The designer has **priorities** $\succsim^D$ on $\mathcal{F}$. A binary relation $\succsim^D$ represents priorities if it is transitive, antisymmetric and **Pareto consistent** in the sense that $M(\mu) \subsetneq M(\mu')$ implies $\mu' \succ^D \mu$. The designer need not rank all possible matchings; $\succsim^D$ may be incomplete. However the rankings he does hold are rational in the sense that $\succsim^D$ is transitive. The designers priorities are aligned with the Pareto preference in the sense that he prefers to match larger sets of agents.

A (deterministic) **mechanism** $\phi : \Omega \to \mathcal{F}$ maps any profile $\succsim \in \Omega$ to a matching $\mu \in \mathcal{F}$. The mechanism ϕ is individually rational and Pareto optimal if $\phi(\succsim) \in \mathcal{F}(\succsim)$ holds for all $\succsim$; it **respects the the priority** $\succsim^D$ if $\phi(\succsim) \in \max_{\succsim^D} \mathcal{F}(\succsim)$ holds for all $\succsim \in \Omega$. The mechanism ϕ is **strategy proof** if $\phi(\succsim) \succsim_i \phi(\succsim_{-i}, \succsim_i')$ holds for all triples $(\succsim, \succsim_i', i)$. An individually rational mechanism ϕ is strategy proof if $\phi(\succsim)(i) = \emptyset$ implies $\emptyset \succsim_i \phi(\succsim_{-i}, \succsim_i')$ for all triples $(\succsim, \succsim_i', i)$.

A mechanism $\phi : \Omega \to \mathcal{F}$ is an f-mechanism if there exists an injective function $f : \mathcal{F} \to \mathbb{R}$ such that for any $\succsim \in \Omega$, $\phi(\succsim)$ maximizes $f(\mu)$ over all $\mu \in \mathcal{F}(\succsim)$. Since f is injective there exists a unique maximizer for every $\succsim$, so ϕ is well-defined. Any f-mechanism ϕ is individually rational and Pareto optimal given that $\phi(\succsim)$ is an element of $\mathcal{F}(\succsim)$ for all $\succsim \in \Omega$.

3 Deterministic Mechanisms

Theorem 1 shows that for any set of feasible matchings and any priorities of the designer there are some Pareto optimal, individually rational and strategy proof mechanisms that respect the designer's preference.

Theorem 1 *Fix a set of feasible matchings $\mathcal{F}$ and a some priorities $\succsim^D$. Then there exists a strategy proof, individually rational and efficient mechanism that respects $\succsim^D$.*

Proof Let $\phi : \Omega \to \mathcal{F}$ be an f-mechanism, where $f : \mathcal{F} \to \mathbb{R}$ is an injective function that represents $\succsim^D$ in the sense that $\mu \succ^D \mu'$ implies $f(\mu) > f(\mu')$ for all $\mu, \mu' \in \mathcal{F}$.[1] As an f-mechanism ϕ is individually rational and Pareto optimal. The mechanism ϕ respects $\succsim^D$ since f represents $\succsim^D$.

To see that ϕ is strategy proof, suppose it was not. That is, suppose that $\phi(\succsim)(i) = \emptyset$ and $\phi(\succsim_{-i}, \succsim'_i) \succ_i \emptyset$ held for some triple $(\succsim, \succsim'_i, i)$. Since $\phi(\succsim_{-i}, \succsim'_i) \succ_i \emptyset$ agent i cannot be single under the matching $\phi(\succsim_{-i}, \succsim'_i)$. Since ϕ is individually rational agent i must also according to $\succsim'_i$ prefer $\phi(\succsim_{-i}, \succsim'_i)$ to $\emptyset$. Given that agent i is single under $\phi(\succsim)$, he prefers $\phi(\succsim_{-i}, \succsim'_i)$ to $\phi(\succsim)$ according to $\succsim_i$ and $\succsim'_i$.

If $\phi(\succsim_{-i}, \succsim'_i) \succsim_j \phi(\succsim)$ holds for all $j \in N \setminus \{i\}$, then $\phi(\succsim_{-i}, \succsim'_i)$ Pareto-dominates $\phi(\succsim)$ at $\succsim$, contradicting $\phi(\succsim) \in \mathcal{F}(\succsim)$. So $\phi(\succsim) \succ_j \phi(\succsim_{-i}, \succsim'_i)$ must hold for some $j \in N \setminus \{i\}$. Given $\phi(\succsim_{-i}, \succsim'_i) \succ_i \phi(\succsim)$ and $\phi(\succsim_{-i}, \succsim'_i) \succ'_i \phi(\succsim)$, the matchings $\phi(\succsim_{-i}, \succsim'_i)$ and $\phi(\succsim)$ are neither Pareto ranked at $\succsim$ nor at $(\succsim_{-i}, \succsim'_i)$. If $\phi(\succsim_{-i}, \succsim'_i) \in \mathcal{F}(\succsim)$ then the definition of $\phi(\succsim)$ implies that $f(\phi(\succsim)) > f(\phi(\succsim_{-i}, \succsim'_i))$. If $\phi(\succsim_{-i}, \succsim'_i) \notin \mathcal{F}(\succsim)$ then there exists a $\mu \in \mathcal{F}(\succsim) \setminus \{\phi(\succsim)\}$ such that $M(\phi(\succsim_{-i}, \succsim'_i)) \subsetneq M(\mu)$. In this case the definition of $\phi(\succsim)$ implies that $f(\phi(\succsim)) > f(\mu)$. Since f represents the

[1] By Szpilrajn's theorem the order $\succsim^D$ has a linear extension $\succsim'$ in the sense that $\succsim'$ is antisymmetric, complete and transitive and $\mu \succ^D \mu'$ implies $\mu \succ' \mu'$ for all $\mu, \mu' \in \mathcal{F}$. Since $\mathcal{F}$ is finite there exists a function $f : \mathcal{F} \to \mathbb{R}$ that represents $\succsim'$. Since $\succsim'$ is antisymmetric f is injective.

- Pareto consistent - priorities $\succsim^D$ we also have $f(\mu) > f(\phi(\succsim_{-i}, \succsim_i'))$.[2] In sum we obtain $f(\phi(\succsim)) > f(\phi(\succsim_{-i}, \succsim_i'))$. Applying the same arguments mutatis mutandis to the definition of $\phi(\succsim_{-i}, \succsim_i')$ we obtain the contradiction $f(\phi(\succsim_{-i}, \succsim_i')) > f(\phi(\succsim))$. In sum, ϕ is strategy proof. □

An f-mechanism automatically satisfies individual rationality, Pareto optimality and respect for the designer's priorities. The fact that preferences are dichotomous is only used to show that $\phi(\succsim_{-i}, \succsim_i') \succ_i \phi(\succsim)$ implies $\phi(\succsim_{-i}, \succsim_i') \succ_i' \phi(\succsim)$ for all triples $(\succsim, \succsim_i', i)$. Any f-mechanism ϕ that satisfies this implication, is strategy proof. Hatfield [4] also uses f-mechanisms in to prove the existence of strategy proof, efficient and individually rational mechanisms. Differently from Hatfield's proof, which relies on some features of kidney exchanges, the present proof applies to all multilateral matching problems.

