# When Pareto-optimal assignments are (not) payoff-equivalent?

Afshin Nikzad*

PRELIMINARY DRAFT

**Abstract**

We prove that in a random matching market, there exists an assignment with an average rank that is independent of the market size (i.e., constant average rank). The proof relies on techniques from random graph theory and the FKG inequality [Fortuin et al. 1971], a fundamental correlation inequality in statistical mechanics and probabilistic combinatorics. The FKG inequality is used here to keep track of the correlations between the objects' ranks on an agent's lists. We elaborate next that the non-existence of such correlations has significant consequences regarding the payoff-equivalency of a large family of assignments including Pareto-optimal assignments.

The existence of an assignment with a constant average rank means that the average rank under the Random Serial Dictatorship mechanism takes a heavy toll compared to the first-best. This is in apparent contrast with recent results that show the payoff-equivalency of Pareto-optimal assignments [Che and Tercieux 2018]. We account for this difference by showing that the equivalence result relies on the assumption that the idiosyncratic components of the utility of an agent for objects are iid. When the iid assumption is replaced with negative correlation, Pareto-optimal assignments can have very different payoff distributions. Strict ordinal preferences indeed encapsulate a notion of correlated idiosyncrasy.

Under the iid assumption, however, we show that a large family of assignments, including Pareto-efficient assignments as well as assignments with large Pareto-inefficiencies, are payoff-equivalent. We introduce *abundance of competent choices*, a likely phenomenon under the iid assumption, as a deriving force of such equivalencies. We show that this phenomenon also occurs in other setups, such as in Gale and Shapley's model of marriage markets [Gale and Shapley 1962].

*Department of Economics, UC Berkeley. Email: nikzad@berkeley.edu

# Contents

# 1 Introduction

We first show that in a random market with $n$ agents and $n$ objects, where each agent ranks all objects independently and uniformly at random, there exists an assignment (bijection) of objects to agents with a constant average rank. That is, there exists a constant $R$, independent of the market size ($n$), such that the object assigned to each agent has an average position of at most $R$ on her list (Theorem 3.1).

On the other hand, a well-known result by [Knuth 1996] shows that in the same market, the average rank of the assignment generated by the Random Serial Dictatorship mechanism, in which agents are ordered randomly and choose objects one by one in that order, is $\log n$. Comparing these two results implies that the average rank of the assignment takes a heavy toll under the Random Serial Dictatorship mechanism compared to the first-best.

We then contrast this gap between the two Pareto-optimal assignments to the payoff-equivalency result of [Che and Tercieux 2018], whose work imply that in the same market, Pareto-optimal assignments are asymptotically payoff-equivalent (up to relabeling of agents) if the utility of an agent from an object is independently and identically distributed (iid) across all agent-object pairs.[1]

One might presume that the difference between these findings stems from the fact that one of them is based on an ordinal notion of preferences and the other on a cardinal notion. We rule this out by showing that the gap between the Pareto-optimal assignments can also persist in cardinal utility models.[2]

What accounts for the difference is, rather, the correlation between the idiosyncratic components of an agent's utilities over the objects. Consider a common utility specification model for an agent $a$ having an object $o$, $u_a(o) = v_o + v_o^a$, which writes the agent's utility as the sum of a common-value component and an idiosyncratic component. When the idiosyncratic components are iid, [Che and Tercieux 2018] show that Pareto-optimal assignments are asymptotically payoff-equivalent under their assumptions. (The iid assumption leads to similar equivalence results in other markets, too, as discussed later.) On the other hand, when the idiosyncratic components are negatively correlated, Proposition 4.1 shows that Pareto-optimal assignments can have very different payoff distributions.

[1] As we will discuss, their result is more general. They show that when an agent's utility for an object is determined by a common-value component and an idiosyncratic component, with the idiosyncratic components being iid, then Pareto-optimal assignments are asymptotically payoff-equivalent up to renaming of agents (and therefore generate the same average utility for agents as well).

[2] This can happen, for instance, when the utility of each agent is determined by the rank of the object assigned to her in her ranking. (Section 4)

Strict ordinal preferences indeed encapsulate a notion of correlated idiosyncrasy: when an object is ranked first, any other object must be ranked second or worse. Such a notion of (negative) correlation is absent from cardinal utility models with iid idiosyncratic components, and its absence causes a phenomenon that we call *abundance of competent choices*.

This phenomenon is discussed in Section 4. To elaborate, we first focus on the case of zero common-value components. Then, abundance of competent choices simply means that each agent has "many choices" that provide her a utility close to her maximum conceivable utility.[3] Proposition 4.3 shows that this leads to asymptotic payoff-equivalency of a large family of assignments which includes Pareto-optimal assignments as well as assignments with large Pareto-inefficiencies.[4] In fact, in all of the assignments in this family, any agent asymptotically attains the maximum conceivable utility. This suggests that models with (negatively) correlated idiosyncratic components can be sharper in distinguishing between the assignments and capturing the assignment inefficiencies than models with iid idiosyncratic components.

The same holds for the case of non-zero common-value components. (Section E in the appendix shows this for the case of iid common-value components.) A brief intuition for why the addition of common-value components does not affect the above observations is that, in a market with $n$ agents and $n$ objects, common-value components always contribute a fixed amount to any assignment. For example, to maximize the assignment's average utility, one can ignore the common-value components and merely maximize the sum of idiosyncratic components.

Finally, in Section 5 we discuss the abundance of competent choices phenomenon in other markets by focusing on two well-studied setups: Gale and Shapley's model of marriage markets [Gale and Shapley 1962] and the *Secretary problem* [Bruss 2000]. In particular, we highlight the results of [Lee and Yariv 2018] which imply that in any stable assignment in a marriage market, any agent asymptotically attains the maximum conceivable utility when the idiosyncratic components are iid. Analogous to our previous findings, we show that this is no longer the case when the idiosyncratic components are negatively correlated (Proposition 5.1).

Theoretical models in economics sometimes leverage the iid idiosyncratic components assumption, possibly for its analytical simplicity (e.g., see [Kanoria and Saban 2015, Arnosti

[3]We say an agent's maximum conceivable utility is $\overline{u}$ when her valuation has support $[\underline{u}, \overline{u}]$.

[4]By an assignment with large Pareto-inefficiencies we mean an assignment in which a large fraction of agents can be made strictly better off without making any other agent worse off.

and Shi 2017, Che and Tercieux 2018, Lee and Yariv 2018], among others.)[5] Our findings suggest that when there is no abundance of competent choices in a marketplace, models with negatively correlated idiosyncratic components are likely to provide a better approximation of reality than models with iid idiosyncratic components.

We conclude this section by summarizing the findings discussed above. First, we establish the existence of Pareto-optimal assignments in random markets which have an average rank that is independent of the market size (i.e., constant average rank). This finding is of independent interest. From a technical perspective, the proof needs to keep track of the correlations between the objects' ranks (which are precisely why the payoff-equivalence result breaks here). To handle these correlations, we use two applications of the FKG inequality [Fortuin et al. 1971], a fundamental correlation inequality in statistical mechanics and probabilistic combinatorics.

The above result also shows that the average rank under the RSD mechanism takes a heavy toll compared to the first-best. We then compare this gap between the two Pareto-optimal assignments with the recent results about the payoff-equivalency of Pareto-optimal assignments [Che and Tercieux 2018]. We explain the apparent contradiction by showing that the payoff-equivalency hinges on the iid idiosyncratic components assumption. When this assumption is dismissed, then Pareto-optimal assignments can have very different payoff distributions. Indeed, some amount of negative correlation exists in ordinal utility models as the objects' ranks on an agent's (strict) preference list are "negatively correlated".

Furthermore, we show that the iid idiosyncratic components assumption leads to the payoff-equivalency of a large family of assignments, which includes Pareto-optimal assignments as well as assignments with large Pareto-inefficiencies. In all of the assignments in this family, any agent attains the maximum conceivable utility asymptotically. We observe similar patterns in other setups, such as in Gale and Shapley's model of marriage markets [Gale and Shapley 1962], and introduce *abundance of competent choices*, a likely phenomenon under the iid assumption, as a deriving force of such equivalencies. Our findings can also guide researchers' modeling choices depending on the presence or the absence of this phenomenon in the marketplaces they study.

The rest of the paper is organized as follows. Section 2 presents the main setup. Section 3 shows that in a random matching market there exists an assignment with a constant average rank. Section 4 compares this finding to the recent payoff-equivalency results and formally

[5] On the other hand, there are empirical work on two-sided matching markets which allow for arbitrary correlations between the idiosyncratic utility components of an agent for objects. E.g., see [Abdulkadiroglu et al. 2017]

introduces the *abundance of competent choices* notion. Section 5 discusses the abundance of competent choices in other setups. The related literature are reviewed throughout the paper.

## 2 Setup

Consider a market with a set of agents $A$ and a set of objects $O$. We use $n, m$ to denote $|A|, |O|$, respectively. Suppose each agent has a complete strict preference relation over objects, which we also call her preference list or ranking. The position of an object $o$ on the preference list of an agent $a$ is called the *rank* of the object for that agent, and is denoted by $r_a(o)$.

Each object has a capacity $c$. Throughout the paper we assume that $n = cm$. An assignment is a function $\mu : A \to O$ that assigns each agent to an object without assigning more than $c$ agents to any object. The average rank of an assignment $\mu$ is defined as

$$\overline{r}(\mu) = \frac{1}{|A|} \cdot \sum_{a \in A} r_a(\mu(a)).$$

The set of preference lists of all agents is called a *preference profile*. When $A, O$ are known from the context, we denote the set of all preference profiles by $\Pi$ and typically denote a member of it by $\pi$. For a preference profile $\pi$, let the *rank-optimal* assignment, $r^\star(\pi)$, be the assignment with the minimum average rank. Observe that, given any $\pi$, $r^\star(\pi)$ is a *Pareto-optimal* assignment, that is, there exists no other assignment which is (weakly) preferred by all agents to $r^\star(\pi)$. We use the terms Pareto-optimality and Pareto-efficiency interchangeably.

A *random market* is a market such as the one above in which agents' preference lists are drawn independently and uniformly at random. One of the notions of interest is the expected average rank of the rank-optimal assignment in a random market (i.e. $\mathbb{E}_{\pi \sim U}[r^\star(\pi)]$ where $U$ is the uniform distribution over $\Pi$.)

The *Random Serial Dictatorship* (RSD) mechanism is an assignment mechanism in which agents are ordered randomly and then choose an item from the remaining items, one by one, in that order. We call the assignment generated by RSD the *RSD assignment*.

**Asymptotic notations**

For two functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$ we write $f = o(g)$ when $\lim_{n\to\infty} f(n)/g(n) = 0$, $f = \omega(g)$ when $g = o(f)$, $f = \Omega(g)$ when $\lim_{n\to\infty} f(n)/g(n) > 0$, $f = O(g)$ when $g = \Omega(f)$, and $f = \Theta(g)$ when $f = O(g)$ and $g = O(f)$.

We say a sequence of events $E_1, E_2, \ldots$ happen *with high probability* (whp) when $\mathbb{P}[E_n]$ approaches 1 as $n$ approaches infinity.

# 3 Average rank of Pareto-optimal assignments

First, we show that the expected average rank of the rank-optimal assignment in a random market is bounded from above by a constant independent of the market size.

**Theorem 3.1.** *There exists a constant $R$ that bounds the expected average rank of the rank-optimal assignment from above in any random market.*

We sketch the proof in Section 3.1, where we also illustrate that the constant $R$ is quite small (less than 2). These findings hold for any capacity parameter $c$.

When the object capacities are 1, [Knuth 1996] shows that the expected average of the RSD assignment is almost equal to $\ln n$, and therefore approaches infinity as $n$ approaches infinity, in contrast to the average rank in the rank-optimal assignment. The gap between the rank-optimal assignment and the RSD assignment persists even when the object capacities are greater than 1; in particular, when $c = o(\log n)$.

**Proposition 3.2.** *The expected average rank of the RSD assignment is at least $\frac{\ln m - 1}{c}$.*

In Section 4, we show that the gap between the rank-optimal assignment and the RSD assignment is also present in cardinal utility models. On the other hand, Pareto-optimal assignments are shown to be asymptotically payoff-equivalent by [Che and Tercieux 2018]. This apparent contradiction has a simple explanation, which we discuss in Section 4.

## 3.1 Discussion of the theorem and proof ideas

To prove Theorem 3.1, we first show that the expected average rank of the rank-optimal assignment is smaller in a random market with $n$ agents and $n/c$ objects each with capacity $c$ than in a random market with $n$ agents and $n$ objects each with capacity 1. (Lemma B.5 in the appendix shows that this holds in a stronger sense: stochastic dominance of the rank distributions.) Given this fact, it suffices to prove the claim of the theorem for unit capacities: any constant $R$ that bounds the expected average rank in markets with unit capacities will also be a valid upper bound in markets with larger capacities.

Before presenting a proof sketch for the case of unit capacities, we discuss how large the constant $R$ can be. Our proof shows that $R < 7\frac{3}{4}$, whereas our simulations demonstrate that $R < 2$. (See Figure 1) The simulations are for markets with unit capacities. As discussed above, the upper bound for the case of unit capacities is also a valid upper bound for the general case.

