Efficiency and Stability in Large Matching Markets *

Yeon-Koo Che[†] Olivier Tercieux[‡]

November 4, 2014

Preliminary and incomplete.

Abstract

We study efficient and stable mechanisms in matching markets when the number of agents is large and individuals' preferences are drawn randomly from a class of distributions allowing for both common and idiosyncratic components. In this context, as the market grows large, all Pareto efficient mechanisms—including top trading cycles, serial dictatorship, and their randomized variants—are asymptotically payoff equivalent "up to the renaming of agents," yielding the utilitarian upper bound in the limit. If agents' priorities with objects are also randomly drawn but their common values for objects are heterogenous, then well-known mechanisms such as deferred acceptance and top trading cycle mechanisms fail either efficiency or stability even in the asymptotic sense. We propose a new mechanism that is asymptotically efficient, asymptotically stable and asymptotically incentive compatible.

Keywords: Large matching markets, Pareto efficiency, Stability, Fairness, Payoff equivalence, Asymptotic efficiency, and asymptotic stability.

*We are grateful to Charles Maurin and Xingye Wu for their research assistance, and to Yash Kanoria, Fuhito Kojima, Scott Kominers, Bobak Pakzad-Hurson, Debraj Ray, Rajiv Sethi, and seminar participants at Columbia, KAEA Conference, Maryland, NYC Market Design Workshop, NYU, PSE Market Design conference, UCL, Warwick Micro theory conference, Wisconsin, and WCU Market Design conference at Yonsei University, for helpful comments.

[†]Department of Economics, Columbia University, USA. Email: yeonkooche@gmail.com.

[‡]Department of Economics, Paris School of Economics, France. Email: tercieux@pse.ens.fr.

1 Introduction

Assigning indivisible resources such as housing, public school seats, employment contracts, branch postings and human organs are important subject for modern market design. Two central goals in designing such matching markets are efficiency and stability. Pareto efficiency means exhausting all gains from trade, a basic desideratum in any allocation problem. Stability means eliminating incentives for individuals to "block"—-or deal around—a suggested assignment. Not only is stability crucial for long-term sustainability of a market, as pointed out by Roth and Sotomayor (1990), but it also guarantees a sense of fairness in eliminating so-called "justified envy."[1] Eliminating justified envy means, for instance in the school choice context, that no student would lose a school seat to another with a lower priority at that school.

Thanks to the recent progress in matching theory, there are by now well-established mechanisms for attaining each of these two goals and for balancing the tradeoffs between them when they are in conflict. For instance, Gale and Shapley's agent-proposing deferred acceptance algorithm (in short, DA) produces a stable matching (Gale and Shapley, 1962), and does so at the minimal efficiency loss.[2] Meanwhile, a range of mechanisms such as serial dictatorship and top trading cycles (henceforth, TTC) attain Pareto efficiency (Shapley and Scarf, 1974). While they do not satisfy stability, a version of TTC allowing agents to trade their priorities minimizes the incidence of instabilities among Pareto-efficient and strategy-proof mechanisms (Abdulkadiroglu, Che, and Tercieux, 2013).[3]

These knowledges are clearly useful. Yet, they leave open several fundamental and practical questions. First, Pareto efficiency is a very weak standard for efficiency, compat-

[1]See Balinski and Sönmez (1999) and Abdulkadiroglu and Sonmez (2003). This fairness property may be more important in applications such as school choice, where the supply side under a public control, so strategic blocking is not a serious concern.

[2]Given strict preferences, DA yields a stable matching that Pareto dominates all other stable matchings for the agents on the proposing side (Gale and Shapley, 1962).

[3]This version of TTC has been proposed by Abdulkadiroglu and Sonmez (2003): In each round, each agent points to the most preferred (acceptable) object that remains in the market, and each object points the agent with the highest priority. A cycle is then formed, and the agents involved in a cycle are assigned the objects they point to, and the same procedure is repeated with the remaining agents and the remaining objects, until the market is all exhausted. Abdulkadiroglu, Che, and Tercieux (2013) shows that this version of TTC is envy minimal in one-to-one matching in the sense that there is no efficient and strategy-proof mechanism that entails a smaller set of blocking pairs than TTC (smaller in the set inclusion sense) for all preferences, strictly so for some preferences. This result does not extend to the many-to-one matching, however.

ible with a range of outcomes vastly different in terms of individual welfare. For instance, a serial dictatorship may assign an individual his most favorite object or the object no other individuals want, depending on his serial order. Likewise, TTC can lead to different outcomes depending on how the individuals' ownership rights are specified. We do not yet know how they differ in terms of utilitarian welfare or in terms of the distribution of agents' payoffs, and the literature has yet to produce a clear prescription on which efficient mechanism should be chosen out of so many.

Second, while the tradeoff between efficiency and stability is well understood, it remains unclear how best to resolve the tradeoff when both goals are important. As noted above, the standard approach is to attain one goal with the minimal sacrifice of the other. Whether this is the best way to resolve the tradeoff is far from clear. For instance, one can imagine a mechanism that is neither stable nor efficient but may be superior to DA and TTC because it involves very little loss on each objective.

The purpose of the current paper is to answer these questions and in the process provide useful insights on practical market design. These questions remain outstanding since our analytical framework is so far driven primarily by the "qualitative" notions of the two goals. To make progress, we therefore need to relax them "quantitatively." And this requires some structure on the model. First, we consider markets that are "large" in the number of participants as well as in the number of object types. Large markets are clearly relevant in many settings. For instance, in the US Medical Match, each year about 20,000 applicants participate to fill positions at 3,000 to 4,000 programs. In NYC School Choice, about 90,000 students apply each year to 500 high school programs. Second, we assume the agents' preferences are generated randomly according to some reasonable distributions. Specifically, we consider a model in which each agent's utility from an object depends on a common component (i.e., that does not vary across agents) and an idiosyncratic component that is drawn at random independently (and thus varies across the agents).

Studying the limit properties of a large market with random preferences generated in this way provides a framework for answering our questions. In particular, this framework enables us to perform meaningful "quantitative" relaxations of the two desiderata: we can look for mechanisms that are **asymptotically efficient** in the sense that, as the economy grows large, with high probability (i.e., approaching one), the proportion of agents who would gain discretely from a Pareto improving assignment vanishes, and mechanisms that are **asymptotically stable** in the sense that, with high probability, the proportion of agents who would have justified envy toward a significant number of agents vanishes in a sufficiently large economy.

Our findings are as follows.

First, all Pareto efficient mechanisms yield aggregate payoffs, or utilitarian welfare, that converge to the same limit—more precisely the utilitarian optimum—as the economy grows large (in the sense described above). This result implies that in large economies the alternative efficient mechanisms become virtually indistinguishable in terms of the aggregate payoff distribution of the participants. In other words, agents' payoffs are asymptotically equivalent across different efficient mechanisms, up to the "renaming" of agents. This result implies that there is no reason to favor one efficient mechanism over another. From a policy perspective, this means that a Pareto efficient allocation favoring or prioritizing a certain group of individuals would not significantly harm utilitarian welfare in a large market.

Second, considering an environment in which the agents' priorities with the objects are drawn at random, we find that the efficiency loss from the DA and the stability loss from TTC do not disappear when the agents' preferences for the objects are significantly correlated. Possible inefficiencies of DA and possible instabilities of TTC are well known from the existing literature; our novel finding here is that they remain "quantitatively" significant (even) in the large market.

The reasons can be explained in intuitive terms. Suppose the objects come in two tiers, high quality and low quality, and every high-quality object is preferred to every low-quality object for each agent regardless of his idiosyncratic preferences. In this case, the (agent-proposing) DA has all agents compete first for every high-quality object before they start proposing to a low-quality object. Such a competition means that in a stable matching—including agent-optimal stable matching—the outcome is dictated largely by how the objects rank the agents not by how the agents rank the objects. In other words, the competition among agents causes the stability requirement to entail a significant welfare loss for the agents.

Meanwhile, under TTC, with non-vanishing probability, a significant proportion of agents who are assigned low-quality objects exhibit justified envy toward a significant number of agents who obtain high quality objects. The reason is that many of these latter agents obtain high-quality objects through the trading of their priorities. Given the nature of TTC, they have high priorities with the objects they are *trading off*, but they could very well have very low priorities with the objects they are *trading in.* This finding is not only an interesting theoretical finding but also has an important practical market design implication, since it suggests that the standard approach of achieving one goal with the minimal sacrifice of the other may not be the best.

These results are consistent with Abdulkadiroglu, Pathak, and Roth (2009), who find that out of about 80,000 eighth grade students assigned to the New York City public high schools in 2006-2007, about 5,800 students would be made better off from a Pareto improving rematching. They also show that if these students were assigned via an efficient mechanism (unlike the practice of DA), then about 35,000 students would have justified envy.[4]

Motivated by these results, we develop a new mechanism that is both asymptotically efficient and asymptotically stable. This mechanism runs a (modified) DA in multiple stages. Specifically, all agents are ordered in some way, and following that order, each agent takes a turn *one at a time* applying to the best object that has not yet rejected him[5] and the proposed object accepts or rejects the applicant, much as in the standard DA. If at any point an agent applies to an object that holds an application, one agent is rejected, and the rejected agent in turn applies to the best object among those that have not rejected him. This process goes on until an agent makes a certain "threshold" number of offers for the first time. Then the stage is terminated at that point, and all the tentative assignments up to that point become final. The next stage then begins with the last agent (who triggered termination of the last stage) applying to the best remaining object. The stages proceed in this way until no rejection occurs.

This "staged" version of DA resembles the standard DA except for one crucial difference: the mechanism periodically terminates a stage and finalizes the tentative assignment up to that point. The event triggering the termination of a stage is the number of offers an agent makes during the stage exceeding a certain threshold. Intuitively, the mechanism turns on a "circuit breaker" whenever the competition "overheats" to a point that puts an agent at the risk of losing an object he ranks highly to an agent who ranks it relatively lowly (more precisely below the threshold rank). This feature ensures that an object assigned at each stage does go to an agent who ranks it relatively highly among those objects available at that stage.

[4]Note that the efficient matching does not coincide with TTC. Instead, Abdulkadiroglu, Pathak, and Roth (2009) produced an efficient matching by first running DA and then running a Shapley-Scarf TTC from the DA assignment. We expect the figures to be comparable if TTC were run to produce an efficient matching.

[5]DA where offers are made according to a serial order was first introduced by McVitie and Wilson (1971).

Given the independent drawing of idiosyncratic shocks, the "right" threshold turns out to be of order between $\log^2(n)$ and n. Given the threshold, the DA with a circuit breaker produces an assignment that is both asymptotically stable and asymptotically efficient. The analytical case for this mechanism rests on the limit analysis, but the mechanism appears to perform well even away from the limit. Our simulation shows that, even for moderately large market and for a more general preference distribution, our mechanism performs considerably better than DA in terms of utilitarian welfare and entails significantly less stability loss than efficient mechanisms such as TTC.

One potential concern about this mechanism is its incentive property. While the mechanism is not strategy proof, the incentive problem does not appear to be severe. A manipulation incentive arises only when an agent is in a position to trigger the circuit breaker since then the agent may wish to apply to some object safer instead of a more popular one with high probability of rejecting him. The probability of this is one out of the number of agents assigned in the current stage, which is in the order of n, so with a sufficient number of participants, the incentive issue is rather small. Formally, we show that the mechanism induces truthful reporting as an ϵ-Bayes Nash equilibrium.

Our DA mechanism with a circuit breaker bears some resemblance to the features that are found in popular real-world matching algorithms. The "staged termination" feature is similar to the school choice program used to assign students to colleges in China (Chen and Kesten (2013)). More importantly, the feature that prohibits an agent from "outcompeting" another over an object that the former ranks lowly but the latter ranks highly is present in the truncation of participants' choice lists, which is practiced in virtually every implementation of the DA in real settings. Our large market result could provide a potential rationale for the practice that is common in actual implementation of DA but has been so far difficult to rationalize (see Haeringer and Klijn (2009), Calsamiglia, Haeringer, and Klijn (2010) and Pathak and Sömez (2013)).

Related Literature

The present paper is connected with several strands of literature. First, it is related to the literature that studies large matching markets, particularly those with large number of object types and random preferences; see Immorlica and Mahdian (2005), Kojima and Pathak (2008), Lee (2012), Knuth (1996), Pittel (1989) and Ashlagi, Kanoria, and Leshno (2013). The first three papers are concerned largely with the incentive issues arising in

DA. The last three papers are concerned with the ranks of the partners achieved by the agents on the two sides of the market under DA, so they are closely related to the current paper whose focus is on the payoffs achieved by the agents. In particular, our asymptotic inefficiency result of DA follows directly from Ashlagi, Kanoria, and Leshno (2013). Unlike these papers, our paper considers not just DA but also other mechanisms and also has broader perspectives dealing with efficiency and stability.

Another strand of literature studying large matching markets considers a large number of agents matched with a finite number of object types (or firms/schools) on the other side; see Abdulkadiroglu, Che, and Yasuda (2008), Che and Kojima (2010), Kojima and Manea (2010), Azevedo and Leshno (2011), Azevedo and Hatfield (2012) and Che, Kim, and Kojima (2013), among others. The assumption of finite number of object types enables one to use a continuum economy as a limit benchmark in these models. At the same time, this feature makes both the analysis and the insights quite different. The two strands of large matching market models capture issues that are relevant in different real-world settings and thus complement each other.[6]

Methodologically, the current paper utilizes the framework developed in the random graph and random mapping theory; see Bollobas (2001) and Dawande, Keskinocak, Swaminathan, and Tayur (2001) for instance.

2 Set-up

We consider a model in which a finite set of agents are assigned a finite set of objects, at most one object for each agent. Since our analysis will involve studying the limit of a sequence of infinite economy, it is convenient to index the economy by its size n. An n-**economy** $E^n = (I^n, O^n)$ consists of **agents** I^n and **object types** O^n, where $|I^n| = n$. For much of the analysis, we shall suppress the superscript n for notational ease.