4 Random matching mechanisms

Fairness can only be achieved using random matching mechanisms. Some further concepts and notation are needed. A **random matching** is a lottery π on the set of feasible matchings $\mathcal{F}$, where $\pi(\mu)$ is the probability of μ under π. The set of all random matchings $\Delta\mathcal{F}$ naturally embeds the set $\mathcal{F}$ as the set of all degenerate lotteries. The degenerate random matching that assigns probability 1 to μ is denoted by μ. Agents are assumed to be expected utility maximizers over $\Delta\mathcal{F}$. The expected utility $U_i : \Delta\mathcal{F} \to \mathbb{R}$ **represents** $\succsim_i$ if $\mu \succsim_i \mu'$ implies $U_i(x) \geq U_i(y)$. Any such U_i is w.l.o.g. normalized to $U_i(\mu) = 1$ if $\mu \succ_i \emptyset$ and $U_i(\mu) = 0$ if $\mu(i) = \emptyset$.

A random matching π is individually rational (ex-post Pareto optimal) at $\succsim$ if all matchings that occur with positive probability according to π are

[2]To see that the Pareto consistency of $\succsim^D$ is necessary for this argument consider $N = \{1, 2, 3, 4, 5\}$ and $\mathcal{F}$ the set of all subsets of N with at least two members. Define a Pareto in-conistent $\succsim^D$ such that $\{3, 4\} \succ^D \{1, 2\} \succ^D \{3, 4, 5\}$. Let f represent $\succsim^D$ and let ϕ be an f-mechanism. Let $\succsim$ be such that $A(\succsim) = (\{1, 2\}, \{3, 4, 5\}, \{3, 4\})$ and therefore $\mathcal{F}(\succsim) = \{\{1, 2\}, \{3, 4, 5\}\}$. Since $\{1, 2\} \succ^D \{3, 4, 5\}$ we have $\phi(\succsim) = \{1, 2\}$; since $\{3, 4\} \succ^D \{1, 2\}$ we have $\phi(\succ_3', \succ_{-3}) = \{3, 4\}$ where $\{3, 4\}$ is the only acceptable team under $\succ_3'$. So we obtain $\phi(\succ_3', \succ_{-3}) \succ_3 \phi(\succ)(3) = \emptyset$ and ϕ is not strategy proof.

individually rational (Pareto optimal). So π is individually rational and ex post Pareto optimal at $\succsim$ if $\pi(\mu) > 0$ implies $\mu \in \mathcal{F}(\succsim)$. The assumption of dichotomous preferences together with the normalization U_i turn out to be convenient for the upcoming analysis. For any profile of preferences $\succsim$ and any individually rational π, $U_i(\pi)$ coincides i's matching probability under π.

An individually rational random matching π is **ex ante Pareto optimal (efficient)** at $\succsim$ if there exists no other individually rational random matching π' such that every agent's matching probability under π' is at least as high as his matching probability under π, and at least one agent faces a higher matching probability under π' than under π.[3] A **random matching mechanism** ψ maps the set of preference profiles Ω to the set of lotteries over feasible matchings $\Delta\mathcal{F}$. The random matching mechanism $\psi : \Omega \to \Delta\mathcal{F}$ is **ordinally strategy proof** if $U_i(\psi(\succsim)) \geq U_i(\psi(\succsim_{-i}, \succsim_i'))$ holds for all triples $(\succsim, \succsim_i', i)$ and all U_i that represent $\succsim_i$. So ψ is ordinally strategy proof if truthfully revealing $\succsim_i$ is a weakly dominant strategy for *any* expected utility over $\Delta\mathcal{F}$ that represents $\succsim_i$.

5 Permutations

Permutations $p : N \to N$ are used to exchange the agents names. The **identity**, $id : N \to N$ is the permutation that changes nothing, so $id(i) = i$. A **transposition**, denoted $(i, j) : N \to N$, is a permutation with $(i, j)(i) = j$, $(i, j)(j) = i$ and $(i, j)(i') = i'$ for all $i' \in N \setminus \{i, j\}$. Any transposition is its own inverse, so $(i, j) \circ (i, j) = id$ holds for all $i, j \in N$. When a permutation is applied to an $i \in N$ or a set $S \subset N$ I write $p(i)$ and $p(S)$ as is standard. The next paragraphs define the application of permutations to matchings, lotteries over matchings and profiles of preferences. In each of

[3] The reason to consider this constrained notion of ex ante Pareto optimality is the presumption that agents can avoid unacceptable teams. Consider a marriage problem with one man and one woman under the unconstrained notion of ex-ante Pareto optimality. Assume that the man would like to marry the woman who would rather stay single. According the unconstrained notion of ex-ante Pareto optimality any lottery over the marriage between the man and the woman and the alternative under which both are single is ex-ante Pareto optimal.

these nonstandard cases I write $p[x]$ when the permutation p is applied to the object x.

For any matching μ the **permuted matching** $p[\mu]$ is such that agent $p(i)$ assumes the role of agent i in the permuted matching; so $p[\mu]$ equals $\{p(\mu(i))\}_{i\in N}$. The definition of permuted matchings directly implies that $(p \circ q)[\mu] = p[q[\mu]]$ holds for all permutations p and q and all matchings μ. For any matching μ, the matching $(i,j)[\mu]$ is obtained from μ by swapping i and j.

Two agents in $i, j \in N$ are **equal (according to $\mathcal{F}$)** if $\mu \in \mathcal{F}$ implies $(i,j)[\mu] \in \mathcal{F}$ for any μ. To see that equality is a transitive relation, fix any three agents i, j, i' such that i and j as well as j and i' are equal. So we have that

$$\mu \in \mathcal{F} \Rightarrow (i,j)[\mu] \in \mathcal{F} \Rightarrow (i,i')[(i,j)[\mu]] \in \mathcal{F} \Rightarrow$$
$$(i,j)[(i,i')[(i,j)[\mu]]] = ((i,j) \circ (i,i') \circ (i,j))[\mu] = (i,i')[\mu] \in \mathcal{F}.$$

The first and third implication follow from i and j being equal, the second from j and i' being equal. The first equality follows from $(p \circ q)[\mu] = p[q[\mu]]$ the second from $(i,j) \circ (j,i') \circ (i,j) = (i,i')$. Ultimately we obtain that $\mu \in \mathcal{F} \Rightarrow (i,i')[\mu] \in \mathcal{F}$, which is none other than i and i' being equal. Equality is also reflexive, since $\mu \in \mathcal{F} \Rightarrow (i,i)[\mu] = id[\mu] = \mu \in \mathcal{F}$ holds for all μ. In sum, equality is an equivalence relation and the set of agents N can be partitioned into a set of n_e subsets $\{E_1, \cdots E_{n_e}\}$ of equal agents. A permutation p is **equality neutral** if any set of equal agents is mapped onto itself according to p, so $p(E_k) = E_k$ holds for all $1 \leq k \leq n_e$. For any $\mu \in \mathcal{F}$ and any equality neutral p, $p[\mu]$ is also an element of $\mathcal{F}$. Letting $\mathcal{E}$ be the set of all equality neutral permutations, define $|\mathcal{E}| :\ = \frac{1}{\epsilon}$.[4]