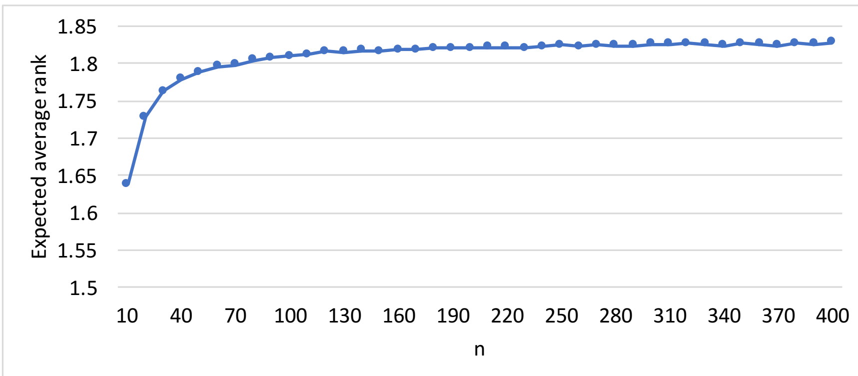


Figure 1: For each $n \in \{10, 20, 30, \ldots, 400\}$, we report the average over 1000 independent samples. The largest reported average is below 1.83. In addition, for $n = 1000$ and $n = 10000$ we take the average over 10 independent draws due to computational limitations; the simulations report average ranks 1.831 and 1.827, respectively.

The proof idea for the case of unit capacities is inspired by a result of [Walkup 1980] in random graph theory. We use two applications of the FKG inequality[6] to make a similar proof approach applicable. [Walkup 1980] shows that in a random bipartite graph with $n$ nodes on each side, where each node is connected to two distinct neighbors independently and uniformly chosen from the other side, there exists a perfect matching, with high probability.

We construct a bipartite graph $G(A, O)$ from a given preference profile $\pi$. Connect each node $a \in A$ to the first 3 items on the preference list of $a$. Also, connect each node in $O$ to "almost" 3 other neighbors in $A$, chosen as follows. For any $o \in O$, among the agents who rank $o$ after the third position, choose the 3 agents who rank $o$ the highest. Connect $o$ to those agents (on the other side of the bipartite graph).

We prove that $G(A, O)$ has a perfect matching with probability at least $1 - n^{-3}$ for all $n \geq 50$. The proof uses the König theorem, which implies that the size of the maximum

[6] We recall the definition of the FKG inequality in the appendix, Section A.1.

matching (number of its edges) plus the size of the maximum independent set[7] in a bipartite graph is equal to the total number of nodes. Therefore, to prove that $G(A, O)$ has a perfect matching, it suffices to show that it has no independent set of size $n+1$. We show that this holds with high probability using a union bound on all the subsets of nodes of size $n+1$.

In the graph that we construct, unlike in the graph in [Walkup 1980], the neighbors of nodes $o \in O$ are not chosen independently. In particular, there are two types of correlations involved: (i) Correlations across the partitions: The "almost" 3 neighbors that are chosen for a node in $O$ are not independent of the 3 neighbors that are chosen for a node in $A$, and (ii) Correlations within each partition: The "almost" 3 neighbors that are chosen for a node $o_1 \in O$ are not independent of the "almost" 3 neighbors that are chosen for a node $o_2 \in O$. The bulk of the proof involves handling these correlations, mainly using the FKG inequality.

These correlations are the reason behind the previously discussed gap between the rank-optimal assignment and the RSD assignment (after Theorem 3.1). In particular, when agents' utilities for objects are iid (for all agent-object pairs), then these correlations vanish, and all Pareto-optimal assignments, including the rank-optimal and RSD assignments, become payoff-equivalent as shown by [Che and Tercieux 2018].

# 4 Abundance of competent choices

First, we show that the gap between the rank-optimal assignment and the RSD assignment is also present in cardinal utility models. The cardinal model that we consider here is similar to our main setup in Section 2. The difference is that in here, each agent $a$ receives utility $u_a(o)$ from having an object $o \in O$. The *utility-optimal assignment* is then defined as the assignment that maximizes the sum of agents' utilities.

The next proposition compares *average utility*, defined as sum of the agents' utilities normalized by the number of agents, in the utility-optimal assignment and in the RSD assignment.

**Proposition 4.1.** *Consider a random market of $n$ agents and $n$ objects with unit capacities. An agent derives utility $v_1$ from having either of her top 3 objects and derives utility $v_2$ from having any other object, where $v_1 > v_2 \geq 0$. Then, the difference between the average utility in the utility-optimal assignment and the average utility in the RSD assignment is at least $\frac{v_1 - v_2}{20}$, with high probability.*[8]

[7]Recall that an independent set in a graph is a subset of its nodes which are pairwise non-adjacent.

[8]An event holds with high probability if it holds with probability that approaches 1 as $n$ approaches

There exists a non-vanishing gap between the utility-maximizing and the RSD assignments in the market considered above. On the other hand, any two Pareto-optimal assignments must be asymptotically payoff-equivalent under the assumptions of [Che and Tercieux 2018]. To explain this difference, we first recall a widely used cardinal utility specification that writes the utility of an agent $a$ from an object $o$ as $u_a(o) = v_o + v_o^a$, with $v_o$ being a common value component and $v_o^a$ being an idiosyncratic component. It is typical in the literature to assume that the idiosyncratic components are iid random variables (e.g., see [Lee and Yariv 2018, Che and Tercieux 2018]). However, in the market considered in Proposition 4.1, the idiosyncratic components are negatively correlated.[9] When the idiosyncratic components are iid, the gap between the utility-optimal assignment and the RSD assignment vanishes: these assignments would be asymptotically payoff-equivalent, as shown by [Che and Tercieux 2018].

To understand the intuition better, first consider the case of no common-value components and iid idiosyncratic components: suppose that $v_o = 0$ for all objects $o$ and that the random variables $v_o^a$ are iid with support $[\underline{u}, \overline{u}]$. In this case, Proposition 4.3 will show that all agents attain the maximum conceivable utility, $\overline{u}$, asymptotically. This does not happen when the idiosyncratic components have a sufficiently large degree of negative correlation between them (e.g., as in Proposition 4.1). Strict ordinal preferences indeed encapsulate a notion of negatively correlated idiosyncrasy as well: when an object is ranked first, any other object must be ranked second or worse. This negative correlation creates the large gap discussed in Section 3 between the average ranks of the RSD and the rank-optimal assignments.

Loosely speaking, under the iid idiosyncratic components assumption, there are many objects that provide a high utility for each agent. We call this effect the *abundance of competent choices*. The next proposition isolates this effect in a simple environment.

**Proposition 4.2.** *Let $u_1, \ldots, u_n$ be a sequence of iid random variables from a strictly increasing continuous CDF $F$ with support $[\underline{u}, \overline{u}]$. Let $u$ denote the $k$-th greatest element of the sequence. Then, as $n$ approaches infinity, $u$ converges in probability to the degenerate distribution centered at $\overline{u}$ as long as $k = o(n)$.*

One might assume that the (heterogeneity of) common-value components might eliminate

inifnity.

[9] In that market, the common-value component are 0. The joint distribution of the idiosyncratic components $v_1^a, \ldots, v_n^a$ is the uniform distribution over the binary vectors $(v_1^a, \ldots, v_n^a)$ such that $v_i^a \in \{0, 1\}$ and $\sum_{i=1}^n v_i^a = 3$. Observe that the random variables $\{v_i^a\}_{i=1}^n$ are negatively correlated; e.g. $\mathbb{P}[v_1^a = 1 | v_2^a = 1] < \mathbb{P}[v_1^a = 1]$ holds.

the abundance of competent choices. While this is true in the setting of the above proposition, this is not the case in our market, as discussed in Section 4.1.

### 4.1 Failure of the large market limit in *ranking* the assignments

In this section, we show that an inefficient version of the RSD mechanism attains the maximum utilitarian upper bound together with many other Pareto-efficient or -inefficient assignment mechanisms (all of which therefore would be payoff-equivalent). First, we present the result for the case of 0 common-value components and then we make a similar observation for the case of positive iid common-value components.

Consider a market with $n$ agents and $n$ objects each with unit capacity. Also, let $k(n) : \mathbb{N} \to \mathbb{N}$ be any increasing function. For notational simplicity, we drop the argument $n$ when it is clearly known from the context.

The *Inefficient Random Serial Dictatorship* (Inefficient-RSD) mechanism assigns agents to objects as follows. The agents are ordered randomly and, in that order, choose objects one by one. The choosing agent chooses her $k$-th top object if it is available, otherwise she chooses her $(k-1)$-th top object if it is available, otherwise she chooses her $(k-2)$-th top object if it is available, and so on. At the end, if the object she ranks first is also not available, the agent chooses her $(k+1)$-th top object if it is available, otherwise she chooses her $(k+2)$-th top object if it is available, and so on.

We say an assignment $\mu : A \to O$ contains a *Pareto-improving cycle of length $l$* when there exist agents $a_0, \ldots, a_{l-1}$ such that, for all $i \in \{0, 1 \ldots, l-1\}$, agent $a_i$ prefers $\mu(a_{i+1})$ to $\mu(a_i)$, where $i+1$ is computed modulo $l$.

**Proposition 4.3.** *Suppose that the utility of an agent $a \in A$ from an object $o \in O$ is drawn iid from a strictly increasing continuous CDF $F$ with support $[\underline{u}, \overline{u}]$. Then:*

*i.* (Abundance of competent choices) *When $k = o(n)$, the utility of any agent in the Inefficient-RSD mechanism converges in probability to the degenerate distribution centered at $\overline{u}$, as $n$ approaches infinity.*

*ii. When $k \geq k_0$, where $k_0$ is a sufficiently large constant,*[10] *the Inefficient-RSD assignment is not Pareto-efficient, whp: it contains a Pareto-improving cycle of length $\Theta(n)$.*

*iii. For all $k$, the average rank in the Inefficient-RSD assignment is $\Omega(\max\{k, \ln n\})$, whereas in the RSD assignment is $O(\ln n)$.*

[10] Our simulations show that $k_0 = 3$ suffices. (Section G in the appendix)

The above observation is not specific to the Inefficient-RSD mechanism; there are other families of assignments with similar properties.[11] Complementing our findings in Section 3, this proposition suggests that (i) in our market, abundance of competent choices is a likely consequence of iid idiosyncratic components, and (ii) under this assumption, cardinal models can fail to capture rank- or Pareto-inefficiencies, asymptotically; however, ordinal models or cardinal models with negatively correlated idiosyncratic components can be sharper in distinguishing between the assignments and capturing the assignment inefficiencies.

In Section E of the appendix, we provide a counterpart for the above proposition in the presence of non-zero iid common-value components. A brief intuition for why a similar observation holds in this case is that, in a market with $n$ agents and $n$ objects, common-value components always contribute a fixed amount to any assignment. For example, to maximize the assignment's average utility, it suffices to maximize the sum of idiosyncratic components.

In Section E, we also show that the Pareto-inefficiency pointed out in part ii can be captured through the alternative notion of *Pareto-improving pairs* (i.e., pairs of agents who prefer to swap their objects). This is done by showing that there exist a family of pairwise-disjoint Pareto-improving pairs that contain almost all agents.

# 5 Abundance of competent choices in other setups

In the absence of negative correlation, *abundance of competent choices* is a likely phenomenon in other environments as well. First, we show that the iid idiosyncratic components assumption can create abundance of competent choices in simple one-sided environments as well. After that, we consider stable assignments in marriage markets and make similar observations.

**Example: The secretary problem**

Consider the *secretary problem* in which an Employer needs to hire precisely one employee from a finite number of applicants, arriving in random order one by one. Employer only knows the number of applicants, $n$, and there is no recall: she either accepts the present applicant or rejects her, waiting to potentially hire one of the the future applicants.

We compare the solutions to Employer's problem in two cases. Case (i) is a well-studied case where Employer derives utility 1 from hiring the best applicant and utility 0 otherwise.

[11] For instance, truncate the lists so they contain only ranks $1, \ldots, k$, and then find the rank-*maximizing* assignment instead of the Inefficient-RSD assignment.

In Case (ii), the utility that Employer derives from hiring each applicant is iid over the interval $[\underline{u}, \overline{u}]$.

[Bruss 2000] provides a short elegant proof that shows the following policy is asymptotically[12] optimal for Case (i): Employer rejects the first $n/e$ applicants (*screens* them), and then hires the first applicant who is preferred to all of the screened applicants. Moreover, the analysis shows that this optimal policy hires the best candidate with probability $1/e$. In Case (ii), however, an asymptotically optimal policy only needs to screen $\omega(1)$ applicants. This guarantees an expected utility of $\overline{u}$ for Employer, asymptotically. The difference between these two cases is, again, due to the abundance of competent choices caused by the iid assumption in Case (ii).[13]

**Example: Stable assignments**

This section demonstrates that iid idiosyncratic components lead to abundance of competent choices in *marriage markets* as well: there, almost all agents attain their maximum conceivable utility in all *stable* assignments [Lee and Yariv 2018]. This, however, would not be the case when the idiosyncratic components are negatively correlated, or in ordinal utility models. These models capture meaningful gaps between different stable assignments, as discussed next.

A *marriage market* consists of $n$ men and $n$ women. Each man (woman) has a complete strict preference list over women (men). A stable assignment is a bijection $\mu$ from men to women with no *blocking pair*, where a blocking pair consists of a man and a woman who are not matched together but prefer each other to their current matches in $\mu$. A *man-optimal* stable assignment in a marriage market is a stable assignment in which every man (weakly) prefers his match in that assignment to any other stable assignment. A *woman-optimal* stable assignment is defined similarly, but for women. Both man-optimal and woman-optimal assignments exist in any marriage market [Gale and Shapley 1962].

A *random marriage market* is a marriage market where the preference lists of all men (women) over women (men) are drawn independently and uniformly at random.