The object types can be interpreted as schools or housing types. Each object type o has $q_o \geq 1$ **copies** or **quotas**. Since our model allows for $q_o = 1$ for all $o \in O^n$,

[6]The latter model is more appropriate for situations in which there are a relatively small number of institutions each with a large number of positions to feel. School choice in some district such as Boston Public Schools could be a suitable application, since only a handful of schools each enroll hundreds of students. The former model is descriptive of settings in which there are numerous participants on both sides of the market. Medical matching and school choice in some district such as New York Public Schools would fit the description.

one-to-one matching is a special case of our model. We assume that total quantity is: $Q^n = \sum_{o\in O^n} q_o = n$. In addition, we assume that the number of copies of each object is uniformly bounded, i.e., there is $\bar{q} \geq 1$ such that $q_o \leq \bar{q}$ for all $o \in O^n$ and all n. The assumption that $Q^n = n$ is only for convenience as long as it grows at order n our results will go through. Similarly, the assumption that the number of copies of each object is uniformly bounded is not necessary as long as it grows at a rate which is low enough.[7]

Throughout, we shall consider a general class of random preferences that allows for a positive correlation among agents on the objects. Specifically, each agent $i \in I^n$ receives utility from obtaining object type $o \in O^n$:

$$U_i(o) = U(u_o, \xi_{i,o}),$$

where u_o is a common value, and the *idiosyncratic shock* $\xi_{i,o}$ is a random variable drawn independently and identically from $[0, 1]$ according to the uniform distribution.[8]

We further assume that the function $U(\cdot,\cdot)$ takes values in $\mathbb{R}_+$, is strictly increasing in the common values and strictly increasing and continuous in the idiosyncratic shock. The utility of remaining unmatched is assumed to be 0 so that all agents find all objects acceptable.[9]

We assume that the agents' common value for object $o \in O$, u_o, takes an arbitrary value in $[0, 1]$ in an n-economy, and its population distribution is given by a cumulative distribution function (CDF):

$$X^n(u) = \frac{\sum_{o\in O^n:u_o\leq u} q_o}{n}$$

interpreted as the fraction of the objects whose common value is less than or equal to u, and by

$$Y^n(u) = \frac{|\{o \in O^n | u_o \leq u\}|}{n},$$

interpreted as the fraction of the object types whose common value is no greater than u.

We assume that these CDFs converge to well-defined limits, X and Y, in the Lévy metric. To be precise, for any two distributions, F and G, consider their distance measured

[7]As will be clear from footnote 28, we can allow $\bar{q}$ to grow in n as long as its growth rate is capped at $O(n/log(n))$.

[8]This assumption is without loss, as long as the type distribution is atom-less and bounded, since one can always focus on the quantile corresponding to the agent's type, as a normalized type, and redefine the payoff function as a function of the normalized type.

[9]This feature does not play a crucial role for our results, which go through as long as a linear fraction of objects are acceptable to all agents.

in Lévy metric:

$$L(F,G) := \inf \{\delta > 0 | F(z-\delta) - \delta \leq G(z) \leq F(z+\delta) + \delta, \forall z \in \mathbb{R}_+\}.$$

According to this measure, any two distributions will be regarded as being close to each other as long as they are uniformly close at all points of continuity.[10] We assume that the limit distributions X and Y are nondecreasing and right-continuous, and satisfy $X(0) = 0, X(1) = 1$ and $X(\cdot) - Y(\cdot) \geq 0$ Note that each object type may possibly have multiple copies, so the model allows for many-to-one matching but also includes as a special case an one-to-one matching with $X(\cdot) = Y(\cdot)$. We allow X and Y to be fairly general, allowing for atoms.

Two special cases of this model are of interest. The first is a **finite-tier model.** In this model, the objects are partitioned into finite tiers, $\{O_1^n, ..., O_K^n\}$, where $\cup_{k \in K} O_k^n = O^n$ and $O_k^n \cap O_j^n = \emptyset$. (With a slight abuse of notation, the largest cardinality K denotes also the set of indexes.) In this model, the CDFs X^n and Y^n are step functions with finite steps. This model offers a good approximation of situations in which the objects have clear tiers, as will be the case in situations in which schools are distinguished in different categories or by regions, and houses may come in clearly distinguishable tiers. For the most part, the finite model serves as an analytical vehicle that will be used to analyze the general model. From this perspective, the finite model is useful to focus on since it brings out, in the most transparent way.

Another special case is the **full-support model** in which the limit distribution Y is strictly increasing in its support. This model is very similar to Lee (2012), who also considers random preferences that consist of common and idiosyncratic terms. One difference is that his framework assumes that the common component of the payoff is also drawn uniform randomly from a positive interval. Our model assumes common values to be arbitrary, but with full support assumption, the values can be interpreted as realizations of random draws (drawn according to the CDF Y). Viewed in this way, the full-support model is comparable to Lee (2012)'s, except that current model also allows for atoms in the distribution of Y.

Unless specified, we are referring to a general model that has these two as special cases. Fix an n-economy. We shall consider a class of matching mechanisms that are Pareto efficient. A **matching** μ in an n-economy is a mapping $\mu : I \to O \cup \{\emptyset\}$ such that $|\mu^{-1}(o)| \leq q_o$ for all $o \in O$, with the interpretation that agent i with $\mu(i) = \emptyset$

[10] Here, convergence of CDFs in Lévy metric is equivalent to weak convergence.

is unmatched. Let M denote the set of all matchings. All these objects depend on n, although their dependence is suppressed for notational convenience.

In practice, the matching chosen by the designer will depend on the realized preferences of the agents as well as other features of the economy that the matching institution may condition on. For instance, if the objects O are institutions or individuals, their preferences on their matching partners will typically impact on what matching will arise. Alternatively, one may wish the matching to respect the existing rights that the individuals may have over the objects; for instance, the objects may be housing, and some units may have existing tenants who may have priority over these units. Likewise, in the school choice context, the matching may favor the students whose siblings already attend the school or those living close to the school. Some of these factors may depend on the features not captured by their idiosyncratic component. We collect all assignment-relevant variables, call its generic realization a "state," and denote it by $\omega = (\{\xi_{i,o}\}_{i\in I, o\in O}, \theta)$, where $\{\xi_{i,o}\}_{i\in I, o\in O}$ is the realized profile of idiosyncratic component of payoffs, and θ is the realization of all other variables that influence the matching, and let Ω denote the set of all possible states.

A **matching mechanism** is a function that maps from a state in Ω to a matching in M. With a slight abuse of notation, we shall use $\mu = \{\mu_\omega(i)\}_{\omega\in\Omega, i\in I}$ to denote a matching mechanism, which selects a matching $\mu_\omega(\cdot)$ in state ω. Let $\mathcal{M}$ denote the set of all matching mechanisms. For convenience, we shall often suppress the dependence of the matching mechanism on ω.

A matching $\mu \in M$ is Pareto efficient if there is no other matching $\mu' \in M$ such that $U_i(\mu'(i)) \geq U_i(\mu(i))$ for all $i \in I$ and $U_i(\mu'(i)) > U_i(\mu(i))$ for some $i \in I$. A matching mechanism $\mu \in \mathcal{M}$ is Pareto efficient if, for each state $\omega \in \Omega$, the matching it induces, i.e., $\mu_\omega(\cdot)$, is Pareto efficient. Let $\mathcal{M}^*_n$ denote the set of all Pareto efficient mechanisms in the n-economy.

3 Payoff Equivalence of Pareto Efficient Mechanisms

A wide variety of mechanisms yield Pareto efficient matchings. Mechanisms such as (deterministic or random) serial dictatorship attain efficiency with no particular regard to agents' property rights or priorities; others such as TTC recognize such rights, and allow agents to trade these rights to achieve efficiency. Market designers can also endow agents with fake money, allowing them to purchase objects efficiently in an artificial market place, as envisioned by Hylland and Zeckhauser (1979). Mechanisms can be further adjusted to meet other social needs, such as "affirmative treatment" of some target groups identified based on their socio-economic backgrounds, for example. Any such adjustments will obviously impact the welfare of the participants at the individual level. But do they impact the total welfare of the agents or their aggregate payoff distribution? If so, how?

These questions have potentially significant market design implications. If accommodating the rights or priorities of some individuals or to satisfy specific social objectives or constraints were to entail significant loss in terms of utilitarian welfare or to have significant distributive impact, this will call into question the merit of the policy interventions. We address these questions below.

3.1 Definitions

To begin, we first define an upper bound for the utilitarian welfare—a highest possible level of total surplus that can be realized under any matching mechanism. To this end, suppose every agent is assigned an object and realizes the highest possible idiosyncratic payoff. Since the common values of the objects are distributed according to X^n, the resulting (normalized) utilitarian welfare is $\int_0^1 U(u,1)dX^n(u)$. This obviously gives the upper bound for the utilitarian welfare in the n-economy. We consider its limit, called the **limit utilitarian upper bound**:

$$U^* := \int_0^1 U(u,1)dX(u).$$

The payoff distribution of an economy, whether it is a finite n-economy or its limit, can be represented by a distribution function, i.e., a nondecreasing right-continuous function F mapping from $[0, U(1,1)]$ to $[0,1]$. The number $F(z)$ is interpreted as the fraction of the agents who realize payoffs no greater than z. We let F^μ denote the payoff distribution induced by mechanism μ.

3.2 Utilitarian efficiency and its implications

We are now in a position to state our first main theorem.

Theorem 1. *Let F^* be the distribution of payoff attaining the limit utilitarian upper bound*

U^*, and recall $\mathcal{M}_n^*$ is the set of Pareto efficient mechanisms in the n-economy. Then,

$$\sup_{\mu^n \in \mathcal{M}_n^*} L(F^{\mu^n}, F^*) \xrightarrow{p} 0.$$[11]

In words, the theorem states that the distance (in Lévy metric) between a payoff distribution resulting from *every* Pareto efficient mechanism and that of the utilitarian upper bound vanishes uniformly in probability as $n \to \infty$. More precisely, the statement is as follows. Fix any $\epsilon, \delta > 0$. Then, with probability at least of $1 - \delta$, the proportion of agents enjoying any payoff u or higher under any Pareto efficient mechanism is within ϵ of the proportion of agents enjoying payoff of $u - \epsilon$ or higher under the utilitarian upper bound, for n sufficiently large. It is remarkable that the rate of convergence is "uniform" with respect to the entire class of the Pareto efficient mechanisms.

The following corollary is immediate:

Corollary 1.

$$\inf_{\mu^n \in \mathcal{M}^*} \frac{\sum_{i \in I} U_i(\mu^n(i))}{|I|} \xrightarrow{p} U^*.$$

The theorem also implies that alternative Pareto efficient mechanisms become payoff equivalent uniformly as the market grows in size—that is, "up to the renaming of the agents":

Corollary 2.

$$\sup_{\mu^n, \tilde{\mu}^n \in \mathcal{M}_n^*} L(F^{\mu^n}, F^{\tilde{\mu}^n}) \xrightarrow{p} 0.$$

These results suggest that, as long as agents are ex ante symmetric in their preferences, there is little ground to discriminate one Pareto efficient mechanism in favor of another in terms of total welfare of participants or aggregate payoff distribution, at least in the large economy. This has important implications for market design. Often designers face extra constraints arising from the existing rights or priorities that some participants have over some objects or there may be a need to treat some target group of participants affirmatively. And there is a concern that accommodating such constraints or needs may sacrifice

[11] We say $Z_n \xrightarrow{p} z$, or Z_n *converges in probability* to z, where both Z_n and z are real-valued random variables, if for any $\epsilon > 0, \delta > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, we have

$$\Pr\{|Z_n - z| > \epsilon\} < \delta.$$

utilitarian welfare or to adversely impact the aggregate distribution of payoffs. Our result implies that accommodating such constraints does not entail any significant loss in these terms in the large economy, as long as Pareto efficiency is maintained.

Remark 1 (Virtual Transferability). That Pareto efficiency implies utilitarian efficiency is surprising since transfers are not allowed in our model. One interpretation is that a large market makes utilities virtually transferable, by creating "rich" opportunities for agents to trade on idiosyncratic payoffs. In other words, objects that are uniformed valued by the participants can be transferred from one set of agents to another set without entailing much loss in terms of the idiosyncratic payoffs. Such a result, while plausible, is neither obvious nor universally true. As will be seen in the second part of the current paper, the strong payoff equivalence does not extend when the welfare of both sides are relevant. Gale and Shapley's deferred acceptance algorithm is Pareto efficient across both sides of the market—i.e., taking the objects as welfare-relevant entities—but does not attain the highest total payoffs across the market, and a different matching (which is Pareto inefficient) yields a higher utilitarian welfare.

Remark 2 (Relationship with other equivalent results). The current equivalence result is reminiscent of a similar equivalence result obtained by Abdulkadiroglu and Sönmez (1998) between two well known mechanisms, random serial dictatorship and TTC with random ownership, and of the large market equivalence result obtained by Che and Kojima (2010) between random serial dictatorship and probabilistic serial mechanism, and their extension by Pathak and Sethuraman (2011). While these results consider arbitrary preferences on the agents, they assume ex ante symmetric random priorities with respect to the objects. By contrast, our equivalence result does not impose any structure on the priorities on the object side, allowing them to be arbitrary, but it does impose a certain structure on the agents' preferences (to consist of common values and iid idiosyncratic preferences). Our result also holds only in the limit as the number of agents and objects becomes large (with the number of object types bounded or growing at a slower rate), whereas the equivalence result by Abdulkadiroglu and Sönmez (1998) and Pathak and Sethuraman (2011) holds for any finite economy.