Now fix a permutation $p \in \mathcal{E}$. Define the **permutation of a probability** π over matchings such that the permuted matching $p[\mu]$ occurs with the same probability under $p[\pi]$ as the original matching does under π, so $p[\pi](p[\mu]) = \pi(\mu)$ holds for all $\mu \in \mathcal{F}$. For any profile of preferences $\succsim$ define a **permuted profile** $p[\succsim]$ such that both profiles describe the same preferences, the only difference being an exchange of names. In particular, $\mu p[\succsim]_i \mu'$ holds if and

[4] Classical roommate problems have only one set of equals, marriage problems have two (men and women). Note that $|\mathcal{E}| = |E_1| \times \cdots \times |E_{n_e}|$.

only if $p[\mu] \succsim_{p(i)} p[\mu']$ holds, so that agent i under $\succsim$ is renamed $p(i)$ under $p[\succsim]$. Two agents i and j have the same preferences over matchings if $(i,j)[\succsim] = \succsim$.

6 Fairness

One of the weakest fairness criteria, **equal treatment of equals**, requires that any two equal agents who truthfully reveal their identical preferences obtain the same utility. In multilateral matching problems individual rationality and Pareto optimality conflict with equal treatment of equals. Take a classic roommate problem with 3 - equal - agents $N = \{1, 2, 3\}$. Let $\succsim$ be such that 2 and 3 have identical preferences: both would (only) like to room with 1, 1 on the other hand only finds 2 acceptable. Given that $\{1, 2\}$ is the only individually rational Pareto optimum at $\succsim$, $\psi(\succsim) = \{1, 2\}$ must hold for any individually rational and ex post Pareto optimal ψ. No such ψ satisfies equal treatment of equals: agents 2 and 3 announce the same preferences but end up with different utilities.[5]

In sum we cannot require that popular agents, such as agent 2 in the preceding example, and unpopular agents, such as agent 3, face the same chances to be matched under an individually rational and Pareto optimal mechanism. What we can require though, is that agents who face the same restrictions by the environment ($\mathcal{F}$) and by their popularity ($\succsim$) are treated equally. Two equal agents i and j are **relatively equal (according to**

[5]This example proves Bogomolnaia and Moulin's [2] claim (on p.272) that their egalitarian and random priority solutions satisfy no-envy wrong. A mechanism ϕ satisfies no envy if no agent i prefers to outcome of any equal agent. The egalitarian and the random priority solution are both Pareto optimal and individually rational. So both mechanisms must map $\succsim$ in the above example to $\{1, 2\}$, meaning that 3 envies 2 under both solutions.

$\mathcal{F}(\succsim)$) if $\mu \in \mathcal{F}(\succsim)$ implies $(i,j)[\mu] \in \mathcal{F}((i,j)[\succsim])$ for any μ.[6] So two agents i, j are relatively equal if they are viewed as exchangeable by any team that they would like to belong to. In terms of the set of acceptable teams, i and j are relatively equal if and only if $T \in A(\succsim) \Rightarrow (i,j)(T) \in A((i,j)[\succsim])$. To see that relative equality is not transitive consider a classic roommate problem with just 3 agents and one room. Suppose agents 1 and 2 are happy to enter into any match. However agent 3 is only willing to room with agent 1. Agents 2 and 3 are relatively equal since 1 treats them equally, by the same token agents 1 and 3 are relatively equal. However agents 1 and 2 are not relatively equal.

To obtain notions of fairness that are applicable to the present context, equal treatment and no envy are only imposed for relative equals. An individually rational random matching mechanism ψ satisfies **relative equal treatment of equals** if for any two relative equals i, j we have $\psi(\succsim)(\{\mu : \mu(i) = T\}) = \psi(\succsim)(\{\mu : \mu(j) = \{j\} \cup T \setminus \{i\}\})$ for any $T, \succsim$ such that $i \in T$ but $j \notin T$ and $(i,j)[\succsim] = \succsim$. So ψ treats relative equals equally, if any two relative equals who announce the same preferences face the same distributions over teams under ψ. The same mechanism satisfies **relative no envy**, if no agent would rather get the assignment of another relatively equal agent, so $U_i(\psi(\succsim)) \geq U_i((i,j)[\psi(\succsim)])$ holds for all $\succsim$, all relatively equal i, j and all U_i that represent $\succsim_i$. We cannot just limit ourselves to matching probabilities in this definition, since agent i might under $(i,j)[\psi(\succsim)]$ be matched in an unacceptable team.

A mechanism ψ is anonymous if the agents' names do not matter. If we take an anonymous mechanism ψ, a profile $\succsim$ and use an equality neutral

[6] Agent i and j's own preferences are used to determine whether they are relatively equal. Even if only one of two women in a marriage problem has received a proposal by man m^*, these two women might be relatively equal if neither one of them is interested in marrying m^*. The stronger notion of relative equality which calls agents i, j relatively equal at $\succsim^*_{-\{i,j\}}$ if $\mu \in \mathcal{F}((\succsim_i, \succsim_j, \succsim^*_{-\{i,j\}}))$ implies $(i,j)[\mu] \in \mathcal{F}((i,j)[(\succsim_i, \succsim_j, \succsim^*_{-\{i,j\}})])$ for all preferences $\succsim_i, \succsim_j$ does not have this flaw. Any two agents who are relatively equal according to this stronger notion are also relatively equal according to the notion defined in the text. Any statement that holds for relatively equal agents also applies to the subset of strongly relatively equal agents. Consequently Theorem 2 also applies to the stronger notion. While the weaker notion might be conceptually less appealing it is easier to state and allows for a more general result on fairness.

permutation p to rename a set of equal agents in this profile to $p[\succsim]$, then the outcome of ψ at the profile $p[\succsim]$ is obtained by using the same permutation p to rename agents in the outcome of ψ at the original profile $\succsim$. In sum ψ is **anonymous** if $\psi(p[\succsim]) = p[\psi(\succsim)]$ holds for all $\succsim$ and all equality neutral permutations p. To save on notation ψ is **fair** if it satisfies anonymity, relative equal treatment of equals and relative no envy. Following the next Lemma we only need to check for two properties.