**Proposition 5.1.** *Consider a random marriage market with $n$ men and $n$ women. Each agent (man or woman) attains utility 1 from being assigned to one of their top $k$ choices,*

[12] In here, asymptotically means as $n$ approaches infinity.

[13] Similar observations hold in other variations. E.g, suppose Employer can choose a subset of the applicants to interview with and hires the best of them. Interviewing $o(n)$ applicants means missing the best applicant with probability 1, asymptotically. However, when Employer's utility from hiring each applicant is drawn iid., interviewing only $\omega(1)$ applicants suffices for attaining utility $\overline{u}$ asymptotically.

*and utility* $0$ *otherwise. When* $k \geq \log^2 n$ *and* $k = o(\frac{n}{\log n})$*, in the man-optimal assignment any man and any woman respectively attain utilities* $1, 0$*, whp, while in the woman-optimal assignment any man and any woman respectively attain utilities* $0, 1$*, whp.*

By the above proposition, when the conditions on $k$ hold, the average utilities of men approaches 1 (the maximum conceivable utility) in the man-optimal assignment as $n$ approaches infinity, while the average utilities of women approaches 0. By symmetry, in the woman-optimal assignment the average utility of women and men approach 1,0, respectively.[14] On the other hand, the results of [Lee and Yariv 2018] imply that both men and women attain average utility approaching 1 (the maximum conceivable utility) when the utilities of men from being matched to women are iid across all pairs drawn from a distribution $F$ and the utilities of women are defined similarly but drawn from a distribution $G$.[15] This is in contrast to the observation made in Proposition 5.1.

These observations follow the patterns observed in the previous sections: iid idiosyncratic components in cardinal utility models can lead to abundance of competent choices. Hence, almost all agents attain their maximum conceivable utility asymptotically. This suggests that models with (negatively) correlated idiosyncratic components can be sharper in distinguishing between the assignments and capturing the inefficiencies than models with iid idiosyncratic components, especially in markets without abundance of competent choices.

# 6 Conclusion

We prove that, in a random matching market, there exists an assignment with an average rank that is independent of the market size (i.e., constant average rank). To prove the result, we need to keep track of the correlations between the objects' ranks, which are the reason that payoff-equivalence results for Pareto-optimal mechanisms such as [Che and Tercieux 2018] do not apply here. To handle these correlations, we use two applications of the FKG inequality [Fortuin et al. 1971], a fundamental correlation inequality in statistical mechanics and probabilistic combinatorics.

One of the implications of the above result is that the average rank under the RSD mechanism takes a heavy toll compared to the first-best. This is in apparent contrast with

[14] We also highlight a related result of [Pittel 1992] that shows in a random matching market, the *average rank* of partners for men and women in the man-optimal assignment is close to $\log n$, $\frac{n}{\log n}$, respectively.

[15] Their findings are more general, e.g. they also allow for (non-zero) common surplus components $c_{m,w}$ for any man $m$ and woman $w$.

the results of [Che and Tercieux 2018] about the payoff-equivalency of Pareto-optimal assignments. We account for this difference by showing that the equivalence result relies on the assumption that the idiosyncratic utility components (of an agent for objects) are iid. When the iid assumption is replaced with negative correlation, Pareto-optimal assignments can have very different payoff distributions.

Under the iid assumption, however, we show that a large family of assignments are payoff-equivalent; the family includes Pareto-optimal assignments as well as assignments with large Pareto-inefficiencies. We introduce *abundance of competent choices*, a likely phenomenon under the iid assumption, as a deriving force of such equivalencies. We show that this phenomenon also occurs in other setups, such as in Gale and Shapley's model of marriage markets [Gale and Shapley 1962].

We hope that our findings can provide more refined guidelines for researchers' modeling choices. In particular, our findings suggest that when there is no abundance of competent choices in a marketplace, models with negatively correlated idiosyncratic components are likely to give a better approximation of reality than models with iid idiosyncratic components.

# References


Atila Abdulkadiroglu, Nikhil Agarwal, and Parag Pathak. 2017. The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match. *American Economic Review* (2017). 3

Nick Arnosti and Peng Shi. 2017. How (Not) to Allocate Affordable Housing. *Working paper (EC2017)* (2017). 2

Itai Ashlagi and Afshin Nikzad. 2018. What matters in school choice tie-breaking? How competition guides design. *Working paper* (2018). 41

F Bruss. 2000. Sum the odds to one and stop. *Annals of Probability* 28 (06 2000). https://doi.org/10.1214/aop/1019160340 2, 11

Yeon-Koo Che and Olivier Tercieux. 2018. Payoff equivalence of efficient mechanisms in large matching markets. *Theoretical Economics* 13, 1 (2018), 239–271. https://doi.org/10.3982/TE2793 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE2793 1, 3, 5, 7, 8, 12, 13

C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. 1971. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics* 22, 2 (01 Jun 1971), 89–103. https://doi.org/10.1007/BF01651330 1, 3, 12, 15

Alan Frieze and Michal Karonski. 2012. Introduction to Random Graphs. (2012). 31

Alan Frieze and Páll Melsted. 2009. Maximum matchings in random bipartite graphs and the space utilization of Cuckoo Hash tables. *Random Structures & Algorithms* 41, 3 (2009), 334–364. https://doi.org/10.1002/rsa.20427 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20427 28

David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15. 1, 2, 3, 11, 13

Yash Kanoria and Daniela Saban. 2015. The Size of the Core in Assignment Markets. *Operations Research* (2015). 2

Donald E. Knuth. 1996. An Exact Analysis of Stable Allocation. *Journal of Algorithms* 20, 2 (1996), 431 – 442. https://doi.org/10.1006/jagm.1996.0020 1, 5, 27, 33

SangMok Lee and Leeat Yariv. 2018. On the Efficiency of Stable Matchings in Large Markets. *Working paper* (2018). https://arxiv.org/abs/1512.07276 2, 3, 8, 11, 12

Boris Pittel. 1992. On likely solutions of a stable marriage problem. *The Annals of Applied Probability* (1992), 358–401. 12, 40, 41, 42

A. E. Roth and Marilda Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Oxford University Press. 40

David W. Walkup. 1980. Matchings in random regular bipartite digraphs. *Discrete Mathematics* 31, 1 (1980), 59 – 64. https://doi.org/10.1016/0012-365X(80)90172-7 6, 7, 23

Douglas West. 1995. Introduction to Graph Theory. (1995). 32, 43

Wikipedia. 2018a. Depth-first search. https://en.wikipedia.org/wiki/Depth-first_search 32

Wikipedia. 2018b. Strongly connected component. https://en.wikipedia.org/wiki/Strongly_connected_component 33

# A Preliminaries

This section defines a few notions which will be used in the appendices.

For a positive integer $n$, we use $[n]$ to denote $\{1, \dots, n\}$.

Consider a market with set of agents $A$ and objects $O$ where each agent has a complete strict preference list over objects. A *preference profile* is a function $\pi : A \to O^{|A|}$ that determines the preference list of each agent. We use a more compact notation of $\pi_a$ instead of $\pi(a)$ to denote the preference list of an agent $a$.

For any lattice $\mathcal{L}$, we denote the set of its element by $V(\mathcal{L})$. For any graph $G$, we denote the set of its nodes by $V(G)$ and the set of its edges by $E(G)$.

## A.1 The FKG inequality

The Fortuin-Kasteleyn-Ginibre (FKG) inequality [Fortuin et al. 1971] is a correlation inequality. Informally, it says that an "increasing event" and a "decreasing event" are negatively correlated, while two "increasing events" are positively correlated.

Let $\mathcal{L}$ be a finite distributive lattice and $\mu : V(\mathcal{L}) \to \mathbb{R}_+$ be a non-negative function on it that satisfies log-supermodularity, i.e. for any two $x, y \in V(\mathcal{L})$,

$$\mu(x \wedge y)\mu(x \vee y) \geq \mu(x)\mu(y),$$

where $\wedge, \vee$ are the meet and join operators of the lattice, respectively.

By the FKG inequality, when $f, g : V(\mathcal{L}) \to \mathbb{R}_+$ are respectively increasing and decreasing functions on the lattice $\mathcal{L}$, it holds that

$$\left(\sum_{x \in X} f(x)g(x)\mu(x)\right)\left(\sum_{x \in X} \mu(x)\right) \leq \left(\sum_{x \in X} f(x)\mu(x)\right)\left(\sum_{x \in X} g(x)\mu(x)\right).$$

The direction of the inequality is reversed when both functions are increasing (or decreasing).

## A.2 Chernoff Bounds

In the following proofs the Chernoff concentration bounds are used as stated below.

Let $X_1, \dots, X_n$ be a sequence of $n$ independent random binary variables such that $X_i = 1$ with probability $p_i$ and $X_i = 0$ with probability $1 - p_i$. Also, let $\mu = \sum_{i=1}^{n} \mathbb{E}[X_i]$. Then for any $\epsilon$ with $0 \le \epsilon \le 1$ we have:

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i > (1+\epsilon)\mu\right] \le e^{-\epsilon^2\mu/(2+\epsilon)}$$
$$\mathbb{P}\left[\sum_{i=1}^{n} X_i < (1-\epsilon)\mu\right] \le e^{-\epsilon^2\mu/2}.$$

Furthermore, the former inequality holds for all $\epsilon \ge 0$.

### A.3 Asymptotic notations

For two functions $f, g : \mathbb{N} \to \mathbb{R}_+$ we write $f = o(g)$ when $\lim_{n\to\infty} f(n)/g(n) = 0$, $f = \omega(g)$ when $g = o(f)$, $f = \Omega(g)$ when $\lim_{n\to\infty} f(n)/g(n) > 0$, $f = O(g)$ when $g = \Omega(f)$, and $f = \Theta(g)$ when $f = O(g)$ and $g = O(f)$.

We say a sequence of events $E_1, E_2, \ldots$ happen *with high probability* (whp) when $\mathbb{P}[E_n]$ approaches 1 as $n$ approaches infinity. We say the sequence happens *with very high probability* (wvhp) if $n^c \cdot (1 - \mathbb{P}[E_n])$ approaches 0 as $n$ approaches infinity for any constant $c > 0$.

## B Proof of Theorem 3.1

First, we prove the claim for the case of unit capacities (i.e. $c = 1$). After that, we present Lemma B.5 which then proves the claim for the case of non-unit capacities (i.e. $c > 1$).

### B.1 Proof for the case of unit capacities ($c = 1$)

We need a few definitions for the proof. We use $\Pi$ to denote the set of all preference profiles in a market of $n$ agents and $n$ objects. Also, let $U_\Pi$ denote the uniform probability measure over $\Pi$.

Let $d$ be a constant independent of $n$ (we will set $d$ to 3). First, we construct a bipartite graph $G[A, O]$. We call $A$ the left side of the graph and $O$ the right side. The set of edges in the graph is $E = E_L \cup E_R$, where $E_L$ and $E_R$ are defined as follows. For any agent $a \in A$, $E_R$ (the set of edges that we draw from left to right) contains $d$ edges that connect $a$ to the first $d$ objects that $a$ ranks on her preference list, i.e. her $d$ most favorite objects.

Define $E_L$ (the set of edges that we draw from right to left) as follows. For any object $o \in O$, let $a^o_1, \ldots, a^o_{n_o}$ denote all agents who have listed object $o$ at position $d+1$ or worse.

Without loss of generality, suppose that $a_1^o, \dots, a_{n_o}^o$ are ordered in the order that object $o$ appears on their list: $a_1^o$ has to earliest appearance of $o$ (the most favorable position) and $a_{n_o}^o$ has the latest appearance of $o$ (the least favorable position). Let $i_o$ denote the smallest index such that $i_o \geq d$ and that $a_{i_o}^o$ and $a_{i_o+1}^o$ assign different ranks to object $o$. If there is no such index $i_o$, then let $i_o = n_o$. $E_l$ contains the edges $(o, a_1^o,), \dots, (o, a_{i_o}^o)$, for all objects $o \in O$.

The proof is done in two steps. In Step A we show that $G[A, O]$ contains a perfect matching whp, and then we use this fact in Step B to provide an upper bound on the expected average rank of the rank-optimal matching.

**Step A: existence of a perfect matching with high probability**

We use $A$ to denote the set of agents and $O$ to denote the set of objects. A $k$-tuple is a pair of sets $(X, Y)$ with $X \subseteq A$ and $Y \subseteq O$ such that $|X| = k$ and $|Y| = n - k + 1$. A $k$-tuple is *independent* if $(X \times Y) \cap E = \varnothing$. We use $p_k$ is the probability that there exists and independent $k$-tuple. Let $p(n) = \sum_{k=1}^{n} p_k$. We will show that $p(n) \leq n^{-3}$ holds for $n \geq 50$.

To provide an upper bound on $p_k$, first we fix a $k$-tuple $(X, Y)$ and provide an upper bound on the chance that that $(X, Y)$ is independent. Without loss of generality, suppose $X = \{1, \dots, k\}$ and $Y = \{1, \dots, l\}$, where $l = n - k + 1$. Let $L, R$ respectively denote the events that $(X \times Y) \cap E_L = \varnothing$ and $(X \times Y) \cap E_R = \varnothing$. (Note that each of the events $L, R$ is just a subset of $\Pi$; therefore, we treat $L, R$ as sets, as well as events.)

Suppose $p_{(X,Y)}$ is the probability that $(X, Y)$ is an independent tuple. First, observe that

$$p_{(X,Y)} = \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot \mathbb{P}\left[L|R\right]. \tag{B.1}$$

The first factor in the RHS of the above equation is just equal to $\mathbb{P}[R]$. We use the next two claims to simplify the RHS of the above equation.