3.3 Sketch of the Proof

Here we sketch the proof of Theorem 1, which is contained in Appendix A. For the current purpose, assume $X(\cdot)$ is degenerate with a single common value u^0, and $X(\cdot) = Y(\cdot)$. In other words, the agents have only idiosyncratic payoffs, and the matching is one-to-one (as opposed to many to one). As will be seen in Appendix A, the same proof argument works for the general case (with some care).

To begin, fix an arbitrary Pareto efficient mechanism $\tilde{\mu}$. We first invoke the fact that any Pareto efficient matching can be implemented by a serial dictatorship[12] with a suitably-chosen serial order (see Abdulkadiroglu and Sönmez (1998)). Let $\tilde{f}$ be the serial order, namely a function that maps each agent in I to his serial order in $\{1, ..., n\}$ that implements $\tilde{\mu}$ under a serial dictatorship. Since $\tilde{\mu}$ induces a Pareto efficient matching that depends on the state, the required serial order $\tilde{f}$ is a random sequence.

Next, for arbitrarily small $\epsilon, \delta > 0$, define the random set:

$$\bar{I} := \left\{ i \in I \,\middle|\, U_i(\tilde{\mu}(i)) \leq U(u^0, 1-\epsilon) \text{ and } \tilde{f}(i) \leq (1-\delta)|O| \right\}.$$

The set $\bar{I}$ consists of agents who are within $1-\delta$ top percentile in terms of their serial order $\tilde{f}$ but fail to achieve payoff ϵ-close to the highest possible payoff.[13] Since $\epsilon, \delta > 0$ are arbitrary, for the proof it will suffice to show that

$$\frac{|\bar{I}|}{n} \xrightarrow{p} 0. \tag{1}$$

To prove this, we exploit a result in random graph theory. It is thus worth introducing the relevant model of random graph. A **bipartite graph** G consists in vertices, $V_1 \cup V_2$, and edges $E \subset V_1 \times V_2$ across V_1 and V_1 (with no possible edges within vertices in each side). An **independent set** is $\bar{V}_1 \times \bar{V}_2$ where $\bar{V}_1 \subseteq V_1$ and $\bar{V}_2 \subseteq V_2$ for which no element in $\bar{V}_1 \times \bar{V}_2$ is an edge of G. A random bipartite graph $B = (V_1 \cup V_2, p)$, $p \in (0,1)$, is a bipartite graph with vertices $V_1 \cup V_2$ in which each pair $(v_1, v_2) \in V_1 \times V_2$ is linked by an edge with probability p independently (of edges created for all other pairs). The following result provides the crucial step for our result.

Lemma 1 (Dawande, Keskinocak, Swaminathan, and Tayur (2001)). *Consider a random bipartite graph $B = (V_1 \cup V_2, p)$ where $0 < p < 1$ is a constant and for each $i \in \{1, 2\}$ and*

[12]A serial dictatorship mechanism specifies an order over individuals, and then lets the first individual – according to the specified ordering – receive his favorite object, the next individual receives his favorite item among remaining objects, etc

[13]Strictly speaking, we should be focusing on individuals receiving payoffs lower than $U(u^0, 1) - \epsilon$. However, given that the utility functions are continuous there is little loss in focusing our attention on agents receiving less that $U(u^0, 1-\epsilon)$. This point will be made clear in the proof.

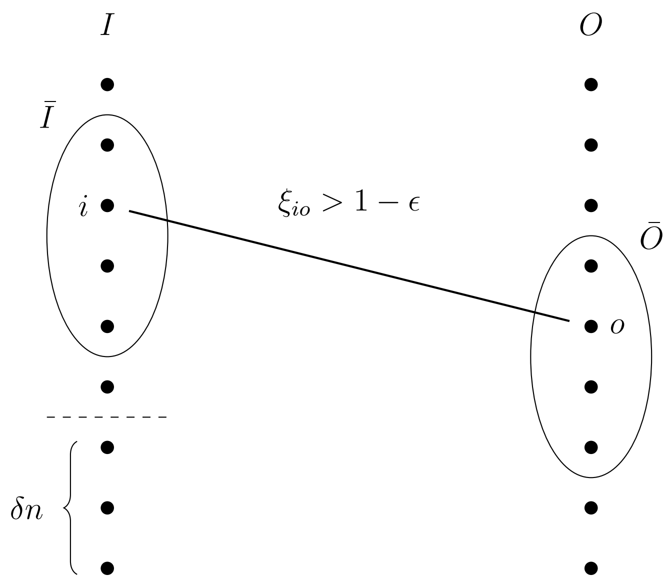

Figure 1: Illustration of a random graph and sets $\bar{I}$ and $\bar{O}$

$|V_1| = n$ *and* $|V_2| = m = O(n)$. *There is* $\kappa > 0$,

$$\Pr\left[\exists \text{ an independent set } \hat{V}_1 \times \hat{V}_2 \text{ with } \min\{|\hat{V}_1|, |\hat{V}_2|\} \geq \kappa \ln(n)\right] \to 0 \text{ as } n \to \infty.$$

This result implies that with high probability, for every independent set, at least one side of that set vanishes in relative size as $n \to \infty$.

To prove our result, then it suffices to show that $\bar{I}$ forms a vanishing side of an independent set in an appropriately-defined random graph. Consider a random bipartite graph consisting of I on one side and O on the other side where an edge is created between $i \in I$ and $o \in O$ if and only if $\xi_{i,o} > 1 - \epsilon$. Let

$$\bar{O} := \left\{ o \in O \,\middle|\, \tilde{f}(\tilde{\mu}(o)) \geq (1 - \delta)|O| \right\}$$

be the (random) set of objects that are assigned to the agents who are at the bottom δ percentile in terms of the serial order $\tilde{f}$.

The key observation is that the (random) subgraph $\bar{I} \times \bar{O}$ is an independent set.

To see this, suppose to the contrary that there is an edge between an agent $i \in \bar{I}$ and an object $o \in \bar{O}$ in some state ω. By construction of $\bar{I}$, agent $i \in \bar{I}$ must realize less than $1 - \epsilon$ of idiosyncratic payoff from $\tilde{\mu}_\omega(i)$. But the fact that there is an edge between i and o

means that i would gain more than $1-\epsilon$ in idiosyncratic payoff from o. So, agent i must prefer o to his match $\tilde{\mu}_\omega(o)$. Yet, the fact that $o \in \bar{O}$ means that o is not yet claimed and is thus available when agent i (who is within top $1-\delta$ of serial order $\tilde{f}_\omega$) picks $\tilde{\mu}_\omega$. This is a contradiction, proving that $\bar{I} \times \bar{O}$ is an independent set.

Next we observe that $|\bar{O}| = \delta n$, meaning that $\bar{O}$ never vanishes in probability. Lemma 1 then implies that the set $\bar{I}$ must vanish in probability. Importantly, this result applies uniformly to all mechanisms in $\mathcal{M}^*$: If we define the sets $\bar{I}(\tilde{\mu})$ and $\bar{O}(\tilde{\mu})$ for each $\tilde{\mu} \in \mathcal{M}^*$ as above, then for each $\tilde{\mu} \in \mathcal{M}^*$, $\bar{I}(\tilde{\mu}) \times \bar{O}(\tilde{\mu})$ forms an independent set of the *same* random graph! This explains the uniform convergence.

Remark 3. If the mechanism $\tilde{\mu}$ were a serial dictatorship with a "deterministic" serial order f, then a simple direct argument proves the result. First, let us note that we can think of each agent as drawing his preferences "along the algorithm," i.e., he draws his preferences for the stage when it is his turn to make a choice. Obviously, the distribution of i's preferences is not affected by the choices of agents ahead of that agent in the serial order. Fix any arbitrary $\epsilon, \delta > 0$ and let E_i be the event that when i gets his turn to make a choice, there remains at least one object o such that $U_i(o) \geq U(u^0, 1-\epsilon)$. Clearly,

$$\begin{aligned}
&\Pr\{U_i(\tilde{\mu}(i)) \geq U(u^0, 1-\epsilon) \text{ for all } i \text{ with } f(i) < (1-\delta)n\} \\
\geq &\Pr\{\cap_{i \in I: f(i) < (1-\delta)n} E_i\} \\
\geq &1 - \Pr\{\cup_{i \in I: f(i) < (1-\delta)n} E_i^c\} \\
\geq &1 - (1-\delta)n(1-\epsilon)^{\delta n} \to 1 \text{ as } n \to \infty.
\end{aligned}$$

Such an argument does not work for an arbitrary Pareto efficient mechanism, however. For a general Pareto efficient mechanism, the serial order implementing the mechanism need not be independent of the agents' preferences (which is required in the last inequality in the above string). Our general proof using random graph theory avoids this difficulty.

4 Resolving the Tradeoff between Efficiency and Stability

As motivated in the introduction, efficiency and stability are important objectives in the designing of markets for allocating indivisible resources to agents. While theses two objectives are generally in conflict with each other, it is unknown how the tradeoff plays out in markets where the number of agents as well as the number of objects are large. In this section, we study whether the two prominent mechanisms—top trading cycles (TTC) and Gale and Shapley's deferred acceptance algorithm (DA)—resolve the tradeoff at least in a suitably-defined asymptotic sense in the large market. Our main result will identify a plausible set of circumstances in which the two mechanisms fail to provide adequate resolution of the tradeoff—namely, the TTC entails a significant loss in stability and the DA entails a significant efficiency loss for the agents. We then propose a new mechanism that attains both objectives in the asymptotic senses and has desirable large market incentive property.

We begin by simplifying the model. First, we focus on the one-to-one matching environment. Formally, we assume $X^n(\cdot) = Y^n(\cdot)$ and $X(\cdot) = Y(\cdot)$.

Second, we assume that the priorities of agents at alternative objects—or objects' "preferences" over agents—are drawn uniform randomly. Formally, we assume that each object $o \in O$ receives utility from getting matched with individual $i \in I$:

$$V_i(o) = V(\eta_{i,o}),$$

where *idiosyncratic shock* $\eta_{i,o}$ is a random variable drawn independently and identically from $[0, 1]$ according to the uniform distribution. This assumption is needed to keep the analysis tractable.[14] Despite being restrictive, the assumption identifies a class of plausible circumstances under which a tradeoff between the two objectives persist and can be addressed more effectively by a novel mechanism. We remark later on how we can generalize the mechanism when there is a significant correlation in the agents' priorities. The function $V(\cdot)$ takes values in $\mathbb{R}_+$, is strictly increasing and continuous in the idiosyncratic shock. The utility of remaining unmatched is assumed to be 0 so that all objects find all individuals acceptable.

Third, we restrict attention to the finite-tier model. Namely, the objects are partitioned into finite sets, $\{O_1, ..., O_K\}$, such that agents realize the same common value u^k from objects in O^k, where $u^1 > ... > u^K$. For each $k = 1, ..., K$, the proportion of objects in O_k, i.e., $\frac{|O_k|}{|O|}$, has a well defined limit denoted by x_k. This simplification, while invoked for tractability of analysis, is not without realism. In the school choice context, schools are often segmented into different tiers based on the geographic districts the schools belong to. Our simulation later will consider a broad class of settings beyond the current structure, and will confirm the robustness of our analytical results.

[14]Without this assumption, the analysis of the mechanisms such as DA and TTC becomes intractable since the preferences of the agents over remaining objects after the first round are not independent of the history, so the principle of "deferred decision"—a key tool in the probabilistic method—cannot be invoked.

We begin with a lemma that will be used throughout.

Lemma 2. *Fix any $\epsilon > 0$, and any $k = 1, ..., K$. There exists $\delta > 0$ small enough such that, with probability going to 1 as $n \to \infty$, for every agent in I, (i) each of his top $\delta|O_k|$ favorite objects in $O_{\geq k}$ yields a payoff greater than $U(u_k, 1) - \epsilon$* and *(ii) and every such object belongs to O_k.*

Proof. See Appendix B. □

We next suggest how efficiency and stability can be weakened in the large market setting.

□ **Asymptotic Notions of Efficiency and Stability**

We say a matching mechanism μ is **asymptotically efficient** if, for any mechanism μ' that weakly Pareto-dominates μ for the agents I,

$$\frac{|I_\epsilon(\mu'|\mu)|}{n} \xrightarrow{p} 0,$$

where

$$I_\epsilon(\mu'|\mu) := \{i \in I | U_i(\mu(i)) < U_i(\mu'(i)) - \epsilon\}$$

is the set of agents who would benefit more than ϵ by switching from μ to μ'. In words, a matching is asymptotically efficient if the fraction of the agents who could benefit discretely from any Pareto improving rematching vanishes in probability as the economy becomes large.

The notion of stability can be weakened in a similar way. We say a matching mechanism μ is **asymptotically stable** if, for any $\epsilon > 0$,

$$\frac{|J_\epsilon(\mu)|}{n(n-1)} \xrightarrow{p} 0,$$

where

$$J_\epsilon(\mu) := \{(i, o) \in I \times O | U_i(o) > U_i(\mu(i)) + \epsilon \text{ and } V_o(i) > V_o(\mu(o)) + \epsilon\}.$$

is the set of ϵ**-block**'s—namely, the set of pairs of unmatched agent and object who would each gain ϵ or more from matching each other rather than matching according to μ. Asymptotic stability requires that for any $\epsilon > 0$ the fraction of these ϵ-blocks out of all $n(n-1)$ "possible" blocking pairs vanishes in probability as the economy grows large. It is possible even in an asymptotically stable matching that some agents may be willing to block with a large number of objects, but the number of such agents will vanish in probability. This can be stated more formally.