Lemma 1 *Fix a set of feasible matchings $\mathcal{F}$ and an individually rational and ex post Pareto optimal random matching mechanism ψ. If ψ is anonymous then it treats relative equals equally.*

Proof Fix a profile of preferences $\succsim$ such that $\succsim = (i,j)[\succsim]$ holds for two relatively equal agents i, j. If ψ is anonymous, then $\psi(p[\succsim]) = p[\psi(\succsim)]$ holds for all equality neutral permutations p, so we have in particular that $\psi((i,j)[\succsim]) = (i,j)[\psi(\succsim)]$. Since $\succsim = (i,j)[\succsim]$ holds we have $\psi(\succsim) = (i,j)[\psi(\succsim)]$, so under ψ agents i and j face the same distribution over teams as required for equal treatment of equals. $\square$

7 A Strategy Proof, Fair and Ex-Ante Pareto optimal Mechanism

Strategy proofness, individual rationality, efficiency and fairness can be reconciled in any multilateral matching problem with dichotomous preferences.

Theorem 2 *For any $\mathcal{F}$ there exists an ordinally strategy proof, individually rational, efficient, and fair random matching mechanism.*

The proof is constructive. I fix an arbitrary strategy proof deterministic mechanism ϕ that maps any profile of preferences to an individually rational matching of maximal size. We know from Theorem 1 that such a mechanism exists. Using the set of all equality neutral permutations I generate a set of strategy proof mechanisms: for any $p \in \mathcal{E}$ I define a mechanism through the permutations of the agents' names in ϕ using p. A random matching mechanism is defined as the uniform randomization over all deterministic mechanisms constructed in the preceding step. Strategy proofness, individual rationality and the maximality of matchings are all robust to randomization, the random matching mechanism therefore satisfies these properties. The mechanism is efficient since all matchings that occur with positive probability are maximal. Fairness is achieved since all equal agents have the same chance to be assigned any one role in the mechanism. The detailed proof is in the Appendix.

8 Bossiness

A mechanism is non-bossy if any agent's change of preference announcement that has an effect on others also has an effect on the agent himself. This basic idea implies two different notions of non-bossiness: we can either consider the effect on matchings or on utilities. A mechanism ϕ is **allocation-non-bossy** if $\phi(\succsim)(i) = \phi(\succsim_{-i}, \succsim_i')(i)$ implies $\phi(\succsim) = \phi(\succsim_{-i}, \succsim_i')$ for all triples $(\succsim, \succsim_i', i)$. The same mechanism is **utility-non-bossy** if $\phi(\succsim) \sim_i \phi(\succsim_{-i}, \succsim_i')$ implies $\phi(\succsim) \sim_j \phi(\succsim_{-i}, \succsim_i')$ for all $(\succsim, \succsim_i', i, j)$. When all agents preferences are linear orders, then the two notions of non-bossiness for deterministic mechanisms coincide. With dichotomous preferences the two notions are markedly different.[7] In the following theorem I show that allocation-non-bossiness is compatible with strategy proofness, efficiency and fairness. However utility-non-bossiness is ruled out be the mere assumption of individual rationality and ex post Pareto optimality.

[7]Neither one of the two notions implies the other. The fact that allocation- does not imply utility-non-bossiness follows from Theorem 3. To see that utility- does not imply allocation-non-bossiness consider $N = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{F} = \{\mu^1, \mu^2, \mu^3\}$ with $\mu^1 := \{\{1, 2\}, \{3, 4, 5, 6\}\}$, $\mu^2 := \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$, and $\mu^3 := \{1, 6\}$. Define ϕ such that $\phi(\succsim) = \mu^1$ if $\mu^3 \succ_1 \emptyset$ and $\mathcal{F}(\succsim) = \{\mu^1, \mu^2\}$, $\phi(\succsim) = \mu^2$ if $\emptyset \succ_1 \mu^3$ and $\mathcal{F}(\succsim) = \{\mu^1, \mu^2\}$, and $\phi(\succsim) = \emptyset$ for all other $\succsim$. The mechanism is utility-non-bossy, since any preference announcement that yields a change in any agent's utility implies a change in all agents' utilities. The mechanism is however not allocation-non-bossy given that agent 1 stays in team $\{1, 2\}$ whether he declares $\{1, 6\}$ acceptable or not, whereas the teams that the other agents belong to depend on agent 1's announcement.

Theorem 3 *Fix a set $\mathcal{F}$ and some priorities $\succsim^D$. There exists a strategyproof, individually rational, allocation-non-bossy and Pareto optimal mechanism that respects $\succsim^D$. If there exist two equal agents i and j and a matching under which i but not j is single, then there exists no individually rational, efficient and utility-non-bossy mechanism.*

Proof To see the first claim, recall from proof of Theorem 1 that any f-mechanism ϕ where f represents $\succsim^D$ is strategyproof, individually rational, Pareto optimal and respects the priorities of the designer. To see that ϕ is allocation-non-bossy suppose that $\phi(\succsim)(i) = \phi(\succsim_{-i}, \succsim'_i)(i)$ and $\phi(\succsim) \neq \phi(\succsim_{-i}, \succsim'_i)$ held for some triple $(\succsim, \succsim_i, i)$. If $\phi(\succsim_{-i}, \succsim'_i) \succ_{j^*} \phi(\succsim)$ held for some $j^* \in N$ and $\phi(\succsim_{-i}, \succsim'_i) \succsim_j \phi(\succsim)$ for all $j \in N$, then $\phi(\succsim_{-i}, \succsim'_i)$ Pareto dominated $\phi(\succsim)$ at $\succsim$ and $(\succsim_{-i}, \succsim'_i)$ contradicting $\phi(\succsim) \in \mathcal{F}(\succsim)$. By the same arguments mutatis mutandis we can rule out that $\phi(\succsim) \succ_{j^*} \phi(\succsim_{-i}, \succsim'_i)$ holds for some $j^* \in N$ and $\phi(\succsim) \succsim_j \phi(\succsim_{-i}, \succsim'_i)$ for all $j \in N$. So either $\phi(\succsim) \sim_j \phi(\succsim_{-i}, \succsim'_i)$ holds for all $j \in N$ or there exist j^* and j° such that $\phi(\succsim_{-i}, \succsim'_i) \succ_{j^*} \phi(\succsim)$ and $\phi(\succsim) \succ_{j^\circ} \phi(\succsim_{-i}, \succsim'_i)$. In either case $\phi(\succsim)$ and $\phi(\succsim_{-i}, \succsim'_i)$ are not Pareto ranked by either $\succsim$ or $(\succsim_{-i}, \succsim'_i)$. The arguments of the last paragraph of the proof of Theorem 1 the yield the contradiction $f(\phi(\succsim)) > f(\phi(\succsim_{-i}, \succsim'_i))$ and $f(\phi(\succsim)) < f(\phi(\succsim_{-i}, \succsim'_i))$. In sum, ϕ is allocation-non-bossy.