**Claim B.1.** $\mathbb{P}[L|R] \leq \mathbb{P}[L]$.

*Proof.* The proof uses the FKG inequality to show that the events $L$ and $R$ are negatively correlated. To use the inequality, we first define a distributive lattice. To define the lattice, we need a few definitions. For any object $i$, let $f(i) = i$ if $i > d$ and let $f(i) = n + i$ otherwise. Define a total order $\preceq$ on the set of objects $O = \{1, \dots, n\}$ as $i \preceq j$ iff $f(i) < f(j)$. With slight abuse of notation, for any two vectors of the same size, namely $u = (u_1, \dots, u_s), v = (v_1, \dots, v_s)$, we write $u \preceq v$ iff $u_j \preceq v_j$ holds for all $j$.

The lattice that we define for applying the FKG inequality is denoted by $\mathcal{L}[l]$. Each element of the lattice, namely $x$, is an ordered list of $n$ $l$-dimensional vectors, namely $x_1, \ldots, x_n$. Since $l$ remains fixed in the proof of this claim, we drop the argument and denote the lattice simply by $\mathcal{L}$. As we will see, $x_a$ corresponds to agent $a$ for any $a \in A$. Each vector, namely $x_a = (x_a^1, \ldots, x_a^l)$, contains $l$ distinct integers belonging to $O$. For $a \leq k$, the integers are ordered in decreasing order with respect to $\preceq$, and for $a > k$, they are ordered in increasing order with respect to $\preceq$. For any two elements of the lattice $x, y$, we have $x \preceq_{\mathcal{L}} y$ iff

$$\begin{cases} y_a \preceq x_a, \ \forall a \in \{1, \ldots, k\} \\ x_a \preceq y_a, \ \forall a \in \{k+1, \ldots, n\}. \end{cases}$$

Observe that the defined partial order $\preceq$ over $V(\mathcal{L})$ is a lattice with the meet ($\wedge$) and join ($\vee$) operators being the component-wise minimum and maximum with respect to $\preceq$, respectively. Figure 2 provides a graphical representation of the lattice elements.

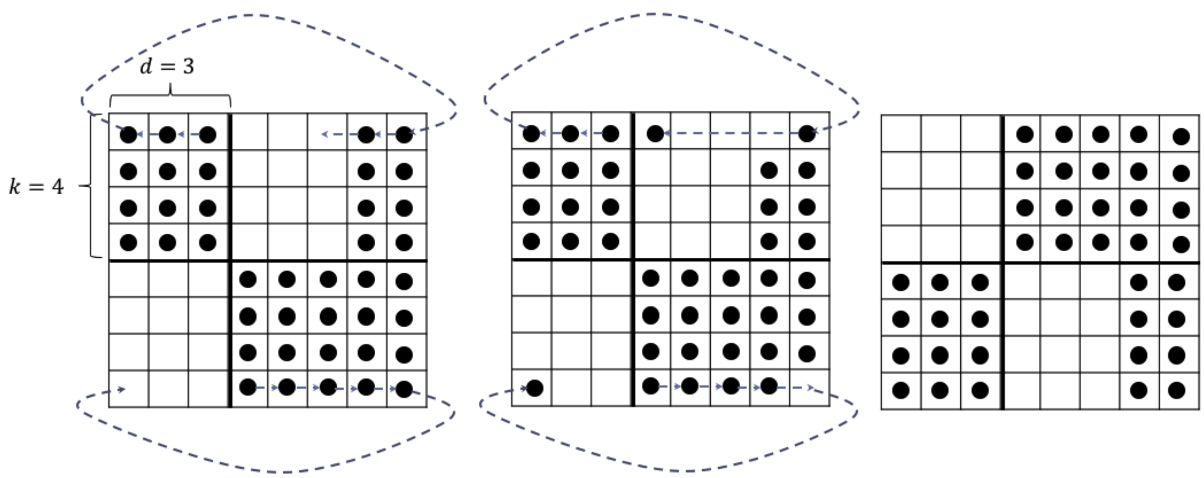


Figure 2: From left to right: the smallest element of the lattice, an element between the smallest and the largest element, and the largest element. We have $A = [8]$ and $O = [8]$. Rows and columns correspond to the elements of $A, O$, respectively. Each row $a \in A$ contains $l = 5$ marked cells which determine the $l$-dimensional vector $x_a$ corresponding to that row. The dotted arrows determine the directions for "moveing" in the lattice from a smaller to a larger element.

Next, we link the definition of the lattice to the set of all preference profiles $\pi$ of the assignment problem. Each preference profile $\pi$ is *represnted* by and element of the lattice, which we denote by $\mathcal{L}(\pi)$, and define it as follows. Given $\pi$, for each agent $a$, construct an $l$-dimensional vector $x_a$ that contains the positions (ranks) of objects $1, \ldots, l$ on $\pi_a$. Order the elements of this vector in decreasing order with respect to $\preceq$ if $a \leq d$, and otherwise, order them in increasing order with respect to $\preceq$. Let $\mathcal{L}(\pi) = \langle x_1, \ldots, x_n \rangle$.

Observe that, given any preference profile $\pi$, all the information that we need to assess whether events $R, L$ happen at $\pi$ are coded in $\mathcal{L}(\pi)$. That is, $\mathcal{L}(\pi)$ is a sufficient statistic for detecting whether events $L, R$ happen in $\pi$. This observation is used in the last step of the proof, as follows.

Define two functions $f_L, f_R : \mathcal{L} \to \{0,1\}$ such that $f_L(x) = 1$ iff the event $L$ holds at element $x$ of the lattice and $f_R(x) = 1$ iff the event $R$ holds at element $x$. A straightforward coupling argument shows that the functions $f_L, f_R$, respectively, are increasing and decreasing with respect to the partial order $\preceq$ defined over $V(\mathcal{L})$. Also, observe that the probability distribution induced by $U_\Pi$ over $\{\mathcal{L}(\pi) : \forall \pi\}$ is a uniform distribution itself, and therefore, it satisfies the log-supermodularity condition required for the FKG inequality. The inequality implies that

$$\mathbb{E}\left[f_L(x) f_R(x)\right] \leq \mathbb{E}\left[f_L(x)\right] \mathbb{E}\left[f_R(x)\right],$$

i.e. the events $L$ and $R$ are negatively correlated. □

**Claim B.2.** $\mathbb{P}[L] \leq \prod_{i=1}^{l} \mathbb{P}[L_i]$, *where* $L_i$ *is the event that there is no edge in* $E_L$ *that connects a node* $i \in Y$ *to a node in* $X$*.*

*Proof.* The proof is by induction. For any $j \leq l$, we will show that

$$\mathbb{P}\left[L_1 \wedge L_2 \wedge \ldots \wedge L_j\right] \leq \prod_{i=1}^{j} \mathbb{P}\left[L_i\right].$$

The induction basis for $j = 1$ is trivial. The induction step supposes that, for some $l' < l$, the claim holds for all $j \leq l'$, and then proves the claim for $j = l' + 1$.

The proof of the induction step is by an application of the FKG inequality. First, we need a few definitions.

**Definitions.** The lattice which will be used in the application of the FKG inequality is $\mathcal{L}[l']$. (Recall the definition from the proof of Claim B.1.) For brevity, we drop the argument $l'$ and denote the lattice simply by $\mathcal{L}$ throughout this proof.

Define a *t-subprofile* to be a partially filled preference profile in the following sense: each row of the preference profile has $n - t$ empty positions. The rest of the positions in the row are filled with numbers $1, \ldots, t$, with each number appearing precisely once. The set of *filled indicies* in a $t$-subprofile is simply the set of all pairs $(a, i)$ where position $i$ of agent $a$'s preference list is filled. Note the correspondence between the notions of an $l'$-subprofile and the lattice $\mathcal{L}$: each element of the lattice corresponds to the set of filled indices of precisely $(l'!)^n$ $l'$-subprofiles.

An *Extension* of an element of the lattice, namely $x$, to an $l'$-subprofile is defined as follows. Recall that $x = (x_1, \ldots, x_n)$ where for each $a \in A$, $x_a = (x_a^1, \ldots, x_a^{l'})$ is a vector that contains $l'$ distinct integers belonging to $O$. Suppose that $\pi$, which is initially empty, denotes the $l'$-subprofile which is to be constructed from $x$. To extend $x$ to an $l'$-subprofile, assign a number from $[l']$ to each preference list position $\pi_a(x_a^i)$ for all $a \in A$ and $i \in [l']$, such that all the numbers assigned to any preference list $\pi_a$ are distinct (i.e. such that for any $a \in A$, the set $\{\pi_a(x_a^i) : i \in [l']\}$ has size $l'$).

An *Extension* of a $j$-subprofile to a complete preference profile is defined in the natural way: by filling out the empty positions in the $j$-subprofile in a way that it creates a valid preference profile.

Let $K_i^j$ denote the set of $j$-subprofiles $\pi'$ such that there is an extension of $\pi'$ to a complete preference profile, namely $\pi''$, with $\pi'' \in L_i$.[16] (Observe that if there exists one such extension of $\pi'$ satisfying this condition, then *any* of its extensions to a complete preference profile also satisfy this condition, so long as $j \geq i$.) Define

$$K^j = K_1^j \cap \ldots \cap K_{l'}^j.$$

For an element $x$ of the lattice, define $f(x)$ to denote the number of extensions of $x$ to an $l'$-subprofile belonging to $K^{l'}$. Also, let $g(x)$ denote the number of extensions of $x$ to an $(l'+1)$-subprofile belonging to $K_{l'+1}^{l'+1}$ divided by $(l'!)^n$. Intuitively, $g(x)$ is the number of extensions of $x$ to a $(l'+1)$-subprofile belonging to $K_{l'+1}^{l'+1}$ in which each of the positions that contain an element of $[l']$ is filled with an '*' instead. (Note that only the position of the object $l'+1$ in the subprofile determines whether it belongs to $K_{l'+1}^{l'+1}$.)

A straight-forward coupling argument shows that the functions $f, g : V(\mathcal{L}) \to \mathbb{R}_+$ are decreasing and increasing functions, respectively.

**The induction step.** Let $U, V$ denote the uniform measures induced over the set of all preference profiles and over $V(\mathcal{L})$, respectively. Since the functions $f, g$ are respectively decreasing and increasing, and since the measure $U$ is uniform (and therefore, log-

[16]Recall the equivalence between events and subsets of $\Pi$.

supermodular), then the FKG inequality implies that

$$\mathbb{E}_U\left[f(x)g(x)\right] \leq \mathbb{E}_U\left[f(x)\right]\mathbb{E}_U\left[g(x)\right]. \tag{B.2}$$

We will complete the proof by rewriting the above inequality. To this end, let

$$\begin{aligned}\alpha &= |V(\mathcal{L})| \cdot (l'!)^n,\\ \beta &= (n-l')^n,\\ \gamma &= (n-l'-1)^n,\end{aligned}$$

and define the events

$$\begin{aligned}A &= L_1 \wedge \ldots \wedge L_{l'},\\ B &= L_{l'+1}.\end{aligned}$$

Observe that

$$\begin{aligned}\mathbb{P}_U\left[A\right] &= \frac{\sum_x \beta f(x)}{\alpha\beta} = \frac{\mathbb{E}_V\left[f(x)\right]}{(l'!)^n},\\ \mathbb{P}_U\left[B\right] &= \frac{\sum_x \gamma \cdot (l')^n \cdot g(x)}{\alpha\beta} = \frac{\gamma \cdot \mathbb{E}_V\left[g(x)\right]}{(n-l')^n},\\ \mathbb{P}_U\left[A \wedge B\right] &= \frac{\sum_x \gamma f(x)g(x)}{\alpha\beta} = \frac{\gamma \cdot \mathbb{E}_V\left[f(x)g(x)\right]}{(l'!)^n(n-l')^n},\end{aligned}$$

where all the sums are taken over $x \in V(\mathcal{L})$. The above equalities together with (B.2) imply that

$$\mathbb{P}_U\left[A \wedge B\right] \leq \mathbb{P}_U\left[A\right] \cdot \mathbb{P}_U\left[B\right], \tag{B.3}$$

which completes the induction step and proves the claim. □

We can provide an upper bound on the RHS of (B.1) using the above two claims.

**Claim B.3.**

$$p_{(X,Y)} \leq \left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k (\omega+\delta)^l, \tag{B.4}$$

*where* $\omega = \left[\binom{n-k}{d}/\binom{n}{d}\right]^{n-k+1}$ *and* $\delta = \frac{d^{n-2d+1}e^d}{n^{n-2d}}$.

*Proof.* By Equation (B.1) and Claims B.1 and B.2, it suffices to show that $\mathbb{P}[L_i] \leq \omega + \delta$ for any $i \in Y$. To this end, let $D$ be the event in which object $i$ has a rank worse than $d$ in at least $d$ of the agents' preference lists. Observe that

$$\mathbb{P}[L_i] = \mathbb{P}[L_i|D] \times \mathbb{P}[D] + \mathbb{P}\left[L_i|\overline{D}\right] \times \mathbb{P}\left[\overline{D}\right].$$

We will show that (i) $\mathbb{P}\left[\overline{D}\right] \leq \delta$ and (ii) $\mathbb{P}[L_i|D] \leq \omega$, which would prove that $\mathbb{P}[L_i] \leq \omega+\delta$.