For any $\epsilon > 0$, let $\hat{O}^i_\epsilon(\mu) := \{o \in O | (i, o) \in J_\epsilon(\mu)\}$ be the set of objects agent i can form an ϵ-block with against μ. Then, a matching is asymptotically stable if and only if the set of agents who can form an ϵ-block with a non-vanishing fraction of objects vanishes; i.e., for any $\epsilon, \delta > 0$,

$$\frac{|I_{\epsilon,\delta}(\mu)|}{n} \xrightarrow{p} 0,$$

where

$$I_{\epsilon,\delta}(\mu) := \left\{ i \in I \middle| |\hat{O}^i_\epsilon(\mu)| \geq \delta n \right\}.$$

If, as is plausible in many circumstances, agents form ϵ-blocks by randomly sampling a finite number of potential partners (i.e., objects), asymptotic stability would mean that only a vanishing proportion of agents will succeed in finding blocking partners in a large market.

A similar implication can be drawn in terms of fairness. Asymptotic stability of matching implies that only a vanishing proportion of agents would have (a discrete amount of) justified envy toward a non-vanishing proportion of agents. If an individual gets aggrieved from justifiably envying say finitely many individuals who she randomly encounters, then the property will guarantee that only a vanishing fraction of individuals will suffer significant aggrievement as the economy grows large.

4.1 Two Prominent Mechanisms

□ Top Trading Cycles (TTC) Mechanism:

Top Trading Cycles algorithm, originally introduced by Shapley and Scarf (1974) and later adapted by Abdulkadiroglu and Sonmez (2003) to the context of strict priorities, has been an influential method for achieving efficiency.[15] The mechanism has some notable applications. For instance, TTC was used until recently in New Orleans school systems for assigning students to public high schools and recently, San Francisco school system announced plans to implement a top trading cycles mechanism. A generalized version of TTC is also used for kidney exchange among donor-patient pairs with incompatible donor kidneys (see Sonmez and Unver (2011)).

[15]The original idea is attributed to David Gale by Shapley and Scarf (1974).

The TTC algorithm (defined by Abdulkadiroglu and Sonmez (2003)) proceeds in multiple rounds as follows: In Round $t = 1, ...,$ each individual $i \in I$ points to his most preferred object (if any). Each object $o \in O$ points to the individual to which it assigns the highest priority. Since the number of individuals and objects are finite, the directed graph so obtained has at least one cycle. Every individual who belongs to a cycle is assigned to the object he is pointing at. Any individuals and objects that are assigned are then removed. The algorithm terminates when all individuals have been assigned; otherwise, it proceeds to Round $t + 1$.

This algorithm terminates in finite rounds. Indeed, there are finite individuals, and at the end of each round, at least one individual is removed. The TTC mechanism then selects a matching via the above algorithm for each realization of individuals' preferences as well as objects' priorities.

As is well-known, the TTC mechanism is Pareto-efficient and strategy-proof, namely making it a dominant strategy for agents to report their preferences truthfully, but is not stable.

□ The Deferred Acceptance (DA) Mechanism

The best-known mechanism for attaining stability is the deferred acceptance algorithm. Since introduced by Gale and Shapley (1962), the mechanism has been applied widely in a variety of contexts. The Medical Match in the US and other countries adopt DA for placing doctors to hospitals for residency programs. The school systems in Boston and New York City use DA to assign eighth-grade students to public high schools (see Abdulkadiroglu, Pathak, and Roth (2005) and Abdulkadiroglu, Pathak, Roth, and Sonmez (2005)). College admissions are organized via DA in many provinces in Australia.

For our purpose, it is more convenient to define DA in a version proposed by McVitie and Wilson (1971), which proceeds in multiple steps as follows:

Step 0: Linearly order individuals in I.

Step 1**:** Let individual 1 make an offer to his most favorite object in O. This object tentatively holds individual 1, and go to Step 2.

Step $i \geq 2$**:** Let agent i make an offer to his most favorite object o in O among the objects to which he has not yet made an offer. If o is not tentatively holding any individual, then o tentatively holds i. whenever $i = n$, end the algorithm; otherwise iterate to Step $i + 1$. If however o is holding an individual tentatively—call him i^*—object o chooses between i and i^* accepting tentatively the one who is higher in its preference list, and rejecting the other. The rejected agent is named i and we go back to the beginning of Step i.

This process terminates in finite time and yields a matching μ. The DA mechanism selects a matching via this process for each realization of individuals' preferences as well as objects' priorities.

As is well known, the (agent-proposing) DA mechanism selects a stable matching that Pareto dominates all other stable matchings, and it is also strategy proof (Dubins and Freedman (1981); Roth (1982)). But the mechanism is not Pareto efficient, meaning the agents may all be better off by another matching (which is not stable).

4.2 Uncorrelated Preferences

We first consider the case in which the participants' preferences for the objects are uncorrelated. That is, the support of the common component of the agents' utilities is degenerate, with a single tier $K = 1$ for the objects. In this case, both DA and TTC involve little tradeoff:

THEOREM 2. *If the support of $Y(\cdot)$ is degenerate, then any Pareto efficient mechanism, and hence TTC, is asymptotically stable and DA is asymptotically efficient.*

PROOF. The asymptotic stability of a Pareto efficient mechanism follows from Theorem 1, which implies that for any $\epsilon > 0$, the proportion of the set $I_\epsilon(\tilde{\mu})$ of agents who realize payoffs less than $U(u^1, 1) - \epsilon$ vanishes in probability as $n \to \infty$, for any $\tilde{\mu} \in \mathcal{M}_n^*$. Since $I_{\epsilon,\delta}(\tilde{\mu}) \subset I_\epsilon(\tilde{\mu})$, asymptotic stability then follows.

The asymptotic efficiency of DA is as follows. Let E_1 be the event that all agents are assigned objects which they rank within $\log^2(n)$. By Pittel (1992), the probability of that event goes to 1 as n goes to infinity. Now, fix $\epsilon > 0$ arbitrarily small and let E_2 be the event that for all agents the objects which they rank within $\log^2(n)$ give them a payoff greater than $U(u_k, 1) - \epsilon$. By Lemma 2, the probability of that event goes to 1 as n tends to infinity. It is clear that whenever both events occur, all agents will get a payoff greater than $U(u_k, 1) - \epsilon$ under DA. Since the probability of both events realizing goes to 1, the DA mechanism is asymptotically efficient.[16] $\square$

[16]Our notion of efficiency focuses on one side of the market: the individuals' side. It is worth noting here that even if we were to focus only on the other side: the objects' side, efficiency would still follow from Pittel (1992) even though we are using DA where individuals are the proposers.

It is worth noting that the tradeoffs of the two mechanisms do not disappear *qualitatively* even in the large markets: DA remains inefficient and TTC remains unstable even as the market grows large. In fact, given random priorities on the objects, the acyclicity conditions that would guarantee efficiency of DA and stability of TTC, respectively,[17] fail almost surely as the market grows large. What Theorem 2 suggests is that the tradeoff disappears *quantitatively*, provided that the agents have uncorrelated preferences. The uncorrelatedness of preferences implies that the conflicts agents may have over the goods disappear as the economy grows large, for each agent is increasingly able to find an object that he likes that others do not like as much. Hence, the diminishing conflicts of preferences means that the agents can attain high payoffs, in fact arbitrarily close to their payoff upper bound as $n \to \infty$. This eliminates (probabilistically) the possibility that a Pareto dominating mechanism can benefit a significant fraction agents significantly, explaining the asymptotic efficiency of DA. Similarly, under TTC, the agents enjoy payoffs that are arbitrarily close to their payoff upper bound as $n \to \infty$, which guarantees that the number of agents who each would justifiably envy a significant number of agents vanishes in the large market.

4.3 Correlated Preferences

As we show below, Theorem 2 no longer holds when the agents' preferences are correlated, in particular, when some objects are perceived by "all" agents to be better than the other objects. This situation is quite common in many contexts, such as school assignment, since schools have distinct qualities that students and parents evaluate in a similar fashion.

To consider such an environment in a simple way, we shall suppose the objects are divided into two tiers O_1 and O_2 such that $|I| = |O_1| + |O_2| = n$. As assumed earlier, $\lim_{n\to\infty} \frac{|O_k|}{n} = x_k > 0$. In addition, we assume that each object in O_1 is considered by every agent to be better than each object in O_2: $U(u^1, 0) > U(u^2, 1)$, where u^1 and u^2 are common values of the objects from tier 1 and tier 2, respectively. The preferences/priorities by the objects are given by idiosyncratic random shocks, as assumed above. In this environment, we shall show that the standard tradeoff between DA and TTC extends to the large markets even in the asymptotic sense — namely, DA is not asymptotically efficient and TTC is not asymptotically stable. This observation runs counter to the common wisdom based on finite market that correlation of preferences on one side usually renders stable allocations

[17]The former is due to Ergin (2002) and the latter is due to Kesten (2006).

efficient or efficient allocations stable.

4.3.1 Asymptotic Instability of TTC

Our first result is that, with correlated preferences, TTC fails to be asymptotically stable.

THEOREM 3. *In our model with two tiers, TTC is not asymptotically stable. More precisely,*

$$\frac{|J_\epsilon(TTC)|}{n(n-1)} \overset{p}{\nrightarrow} 0.$$

PROOF. See Appendix C. □

We provide the main idea of the proof here. In a nutshell, the asymptotic instability arises from the key feature of TTC. In TTC, agents attain efficiency by "trading" among themselves the objects with which they have high priorities. This process entails instabilities since some agent may gain an object despite having a low priority, ahead of an agent who has a higher priority but whose priority is still lower than the agent who traded off his priority to the gaining agent. This insight is well known but informs little about the magnitude of the instabilities.

The results in the previous subsection suggest that instabilities are not significant in case agents' preferences are uncorrelated. In that case, the agents' preferences do not conflict with each other, and they all attain close to their "bliss" payoffs in TTC, resulting in only a vanishing number of agents with justifiable envy toward any significant number of agents. The situation is different, however, when their preferences are correlated significantly. In the two-tier case, for instance, a large number of agents are assigned objects in O_2, and they would all envy the agents who are assigned objects in O_1. The asymptotic stability of the mechanism then depends on whether a significant number of the latter agents (those assigned objects in O_1) would have significantly lower priorities (with the objects they are obtaining) relative to the former agents who envy them.

This latter question boils down to the length of the cycles through which the latter agents (who are assigned the objects in O_1) are assigned in the TTC mechanism. Call a cycle of length two—namely, an agent top-ranks an object, and also has the highest priority with that object, among those remaining in each round—-a **short cycle**, and any cycle of length greater than two a **long cycle**.

It is intuitive that agents who are assigned via short cycles are likely to have high

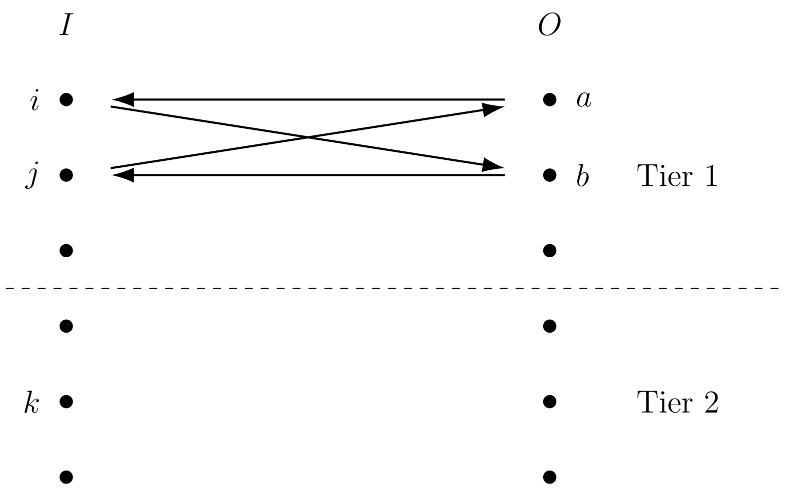

Figure 2: Possible justified envy by k toward j

priorities (with the objects they are assigned).[18] By contrast, the agents who are assigned via long cycles are unlikely to have high priorities. Agents in the long cycles tend to have high priorities with the objects they *trade up* (since the objects must have pointed to them), but they could have very low priorities with the objects they *trade in*. In fact, their priorities with the objects they are assigned play no (contributory) role for such a cycle to form.[19] Hence, their priorities with the objects they are assigned (in O_1) are at best simple iid draws, with one half chance of them being higher than the priorities of those agents assigned objects in O_2. This suggests that any agent assigned objects in O_2 will have on average a significant amount of justified envy toward one half of those agents who are assigned objects in O_1 via long cycles.

The crucial part of the proof of Theorem 3, provided in Appendix C, is to show that the number of agents assigned O_1 via long cycles is significant—i.e., the number does not vanish in probability as $n \to \infty$. While this result is intuitive, its proof is not trivial. By an appropriate extension of "random mapping theory," we can compute the expected number

[18]This is obvious for the agents assigned in the first round, for they have the highest priorities. But even those assigned in later rounds are likely to have high priorities as long as they are assigned via short cycles: Theorem 1 implies that almost all agents are assigned within the number of steps in TTC that is sub linear—i.e., small relative to n, meaning that those assigned via short cycles tend to have relatively high priorities.

[19]If any, the role is negative. That an agent is assigned via a long cycle, as opposed to a short cycle, means that she does *not* have the highest priority with object he is getting in that round.

of objects in O_1 that are assigned via long cycles in the first round of TTC. But, this turns out to be insufficient for our purpose since the number of objects that are assigned in the first round of TTC (which turns out to be in the order of $\sqrt{n}$) comprises a vanishing fraction of n as the market gets large. But extending the random mapping analysis to the subsequent rounds of TTC is difficult since the preferences of the agents and objects remaining after the first round are no longer i.i.d. The extension requires us to gain a deeper understanding about the precise random structure of the algorithm evolves over time. Appendix D does this. In particular, we establish that the *number* of objects (and thus agents) assigned in each round of TTC follows a simple Markov chain, implying that the number of agents cleared in each round is not subject to the conditioning issue. The *composition* of the cycles, in particular short versus long cycles, is subject to the conditioning issue, however. Nevertheless, we manage to bound the number of short cycles formed in each round of TTC, this bound, combined with the Markov property of the number of objects assigned in each round, produces the result.