To see the second claim, fix an individually rational and efficient mechanism ϕ,[8] a matching μ^* with $3 \in \mu^*(1)$ and $\mu^*(2) = \emptyset$ where 1 and 2 are equal, and a profile of preferences $\succsim$ where $\mu \succ_i \emptyset$ holds if and only if $\emptyset \neq \mu(i) \in \{\mu^*(i), (1,2)[\mu^*](i)\}$. So agents 1 and 2 respectively only find teams $\mu^*(1)$ and $(1,2)[\mu^*](2)$ acceptable. The remaining agents do not distinguish between agents 1 and 2: any such agent finds teams T and $(1,2)(T)$ either both acceptable or both unacceptable. Since $\mathcal{F}(\succsim) = \{\mu^*, (1,2)[\mu^*]\}$, $\phi(\succsim) \in \{\mu^*, (1,2)[\mu^*]\}$. Suppose w.l.o.g. that $\phi(\succsim) = (1,2)[\mu^*]$. Define $\succsim'_3$ such that $\mu \succ'_3 \emptyset$ if and only if $\mu(3) = \mu^*(3)$. Since $\mathcal{F}(\succsim_{-3}, \succsim'_3) = \{\mu^*\}$, $\phi(\succsim_{-3}, \succsim'_3) = \mu^*$ must hold. Since 3 finds μ^* and $(1,2)[\mu^*]$ acceptable we have that $U_3(\phi(\succsim)) = U_3(\phi(\succsim_{-3}, \succsim'_3)) = 1$. However $(1,2)[\mu^*] \succ_2 \emptyset$ and $\mu^*(2) = \emptyset$ imply $U_2(\phi(\succsim)) > U_2(\phi(\succsim_{-3}, \succsim'_3)) = 0$ in violation to utility-non-bossiness.

[8]This part of the proof can be extended to random matching mechanisms.

□

Any marriage problem with at least one man and two women satisfies the assumptions of the second part of Theorem 3. This part therefore generalizes Bogomolnaia and Moulin's [2] observation that no strategy proof, utility-non-bossy, individually rational, and efficient mechanism exists for marriage problems.

The proofs of Theorem 1 and Theorem 3 establish that any f-mechanism $\phi : \Omega \to \mathcal{F}$ is strategy proof, efficient, individually rational and allocation-non-bossy. So one might ask whether the converse also holds. Is any strategyproof, individually rational, allocation-non-bossy and efficient mechanism an f-mechanism?

To see that the answer is negative, let $N = \{1,2,3,4\}$ and let $\mathcal{F}$ consist of all teams containing exactly three agents, so $\mu \in \mathcal{F} \Leftrightarrow |\mu| = 3$. Let f' and f'' be two injections from $\mathcal{F}$ to $\mathbb{R}$ with

$$f'(\{1,2,3\}) > f'(\{1,2,4\}) > f'(\{1,3,4\}) > f'(\{2,3,4\})$$
$$f''(\{1,2,4\}) > f''(\{1,2,3\}) > f''(\{1,3,4\}) > f''(\{2,3,4\}).$$

Define the mechanism ϕ^* through $\phi^*(\succsim) = argmax_{\mu \in \mathcal{F}(\succsim)} f'(\mu)$ if $\{1,3,4\} \succ_1 \emptyset$ and $\phi^*(\succsim) = argmax_{\mu \in \mathcal{F}(\succsim)} f''(\mu)$ if $\emptyset \succ_1 \{1,3,4\}$. Since ϕ^* maps any $\succsim$ to some matching in $\mathcal{F}(\succsim)$ it is Pareto optimal and individually rational.

To see that ϕ^* is strategy proof consider a profile $\succsim$ with $\{1,3,4\} \succ_1 \emptyset$. Since the mechanism that maps any $\succsim$ to $argmax_{\mu \in \mathcal{F}(\succsim)} f'(\mu)$ is strategy proof and since $\phi^*(\succsim_{-i}, \succsim_i') = argmax_{\mu \in \mathcal{F}(\succsim_{-i}, \succsim_i')} f'(\mu)$ holds for any $i \neq 1$, no agent $i \neq 1$ has an incentive to deviate. If $\phi^*(u) \sim_1 \emptyset$, then 1 is single according to all matchings in $\mathcal{F}(\succsim)$, and agent 1 cannot improve his utility by announcing some $\succsim_1'$ given that all other agents announce $\succsim_{-1}$. If $\phi^*(\succsim) \succ_1 \emptyset$ agent 1 is already maximally satisfied, and does not have an incentive to deviate. The case of a profile $\succsim$ with $\emptyset \succ_1 \{1,3,4\}$ is dealt with in parallel. To see that ϕ^* is allocation-non-bossy, observe that $\mu \neq \mu'$ implies $\mu(i) \neq \mu'(i)$ for all $\mu, \mu' \in \mathcal{F}$ and all $i \in N$. Therefore $\phi^*(\succsim) \neq \phi^*(\succsim_{-i}, \succsim_i')$ must imply $\phi^*(\succsim)(i) \neq \phi^*(\succsim_{-i}, \succsim_i')(i)$ for any $(\succsim, \succsim_i, i)$. Consequently ϕ^* is allocation-non-bossy.

Now suppose there existed an injection f such that $\phi^*(\succsim) = argmax_{\mu \in \mathcal{F}(\succsim)} f(\mu)$. Let $\succsim$ be such that $\{1,2,3\} \succ_i \emptyset$ for $i = 1,2,3$, $\{1,2,4\} \succ_i \emptyset$ for $i = 1,2,4$, $\{1,3,4\} \succ_1 \emptyset$ and $\emptyset \succ_i \mu$ for all other i, μ. Let $\{1,2,4\} \sim'_1 \{1,3,4\} \succ'_1 \emptyset \succ'_1 \{1,3,4\}$. Given that $\mathcal{F}(\succsim) = \mathcal{F}(\succsim_{-1}, \succsim'_1)$ we have $argmax_{\mu \in \mathcal{F}(\succsim)} f(\mu) = argmax_{\mu \in \mathcal{F}(\succsim_{-1}, \succsim'_1)} f(\mu)$. However ϕ^* was defined such that $\phi^*(\succsim) = \{1,2,3\}$ and $\phi^*(\succsim_{-1}, \succsim'_1) = \{1,2,4\}$ a contradiction.[9]

9 The Special Case of Bilateral Matching

Bogomolnaia and Moulin [2] as well as Roth, Sönmez, and Ünver [7] draw on the Gallai-Edmonds theorem for bipartite graphs to prove some stronger results for the case that $\mathcal{F}$ is a marriage, classical roommate or housing problem. The Gallai Edmonds theorem implies that that all individually rational Pareto optima in a classical roommate problem have the same size. In the next Theorem I show that generic matching problems $\mathcal{F}$ do not have this feature. A problem $\mathcal{F}$ is **based** on a set of teams $\mathcal{T}$ if a matching is feasible if and only if it is composed of teams in $\mathcal{T}$, so $\mathcal{F} = \{\mu \mid T \in \mu \Rightarrow T \in \mathcal{T}\}$. Marriage and classical roommate problems are based on teams. The matching problem in which a team of children may only form if there is also a team of adults is not based on teams.