**Step (i) $\mathbb{P}\left[\overline{D}\right] \leq \delta$.** Let $\overline{D}_j$ be the event in which object $i$ a rank worse than $d$ in precisely $j$ of the agents' preference lists. Observe that

$$\mathbb{P}\left[\overline{D}_j\right] = \binom{n}{j}\left(\frac{d}{n}\right)^{n-j}\left(1-\frac{d}{n}\right)^{j}.$$

Therefore, we can write

$$\mathbb{P}\left[\overline{D}\right] = \sum_{j=0}^{d-1}\mathbb{P}\left[\overline{D}_j\right] \leq d\binom{n}{d}\left(\frac{d}{n}\right)^{n-d} \leq \frac{d^{n-2d+1}e^{d}}{n^{n-2d}} = \delta, \tag{B.5}$$

where in the last inequality we have used the bound $\binom{n}{d} \leq \left(\frac{ne}{d}\right)^{d}$.

**Step (ii) $\mathbb{P}[L_i|D] \leq \omega$.** let $A_j$ be the set of agents who rank object $i$ on the $j$-th position of their preference list, and let $A^h = \cup_{j=d+1}^{h} A_j$ such that $h$ is the smallest number for which $|A^h| \geq d$. Observe that $A^h$ is a random variable whose distribution, conditional on its size being equal to $x$, is the uniform distribution over the set of all subsets of $A$ with size $x$. (This holds by symmetry.) Therefore, we can write

$$\mathbb{P}\left[L_i \middle| g = |A^h|\right] \leq \frac{\binom{n-k}{g}}{\binom{n}{g}}.$$

Since the RHS of the above equality is decreasing in $g$, and since $|A_h| \geq d$, therefore we can write

$$\mathbb{P}[L_i|D] \leq \frac{\binom{n-k}{d}}{\binom{n}{d}} = \omega$$

□

We now use Claim B.3 and a union bound to write

$$p_k \leq \binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k (\omega+\delta)^l$$
$$\leq \binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \left(\omega^l+2^l\delta\right),$$

where recall that $l = n-k+1$, and the last inequality holds since $\omega, \delta \leq 1$. Using the above inequality, we can bound $p(n)$ as follows:

$$p(n) = \sum_{k=1}^{n} p_k \leq \sum_{k=1}^{n}\binom{n}{k}\binom{n}{l}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \left(\omega^l+2^l\delta\right)$$
$$= \sum_{k=1}^{n}\binom{n}{k}\binom{n}{k-1}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot \omega^l + \sum_{k=1}^{n}\binom{n}{k}\binom{n}{k-1}\left[\frac{\binom{k-1}{d}}{\binom{n}{d}}\right]^k \cdot 2^l\delta \quad \text{(B.6)}$$

Observe that (B.6) has two summands. Let $S_1, S_2$ denote the first and the second summand, respectively. [Walkup 1980] shows that

$$S_1 \leq \frac{1}{122}\left(\frac{d}{n}\right)^{(d+1)(d-2)}.$$

We finish the proof by providing an upper bound on $S_2$. Observe that

$$S_2 \leq 2^{3n}\delta = 2^{3n}\cdot\frac{d^{n-2d+1}e^d}{n^{n-2d}} \quad \text{(B.7)}$$

In particular, letting $d = 3$ and summing up $S_1$ and $S_2$ implies that for sufficiently large $n$, we have $p(n) \leq n^{-3}$. (While the crude bounds we have used provide that $n \geq 50$ is sufficiently large, the right-hand side 50 can be reduced with a more careful analysis.)

#### Step B: Bounding the average rank in the perfect matching

Recall that $G[A, O]$ denotes the bipartite graph constructed in Step i. We showed that, for sufficiently large $n$, $G[A, O]$ has a perfect matching with probability at least $1 - n^{-3}$. To bound the expected sum of the ranks in the matching, we provide an upper bound on $\sum_{o\in O} w_o$, where $w_o$ is the weight of the maximum-weight edge adjacent to $o \in O$. (Define $w_o = 0$ if there are no edges incident to $o$.)

Recall that $U_\Pi$ denotes the uniform distribution over $\Pi$. We can show that, for any $o$, $\mathbb{E}_{U_\Pi}[w_o]$ is bounded by $d + dz$, where $z = e/(e-1)$. This will imply that, for $d = 3$, the expected average rank is at most

$$(1 - n^{-3}) \cdot \frac{n(d + 3e/(e-1))}{n} + n^{-3} \cdot n^2 < 7.75$$

for $n \geq 50$.

The formal proof is presented below. For notational simplicity, we drop the subscript $U_\Pi$ from the notations $\mathbb{E}_{U_\Pi}[\cdot]$ and $\mathbb{P}_{U_\Pi}[\cdot]$. Also, all the inequalities that we write below hold for sufficiently large $n$ (i.e. $n \geq 50$). Let $w = \sum_{o \in O} w_o$. Let $\mathfrak{m}$ denote the event that $G[A, O]$ has a perfect matching. By our analysis in Step A, we have

$$\mathbb{E}[w|\mathfrak{m}] \leq \frac{\mathbb{E}[w]}{1 - n^{-3}}.$$

Therefore, to complete Step B, it suffices to bound $\mathbb{E}[w]$. The following claim implies that $\mathbb{E}[w] \leq d + 3e/(e-1)$, would imply that

$$\mathbb{E}[w|\mathfrak{m}] \leq \frac{d + 3e/(e-1)}{1 - n^{-3}} < 7.75$$

for $d = 3$ and $n \geq 50$.

**Claim B.4.** *For any object $o \in O$, $\mathbb{E}[w_o] \leq d + 3e/(e-1)$.*

*Proof.* Let $\mathcal{M}$ be an $|A| \times |O|$ matrix that contains $\pi_a$ as its $a$-th row for each agent $a \in A$. For any integer $i > d$, let $H_i^j$ denote the event that object $o$ appears exactly $j$ times in columns $d+1, \ldots, i$ of the matrix $\mathcal{M}$.

Observe that

$$\mathbb{P}\left[H_{d+1}^0\right] = (1 - 1/n)^n \leq 1/e.$$

Also, observe that for all $i \geq d+1$,

$$\mathbb{P}\left[H_{i+1}^0 | H_i^0\right] < (1 - \frac{1}{n})^n \leq 1/e,$$
$$\mathbb{P}\left[H_{i+1}^2 | H_i^1\right] < (1 - \frac{1}{n})^n \leq 1/e,$$

and that

$$\mathbb{P}\left[H_{i+1}^3 | H_i^2\right] \leq \begin{cases} (1 - \frac{1}{n-2})^{n-2} \leq 1/e, & \text{for } i > d+1 \\ (1 - \frac{1}{n-1})^{n-2} \leq (1/e) \cdot \frac{n-1}{n-2}, & \text{for } i = d+1. \end{cases}$$

Define the random variable $H$ as follows: let $H = i$ when $i$ is the smallest index for which the event $H_i^3 \vee \ldots \vee H_i^n$ holds. If such $i$ does not exist, then let $H = 0$. Observe that $H = w_o$, by definition. The above four inequalities imply that $H - d$ is stochastically dominated by the sum of three independent geometric random variables with means $\frac{1}{1-1/e}$, $\frac{1}{1-1/e}$, and $\frac{1}{1-\frac{n-1}{n-2}/e}$. In fact, with a straight-forward coupling argument it is possible to slightly refine this bound and eliminate the coefficient $\frac{n-1}{n-2}$, i.e. it is possible to show that $H - d$ is stochastically dominated by sum of three independent geometric random variables each with mean $\frac{1}{1-1/e}$. Therefore, $\mathbb{E}[w_o] = \mathbb{E}[H] \leq d + \frac{3e}{e-1}$. □

## B.2 Proof for the case of non-unit capacities ($c > 1$)

First, we need a few definitions. The *rank distribution* of an assignment $\mu : A \to O$ is a function $\mathcal{R} : \mathbb{N} \to \mathbb{R}_+$ where $\mathcal{R}(i)$ denotes the fraction of agents who are assigned to their $i$-th favorite object (i.e. the object that they rank $i$-th). Given a random market $M$, the *expected rank distribution* of the rank optimal assignment in $M$ is a function $\mathcal{R} : \mathbb{N} \to \mathbb{R}_+$ where $\mathcal{R}(i)$ denotes the expected fraction of agents who are assigned to their $i$-th favorite object in the rank-optimal assignment.

We say a rank distribution $\mathcal{R}' : \mathbb{N} \to \mathbb{R}_+$ stochastically dominates another rank-distribution $\mathcal{R} : \mathbb{N} \to \mathbb{R}_+$ when for any $r \in \mathbb{N}$,

$$\sum_{i=1}^{r} \mathcal{R}'(i) \geq \sum_{i=1}^{r} \mathcal{R}(i).$$

The next lemma completes the proof of Theorem 3.1, as it shows that an upper bound on the expected average rank for the case of $c = 1$ is a valid upper bound for the general case ($c \geq 1$) as well.

**Lemma B.5.** *Let $M, M'$ be two random markets both with $n$ agents. Suppose that $M$ has $n$ objects each with capacity 1 and $M'$ has $n/c$ objects each with capacity $c \geq 1$, respectively. Let $\mathcal{R}, \mathcal{R}'$ denote the expected rank distributions of the rank-optimal assignments in $M, M'$, respectively. Then, $\mathcal{R}'$ stochastically dominates $\mathcal{R}$.*

*Proof.* For any preference profile, $\pi$, let $r_\pi$ denote the average rank in the rank-optimal assignment when the preference profile is $\pi$. Let $\Pi, \Pi'$ denote the set of possible preference profiles in $M, M'$. The proof works by defining a function $f : \Pi \to \Pi'$ that maps every $\frac{|\Pi|}{|\Pi'|}$ elements of $\Pi$ to precisely one element of $\Pi$. Moreover, this function is defined such that, for any $\pi \in \Pi$, the rank distribution of the rank-optimal assignment for $\pi$ is stochastically dominated by the rank distribution of the rank-optimal assignment for $f(\pi)$. The existence of such a function will prove the claim.

In the rest of the proof, we define the function. Let $n$ be the number of agents in both markets. In the market $M$, relabel the objects as $O = \{o_t^j : j \in [c], t \in [n/c]\}$. We say an object $o$ is of type $t$ if $o = o_t^j$ for some $j, t$. With slight abuse of notation, we use $t(o)$ to denote the type of an object $o$. Given a preference list $\sigma$ over $O$, define $\overline{\sigma}$ to be a list in which $\overline{\sigma}(i) = t(\sigma(i))$. (That is, each object is replaced with its type.)

Let $g(\sigma)$ be the preference profile over $[n/c]$ defined as follows: in its $i$-th position, $g(\sigma)$ contains the $i$-th distinct number that appears in $\overline{\sigma}$, for $i \in [n/c]$. In other words, the function $g$ removes the second, third, and the higher appearances of a number in $\overline{\sigma}$ and outputs the resulting list.

For each preference profile $\pi \in \Pi$, define $f(\pi)$ to be the preference profile $\pi'$ where $\pi'_a = g(\overline{\pi_a})$ for all $a \in A$. Observe that, by symmetry, $|f^{-1}(\pi')|$ does not depend on $\pi'$, i.e. $|f^{-1}(\pi')| = \frac{|\Pi|}{|\Pi'|}$. To complete the proof, it remains to show that the rank distribution of the rank-optimal assignment for $\pi$ is stochastically dominated by the rank distribution of the rank-optimal assignment for $f(\pi)$, for all preference profiles $\pi \in \Pi$. Let $\mu$ be the rank-optimal assignment for the preference profile $\pi$. We define the assignment $\mu'$ in the market $M'$ for the preference profile $\pi'$ as follows: for each agent $a$, let $\mu'(a) = t(\mu(a))$. Observe that $\mu'$ is a feasible assignment, and that $\mu'(a)$ does not have a worse rank in $\pi'_a$ than $\mu(a)$ has in $\pi_a$. Therefore, the rank distribution of $\mu'$ stochastically dominates the rank distribution of $\mu$. This finishes the proof.

□

This completes the proof of Theorem 3.1. The above lemma together with our analysis in Section B.1 immediately imply the following corollary.

**Corollary B.6** (of Theorem 3.1)**.** *The expected average rank in the rank-optimal assignment is at most* $7\frac{3}{4}$ *for all* $n \geq 50$*.*

# C Proof of Proposition 3.2

Consider the last $m$ agents who choose (i.e. the $m$ agents with lowest priority numbers). Let them be indexed by $a_0, \ldots, a_{m-1}$, ordered with respect to their priority numbers with $a_0$ having the best priority number and $a_{m-1}$ having the worst. Also, let $R_i$ denote the average rank of agent $a_i$, and $R = \frac{1}{m} \cdot \sum_{i=0}^{m-1} R_i$.

To provide a lower bound on $R_i$, we define an auxiliary problem instance, which is just running RSD on a market with $m$ agents, namely $a'_0, \ldots, a'_{m-1}$, and $m$ objects with unit capacities. Suppose the agents rank objects independently and uniformly at random, and that agents choose objects in the same order as their indices: agent $a'_0$ chooses the first object. Let $R'_i$ denote the average rank of agent $a'_i$, and $R' = \frac{1}{m} \cdot \sum_{i=0}^{m-1} R'_i$. [Knuth 1996] shows that $R' \geq \ln m - 1$. The next claim states that $R_i \geq R'_i$, which would imply $R \geq R'$. That would complete the proof, as it shows that the expected average rank in the original instance is at least $R' \cdot \frac{m}{n} \geq \frac{\ln m - 1}{c}$.