Remark 4 (Markov Property of TTC). The Markov property we establish in Appendix D is of independent interest and is likely to be of use beyond the current paper. It means that the number of agents (and thus objects) assigned in each round depends only on the number of agents and objects remaining at the beginning of that round, and importantly does not depend on the history before that round. Furthermore, we explicitly derive in Theorem 8 of Appendix D the formula for the distribution of these variables. This formula can be used to analyze the welfare of agents under TTC even for the finite economy.

4.3.2 Asymptotic Inefficiency of DA

Given correlated preferences, we also find the inefficiency of DA to remain significant in the large market.

Theorem 4. *In our two tier model, DA is not asymptotically efficient. More precisely, there exists a matching μ that Pareto dominates DA and*

$$\frac{|I_\epsilon(\mu|DA)|}{|I|} \overset{p}{\not\to} 0.$$

Proof. See Appendix E. □

Corollary 3. *Any stable matching mechanism fails to be asymptotically efficient in our two-tier model.*

Proof. The DA matching Pareto dominates all other stable matching, as is shown by Gale and Shapley (1962). Hence, any matching μ that Pareto dominates and satisfies the property stated in Theorem 4 will Pareto dominate any stable matching and satisfy the same property. $\square$

The intuition behind Theorem 4 is as follows. When the agents' preferences are correlated, they tend to compete excessively for the same set of objects, and this competition results in a significant welfare loss under a stable mechanism. To see this intuition more clearly, recall that all agents prefer every object in O_1 to any object in O_2. This means that in the DA they all first propose to objects in O_1 before they ever propose to any object in O_2. The first phase of the DA (in its McVitie-Wilson version) is then effectively a sub-market consisting of I agents and O_1 objects with random preferences and priorities. Given that there are excess agents of size $|I| - |O_1|$, which grows linearly in n, even those agents lucky enough to be assigned objects in O_1 must compete so much so that their payoffs will be bounded away from $U(u^1, 1)$.[20]

This result is quite intuitive. Note that all agents who are eventually assigned objects in O_2 must have made each $|O_1|$ offers to the objects in O_1 before they are rejected by all of them. This means each object in O_1 must receive at least $|I| - |O_1|$ offers. Then, from an agent's perspective, to get assigned an object in O_1, he must survive competition from at least $|I| - |O_1|$ other agents. The odds of this is $\frac{1}{|I|-|O_1|}$, since the agents are all ex ante symmetric. Hence, the odds that an agent gets rejected by his top δn choices, for small enough $\delta > 0$, is at least

$$\left(1 - \frac{1}{|I| - |O_1|}\right)^{\delta n} \to \left(\frac{1}{e}\right)^{\frac{\delta}{(1-x_1)}}, \tag{2}$$

since $|I| - |O_1| \to (1 - x_1)n$ as $n \to \infty$. Note that this probability gets close to one, for δ sufficiently small. This probability is not conditional on whether an agent is assigned an object in O_1, and surely the probability will be large (in fact, approach one) conditional on an agent being unassigned any object in O_1. But an agent *does* get assigned an object in O_1 with positive probability (i.e., approaching $x_1 > 0$), so for the unconditional probability of an agent making at least δn offers to be close to one, the same event must occur with positive probability even conditional on being assigned an object in O_1. As shown more

[20] This result is obtained by Ashlagi, Kanoria, and Leshno (2013) and Ashlagi, Braverman, and Hassidim (2011) building on the algorithm originally developed by Knuth, Motwani, and Pittel (1990) and Immorlica and Mahdian (2005). Here we provide a direct proof which is much simpler. This proof is sketched here and detailed in Appendix E.

precisely in Appendix E, therefore, even the agents who are lucky enough to be assigned objects in O_1 have a non-vanishing chance of suffering a significant number of rejections before they are assigned. These agents will therefore attain payoffs that are on average bounded away from $U(u^1, 1)$.

This outcome is inconsistent with asymptotic efficiency. To see this, suppose that, once objects are assigned through DA, the Shapley-Scarf TTC is run with their DA assignment serving as the agents' initial endowment. The resulting reassignment Pareto dominates the DA assignment. Further, it is Pareto efficient. Then by Theorem 1, all agents assigned to O_1 enjoy payoffs arbitrarily close to $U(u^1, 1)$ when the market grows large. This implies that a significant number of agents will enjoy a significant welfare gain from a Pareto dominating reassignment.

It is worth emphasizing that in the presence of systematic correlation in agents' preferences, DA, or equivalently stability, forces the agents to compete one another so intensively as to entail significant welfare loss. This observation serves as a key motivation for designing a new mechanism that we will show next is asymptotically efficient as well as asymptotically stable.

4.4 DA with Circuit Breaker

As we just saw, two of the most prominent mechanisms fail to find matchings which are asymptotically efficient and asymptotically stable. Is there a mechanism that attains both properties? In the sequel, we propose a new mechanism which finds such matchings.[21] To be more precise, we define a class of mechanisms indexed by some integer κ (allowed to be ∞ as well). We will show how an appropriate value of κ can be chosen in order to achieve our goal.

Given a value κ, the DA with Circuit Breaker algorithm (DACB) is defined recursively on triplets: $\hat{I}$ and $\hat{O}$, the sets of remaining agents and objects, respectively, and a counter for each agent that records the number of times the agent has made an offer. We first initialize $\hat{I} = I$ and $\hat{O} = O$, and set the counter for each agent to be zero.

Step 0: Linearly order individuals in $\hat{I}$.

Step 1**:** Let the individual with the lowest index in $\hat{I}$ make an offer to his most favorite

[21]As we discuss in Remark 5, existence of such a mechanism can be established by appealing to Erdos-Renyi theorem. The implied mechanism is not practical and unlikely to have a good incentive property. By contrast, the mechanism that is proposed here does have a good incentive property, as we show below.

object in $\hat{O}$. The counter for that agent increases by one. This object tentatively holds that individual, and go to Step 2.

Step $i \geq 2$**:** The individual with index i (i.e., i-th lowest index) in $\hat{I}$ makes an offer to his most favorite object o in $\hat{O}$ among the objects to which he has not yet made an offer. The counter for that agent increases by one. If o is not tentatively holding any individual, then o tentatively holds that individual. Whenever the index of the agent who made an offer is equal to $|\hat{I}|$, end the algorithm; otherwise iterate to Step $i+1$. If however o is holding an individual tentatively, he accepts tentatively the one who is higher in its priority list, and rejects the other. There are two cases to consider:

1. If the counter for the agent who has made an offer is greater than or equal to κ, then each agent who is assigned tentatively an object in Steps $1, ..., i$ is assigned that object. Reset $\hat{O}$ to be the set of unassigned objects and $\hat{I}$ to be the set of unassigned individuals. Reset the counter for the agent rejected at step i to be zero. If $\hat{I}$ is non-empty, go back to Step 1, otherwise, terminate the algorithm.

2. If the counter for the agent who has made an offer is strictly below κ, we return to the beginning of Step i.

This process terminates in finite time and yields a matching μ. This algorithm modifies the McVitie and Wilson (1971) version of DA where the tentative assignments are periodically finalized. We say that a **stage** begins whenever $\hat{O}$ is reset, and the stages are numbered $1, 2, ...$ serially.

The DACB mechanism encompasses a broad spectrum of mechanisms depending on the value of κ. If $\kappa = 1$, then each stage consists of one step, wherein an agent acts as a dictator with respect to the objects remaining at that stage. Hence, with $\kappa = 1$, the DACB reduces to a serial dictatorship mechanism with the ordering over agents given in Step 0. A serial dictatorship is efficient, but obviously fails (even asymptotic) stability since it completely ignores the agents' priorities at the objects. By contrast, if $\kappa = +\infty$, then the DACB mechanism coincides with the DA mechanism. As was seen already, DA is stable but fails to be asymptotically efficient. So intuitively, κ should be large enough to allow agents to make enough offers (or else, we will not achieve asymptotic stability), but should be small enough to avoid excessive competition by the agents (or else, the outcome would not be asymptotically efficient).

The next theorem provides the relevant lower and upper bounds on κ to ensure that the DACB mechanism attains both asymptotic efficiency and asymptotic stability.

THEOREM 5. *If $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$ then DACB is asymptotically efficient and asymptotically stable.*

In the sequel, we assume that $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$. The Theorem directly follows from the proposition below.

PROPOSITION 1. *Fix any $k \geq 1$. As $n \to \infty$, with probability approaching one, stage k of the DACB ends at step $|O_k| + 1$ and the set of assigned objects at that stage is O_k. In addition, for any $\epsilon > 0$ and γ*

$$\frac{|\{i \in I_k | U_i(DACB(i)) \geq U(u_k, 1) - \epsilon\}|}{|I_k|} \xrightarrow{p} 1$$

as $n \to \infty$; where $I_k := \{i \in I | DACB(i) \in O_k\}$. Similarly,

$$\frac{|\{o \in O_k | V_o(DACB(o)) \geq V(1) - \epsilon\}|}{|O_k|} \xrightarrow{p} 1$$

as $n \to \infty$.

PROOF OF PROPOSITION 1. We focus on $k = 1$, as will become clear, the other cases can be treated exactly in the same way. In the sequel, we fix ϵ and $\gamma > 0$.

First, consider the submarket that consists of the $|O_1|$ first agents (according to the ordering given in the definition of DACB) and of all objects in O_1 objects. If we were to run standard DA just for this submarket, then because preferences are drawn iid, by Pittel [Theorem 6.1., (b) 1992], with probability approaching 1 as $n \to \infty$ *all* agents have made less than $\log(n)^2$ offers at the end of (standard) DA.

Consider now the original market. By Lemma 2 (and the fact that $\kappa(n) = o(n)$ implies that $\kappa(n) \leq \delta |O_1|$ for any n large enough), the event that all agents' $\kappa(n)$ favorite objects are in O_1 has probability approaching 1 as $n \to \infty$. Let us condition on this event, labeled $\mathcal{E}$. Given this conditioning event $\mathcal{E}$, no object outside O_1 would receive an offer before somebody reaches his $\kappa(n)$-th offer. Also since not all of the first $|O_1| + 1$ agents can get assigned objects in O_1, given $\mathcal{E}$, one of these agents must reach his $\kappa(n)$-th offer, having made offers only to objects in O_1. We thus conclude that Stage 1 will end at Step $|O_1| + 1$ or before, with *only* the objects in O_1 being assigned by the end of that stage, conditioning on the event $\mathcal{E}$.

We now show that *all* objects in O_1 are assigned by the end of Stage 1, and that Stage 1 indeed ends at Step $|O_1| + 1$. Note that under our conditioning event $\mathcal{E}$, the distribution of individuals' preferences over objects in O_1 is the same as the unconditional one (of course, this is also true for the distribution of objects' priorities over individuals). Given event $\mathcal{E}$, as long as all agents have made fewer than $\kappa(n)$ offers, the $|O_1|$ first steps of DACB proceed exactly in the same way as DA in the submarket composed of the $|O_1|$ first agents (according to the ordering used in DACB) and of all objects in O_1 objects. Applying the result by Pittel mentioned above, with probability going to 1 as $n \to \infty$, we then reach the end of Step $|O_1|$ of DACB before Stage 1 ends (i.e., before any agent has applied to his $\log(n)^2 \leq \kappa(n)$ most favorite object). Thus, with probability going to 1, the outcome so far coincides with the one attained in DA in the submarket composed of the $|O_1|$ first agents and of all objects in O_1 objects. This implies that, conditional on $\mathcal{E}$, with probability going to 1, all objects in O_1 are assigned and thus step $|O_1| + 1$ must be triggered. In addition, given that $\Pr(\mathcal{E}) \to 1$ as $n \to \infty$, with unconditional probability going to 1, Step $|O_1| + 1$ will be triggered. This completes the proof of the first part of Proposition 1.

In sum, with probability going to 1, the first $|O_1|$ Steps (i.e., Stage 1) of DACB proceed exactly the same way as DA in the submarket that consists of the $|O_1|$ first agents and of all objects in O_1 objects. Thus, appealing again to Pittel, with probability going to 1 by the end of Stage $|O_1|$, the proportion of objects in O_1 for which $V_o(DACB(o)) \geq 1 - \epsilon$ is above any γ arbitrarily close to 1. Since objects in O_1 will have received even more offers at the end of Stage 1, it must still be that, with probability going to 1, the proportion of objects in O_1 for which $V_o(DACB(o)) \geq 1 - \epsilon$ is above γ when n is large enough. Finally, by construction, all matched individuals obtain an object within their $\kappa(n)$ most favorite objects which by Lemma 2 implies that with probability going to 1, they enjoy a payoff above $U(u_1, 1) - \epsilon$.[22] Thus, for $k = 1$, the second statement in Proposition 1 is proved provided that our conditioning event $\mathcal{E}$ holds. Since, again, this event has probability going to 1 as $n \to \infty$, the result must hold even without the conditioning. Thus, we have proved Proposition 1 for the case $k = 1$.

Consider next Stage $k > 1$. The objects remaining in Stage k have received no offers in Stages $j = 1, ..., k-1$ (or else the objects would have been assigned in those stages). Hence, by the principle of deferred decisions, we can assume that the individuals' preferences over those objects are yet to be drawn in the beginning of Stage k. Similarly, we can assume

[22]Note that this implies that

$$\Pr\left[\{i \in I_k | U_i(DACB(i)) \geq U(u_k, 1) - \epsilon\} = I_k\right] \to 1$$

as $n \to \infty$. Hence, part of the statement of Proposition 1 can be strengthened.

that priorities of those objects are also yet to be drawn. Put in another way, conditional on Stage $k-1$ being over, we can assume without loss that the distribution of preferences and priorities is the same as the unconditional one. Thus, we can proceed inductively to complete the proof. $\square$

Theorem 5 shows that DACB is superior to DA or TTC in large markets when the designer cares about both (asymptotic) efficiency and (asymptotic) stability. One potential drawback of DACB is that it is not strategy-proof.[23] In particular, at each stage k, the agent who makes applications but is eventually unassigned at that stage may want to misreport his preferences by including in his top κ favorite objects a "safe" item which is outside his top κ favorite objects but is unlikely to be popular among other agents. Such a misreporting could benefit the agent since the safe item would not have received any other offer and thus would accept him whereas truthful reporting could let him unassigned at that stage and result in the agent receiving a worse object. But the chance of becoming such an agent is one out of the number of agents assigned in the stage, so for an appropriate choice of κ, it is very small in a large economy. Hence, the incentive problem with the DACB is not very serious.