Theorem 4 *Fix a problem $\mathcal{F}$ that is based on a set of teams $\mathcal{T}$. Suppose there exist three teams T', T° and T^* in $\mathcal{T}$ such that $T' \cap T^* = \emptyset$ and T° intersects with T', T^* and $N \setminus (T', T^*)$. Then there exists a profile of preferences $\succsim$ with individually rational Pareto optima of different sizes.*

Proof First assume that $\mid T^\circ \mid \neq \mid T' \mid$. Define $\succsim$ such that $A(\succsim) = \{T', T^\circ\}$. Since $T' \cap T^\circ \neq \emptyset$ there are exactly two individually rational Pareto optima at $\succsim$: T' and T°, which are in turn by definition of different size. Next

[9]The current claim that there exist strategyproof, individually rational, allocation-non-bossy and efficient mechanisms that are not f-mechanisms contrasts with Hatfield's [4] Theorem 4 which seems to make the diametrically opposed claim. The different results arise out of two different definitions of non-bossiness. Hatfield's definition does not directly refer to the agents' preferences but only to the set of acceptable matchings ($A(\succsim)$ in the terminology of the present paper).

assume that $\mid T^\circ \mid=\mid T' \mid$. Define $\succsim$ such that $A(\succsim) = \{T', T^\circ, T^*\}$. Since $T^* \cap T' = \emptyset$, $\{T^*, T'\}$ is a feasible matching. Since $T^*, T' \in A(\succsim)$, $\{T^*, T'\}$ is individually rational. Since T° intersects with T' and T^* $\{T^*, T'\}$ is Pareto optimal. Since $T^\circ \in A(\succsim)$ and since $T^\circ \setminus (T^* \cup T') \neq \emptyset$, T° is also an individually rational Pareto optimum. In sum we have that $\mathcal{F}(\succsim) = \{\{T^*, T'\}, T^\circ\}$, where $\mid \{T^*, T'\} \mid=\mid T^* \mid + \mid T' \mid>\mid T' \mid=\mid T^\circ \mid$. □

It is crucial that the team T° has at least three members; one that also belongs to T' another that also belongs to T^* and yet another that belongs to neither T' nor T^*. Theorem 4 does not impose much else. The number of sets of equals does not matter. Whether all teams have the same size or not is irrelevant. Still team sizes alone do not determine whether all matchings in $\mathcal{F}(\succsim)$ are of the same size. Take the problem of allocating some boys and girls to one apartment with two double rooms. There are two sets of equals 4 girls $\{1, 2, 3, 4\}$ and 2 boys $\{5, 6\}$. Assume that it is not feasible for two children of opposite gender to occupy the apartment. Otherwise any combination of two children as roommates is feasible. If all children are compatible according to $\succsim$, then there are individually rational Pareto optima of different sizes: $\{\{1, 2\}, \{3, 4\}\}$ and $\{5, 6\}$. Conversely, if $\mathcal{T}$ consists of a set of nested teams, then all individually rational Pareto optima have the same size, even though there are some teams with more than two members.[10]

Problems in which any individually rational Pareto optimum is of the same size have the advantage that any randomization over Pareto optimal and individually rational mechanisms yields an efficient random matching mechanism. Bogomolnaia and Moulin [2] as well as Roth, Sönmez and Ünver [7] use this fact to define two mechanisms that elegantly combine strategy proofness, efficiency and fairness: the random priority and the egalitarian mechanism. However, in multilateral matching problems the random priority mechanism is not ex ante Pareto optimal, the egalitarian mechanism is not ordinally strategy proof.

To define the random priority mechanism, fix any permutation p on N, and let $\min p(\emptyset) := \infty$. Define an injective priority-function $f^p : \mathcal{F} \to \mathbb{R}$

[10]For example $N = \{1, 2, 3, 4, 5\}$ and $\mathcal{F} = \{\{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4, 5\}\}$. In this case $\mathcal{F}(\succsim)$ is a singleton for all $\succsim$. Similarly, if any feasible matching consists of exactly one team of some fixed size s then all individually rational Pareto optima have the same size.

such that $f^p(\mu) > f^p(\mu')$ holds if $\min p(M(\mu) \setminus M(\mu')) < \min p(M(\mu') \setminus M(\mu))$. If the two sets of agents matched by μ and μ' are not nested then the highest priority agent who is not matched by both μ and μ' is decisive. For any p let $\rho^p : \Omega \to \mathcal{F}$ be the corresponding f^p-mechanism. The **random priority mechanism** $\alpha : \Omega \to \Delta\mathcal{F}$ results from drawing a priority ranking p from the uniform distribution over all permutations on N to then determine a matching as the outcome of ρ^p at $\succsim$. So the probability of $\mu \in \mathcal{F}$ under $\alpha(\succsim)$ is $\alpha(\succsim)(\mu) = \frac{1}{n!}|\{p|\rho^p(\succsim) = \mu\}|$.

By the proof of Theorem 1 ρ^p is strategy proof, individually rational, and Pareto optimal for any p. As a randomization over strategy proof, individually rational and Pareto optimal mechanisms, α is ordinally strategy proof, individually rational, and ex post Pareto optimal. To see that α is not efficient consider a problem $\mathcal{F}^\alpha$ with six equal agents and one apartment that can be occupied by two or three of them. So $\mathcal{F}^\alpha$ is the set of all two or three agent subsets of $\{1, 2, \cdots, 6\}$. Now fix $\succsim$ such that $\mathcal{F}^\alpha(\succsim) = \{\{1, 2, 3\}, \{4, 5, 6\}, \{1, 4\}, \{2, 5\}, \{3, 6\}\}$. The symmetry of $\mathcal{F}^\alpha(\succsim)$ implies that the matchings $\{1, 2, 3\}$ and $\{4, 5, 6\}$ are equally likely under the random priority mechanism, so $\alpha(\succsim)(\{1, 2, 3\}) = \psi(\succsim)(\{4, 5, 6\}) \colon = x$. By the same logic we have that $\alpha(\succsim)(\{1, 4\}) = \alpha(\succsim)(\{2, 5\}) = \alpha(\succsim)(\{3, 6\}) \colon = y$. Since $\rho^p(\succsim)$ equals $\{1, 4\}$ for any p that assigns the highest and second highest priority to agents 1 and 4 we have that $\alpha(\succsim)(\{1, 4\}) = y > 0$. Any agent's probability of being matched is $x + y$. Since $2x + 3y = 1$, any agent is matched with a probability less than $\frac{1}{2}$ under $\alpha(\succsim)$. This means that at $\succsim$ every single agent prefers the lottery π which assigns probability $\frac{1}{2}$ to $\{1, 2, 3\}$ and to $\{4, 5, 6\}$ to $\alpha(\succsim)$.