**Claim C.1.** *For any* $i \in \{0, \ldots, m-1\}$*,* $R_i \geq R'_i$*.*

*Proof.* When agent $a'_i$ is choosing in the auxiliary problem, there are exactly $i$ objects allocated by the agents before her, and therefore $m - i$ possible choices remain. In the original problem, when agent $a_i$ is choosing, at most $i$ objects have a positive number of copies remaining; therefore, agent $a_i$ has at most $m - i$ possible choices. This implies that the rank distribution for agent $a'_i$ stochastically dominates the rank distribution for agent $a_i$, which implies that $R_i \geq R'_i$. □

# D Proofs from Section 4

## D.1 Proof of Proposition 4.1

We prove that the *expected* difference between the average utility in the utility-optimal assignment and the average utility in the RSD assignment is at least $\frac{v_1 - v_2}{20}$. The high probability result follows from a straight-forward application of Chernoff bounds.[17]

First, we compute an upper bound for average utility under the RSD mechanism. For any $i \geq 3$, the $i+1$-th agent who chooses in the RSD mechanism attains utility $v_2$ with

[17] The presented analysis for the utility-optimal assignment is a high-probability statement. The application of Chernoff bounds is required for the RSD analysis.

probability

$$\frac{i}{n} \cdot \frac{i-1}{n-1} \cdot \frac{i-2}{n-2} \geq (\frac{i-2}{n-2})^3.$$

Therefore, the expected number of agents who attain utility $v_2$ is at least

$$\sum_{i=1}^{n-2} (\frac{i}{n-2})^3 = (n-2)^{-3} \cdot \left(\frac{(n-2)(n-1)}{2}\right)^2 \geq \frac{n}{4}.$$

This shows that the expected average utility under the RSD mechanism is at most $\frac{3}{4} \cdot v_1 + \frac{1}{4} \cdot v_2$.

Next, we show that, whp, there exists a Pareto-optimal assignment with average utility at least $\frac{4}{5} \cdot v_1 + \frac{1}{5} \cdot v_2$.[18] To this end, construct the bipartite graph $G[A', O]$ where $A' = [\lceil \frac{4n}{5} \rceil]$. For any agent $a \in A'$, add 3 edges to $G[A', O]$ that connect $a$ to the first, second, and the third object that $a$ lists on her preference list. Since the preference lists of agents are drawn independently and uniformly at random, then Theorem 3 of [Frieze and Melsted 2009] implies that $G[A', O]$ has a matching of size $|A'|$, whp. Therefore, whp, there exists an assignment in the original random market with average utility at least $\frac{4}{5} \cdot v_1 + \frac{1}{5} \cdot v_2$. This completes the proof.

## D.2 Proof of Proposition 4.2

Fix an arbitrary small $\epsilon > 0$, and let $I$ denote the interval $[\overline{u} - \epsilon, \overline{u}]$. Also, let $\delta = 1 - F(\overline{u} - \epsilon)$. Observe that

$$\begin{aligned}
\mathbb{P}[v \notin I] &= (1-\delta)^n + \sum_{i=1}^{k-1} \binom{n}{i} \cdot \delta^i (1-\delta)^{n-i} \\
&\leq (1-\delta)^n + \sum_{i=1}^{k-1} \delta^i \cdot e^{i \log(\frac{ne}{i})} (1-\delta)^{n-i} \\
&\leq (1-\delta)^n + \sum_{i=1}^{k-1} \delta^i \cdot e^{i \log(\frac{ne}{i}) - (n-i) \log(\frac{1}{1-\delta})} \\
&= (1-\delta)^n + \sum_{i=1}^{k-1} \delta^i \cdot s_i \qquad \text{(D.1)}
\end{aligned}$$

where $s_i = i \log(\frac{ne}{i}) - (n-i) \log(\frac{1}{1-\delta})$. (In the first inequality above, we have used the fact that $\binom{n}{i} \leq (\frac{en}{i})^i$.)

[18] In fact, we can derive a tighter bound of $(\frac{4}{5} + \epsilon) \cdot v_1 + (\frac{1}{5} - \epsilon) \cdot v_2$ for some $\epsilon > 0.01$.

**Claim D.1.** *For any fixed $i > 0$ (i.e. $i$ can depend on $\epsilon$ but not on $n$), $\lim_{n\to\infty} s_i = 0$.*

*Proof.* It suffices to show that

$$\lim_{n\to\infty} i\log(\frac{en}{i}) - (n-i)\cdot\log(\frac{1}{1-\delta}) = -\infty.$$

To see why the above equation holds, note that $i \le k$ and $k = o(n)$, which imply that

$$i\log(\frac{en}{i}) = o(n),$$
$$(n-i)\cdot\log(\frac{1}{1-\delta}) = \Theta(n),$$

where the first equation holds by L'Hospital's rule. □

Claim D.1 together with (D.1) imply that, for any sufficiently large fixed $j > 0$,

$$\lim_{n\to\infty} \mathbb{P}[v \notin I] \le \delta^j.$$

Since $j$ can be any sufficiently large fixed number, the above inequality implies that

$$\lim_{n\to\infty} \mathbb{P}[v \notin I] = 0. \quad \text{(D.2)}$$

Observe that the above equation holds for any arbitrary small $\epsilon$ (because we imposed no restrictions on $\epsilon$ and the CDF $F$ is strictly increasing and continuous). Therefore, by (D.2), we have shown that $v$ converges in probability to $\overline{u}$, i.e. for all $\epsilon > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}[|v - \overline{u}| > \epsilon] = 0.$$

## D.3 Proof of Proposition 4.3

Let $a_i$ denote the $i$-th agent who chooses an object. Suppose $\pi_i$ is the ordinal preference list of $a_i$, i.e. the list of the objects ordered in a decreasing order with respect to their utilities for agent $a_i$. Let $r_i$ be a random variable denoting the rank of the object assigned to $a_i$ on $\pi_i$.

**Lemma D.2.** *In expectation, at least $n/2$ agents are assigned to their $k$-th ranked object. Furthermore, for any fixed $\epsilon > 0$, at least $n(1-\epsilon)/2$ and at most $n(1-\epsilon)/2$ agents are assigned to their $k$-th ranked object, wvhp.*

*Proof.* Observe that

$$\mathbb{P}\left[r_{i+1} \neq k\right] = \frac{i}{n},$$

which implies that the expected number of students not assigned to their $k$-th ranked object is $\frac{n-1}{2}$. This proves the first part of the claim. A standard application of Chernoff concentration bounds proves the second part. □

### D.3.1 Proof of Proposition 4.3, part (i)

Let $r^* = \max\{k(n), \sqrt{n}\}$, where we have suppressed the argument of $r^*(n)$ for notational brevity. Observe that $r^* = o(n)$. Let $p_i$ denote the probability that $r_i > r^*$. First, we provide an upper bound on $p_i$. Observe that for $i \leq r^*$, $p_i = 0$. For $i \geq r^*$,

$$p_{i+1} = \frac{i}{n} \cdot \frac{i-1}{n-1} \cdots \frac{i-r^*}{n-r^*} \leq \left(\frac{i}{n}\right)^{r^*} \leq e^{-\frac{r^* j}{n}}, \tag{D.3}$$

where $j = n - i$. Observe that, for any $j$ such that $j = \omega(n/r^*)$, we have $p_{i+1} = o(1)$. Choose such arbitrary $j$, namely $j^*$. Therefore, (D.3) implies that for all $i \leq n - j^*$, we have $p_{i+1} = o(1)$.

This fact, together with Proposition 4.2 imply that that the expected utility of any agent under the Inefficient-RSD mechanism is at least $\frac{n-j^*}{n} \cdot (1 - o(1)) \cdot \overline{u}$, which approaches $\overline{u}$ as $n$ approaches infinity.

### D.3.2 Proof of Proposition 4.3, part (ii)

Let $\mu : A \times O$ denote the assignment generated by the Inefficient-RSD mechanism. Let $A'$ denote the subset of agents who are assigned to the object ranked $k$-th on their list. Without loss of generality, suppose that $A' = \{a_1, \ldots, a_N\}$. Recall that, by Lemma D.2, $N \geq \frac{n(1-\epsilon)}{2}$ holds whp, for any $\epsilon > 0$. Let $o_i = \mu(a_i)$ for all $i \in [N]$, and let $O' = \{o_1, \ldots, o_N\}$.

Construct a directed graph $G$ as follows. Let the set of nodes of $G$ be $O'$. For any $o_i \in O'$, add a directed edge to $G$ from $o_i$ to $o$ where $o$ is any object belonging to $O'$ that $a_i$ prefers to $o_i$. Observe that any node in $G$ has out-degree $k-1$ if it belongs to $O'$, and otherwise it has out-degree 0. In the rest of the proof, we show that there exists a sufficiently large constant $k_0$ such that, when $k \geq k_0$, $G$ contains a *large* cycle whp, i.e. a cycle of length $\Theta(n)$. This will complete the proof since any cycle in $G$ corresponds to a Pareto-improving cycle of the same length in $\mu$.

The existence result is based on the probabilistic method, and is inspired from the techniques used in [Frieze and Karonski 2012]. The proof has two steps. In the first step, we show that $G$ does not contain a certain subgraph, whp. (Lemma D.4) In the second step, we use this result to complete the proof. In the rest of the proof, suppose $k_0$ is a sufficiently large constant which will be fixed at the end.

**Definition D.3.** *An* independent pair $(X, Y)$ *is a pair of subsets* $X, Y \in V(G)$ *such that* $(X \times Y) \cap E(G) = \varnothing$. *Size of an* independent pair *is* $\min\{|X|, |Y|\}$.

**Lemma D.4.** *Let* $\beta \in (0, 1)$ *be a constant such that* $\left(\frac{1}{1-\beta}\right)^{k_0-1} \geq \frac{e^2}{4\beta^2}$. *Then, there is no independent pair* $(X, Y)$ *with size at least* $\beta n$, *whp.*

*Proof.* It suffices to show that there exists no independent pair with size exactly $\lfloor \beta n \rfloor$. To avoid notational complexity, we suppose $\beta n = \lfloor \beta n \rfloor$. This will not change any step of the proof.

Fix two subsets $X, Y \subseteq V(G)$ of size $\beta n$. Let $p_{X,Y}$ denote the chance that $(X, Y)$ is an independent pair. First, First, we compute an upper bound on $p_{X,Y}$, and then we use a union bound over all such subsets $X, Y$ to prove the lemma. In the rest of the proof we suppose that $n(1-\epsilon)/2 \leq N \leq n(1+\epsilon)/2$ holds for any arbitrarily small constant $\epsilon > 0$, since this condition holds whp by Lemma D.2.

Observe that

$$p_{X,Y} = \left(\frac{\binom{n-\beta n}{k0-1}}{\binom{n}{k0-1}}\right)^{|X|} \leq (1-\beta)^{n\beta(k_0-1)}.$$

A union bound implies that

$$\begin{aligned}
\sum_{X,Y} p_{X,Y} &\leq \binom{N}{n\beta}^2 (1-\beta)^{n\beta(k_0-1)} \\
&\leq \binom{N}{n\beta}^2 (1-\beta)^{n\beta(k_0-1)} \\
&\leq \left(\frac{e^2(1+\epsilon)^2}{(2\beta)^2}\right)^{n\beta} \left((1-\beta)^{k_0-1}\right)^{n\beta}.
\end{aligned}$$

Therefore, so long as

$$\left(\frac{1}{1-\beta}\right)^{k_0-1} \geq \frac{e^2}{(2\beta)^2},$$

we have $\sum_{X,Y} p_{X,Y} = o(1)$ (since we can take $\epsilon$ to be any arbitrarily small constant). This completes the proof.

□

Set $\alpha = 1/25$ and choose a sufficiently large constant $k_0$ such that $\left(\frac{1}{1-\alpha}\right)^{k_0-1} \geq \frac{e^2}{4\alpha^2}$. (Choosing $k_0 = 175$ satisfies this condition. Our simulations in Section G show that $k_0 = 3$ suffices for the main claim to hold.)

**Corollary D.5** (Corollary of Lemma D.4)**.** *There is no independent pair of size at least $\alpha n$.*

**Lemma D.6.** *Suppose $H$ is a strongly connected induced subgraph of $G$. Then, $H$ contains a cycle of length at least $|V(H)| - 4\alpha n$.*

*Proof.* The proof uses the Depth-First Search (DFS) algorithm. (The reader may recall the definition of DFS from [West 1995] or [Wikipedia 2018a], among many other sources.)

Run DFS on the graph $H$ starting from an arbitrary node. We keep track of two sets of vertices in the course of DFS: the set of nodes that the algorithm has not visited yet, namely $S$, and the set of nodes that the algorithm has finished visiting them, namely $T$. (Recall that when DFS visits a node $v$, it runs DFS recursively from all of its unvisited neighbors. The DFS at node $v$ is finished when all the recursive calls are finished.)

When DFS starts, $|S| = |V(H)|$ and $|T| = 0$. When DFS is finished, $|S| = 0$ and $T = |V(H)|$. On the other hand, at each step (recursive call) of DFS, either $|S|$ decreases by 1 and $|T|$ is unchanged, or $|S|$ is unchanged and $|T|$ increases by 1. Therefore, at some point in the course of the algorithm, we must have $|S| = |T|$. Fix the values of $S, T$ to be their values at that point, and let $P = V(H) \backslash (S \cup T)$. Note that $P$ must form a directed path by the definition of DFS, and that $(S \times T) \cap E(H) = \varnothing$. The proof is almost complete using these two facts: the latter one implies that $(S, T)$ is an independent pair. Corollary D.5 implies that the size of $(S, T)$ is at most $\alpha n$, and therefore, $|S|, |T| \leq \alpha n$. This means $|P| \geq |V(H)| - 2\alpha n$.