To formalize this idea, we study the Bayesian game induced by DACB. In this game, the set of types for each individual corresponds to his vector of cardinal utilities, i.e., $\{U_i(o)\}_{o\in O}$, or equivalently, $\xi_i := \{\xi_{i,o}\}_{o\in O}$. These values are drawn according to the distributions assumed so far. The underlying informational environment is Bayesian: each individual only knows his own preferences, labeled his "type," and knows the distribution of others' preferences and the distribution of priorities (including its own).

[23] Similarly, truthful-reporting may not be Bayesian incentive compatible under DACB in a finite economy. To see this, suppose there are three individuals and three tiers of objects each containing only one object. Let us note these objects o_1, o_2 and o_3. Further assume that the differences in common values are so large that irrespective of the idiosyncratic shocks, individuals all agree on their ordinal ranking: o_1 is ranked first, then o_2 and finally o_3 is the less desirable object. Assume that $\kappa = 2$. In that case if all individuals report truthfully, under DACB, individual 1 gets object o_1 with probability 1/2 (in the event o_1 ranks 1 above 2 in which case 1 is the individual matched in the first stage), object o_2 with probability 1/4 (in which case 1 is the individual matched in the second stage and o_2 ranks 1 above 3) and object o_3 with probability 1/4 (in which case 1 is the individual matched in the third stage). However, if individual 1 lies and reports that o_2 is his most favorite object, then, provided that the other individuals report truthfully, 1 must believe that he will get o_2 with probability 1 (either 2 or 3 end stage 1 and individual 1 is matched in the fist stage for sure). With an appropriate choice of common values and of the upper bound on the idiosyncratic shock (i.e., 1), the misreport will be profitable. As we argue below, however, in the large economy, truthful reporting is ϵ-Bayesian Nash equilibrium.

DACB maps profiles of ordinal preferences reported by the agents and their priorities with objects into matchings. In the game induced by DACB, the set of actions by individual i of a given type ξ_i is the set of all possible ordinal preferences the agent may report. A typical element of that set will be denoted P_i. Each type ξ_i induces an ordinal preference which we denote $P_i(\xi_i)$. This is interpreted as the truthful report of individual i of type ξ_i. Given any $\epsilon > 0$, say that truthtelling is an interim ϵ-Bayes-Nash equilibrium if for each individual i, each type ξ_i and any possible report of ordinal preferences P_i', we have

$$\mathbb{E}\left[U_i(DACB_i(P_i(\xi_i), \cdot)) \mid \xi_i\right] \geq \mathbb{E}\left[U_i(DACB_i(P_i', \cdot)) \mid \xi_i\right] - \epsilon,$$

where $U_i(DACB(P_i, .))$ denotes the random utility that i gets given that he reports P_i. We state the following result which provides a sense in which DACB performs well from incentives perspectives.

THEOREM 6. *Let us assume that $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$. Fix any $\epsilon > 0$. Under DACB, there exists $N > 0$ such that for all $n > N$, truthtelling is an interim ϵ-Bayes-Nash equilibrium.*

PROOF. See Appendix F. □

So far the informational environment assumes that each agent only knows his own preferences. One could assume further that the agent's private information contains some additional pieces of information like, for instance, the information on his priorities. In such a case, agent i's type would be the pair $(\xi_i, \eta_i) := (\{\xi_{i,o}\}_{o\in O}, \{\eta_{i,o}\}_{o\in O})$. It is worth pointing out that DACB still has nice incentive properties even in this richer context. Indeed, for the relevant choice of $\kappa(n)$ (i.e., $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$), it is easy to see that, for any $\epsilon > 0$, it is an ex ante ϵ-Bayes-Nash equilibrium to report truthfully when the number of agents is large enough.[24] Indeed, to see this, let us fix $k = 1, ..., K$ and $i \in \{|O_{\leq k-1}|+2, ..., |O_{\leq k}|+1\}$ (with the convention that $|O_{\leq 0}|+2 = 1$ and $|O_{\leq K}|+1 = n$). By Proposition 1, under truthful behavior, the ex ante probability that i gets a payoff arbitrarily close to $U(u_k, 1)$ converges to 1 as n goes to infinity. Hence, i's ex ante payoffs when agents report truthfully must be above $U(u_k, 1) - \frac{\epsilon}{2}$ when the number of agents is large enough. In addition, we know that, by the argument given in the proof of Proposition 1, irrespective of agent i's behavior, provided that the other agents are truthful, the probability that i is assigned in a stage prior to k goes to 0 as n gets large. Hence, for any possible report of individual i, when the number of agents is large enough, i's ex ante payoffs must

[24] A truthful report consists in reporting the true preferences irrespective of the information on priorities.

be bounded above by $U(u_k, 1) + \frac{\epsilon}{2}$. Thus, a deviation from truthful behavior cannot make the deviating agent ϵ-better off in ex ante terms. The following proposition summarizes the above discussion.

THEOREM 7. *Let us assume that $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$. Fix any $\epsilon > 0$. Under DACB, there exists $N > 0$ such that for all $n > N$, truthtelling is an ex ante ϵ-Bayes-Nash equilibrium.*

REMARK 5 (Feasibility of asymptotic efficiency and asymptotic stability). The feasibility of attaining both asymptotic efficiency and asymptotic stability can be seen directly by appealing to the Erdös-Renyi theorem. The theorem states that a random bipartite graph that consists of vertices I and O and of edges created at random with fixed probability $p > 0$ for each $(i, o)_{i \in I, o \in O}$ admits a *perfect bipartite matching* with probability approaching one as $n = |I| = |O|$ tends to ∞.[25] Well-known algorithms such as the *augmenting path algorithm* would find a maximal matching and thus would find a perfect matching whenever it exists.

Exploiting the Erdös-Renyi theorem, one can construct a mechanism in which (1) agents and objects (more precisely their suppliers) report their idiosyncratic shocks, (2) an edge is created between an agent and an object if and only if $\xi_{i,o} > 1 - \epsilon$ and $\eta_{i,o} > 1 - \epsilon$, for any arbitrary $\epsilon > 0$, and (3) a maximal bipartite matching is found. The matching obtained in this way is asymptotically efficient and asymptotically stable, since all objects attain arbitrarily high payoffs and all agents realize arbitrarily high idiosyncratic payoffs. But this mechanism would not be desirable for several reasons. First, it would not work if the agents cannot tell apart common values from idiosyncratic values. More important, the mechanism would not have a good incentive property. An agent will be reluctant to report the objects in lower tiers even though they have high idiosyncratic preferences. Indeed, if he expects that with significant probability, he will not get any object in the highest tier, he will have incentives to claim that he enjoys high idiosyncratic payoffs with a large number of high tier objects and that all his idiosyncratic payoffs for the other tiers are low. It is very likely that there is a perfect matching even under this misreport and this will ensure him to get matched with a high tier object.

[25] A perfect bipartite matching is a bipartite graph in which each vertex is involved in exactly one edge.

5 Simulations

The analyses so far have focused on the limiting situation in which the size of market becomes arbitrarily large. Furthermore, part of the analysis has assumed that the common values have a tier structure. To understand the outcomes for broader settings, we perform a number of simulations varying the market sizes as well as the agents' preference distributions.

Figure 3 shows the utilitarian welfare—more precisely the average idiosyncratic utility enjoyed by the agents—under the alternative algorithms for varying market sizes ranging from $n = 10$ to $10,000$, where idiosyncratic and common values are drawn uniformly from $[0, 1]$.[26]

First, the figure shows that agents' payoffs under TTC converge fast to the utilitarian upper bound, attaining 97% for a moderate market size of $n = 1,000$, and reaching 99% when $n = 10,000$. This result reinforces Theorem 1, showing that the payoff implications of Pareto efficiency hold even for a moderate market size. More important, the same figure shows the relative performances of DA and DACB. As expected, the agents attain lower payoffs under these mechanisms than under TTC. There are significant differences between them, however. The agents attain 90% and 96% of payoff upper bound under DACB respectively for $n = 1,000$ and $n = 10,000$, whereas they achieve 80% and 81% respectively under DA. As n rises in this range, the DACB's gap relative to TTC narrows to 3%, while its gap relative to DA widens to 15%.

Figure 4 shows the fraction of ϵ-blocking pairs under DA, DACB and TTC. Clearly, DA admits no blocking pairs, so the fraction is always zero. Between TTC and DACB, there is a substantial difference. Blocking pairs admitted by TTC comprise almost 9% of all possible pairs, whereas DACB admits blocking pairs that are less than 1% of all possible pairs, and these proportions do not vary much with the market size.

When we weaken the notion of stability to $\epsilon-$stability, focusing only on the pairs of the agent and object that would benefit from blocking more than ϵ percentile improvement in rankings, the fractions decline (not surprisingly), but the differences between the two

[26] We also performed several simulations in which the common values are distributed over two values, as would be in the two tier case. The simulations outcomes in this case are largely similar to the ones reported here. The mechanisms were simulated under varying number of random drawings of the idiosyncratic and common utilities: $1000, 1000, 500, 500, 200, 200, 100, 100, 20, 10$ for the market sizes $n = 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000$.

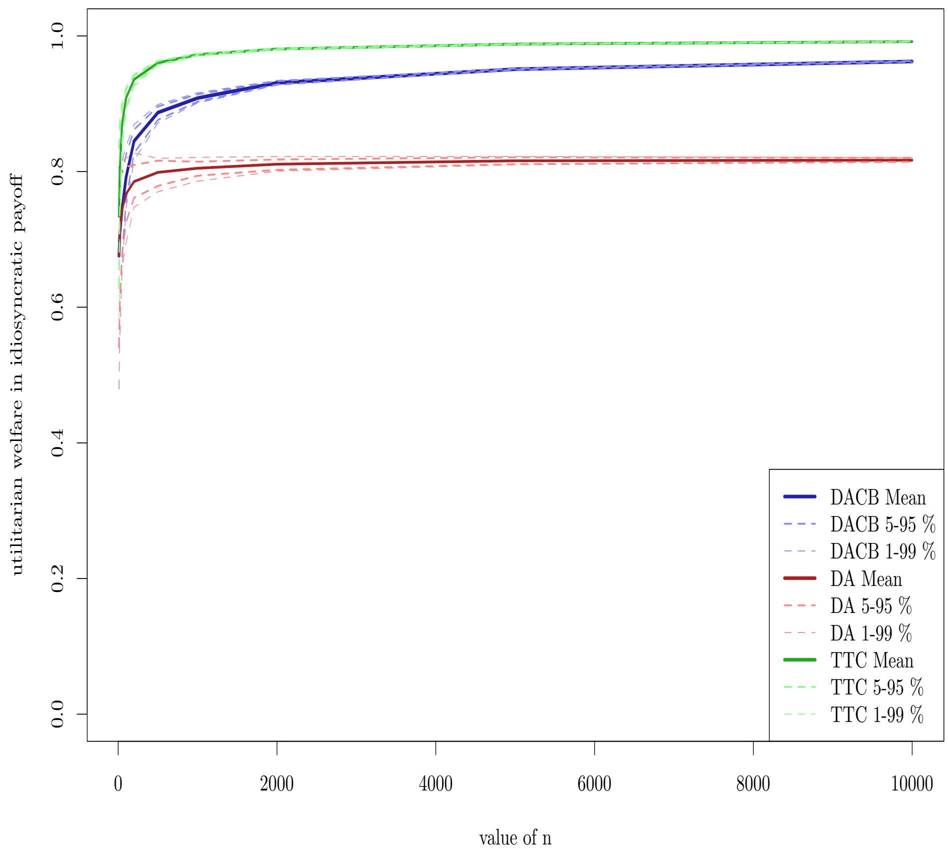

Figure 3: Utilitarian welfare under alternative mechanisms

mechanisms remain significant. Under $\epsilon = 0.05$, the fractions of $\epsilon-$blocking pairs are around 7% for TTC, but the fraction is close to zero under DACB, again largely irrespective of the market size.