The **egalitarian mechanism** $\beta : \Omega \to \Delta\mathcal{F}$ chooses the lottery over individually rational and Pareto optimal matchings that maximizes the minimal matching probability. If there are multiple such lotteries it picks the one that maximizes the second lowest matching probability and so forth. For any random matching π define $V(\pi)$ as the n-vector listing all agents' matching probabilities under π in increasing order, so V_1^π is the lowest matching probability of any agent under π, V_2^π is the second lowest (possibly equal) matching probability and so forth. Define a linear order $\succsim^R$ on the set of all lotteries $\Delta\mathcal{F}$ such that $k_d = \min\{k : V(\pi)_k \neq V(\pi')_k\}$ and $V(\pi)_{k_d} > V(\pi')_{k_d}$

imply $\pi \succ^R \pi'$.[11] The egalitarian mechanism $\beta : \Omega \to \Delta\mathcal{F}$ is defined by $\alpha(\succsim) = \max_{\succsim^R}(\Delta\mathcal{F}(\succsim))$.

Given that the egalitarian mechanism maps any $\succsim$ to a lottery in $\Delta\mathcal{F}(\succsim)$ it is individually rational. Its ex ante Pareto optimality follows from its definition via the maximization of utility vectors. To see that β is not ordinally strategy proof, consider a problem $\mathcal{F}^\beta$ with six equal agents and one apartment that may be occupied by 2, 3 or 4 agents. Let $\succsim$ be such that $\mathcal{F}^\beta(\succsim) = \{\{1,5\},\{2,3\},\{3,4\}\}$ and $\{1,2,3,4\} \succ_i \emptyset$ for $i = 2,3,4$. According to $\beta(\succsim)$ each of the Pareto optimal matchings is chosen with probability $\frac{1}{3}$. If agent 1 declares $\succsim'_1$ instead, where $\{1,2,3,4\} \succ'_1 \emptyset$ keeping all else equal, then we have $\mathcal{F}(\succsim_{-1},\succsim'_1) = \{\{1,5\},\{1,2,3,4\}\}$ and the egalitarian mechanism assigns probability $\frac{1}{2}$ to either one of these two matchings. The mechanism is not ordinally strategy proof given that agent 1 prefers the lottery $\frac{1}{2} : \{1,5\}$, $\frac{1}{2} : \{1,2,3,4\}$ to the lottery the lottery $\frac{1}{3} : \{1,5\}$ and $\frac{2}{3}$: single when he assigns utility $-.1$ to $\{1,2,3,4\}$.

References

[1] Alcalde J. and S. Barbera : "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems" *Economic Theory* , 4, (1994), 417-435.

[2] Bogomolnaia, A. and H. Moulin: " Random matching under dichotomous preferences," *Econometrica*, 72, (2004), 257-279.

[3] Bogomolnaia, A., H. Moulin and R. Stong: " Collective choice under dichotomous preferences," *Journal of Economic Theory*, 122, (2005), pp. 165-184.

[4] Hatfield, J.: "Pairwise Kidney Exchange: Comment," *Journal of Economic Theory*, 125, (2005), pp. 289-193.

[5] Maniquet, F. and P. Mongin: "Approval Voting and Arrow's Impossibility Theorem," *Social Choice and Welfare*, forthcoming.

[11] Szpilrajn's theorem guarantees the existence of such a linear order.

[6] Mishra, Debasis and Souvik Roy: "Implementation in Multidimensional Dichotomous Domains", *Theoretical Economics*, 8, (2013), 431-466.

[7] Roth, Tayfun Sönmez, and Alvin, Utku Ünver: "Pairwise Kidney Exchange" *Journal of Economic Theory*, 125, (2005), 151-188.

[8] Vorsatz, M.: "Approval voting on dichotomous preferences," *Social Choice and Welfare* 28, (2007), pp. 127-141.

[9] Vorsatz, M.: "Scoring rules on dichotomous preferences," *Social Choice and Welfare* 31, (2008), pp. 151-162.

[10] Yilmaz, O.: " Kidney Exchange: An Egalitarian Mechanism," *Journal of Economic Theory*, 146, (2011), pp. 592-618.

Appendix: Proof of Theorem 2

Fix an injection $f : \mathcal{F} \to \mathbb{R}$ that is increasing in the size of matchings so that $|\mu| > |\mu'|$ implies $f(\mu) > f(\mu')$. For any equality neutral permutation p define a mechanism $\phi^p : \Omega \to \mathcal{F}$ through $f(p[\phi^p(\succsim)]) \geq f(p[\mu])$ for all $\mu \in \mathcal{F}(\succsim)$. The mechanism ϕ^p is well-defined as the equality neutrality of p implies that that $p[\mu] \in \mathcal{F}$ holds for any $\mu \in \mathcal{F}$.

Recalling that there are $\frac{1}{\epsilon}$ different equality neutral permutations, define a random matching mechanism $\psi : \Omega \to \Delta\mathcal{F}$ through

$$\psi(\succsim)(\mu) = \epsilon|\{p|\phi^p(\succsim) = \mu\}|.$$

So the probability of any matching μ can be calculated as the probability of μ being the outcome of ϕ^p at $\succsim$ where p is drawn from a uniform distribution on the set of all equality neutral permutations $\mathcal{E}$.

Fix any equality neutral permutation p. The construction in the proof of Theorem 1 implies hat ϕ^p is individually rational and strategy proof. As a randomization over individually rational and strategy proof mechanisms, ψ is individually rational and ordinally strategy proof. Since f is increasing in the size of matchings and since ϕ^p is individually rational, $\phi^p(\succsim)$ maximizes the sum of all agents utilities $U_i(\mu)$ over all individually rational matchings μ. Since ψ maps any $\succsim$ to a lottery over welfare maximal matchings it is efficient.

To see that ψ satisfies relative no envy fix a profile of utilities $\succsim$ such that agents 1 and 2 are relatively equal according to $\mathcal{F}(\succsim)$. Then ψ satisfies no envy if the following inequality holds for any U_1 that represents $\succsim_1$:

$$U_1(\psi(\succsim)) \geq U_1((1,2)[\psi(\succsim)]).$$

Partition the set $\mathcal{E}$ of all neutrality equal permutations into the sets A, B, C, D, E, and F, defined by the following table, with the understanding that p belongs to set X if and only if $U_1(\phi^p(\succsim))$ and $U_1((1,2)[\phi^p(\succsim)])$ equal the values in the column for set X.[12]

	A	B	C	D	E	F
$U_1(\phi^p(u))$	1	0	1	0	0	1
$U_1((1,2)[\phi^p(u)])$	0	1	1	0	< 0	< 0

Since $U_1(\psi(\succsim)) = \epsilon(|A| + |C| + |F|)$ and $U_1((1,2)[\psi(\succsim)]) \leq \epsilon(|B| + |C|)$, $U_1(\psi(\succsim)) \geq U_1((1,2)[\psi(\succsim)])$ holds if $|A| \geq |B|$. To prove that $|A| \geq |B|$ holds I show that the function $g : B \to A$ with $g(p) = p \circ (1,2)$ is an injection from B to A. Since $p \in \mathcal{E}$ and since 1 and 2 are equal, $g(p) = p \circ (1,2)$ is an element of $\mathcal{E}$ too. Since two permutations $p, p' \in B$ are identical if and only if $p \circ (1,2) = p' \circ (1,2)$, the function g from B to $\mathcal{E}$ is injective. All that remains to be shown is that $p \circ (1,2)$ is an element of A for any $p \in B$.