Suppose that $|P| \geq 2\alpha n$, otherwise the claim is trivial. Let the $P_1, P_2$ be subsets denoting the first and the last $\alpha n$ nodes of $P$, respectively. $(P_2, P_1)$ cannot be an independent pair by Corollary D.5. Therefore, there must be an edge from some node in $P_1$ to some node in $P_2$. This creates the promised cycle.

□

**Lemma D.7.** *$G$ contains a strongly connected subgraph that contains at least $N/3$ nodes.*

*Proof.* The proof is by contradiction. Suppose the claim is false. Let $C_1, \ldots, C_l$ denote the maximal strongly connected components of $G$ that partition its vertices. Without loss of generality, suppose $G$ contains no edge from any node in $C_i$ to any node in $C_j$, for any $i < j$.

(This is without loss of generality since the condensation of $G$ is an acyclic directed graph [Wikipedia 2018b].) Let $i$ be the smallest integer for which $\sum_{j=1}^{i} |V(C_j)| \geq N/3$. Then, we must have that $\sum_{j=i+1}^{l} |V(C_j)| \geq N/3$. Let $X = \cup_{j=i+1}^{l} C_j$ and $Y = \cup_{j=1}^{i} C_j$. Observe that $(X, Y)$ is an independent pair of size at least $N/3$. This contradicts Corollary D.5.

□

Lemmas D.6 and D.7 together imply that, whp, $G$ contains a cycle of length at least $N/3 - 4\alpha n \geq n(\frac{4(1-\epsilon)}{24} - \frac{4}{25})$, for any arbitrary small constant $\epsilon > 0$. This completes the proof.

### D.3.3 Proof of Proposition 4.3, part (iii)

First, observe that the rank distribution under the Inefficient-RSD mechanism is stochastically dominated by the rank distribution under RSD. Therefore, the average rank under the Inefficient-RSD mechanism is higher than the average rank under RSD, which is at least $\ln n - 1$, as shown by [Knuth 1996]. On the other hand, Lemma D.2 says that, in expectation, at least $n/2$ agents are assigned to their $k$-th ranked object, which means the expected average rank under the Inefficient-RSD mechanism is at least $k/2$. The two latter facts imply that the expected average rank under Inefficient-RSD is at least $\max\{\ln n - 1, k/2\}$.

# E The case of positive common-value components

The main finding of this section is Proposition E.2, a counterpart for Proposition 4.3 in the presence of non-zero common-value components. In part ii of Proposition E.2 we will use the notion of Pareto-improving pairs for capturing Pareto-inefficiencies, whereas in Proposition 4.3 we used the notion of Pareto-improving cycles. We can similarly prove similar findings about Pareto-improving pairs for Proposition 4.3.

First, we need a few definitions. Consider a market with an equal number of agents and objects where objects have unit capacities. An assignment is a function from the set of agents to the set of objects that assigns no two agents to the same object. The utility that an agent $a$ attains from an object $o$ is denoted by $u_a(o)$. An *assignment mechanism* is an algorithm that takes all the utilities $\{u_a(o)\}_{o \in O, a \in A}$ as input and outputs an assignment of agents to objects.

For any assignment $\mu$ the *utility distribution of* $\mu$ is a CDF $D_\mu$ where $D_\mu(u)$ denotes the fraction of agents who attain utility at most $u$ in $\mu$. We say a utility distribution $D_{\mu_1}$

*stochastically dominates* $D_{\mu_2}$ if $D_{\mu_1}(u) \le D_{\mu_2}(u)$ holds for all $u \ge 0$.

Suppose that the utility of an agent $a \in A$ from an object $o \in O$ is given by

$$u_a(o) = v_o + v_o^a, \tag{E.1}$$

where all the random variables $\{v_o\}_{o\in O}$ are drawn independently from a CDF $F$ with finite mean and all the random variables $\{v_o^a\}_{o\in O, a\in A}$ are drawn independently from a strictly increasing CDF $G$ with supportl $[\underline{u}, \overline{u}]$.

Define the CDF $D^*$ by $D^*(x) = F(x - \overline{u})$ for all $x \in [\underline{u} + \overline{u}, 2\overline{u}]$.

**Fact E.1.** *Let $M_1, M_2, \ldots$ be a sequence of markets such that market $M_n$ has $n$ agents and $n$ objects where the agents' utilities are defined by* (E.1)*. Let $\mu_n$ be an assignment of objects to agents in $M_n$. Then, when $\lim_{n\to\infty} D_{\mu_n}$ exists, it is stochastically dominated by $D^*$.*

According to this fact, $D^*$ is an asymptotic "upper bound" on the utility distribution of any assignment; i.e., no assignment mechanism can, asymptotically, attain a "better" utility distribution than the distribution $D^*$. The proof is straight-forward using the law of large numbers; we omit the proof.

To provide the counterpart for Proposition 4.3, we first need to adapt the the definition of the Inefficient-RSD mechanism from Section 4.1 to the setting with non-zero common-value components.

**The Inefficient-RSD mechanism.** Let the preference list $\pi'(a)$ for any agent $a$ be a list of the objects in which object $p$ appears before object $q$ if $v_p^a \ge v_q^a$. The preference profile $\pi'$ is the collection of all these preference lists. Given an increasing function $k(n) : \mathbb{N} \to \mathbb{N}$, run the Inefficient-RSD mechanism, as defined in Section 4.1, with the preference profile $\pi'$ given as its input. This generates an assignment which we will call the *Inefficient-RSD assignment* through out this section.

A pair of agents $a, a'$ is called a Pareto-improving pair in an assignment $\mu : A \to O$ if $a$ prefers $\mu(a')$ to $\mu(a)$ and $a'$ prefers $\mu(a)$ to $\mu(a')$.

**Proposition E.2.** *Let the distributions $F, G$ be strictly increasing CDFs with support $[\underline{u}, \overline{u}]$.*

*i. When $k = o(n)$, the utility distribution of the Inefficient-RSD assignment converges in probability to $D^*$.*

*ii. When $k = \omega(n^{1/2})$, any fixed agent is involved in at least one Pareto-improving pair in the Inefficient-RSD assignment, whp. Furthermore, when $F, G$ are the uniform distribution* *over $[\underline{u}, \overline{u}]$, at least $n - o(n)$ agents are members of pairwise-disjoint Pareto-improving pairs, in expectation.*

The uniformity assumption on $F, G$ in the second part of part (ii) can be replaced with a much more general assumption: the distributions have strictly positive support and are continuous. The proof remains almost identical. The uniformity assumption, however, simplifies the algebraic expressions. We sketch the proof for the more general case in Section E.1.

**Proof of Proposition E.2, Part i.**

The proof is an immediate corollary of Proposition 4.3, part i. Since, in here, the mechanism assigns the objects solely based on the idiosyncratic components, Proposition 4.3 implies that each agent asymptotically attains utility approaching $\overline{u}$ from the idiosyncratic component. Since the common-value component of the object owned by any agent is drawn independently, its distribution is equal to $F$. This proves the claim.

**Proof of Proposition E.2, Part ii.**

For this proof we suppose that the support of the distributions is the unit interval, i.e. $[\underline{u}, \overline{u}] = [0, 1]$. This is just a normalization that simplifies notation; the same proof works for the general case. Let $\mu$ denote the Inefficient-RSD assignment. Let $a_1, \ldots, a_n$ denote the agents in the order they choose objects, and let $r_a$ denote the rank of the object of agent $a$ on her list. Let $l = k/3$ and define the interval $I = (2l, 3l]$.

**Claim E.3.** *For any agent $a \in A$, $r_a \in I$ holds, whp.*

*Proof.* Let $p_i$ denote the probability that $r_i \notin I$. First, we provide an upper bound on $p_i$. Observe that for $i \leq l$, $p_i = 0$. For $i \geq l$,

$$p_{i+1} = \frac{i}{n} \cdot \frac{i-1}{n-1} \cdots \frac{i-l}{n-l} \leq \left(\frac{i}{n}\right)^l \leq e^{-\frac{lj}{n}}, \tag{E.2}$$

where $j = n - i$. Observe that, for any $j$ such that $j = \omega(n/l)$, we have $p_{i+1} = o(1)$. Choose such arbitrary $j$, namely $j^*$. Therefore, the above inequality implies that for all $i \leq n - j^*$, we have $p_{i+1} = o(1)$. This concludes the proof since the order of agents is chosen uniformly at random. □

We need a few definitions for the rest of the proof. Let $N = |\{a : r_a \in I\}|$. Claim E.3 implies that $N = n - o(n)$. Let the set $B = \{b_1, \ldots, b_N\}$ denote the set of agents $\{a : r_a \in I\}$.

Also, let $o_i = \mu(b_i)$, for all $b_i \in B$. Denote the common-value component corresponding to objects $o_i$ by $v_i$. Without loss of generality suppose that $v_1 \leq \ldots \leq v_N$. Furthermore, by relabeling the agents we can assume that $b_i = i$ for all $i \in [N]$.

**Claim E.4.** *Any agent $b \in B$ is involved in at least one Pareto-improving pair, whp.*

*Proof.* The proof uses the Principle of Deferred Decisions. While running the Inefficient-RSD mechanism, suppose that the preference lists of agents are filled gradually in the course of the algorithm: whenever an agent goes to the next position on her list, a random draw from the remaining objects will fill out that position.

From the proof of Claim E.3, it implies that the positions $1, \ldots, 2l$ will be unfilled on the preference lists of at least $N = n - o(n)$ agents, i.e. the agents in the set $B$. For any two distinct $i, j \in [N]$, we let $p_{i,j}$ denote the chance that agents $b_i, b_j$ form a Pareto-improving pair. By the Principle of Deferred decisions, the chance that agent $b_j$ prefers object $o_i$ to $o_j$ is at least $l/n$. Similarly, the chance that $b_i$ prefers $o_j$ to $o_i$ is at least $l/n$. Therefore, the chance that $(b_i, b_j)$ is a Pareto-improving pair is at least $l^2/n^2$.

It is straight-forward to verify that the chance that agent $b_i$ is not involved in any Pareto-improving pair is at least

$$\prod_{j \in B, j \neq i} (1 - p_{i,j}) \leq (1 - l^2/n^2)^{n-1} \leq e^{\frac{l^2(n-1)}{n^2}} = o(1)$$

□

In the rest of the proof, we show that there exists $\Theta(n)$ pairwise-disjoint Pareto-improving pairs in the Inefficient-RSD assignment. This requires a few definitions.

Let $x$ be a positive number which we will fix later. Partition the unit interval to segments of length $x$: the partition would be

$$(0, x], (x, 2x], (2x, 3x] \ldots,$$

where $x$ is chosen so that $x = \omega(\frac{\log n}{n})$ and $x = o(\frac{k}{n})$. (The last element of the partition may be an interval with length smaller than $x$, in which case we will safely ignore the last element in the analysis below.)

For any integer $i \geq 0$, let $A_i$ be the set of indices $j$ such that $v_j \in (ix, ix + x]$. (Recall that, since the object with common-value $v_i$ is owned by the agent $b_i$, and since we supposed $b_i = i$ for all $i \in [N]$, we can also interpret $A_i$ as a set of agents.)

**Claim E.5.** *For any constant $\epsilon > 0$ and any $i \geq 0$, wvhp*

$$|A_i| \in [(1-\epsilon)xN, (1+\epsilon)xN]$$

*Proof.* Recall that the common-value components of the objects are drawn iid from the uniform distribution over the unit interval. Therefore, the expected number of objects in $B$ whose common-value component falls into a subinterval of the unit interval with length $\alpha$ is $\alpha N$. Since $x = \omega(\frac{\log n}{n})$, the length of the subinterval defining $A_i$ (i.e. the interval $(ix, ix+x]$) would be $\omega(\frac{\log n}{n})$, and therefore the expected size of $A_i$ would be $\omega(\log n)$. The rest of the proof follows from a standard application of Chernoff concentration bounds. □

Next, we show that the set of agents $A_i$ contains at least $|A_i|/2 - o(|A_i|)$ pairwise-disjoint Pareto-improving pairs, whp. This would imply that the Inefficient-RSD assignment contains at least $n/2 - o(n)$ pairwise-disjoint Pareto-improving pairs, in expectation, which is the claim. The next lemma completes the proof.

**Lemma E.6.** *For any $i$, the set of agents $A_i$ contains at least $|A_i|/2 - o(|A_i|)$ pairwise-disjoint Pareto-improving pairs, whp.*

*Proof.* First, we construct an undirected graph $G$ as follows. Let $V(G) = A_i$, and connect a node $z$ to node $z'$ iff the rank of $z$ on $\pi(z')$ is at most $l$ and the rank of $z'$ on $\pi(z)$ is at most $l$. Let $t$ denote the number of nodes in $G$. The proof contains two steps: we will show that (i) any edge in $G$ corresponds to a Pareto-improving pair, wvhp, and (ii) for any constant $\beta < 1$, $G$ contains a matching that covers at least $\beta|V(G)|$ nodes, whp. These two steps, together with a union bound, conclude the lemma.

**Step (i)** We can assume that the positions $1, \ldots, l$ on the preference lists of all agents in $B$ are unfilled (by the Principle of Deferred Decisions). Let $E_{b,b'}$ denote the event that $\mu(b)$ has rank $l$ or better on $\pi(b')$ and $\mu(b')$ has rank $l$ or better on $\pi(b)$. We will show that that if $E_{b,b'}$ holds, then $b, b'$ must form a Pareto-improving pair, wvhp.