Recall our analysis focuses on one-to-one matching. In order to test the robustness of our results to a realistic many-to-one matching setting, we consider a model in which 100,000 students are to be assigned to 500 schools each with 200 seats. We assume that common and idiosyncratic values are both drawn uniformly from $[0, 1]$. We simulate DACB under a variety of values of κ's. Figure 5 shows the sense in which DACB with different κ's span different ways to compromise on efficiency (left panel) and stability (right panel). For a low value of κ, DACB performs similarly to TTC, attaining about 97.5% of the utilitarian

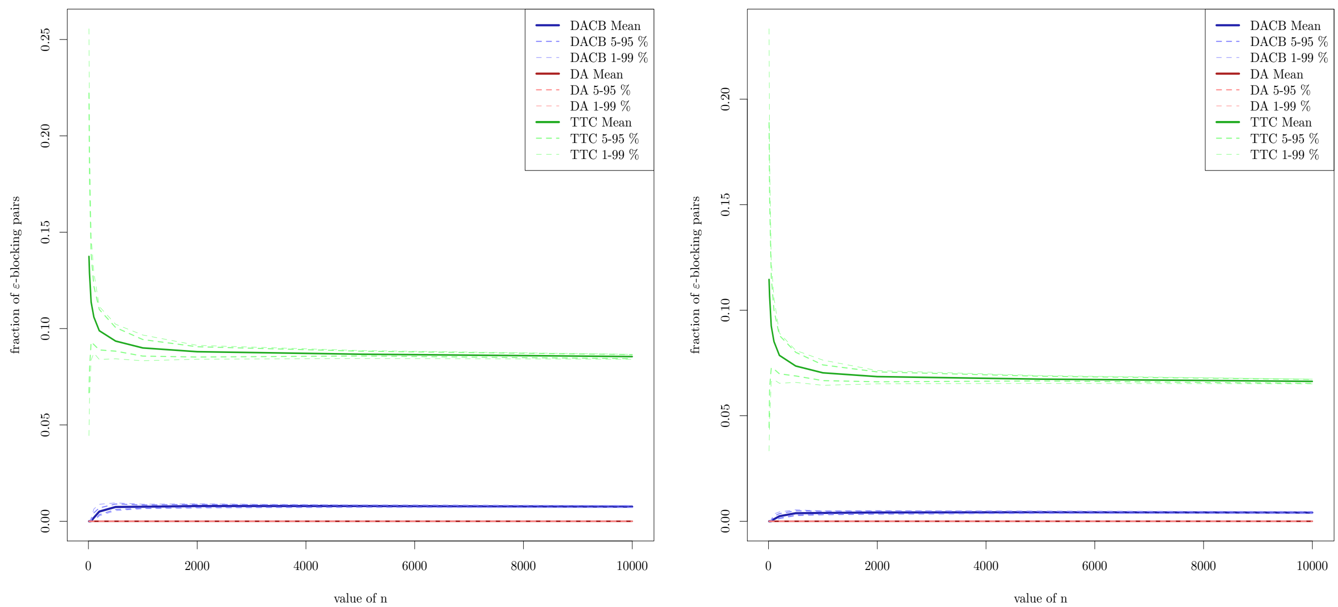

Figure 4: The fraction of ϵ-blocking pairs under alternative mechanisms with $\epsilon = 0$ (left panel) and $\epsilon = 0.05$ (right panel)

welfare upper bound and admitting about 18% of blocking pairs, whereas DA attains close to 80% of the utilitarian welfare upper bound. For a low κ, DACB resembles TTC with high welfare and high incidences of blocking pairs. As κ rises, the utilitarian welfare under DACB falls, but the number of blocking pairs declines as well. For instance, DACB with $\kappa = 35$ keeps the utilitarian welfare still very high at 97% but the proportion of blocking pairs below 10%. For DACB with $\kappa = 70$, the utilitarian welfare remains still high at 95% but the fraction of blocking pairs drops down to almost 5%.

The simulations show that the results obtained in the paper hold broadly in terms of the market sizes and in terms of the distribution of the common values.

6 Concluding Remarks

The current paper has studied the tradeoff between efficiency and stability—two desiderata in market design—in large markets. The two standard design alternatives, Gale and

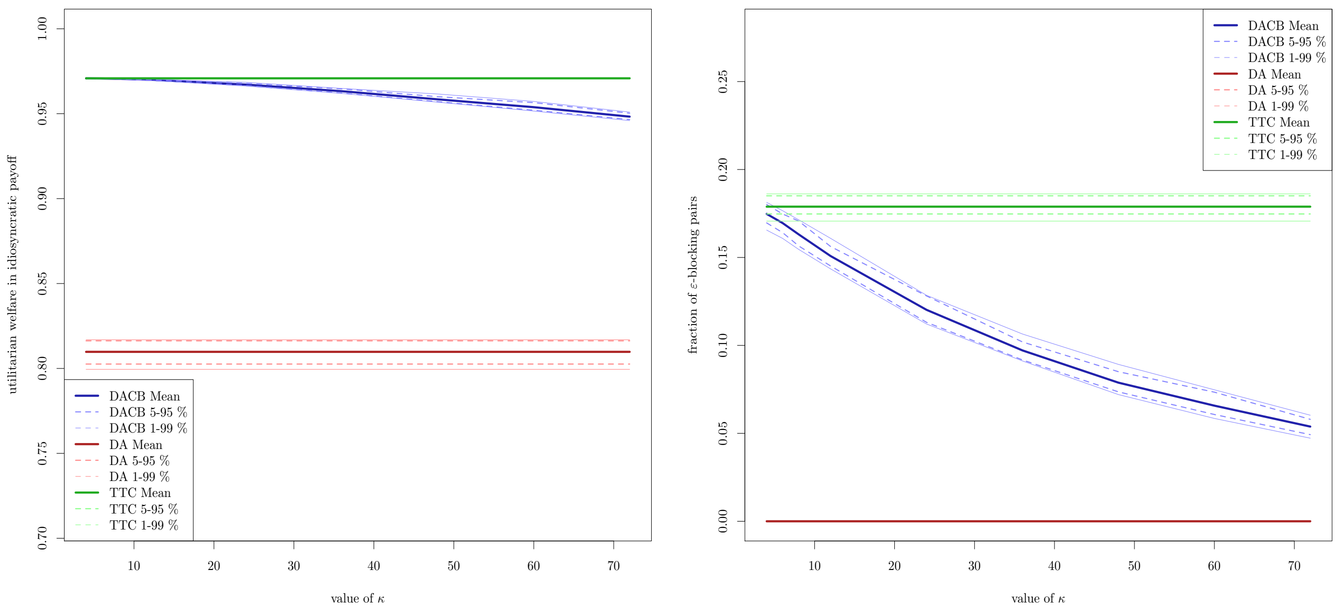

Figure 5: Utilitarian welfare (left panel) and the fraction of 0-blocking pairs (right panel) under DACB with varying κ's

Shapley's deferred acceptance algorithm (DA) and top trading cycles (TTC), satisfy one property but fails the other. Considering a plausible class of situations in which individual agents have random preferences for objects that contain both common and idiosyncratic components and their priorities at the objects are independently and identically drawn, we show that theses failures—the inefficiency of DA and instability of TTC—remain significant even in large markets.

We then propose a new mechanism, deferred acceptance with a circuit breaker (DACB), which modifies the DA to keep agents from competing excessively for over-demanded objects—a root cause of its significant efficiency loss in the large market. Specifically, the proposed mechanism builds on the the McVitie and Wilson's version of DA in which agents make offers one at a time following a serial order, but involves a "circuit breaker" which keeps track of the number of offers each agent makes (or equivalently, the number of rejections he/she suffers) in the serialized offer/acceptance process, and periodically stops the process and finalizes the assignment up to that point when the number of offers for some agent reaches a pre-specified threshold ("trigger"), and restarts the serialized offer/acceptance process with the remaining agents and objects, again with a circuit breaker, and so on. We show in our model with finite common values that DACB, with the trigger set appropriately, achieves both efficiency and stability in an approximate sense as the economy grows large, and it induces truth-telling in an ϵ-Bayes Nash equilibrium.

Our results, along with a number of simulations performed, suggest that DACB will perform well on both objectives in broad and realistic settings. For a practical application of DACB, the two design parameters, namely (1) the serial order of the agents and (2) the threshold number of offers that triggers assignment, can be optimized relative to the detailed features of the market in question. For instance, the serial order can be chosen to reflect the priorities of the schools that are demanded most and the trigger can be set to reflect the common value structure of the market.[27] In practice, both instruments can be fine-tuned toward the specifics of a given market in question.

Finally, it is worth noting that our proposed mechanism shares several features of mechanisms that are already in use. The mechanism used by Shanghai school system for assigning university seats to students employs the "staged" version of DA: the procedure runs DA in each stage to assign students based on the rank-order list of schools submitted by students, which is truncated by a pre-specified number; and those who are unassigned are assigned following the same procedure in the subsequent stages. Truncation of the choice used in each stage performs a similar function as our circuit breaker. The difference is that the students apply simultaneously in the Chinese mechanism whereas in DACB offers are made sequentially one at a time. The simultaneous offer process means that each student faces a significant chance of being unassigned after each stage; this creates a significant incentive problem since a student will find it optimal to insert a safe choice within her truncated list of choice to avoid not being matched in that stage. By contrast, the sequential offer process of DACB makes it very unlikely for each participant active in a given stage to be unassigned in a reasonably-sized market, so the agent has no serious incentive to manipulate his preference. Nevertheless, the Chinese mechanism may exhibit the similar desirable properties as DACB in equilibrium. In addition, the simultaneous offer process of the Chinese mechanism has the desirable feature that students ranked highly by popular schools will likely claim these schools in case they prefer them, a feature that is conducive to minimizing justified envy. This suggests that Chinese mechanism may be a potentially

[27] For instance, if the priorities of the schools are largely given by the standardized test scores, the test scores will form a natural serial order. In the absence of such a universal priority criterion, one could average a student's priorities at alternative schools and use that average priority to determine a serial order.

promising way to implement the outcome of DACB; It remains an open question both theoretically and empirically how students would react to the strategic environment of the Chinese mechanism.

A Proof of Theorem 1

A.1 Preliminaries

We first partition the set of objects in each n-economy based on their common values into finite tiers. That is, let $\sup 1 : u^1 > u^2 > ... > u^K = 0$. In the finite-tier economy, the tiers here can be defined to correspond to the common values (since there is a finite number of them). In the general model, any such tiers will induce a CDF which will approximate the true distribution Y^n from below (as K increases). For each $k = 1, \dots, K$, define $O^n_{\leq k} := \{o \in O^n | u_o \geq u^k\}$ be the set of objects in tier k or better, and let $Y^n_{\leq k} := Y^n(1) - \lim_{u' \uparrow u^k} Y^n(u')$ and $X^n_{\leq k} := X^n(1) - \lim_{u' \uparrow u^k} X^n(u')$, denote the associated mass of objects and the associated mass of copies of objects. Define similarly $Y_{\leq k} := Y(1) - \lim_{u' \uparrow u^k} Y(u')$ and $X_{\leq k} := X(1) - \lim_{u' \uparrow u^k} X(u')$ for the limit economy. From now on, for notational ease, we shall suppress n except for X^n and Y^n to avoid confusion with their limit counterparts.

Now, consider any Pareto efficient mechanism $\mu \in \mathcal{M}^*$. By a well known result (e.g., Abdulkadiroglu and Sönmez (1998)), any Pareto efficient matching can be equivalently implemented by a serial dictatorship mechanism with a suitably chosen serial order. Let SD^{f_μ} be the serial dictatorship mechanism where for each state ω a serial order $f_\mu(\omega) : I \to I$, a bijective mapping, is chosen so as to implement $\mu_\omega(\cdot)$. That is, for each state $\omega \in \Omega$, the serial order f_μ is chosen so that $SD^{f_\mu(\omega)}_\omega(i) = \mu_\omega(i)$ for each $i \in I$. Since the matching μ arising from the mechanism depends on the random state ω, so is the serial order f implementing μ. In the sequel, we shall study a Pareto efficient matching mechanism μ via the associated SD^{f_μ}. To avoid clutter, we shall now suppress the dependence of f on μ.

Given an n-economy, for any Pareto efficient mechanism μ and the associated serial order f, let

$$I_{\leq k}(\mu) := \{i \in I | f(i) \leq Q^n X^n_{\leq k}\}.$$

be the set of agents who have a serial order within the total supply of objects in tiers k or better (in the equivalent serial dictatorship implementation). For any ϵ, the set

$$\bar{I}_{\leq k}(\mu) = \{i \in I_{\leq k}(\mu) | U_i(SD^f(i)) \leq U(u^k, 1 - \epsilon)\},$$

consists of the agents who realize payoff no greater than $U(u^k, 1-\epsilon)$ while having a serial order within $Q^n X^n_{\leq k}$. The following lemma will be crucial for the main result.

Lemma 3. *For any ϵ and $\gamma > 0$,*

$$\Pr\left[\exists \mu \in \mathcal{M}^* \text{ such that } \frac{|\bar{I}_{\leq k}(\mu)|}{|I|} \geq \gamma\right] \to 0$$

as $n \to \infty$.

PROOF. For each k such that $Y_{\leq k} = 0$, the result holds trivially since in that case, $\frac{|\bar{I}_{\leq k}(\mu)|}{|I|} \leq \bar{q} Y^n_{\leq k} \to \bar{q} Y_{\leq k} = 0$. So let us consider k s.t. $Y_{\leq k} > 0$. Fix any $\epsilon > 0$ and $\gamma > 0$. We first build a random graph on $I \cup O$ where an edge (i, o) is added if and only if $\xi_{i,o} > 1 - \epsilon$.

Now choose any $\delta \in (0, 1)$. For each $\mu \in \mathcal{M}^*$, define random sets $I^{\delta}_{\leq k}(\mu) := \{i \in I \,\big|\, f(i) \leq Q^n X^n_{\leq k}(1 - \delta)\}$, $\bar{I}^{\delta}_{\leq k}(\mu) := \{i \in I^{\delta}_{\leq k} \,\big|\, U_i(SD^f(i)) \leq U(u^k, 1 - \epsilon)\}$, and

$$\bar{O}^{\delta}_{\leq k}(\mu) := \{o \in O_{\leq k} \,\big|\, \exists i \in \mu^{-1}(o) \text{ s.t. } f(i) > Q^n X^n_{\leq k}(1 - \delta)\},$$

which consists of objects in $O_{\leq k}$ assigned to the agents with serial order worse than $Q^n X^n_{\leq k}(1 - \delta)$.