Fix an arbitrary $p \in B$. To see that $\phi^{p\circ(1,2)}(\succsim)$ equals $(1,2)[\phi^p(\succsim)]$, one needs to recognise that $f((p \circ (1,2))[(1,2)[\phi^p(\succsim)]]) \geq f((p \circ (1,2))[\mu])$ holds for all $\mu \in \mathcal{F}(\succsim)$ and that $(1,2)[\phi^p(\succsim)]$ is an element of $\mathcal{F}(\succsim)$. To see the first part, observe that

[12] Since ϕ^p is individually rational, $U_1(\phi^p(\succsim)) \in \{1, 0\}$ holds for all p and we have that $A \cup \cdots \cup F = \mathcal{E}$. Since either $U_1(\phi^p(\succsim)) \neq U_1(\phi^{p'}(\succsim))$ or $U_1((1,2)[\phi^p(\succsim)]) \neq U_1((1,2)[\phi^{p'}(\succsim)])$ holds for any two p, p' belonging to different sets, we must have $p \neq p'$ for any such two permutations. In sum, the sets A, B, C, D, E, and F indeed partition $\mathcal{E}$.

$$
\begin{aligned}
f(p[\phi^p(\succsim)]) \geq f(p[\mu]) \text{ for all } \mu \in \mathcal{F}(\succsim) \Leftrightarrow \\
f((p \circ (1,2) \circ (1,2))[\phi^p(\succsim)]) \geq f((p \circ (1,2) \circ (1,2))[\mu]) \text{ for all } \mu \in \mathcal{F}(\succsim) \Leftrightarrow \\
f((p \circ (1,2))[(1,2)[\phi^p(\succsim)]]) \geq f((p \circ (1,2))[(1,2)[\mu]]) \text{ for all } \mu \in \mathcal{F}(\succsim) \Leftrightarrow \\
f((p \circ (1,2))[(1,2)[\phi^p(\succsim)]]) \geq f((p \circ (1,2))[\mu]) \text{ for all } (1,2)[\mu] \in \mathcal{F}((1,2)[\succsim]) \Leftrightarrow \\
f((p \circ (1,2))[(1,2)[\phi^p(\succsim)]]) \geq f((p \circ (1,2))[\mu]) \text{ for all } \mu \in \mathcal{F}(\succsim).
\end{aligned}
$$

The first line follows from the definition of ϕ^p. The second follows from $(1,2)$ being its own inverse. The third follows from $(p \circ q)[\mu] = p[q[\mu]]$ holding for all permutations p, q and all matchings μ. The fourth uses the fact that the set of all $\{\mu' : \mu' = (1,2)[\mu] \text{ for } \mu \in \mathcal{F}(\succsim)\}$ can be rewritten as $\{\mu' : (1,2)[\mu'] \in \mathcal{F}((1,2)[\succsim])\}$. The the relative equality of 1 and 2 (that holds if and only if $\mu \in \mathcal{F}(\succsim) \Leftrightarrow (1,2)[\mu] \in \mathcal{F}((1,2)[\succsim])$) implies the fifth line.

To see that $(1,2)[\phi^p(\succsim)]$ is an element of $\mathcal{F}(\succsim)$, observe that the relative equality of 1 and 2 together with the individual rationality of ϕ^p imply $\phi^p(\succsim) \sim_i (1,2)[\phi^p(\succsim)] \succsim_i \emptyset$ for all $i \neq 1, 2$. Since $p \in B$ we have that $U_1(\phi^p(\succsim)) = 0$, so agent 1 is unmatched under $\phi^p(\succsim)$. This in turn implies that agent 2 is unmatched under $(1,2)[\phi^p(\succsim)]$ and therefore assigns utility 0 to $(1,2)[\phi^p(\succsim)]$. At the same time $p \in B$ implies that $U_1((1,2)[\phi^p(\succsim)]) = 1$. In sum $(1,2)[\phi^p(\succsim)]$ is individually rational. Since f is increasing in the size of matchings, $\phi^p(\succsim)$ is a maximal individually rational matching. Since $(1,2)[\phi^p(\succsim)]$ has the same size as $\phi^p(\succsim)$ it is also a maximal individually rational matching and therefore Pareto optimal. In sum we have that $\phi^{p \circ (1,2)}(\succsim) = (1,2)[\phi^p(\succsim)]$. Since $U_1(\phi^{p \circ (1,2)}(\succsim)) = U_1((1,2)[\phi^p(\succsim)]) = 1$ and $U_1((1,2)[\phi^{p \circ (1,2)}(\succsim)]) = U_1(((1,2) \circ (1,2))[\phi^p(\succsim)]) = U_1(\phi^p(\succsim)) = 0$ we have $p \circ (1,2) \in A$. The mechanism ψ indeed satisfies relative no envy.

To see that ψ is anonymous, fix a profile of utilities $\succsim$, a matching μ^*, and two equality neutral permutations p, q such that $\phi^p(q[\succsim]) = q[\mu^*]$ holds. Note that $\phi^p(q[\succsim]) = q[\mu^*]$ implies $q[\mu^*] \in \mathcal{F}(q[\succsim])$ and thereby $\mu^* \in \mathcal{F}(\succsim)$ as well as

$$f(p[q[\mu^*]]) \geq f(p[\mu]) \text{ for all } \mu \in \mathcal{F}(q[\succsim]) \Rightarrow$$
$$f(p[q[\mu^*]]) \geq f(p[q[\mu]]) \text{ for all } q[\mu] \in \mathcal{F}(q[\succsim]) \Rightarrow$$
$$f((p \circ q)[\mu^*]) \geq f((p \circ q)[\mu]) \text{ for all } \mu \in \mathcal{F}(\succsim).$$

In sum we obtain that $\phi^p(q[\succsim]) = q[\mu^*]$ implies $\phi^{p \circ q}(\succsim) = \mu^*$. Since $p \mapsto p \circ q$ defines a bijection from $\mathcal{E}$ to $\mathcal{E}$ we have that

$$\psi(q[\succsim])(q[\mu^*]) = \epsilon|\{p \in \mathcal{E} \mid \phi^p(q[\succsim]) = q[\mu^*]\}| =$$
$$\epsilon|\{p \in \mathcal{E} \mid \phi^{p \circ q}(\succsim) = \mu^*\}| = \epsilon|\{p \in \mathcal{E} \mid \phi^p(\succsim) = \mu^*\}| = \psi(\succsim)(\mu^*).$$

The definition of $q[\psi(\succsim)]$ implies that $q[\psi(\succsim)](q[\mu^*]) = \psi(\succsim)(\mu^*)$. Since q and μ^* were arbitrarily chosen, $\psi(q[\succsim])(\mu) = q[\psi(\succsim)](\mu)$ holds for all equality neutral permutations q and all μ. So $\psi(q[\succsim]) = q[\psi(\succsim)]$ holds for any equality neutral q, which is none other than ψ being anonymous.