**Claim E.7.** *For any two agents $b, b' \in A_i$, conditioned on $E_{b,b'}$, agents $b, b'$ form a Pareto-improving pair, wvhp.*

*Proof.* Recall that for any agent $a$ and object $o$, the utility of $a$ from $o$ is

$$u_a(o) = v_o + v_o^a.$$

Observe that $|v_{\mu(b)} - v_{\mu(b')}| \leq x$, by the definition of $A_i$. Then, if we show that We have

$$v^b_{\mu(b)} + x < v^b_{\mu(b')} \tag{E.3}$$
$$v^{b'}_{\mu(b')} + x < v^{b'}_{\mu(b)}, \tag{E.4}$$

To complete the proof we will show that the above two inequalities hold wvhp. To prove (E.3), it suffices to show that

$$v^b_{\mu(b)} < 1 - 2l(1-\epsilon)/n,$$
$$v^b_{\mu(b')} > 1 - l(1+\epsilon)/n,$$

hold wvhp. ((E.4) is proved similarly.) The two equations above hold whvp because of Chernoff concentration bounds: Fix an arbitrary small constant $\epsilon > 0$. Since the idiosyncratic components are drawn iid, then, for any agent there are at least $l$ of the objects that give her and idiosyncratic utility component at least $1 - \frac{l(1+\epsilon)}{n}$, wvhp. This proves the latter inequality. Similarly, for any agent there are at most $2l$ of the objects that give her an idiosyncratic utility component at least $1 - \frac{2l(1-\epsilon)}{n}$, wvhp. This proves the former inequality.

□

**Step(ii)** In this step, we will show that for any constant $\beta < 1$, $G$ contains a matching that covers at least $\beta|V(G)|$ nodes, whp. First, we need a definition. An independent set in $G$ is a set of nodes which are pairwise non-adjacent. To prove the lemma, we show that, whp, $G$ does not contain an independent set of size $\delta|V(G)|$, for any arbitrary small constant $\delta > 0$. This completes this step since if the maximum size of a matching in $G$ is $M$, then $G$ must contain an independent of size $|V(G)| - M$.

The proof uses the Principle of Deferred Decisions. While running the Inefficient-RSD mechanism, suppose that the preference lists of agents are filled gradually in the course of the algorithm: whenever an agent goes to the next position on her list, a random draw from the remaining objects will fill out that position.

For any two distinct $b, b' \in B$, we let $p_{b,b'} = \mathbb{P}[E_{b,b'}]$. By the Principle of Deferred decisions, the chance that agent $b$ prefers object $\mu(b')$ to $\mu(b)$ is at least $l/n$. Similarly, the chance that $b'$ prefers $\mu(b)$ to $\mu(b')$ is at least $l/n$. Therefore, $p_{b,b'} \geq l^2/n^2$. It is straight-forward to verify that, for any subset $B' \subseteq B$

$$\mathbb{P}\left[\wedge_{b,b' \in B'} \overline{E_{b,b'}}\right] \leq \prod_{b,b' \in B} (1 - p_{b,b'})$$
$$\leq (1 - l^2/n^2)^{\binom{|B'|}{2}}.$$

When $|B'| > \delta|B|$ for a constant $\delta > 0$, the above inequality implies that

$$\mathbb{P}\left[\wedge_{b,b' \in B'} \overline{E_{b,b'}}\right] \leq (1 - l^2/n^2)^{\frac{\delta^2 |B|^2}{2}}$$
$$\leq e^{-\left(\frac{\delta l |B|}{\sqrt{2}n}\right)^2}$$

A union bound then implies that $G$ contains an independent set of size at least $\delta|B|$ with probability at most

$$2^{|B|} e^{-\left(\frac{\delta l |B|}{\sqrt{2}n}\right)^2} = o(1),$$

where the above equation holds because $|B| = n - o(n)$ and $l = \omega(n^{1/2})$. This completes the proof of the lemma.

□

## E.1 Dismissing the uniformity assumption in part ii

The uniformity assumption is used in the latter part of part ii in Proposition E.2. Below we dismiss the uniformity assumption.

**Proposition E.8.** *Let the distributions $F, G$ have continuous and bounded PDFs with strictly positive support over $[\underline{u}, \overline{u}]$. Then, at least $n - o(n)$ agents are members of pairwise-disjoint Pareto-improving pairs, in expectation.*

For the proof, without loss of generality, we can suppose that the support of the distributions is the unit interval, i.e. $[\underline{u}, \overline{u}] = [0, 1]$. This is just a normalization that simplifies notation. By assumption, there exists a constant $\Delta > 0$ that bounds the PDFs of $F, G$ from above.

The proof follows the same steps as the proof for the case of uniform distributions. We start by partitioning the interval $[1, k]$ to 3 intervals by choosing two numbers $l_1, l_2$. The partition would be $[1, l_1]$, $(l_1, l_2)$, $[l_2, k]$. In the previous proof, we chose $l_1 = k/3$ and $l_2 = \frac{2k}{3}$. In this proof, we choose $l_2 = \frac{2k}{3}$, but we choose a different $l_1$, as follows. Since the PDFs of $F, G$ have strictly positive support, then there also exists a positive constant $\delta > 0$ which bounds the PDFs from below. We set $l_1 = \frac{\delta k}{3\Delta}$. As in previous proof, we denote the interval $[l_2, k]$ by $I$. As in the previous proof, we also choose a parameter $x$ such that $x = \omega(\frac{\log n}{n})$ and $x = o(\frac{k}{n})$

We have chosen the numbers $l_1, l_2$ and $x$ so that some crucial properties, which make the previous proof work, still hold. Loosely speaking, the properties are as follows.

i. The interval $[l_2, k]$ should be large enough so that are for almost all agents $a \in A$, $r_a \in I$ holds. Almost the identical proof for Claim E.3 guarantees this property.

ii. The number $x$ should be large enough so that there are sufficiently many nodes in the graph $G$ (which is defined similar as before). Almost the same proof for Claim E.5 guarantees this property, with a minor difference that Claim E.5 now changes to $|A_i| \in [(1-\epsilon)\delta x N, (1+\epsilon)xN]$.

iii. The number $x$ should be sufficiently small so that Step i in the proof of Lemma E.6 holds similarly.

iv. The interval $[1, l_1]$ should be large enough so that each node in $G$ has sufficiently many neighbors in $G$. By sufficiently many neighbors, we mean sufficiently many so that Step ii in the proof of Lemma E.6 holds similarly.

All these conditions are satisfied by our choices of $l_1, l_2, x$. This is all the previous proof needs to be applicable here.

# F Proof of Proposition 5.1

In the proof we use the notions of the Deferred Acceptance (DA) algorithm, the man-proposing DA, and the man-optimal assignment. The reader may recall the definitions from [Roth and Sotomayor 1990], among other possible sources.

The proof directly follows from the next lemma together with a result of [Pittel 1992] that shows in the man-optimal matching, all men are matched to women that are ranked $\ln^2 n$ or better on the list of their match, whp.

**Lemma F.1.** *Consider a marriage market consisting $n$ men and $n$ women. Let $r(n) : \mathbb{N} \to \mathbb{R}_+$ be any positive function such that $r(n) = o(\frac{n}{\ln n})$. Then, there exists a stable matching in which each women, whp, is assigned to a man ranked $r(n)$ or worse on her list.*

*Proof.* The proof adapts the proof of of Lemma C.1. from [Ashlagi and Nikzad 2018]. We prove the lemma in two steps. In the Step (i), we show that under the man-proposing differed acceptance, each woman receives at most $(1+\epsilon)\ln n$ proposals whp, where $\epsilon > 0$ is an arbitrary small constant independent of $n$. In Step (ii) we prove the claim of the lemma using the result of Step (i).

**Step (i)** The proof idea is defining another stochastic process that we denote by $\mathcal{B}$. Process $\mathcal{B}$ is defined by a sequence of binary random variables $X_1, \ldots, X_k$, where $k = (1+\delta)n\ln n$ for some arbitrary small constant $\delta > 0$. Each random variable in this sequence takes the value 1 with probability $\frac{1}{n-\ln^2 n}$, and 0 otherwise. For convenience, we also refer to these random variables by *coins*, and the process that determines the value of a random variable by *coin-flip*.

Fix a woman $w$. Define $X = \sum_{i=1}^{k} X_i$. The goal is to show that $X$ is a good upper bound on the number of proposals that are received by $w$. The high-level idea is based on two facts: First, the total number of proposals made by all men is stochastically dominated by the solution to the coupon-collector problem, and so, whp is at most $k$. Second, by [Pittel 1992], there is no man who makes more than $\ln^2 n$ proposals, whp. So, each proposal is made to $w$ with probability at most $\frac{1}{n-\ln^2 n}$. Consequently, the number of proposals made to $w$ cannot be more than $\frac{k(1+\delta')}{n-\ln^2 n}$ whp, for any constant $\delta' > 0$. (The latter fact is a direct consequence of the Chernoff bound which is applicable since the coin flips are independent).

To formalize the above argument, we couple the stochastic process governing DA with a new random process, $\mathcal{B}$, which is a simple coin-flipping process: it flips $k$ coins independently, all with success probabilities $\frac{1}{n-\ln^2 n}$. The coupled process, $(\mathrm{DA}, \mathcal{B})$, has two components, one for each of the original random processes. For each of the components, the marginal distribution induced on its sample paths is identical to the distribution of the sample paths in the original process, but there is no restriction on the joint distribution of the sample paths in the coupled process. Next, we define a simple coupling in which in almost all sample paths (i.e. whp), the number of successful coin flips is an upper bound on the number of proposals made to $w$. Whenever a man $m$ wants to make a proposal during DA, process $\mathcal{B}$ flips the next coin. Then:

1. If $m$ has made a proposal to $w$ before, ignore the coin flip, and let $m$ make a proposal to a woman picked uniformly at random from the set of women to whom he has not proposed yet.

2. If $m$ has not made a proposal to $w$ before, then

(a) Suppose $m$ has made $d \leq \ln^2 n$ proposals so far. (Otherwise, ignore this sample path)

i. If the coin flip was successful: with probability $\frac{n-\ln^2 n}{n-d}$ let $m$ make a proposal to $w$, and otherwise (with probability $1 - \frac{n-\ln^2 n}{n-d}$ ) let $m$ make a proposal to a woman picked uniformly at random from the set of women to whom he has not proposed yet.

ii. If the coin flip was not successful: let $m$ make a proposal to a woman picked uniformly at random from the set of women to whom he has not proposed yet, excluding $w$.

It is straight-forward to verify that this defines a valid coupling of DA, $\mathcal{B}$. Note that the total number of successful coin flips in $\mathcal{B}$ is an upper bound on the total number of proposals made to $w$ in the coupled DA process, in almost all sample paths (i.e. whp). Therefore, as we explain next, we can apply the argument mentioned in the beginning of the proof to conclude the lemma.

[Pittel 1992] shows that, whp, no man makes more than $\ln^2 n$ proposals. Therefore, whp, a sample path is not ignored in line (2-a); that is, only a vanishing fraction of sample paths are ignored. On the other hand, we mentioned earlier that a standard application of Chernoff bounds implies that $X$ is at most $\frac{k(1+\delta')}{n-\ln^2 n}$ whp, for any constant $\delta' > 0$. A union bound then implies that, whp, woman $w$ receives at most

$$\frac{k(1+\delta')}{n-\ln^2 n} = \frac{(1+\delta)(1+\delta')n\ln n}{n-\ln^2 n}$$

proposals. Setting $\epsilon$ to be a constant larger than $\delta + \delta' + \delta\delta'$ completes the first step.

**Step (ii)** Fix a woman $w$. For any $k \leq (1+\epsilon)\ln n$, conditioned on $w$ receiving $k$ proposals, the chance that she is assigned to a man ranked $r(n)$ or worse on her list is

$$\left(1 - \frac{r(n)}{n}\right)^k \geq 1 - \frac{k \cdot r(n)}{n} = 1 - o(1).$$

Also, note that $w$ receives no more than $(1+\epsilon)\ln n$ proposals, whp. Therefore, a union bound implies that $w$ is assigned to a man ranked $r(n)$ or worse on her list, whp.

□

# G Simulations for Proposition 4.3

We provide some computational experiments which show that setting $k_0 = 3$ makes $k_0$ sufficiently large for the purpose of Proposition 4.3, part ii.

In each experiments, after fixing $n$, we generate 1000 random markets, each market involving $n$ agents and $n$ objects. We run the Inefficient-RSD mechanism with $k = 3$ on each market. In Table 1, we report the fraction of random markets in which the Inefficient-RSD assignment contains a Pareto-improving cycle. We observe that the empirical probability of the existence of a Pareto-improving cycle is essentially 1.

Table 1 also reports another statistic in its last column, which is the average length of a "large" cycle that we can find (conditional on the existence of a cycle). Since the algorithmic problem of finding the largest cycle in a graph is NP-Complete, finding the largest Pareto-improving cycle in our problem is also quite difficult (computationally). We therefore use a heuristic algorithm using the Depth-First Search (DFS) method [West 1995], which does not necessarily find the largest possible cycle. Therefore, the statistics reported in the last column of the table is a lower bound on the length of the largest possible Pareto-improving cycle. Observe that the lower bound grows roughly proportional to $n$.

| $n$ | Probability | Average length |
|---|---|---|
| $2 \times 10^2$ | 0.998 | 19.09 |
| $2 \times 10^3$ | 1 | 162.22 |
| $2 \times 10^4$ | 1 | 1584.07 |
| $2 \times 10^5$ | 1 | 15779.3 |

Table 1: The second column reports the empirical probability that a cycle exists and the third column reports the average length of the cycle found by our heuristic.