Then, the set $\overline{I}^{\delta}_{\leq k}(\mu) \cup \bar{O}^{\delta}_{\leq k}(\mu)$ must be an independent set. If not, there would exist an edge $(i, o) \in \overline{I}^{\delta}_{\leq k} \times \bar{O}^{\delta}_{\leq k}$. Then,

$$U_i(o) > U(u^k, 1 - \epsilon) \geq U_i(SD^f(i))$$

where the strict inequality holds since $\xi_{i,o} > 1 - \epsilon$ (i.e., (i, o) is an edge), $o \in O_{\leq k}$, and since $U(\cdot, \cdot)$ is monotonic (in particular strictly increasing in idiosyncratic component). The weak inequality holds because $i \in \overline{I}^{\delta}_{\leq k}$. In addition, we must have

$$f(i) \leq Q^n X^n_{\leq k}(1 - \delta) < f(i'), \text{ for some } i' \in \mu^{-1}(o)$$

where the first inequality comes from the fact that $i \in I^{\delta}_{\leq k}(\mu)$ while the second from the fact that $o \in \bar{O}^{\delta}_{\leq k}(\mu)$. Thus, this means that when i becomes the dictator under SD^f, object o is still available, and the agent does not choose it. But $U_i(o) > U_i(SD^f(i))$ means that i chooses an object worse than o, which yields a contradiction.

Thus, for each $\mu \in \mathcal{M}^*$, $\overline{I}^{\delta}_{\leq k}(\mu) \cup \bar{O}^{\delta}_{\leq k}(\mu)$ contains a balanced independent set with size $\min\left\{|\overline{I}^{\delta}_{\leq k}(\mu)|, |\bar{O}^{\delta}_{\leq k}(\mu)|\right\}$. Since $|I| = n$ and $|O|$ is in the order of n, applying Lemma 1, we get that, for any $\tilde{\gamma} > 0$:

$$\Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } \min\left\{|\overline{I}^{\delta}_{\leq k}(\mu)|, |\bar{O}^{\delta}_{\leq k}(\mu)|\right\} \geq \tilde{\gamma} n\right] \to 0 \tag{3}$$

as n goes to infinity.

Since $\left|\bar{O}^{\delta}_{\leq k}(\mu)\right| \bar{q} \geq \sum_{o \in \bar{O}^{\delta}_{\leq k}(\mu)} q_o \geq \left\lfloor \delta Q^n X^n_{\leq k} \right\rfloor = \left\lfloor \delta n X^n_{\leq k} \right\rfloor$ for each $\mu \in \mathcal{M}^*$, and since $X^n_{\leq k} \to X_{\leq k} \geq Y_{\leq k} > 0$ as $n \to \infty$, one can find $\beta > 0$ and $N_1 \in \mathbb{N}$ such that for all $n > N_1$,

$\left|\bar{O}^{\delta}_{\leq k}(\mu)\right| \geq \beta n$ for each $\mu \in \mathcal{M}^*$.[28] Hence, for any $\gamma' > 0$ and for any $n > N_1$:

$$\begin{aligned}\Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } |\overline{I}^{\delta}_{\leq k}(\mu)| \geq \gamma' n\right] &\leq \Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } |\overline{I}^{\delta}_{\leq k}(\mu)| \geq \min\{\gamma', \beta\} n\right] \\ &= \Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } \min\left\{|\overline{I}^{\delta}_{\leq k}(\mu)|, |\bar{O}^{\delta}_{\leq k}(\mu)|\right\} \geq \min\{\gamma', \beta\} n\right] \\ &\to 0,\end{aligned}$$

as n goes to infinity, where the equality comes from the choice of β and N_1 while the convergence to 0 holds by (3).

Finally, by construction, $|\bar{I}^{\delta}_{\leq k}(\mu)| \geq |\bar{I}_{\leq k}(\mu)| - \lfloor \delta Q^n X^n_{\leq k} \rfloor \geq |\bar{I}_{\leq k}(\mu)| - \delta Q^n X^n_{\leq k}$. Since $Q^n = |I|$ and $X^n_{\leq k} \leq 1$, we get that

$$\frac{\left|\bar{I}^{\delta}_{\leq k}(\mu)\right|}{|I|} \geq \frac{\left|\bar{I}_{\leq k}(\mu)\right|}{|I|} - \delta$$

for each $\mu \in \mathcal{M}^*$. Hence, it follows that

$$\Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } \frac{\left|\bar{I}_{\leq k}(\mu)\right|}{|I|} \geq \gamma' + \delta\right] \leq \Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } \frac{\left|\bar{I}^{\delta}_{\leq k}(\mu)\right|}{|I|} \geq \gamma'\right] \to 0.$$

Set δ and γ' such that $\delta + \gamma' = \gamma$. Then,

$$\Pr\left[\exists \mu \in \mathcal{M}^* \text{ s.t. } \frac{\left|\bar{I}_{\leq k}(\mu)\right|}{|I|} \geq \gamma\right] \to 0.$$

□

We are now ready to prove Theorem 1.

A.2 Proof of Theorem 1

To prove the statement, we will show that the payoff distributions induced by Pareto efficient mechanisms converge to F^* in the sense defined earlier.

Fix any $\epsilon' > 0$ and $\upsilon' > 0$. We shall show that there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\Pr\left[\sup_{\mu \in \hat{\mathcal{M}}^*} \sup_{z \in Z} \inf_{\hat{z} \in [z-\epsilon', z+\epsilon']} \{|F^{\mu}(\hat{z}) - F^*(z)|\} \geq \epsilon'\right] < \upsilon', \tag{4}$$

[28] Here we use the assumption that $\bar{q}$ does not increase in n. If it were to depend on n and we further assume that it is $o(n/log(n))$, one can check that $\left|\bar{O}^{\delta}_{\leq k}(\mu)\right|$ is $\omega(log(n))$. Using Lemma 1, one can show that Theorem 1.

where F^* and F^μ are respectively the CDF of the payoff induced by the limit utilitarian upper bound and the CDF induced by mechanism μ in $\mathcal{M}^*$. We later show that this is sufficient for the proof.

To this end, we partition the common value space $[0,1]$ into intervals $\cup_{k=1}^K [u^k, u^{k-1})$, where $1 =: u^1 > u^2 > ... > u^K := 0)$ are such that $\sup_k \sup_{u,u' \in [u^k, u^{k-1})} |X(u) - X^n(u')| < \frac{\epsilon'}{2}$, for any $n > \hat{N}$ for some $\hat{N} \in \mathbb{N}$. Such a partition exists since $X^n \to X$ in Lévy metric and since one can select all points of discontinuity of X to be a subset of the threshold values for the partition. Define as before $O_{\leq k} = \{o \in O | u_o \geq u^k\}$ and $Y^n_{\leq k} := Y^n(1) - \lim_{u' \uparrow u^k} Y^n(u')$, $X^n_{\leq k} := X^n(1) - \lim_{u' \uparrow u^k} X^n(u')$.

The partition induces a corresponding partition of the payoff space $Z := [U(u_K, 1), U(1,1)]$ into intervals $Z_k := [U(u^k, 1), U(u^{k-1}, 1))$, $k = 1, ..., K$. Next, let $\epsilon > 0$ be such that $U(u^{k-1}, 1-\epsilon) > \max\{U(u^{k-1}, 1) - \epsilon', U(u^k, 1)\}$, for all $k = 1, ..., K$. Then, for each Pareto efficient mechanism $\mu \in \mathcal{M}^*$ and for each $z \in Z_k$, $k \in \{1, ..., K\}$, there exists $z' \in [z - \epsilon', z + \epsilon'] \cap Z_k$. Specifically, let $z' := \min\{z, U(u^{k-1}, 1-\epsilon)\}$. Clearly, given our choice of ϵ, we have that $z' \in Z_k$. In addition, given this choice of ϵ, $z' \in [z - \epsilon', z + \epsilon']$. Indeed, this is trivially true if $z \leq U(u^{k-1}, 1-\epsilon)$ (in which case $z' = z$) and if $z > U(u^{k-1}, 1-\epsilon)$, we have that $z - \epsilon' < U(u^{k-1}, 1) - \epsilon' < U(u^{k-1}, 1-\epsilon) = z' (< z)$ where the first inequality comes from the fact that $z \in Z_k$ while the second is by the choice of ϵ.

Define

$$J^\mu(z) = \{i \in I | U_i(\mu(i)) \leq z\}.$$

be the set of agents enjoying payoff of at most z under matching mechanism μ. Let u_z be such that $U(u_z, 1) = z$. (Given that we are interested in $z \in [U(0,1), U(1,1)]$, such a u_z is well defined since $U(\cdot, 1)$ continuous.) Clearly, any agent matched with an object having common value no greater than u_z must be in $J^\mu(z)$. This means that $|J^\mu(z)| \geq Q^n X^n(u_z)$ for all $\mu \in \mathcal{M}^*$.

By definition, for each z, $F^*(z) = X(u_z)$. Now, since $Q^n = n = |I|$, for each z, $\frac{|J^\mu(z')|}{|I|} - F^*(z) \geq X^n(u_{z'}) - X(u_z)$, for any $z' \in [z - \epsilon', z + \epsilon'] \cap Z_k$. (Hence, one can set z' as above.)

Fix $k = 1, ..., K$. Then, for all $n > \hat{N}$,

$$
\begin{aligned}
&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z\in Z_k}\inf_{z'\in[z-\epsilon',z+\epsilon']}\left(-\frac{|J^\mu(z')|}{|I|}+F^*(z)\right)\geq\epsilon'\right]\\
\leq&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z,z'\in Z_k}\left(-X^n(u_{z'})+X(u_z)\right)\geq\epsilon'\right]\\
=&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{u,u'\in[u^k,u^{k-1})}\left(-X^n(u')+X(u)\right)\geq\epsilon'\right]=0,
\end{aligned}
\tag{5}
$$

where the equality to 0 is from the definition of $\hat{N}$.

For each k, recall $I_{\leq k}(\mu) := \left\{i \in I | f(i) \leq Q^n X^n_{\leq k}\right\}$ and $\bar{I}_{\leq k}(\mu) := \left\{i \in I_{\leq k}(\mu) | U_i(SD^f(i)) \leq U(u^k, 1-\epsilon\right.$ where SD^f is the SD rule implementing μ.

Then, any agent who obtains a payoff weakly less than $U(u^{k-1}, 1-\epsilon)$ must be in the set $(I \setminus I_{\leq k-1}(\mu)) \cup \bar{I}_{\leq k-1}(\mu)$. As shown above, for any $z \in Z_k$, there exists $z' \in [z-\epsilon', z+\epsilon'] \cap Z_k$ such that $z' \leq U(u^{k-1}, 1-\epsilon)$. For such a z', we have $J^\mu(z') \subset (I \setminus I_{\leq k-1}(\mu)) \cup \bar{I}_{\leq k-1}(\mu)$. Hence, there exists $N_k \in \mathbb{N}$, with $N_k \geq \hat{N}$, such that for all $n > N_k$,

$$
\begin{aligned}
&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z\in Z_k}\inf_{z'\in[z-\epsilon',z+\epsilon']}\left(\frac{|J^\mu(z')|}{|I|}-F^*(z)\right)\geq\epsilon'\right]\\
\leq&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z\in Z_k}\left(\frac{|\bar{I}_{\leq k-1}(\mu)|+|I|-|I_{\leq k-1}|}{|I|}-X(u_z)\right)\geq\epsilon'\right]\\
=&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z\in Z_k}\left(\frac{|\bar{I}_{\leq k-1}(\mu)|}{|I|}+\lim_{u'\uparrow u^{k-1}}X^n(u')-X(u_z)\right)\geq\epsilon'\right]\\
\leq&\Pr\left[\sup_{\mu\in\mathcal{M}^*}\frac{|\bar{I}_{\leq k-1}(\mu)|}{|I|}+\sup_{u,u'\in[u^k,u^{k-1})}|X^n(u')-X(u)|\geq\epsilon'\right]\\
<&\upsilon'/K,
\end{aligned}
\tag{6}
$$

where the last inequality follows from Lemma 3 (with $\gamma = \epsilon'/2$) and an appropriate choice of N_k together with the definition of $\hat{N}$.

Combining (5) and (6), we get that for each $k = 1, \ldots, K$, and $n > N_k$,

$$
\Pr\left[\sup_{\mu\in\mathcal{M}^*}\sup_{z\in Z_k}\inf_{z'\in[z-\epsilon',z+\epsilon']}\left|\frac{|J^\mu(z')|}{|I|}-F^*(z)\right|\geq\epsilon'\right]<\upsilon'/K,
$$

Since $F^\mu(z') = \frac{|J^\mu(z')|}{|I|}$, for all $n > \max_k N_k$,

$$
\begin{aligned}
&\Pr\left[\sup_{\mu\in\mathcal{M}^*} L(F^\mu, F^*) \geq \epsilon'\right] \\
\leq &\Pr\left[\sup_{\mu\in\mathcal{M}^*}\left\{\sup_{z\in Z}\inf_{z'\in[z-\epsilon',z+\epsilon']}|F^\mu(z') - F^*(z)|\right\} \geq \epsilon'\right] \\
\leq &\sum_{k=1}^{K}\Pr\left[\sup_{\mu\in\mathcal{M}^*}\left\{\sup_{z\in Z_k}\inf_{z'\in[z-\epsilon',z+\epsilon']}|F^\mu(z') - F^*(z)|\right\} \geq \epsilon'\right] \\
< &K\upsilon'/K = \upsilon',
\end{aligned}
$$

where the first inequality holds since if $L(F^\mu, F^*) \geq \epsilon'$, then there exists z such that $\inf_{z'\in[z-\epsilon',z+\epsilon']}|F^\mu(z') - F^*(z)| \geq \epsilon'$; the second follows from the union bound, and the last follows from the above argument. This completes the proof.

B Proof of Lemma 2

We first prove the following claim:

Claim: *Fix any $\tilde{\epsilon} > 0$. Let $\hat{I}$ and $\hat{O}$ be two sets such that both $|\hat{I}|$ and $|\hat{O}|$ are in between αn and n for some $\alpha > 0$. For each $i \in \hat{I}$, let X_i be the number of objects in $\hat{O}$ for which $\xi_{io} \geq 1 - \tilde{\epsilon}$. For any $\delta < \tilde{\epsilon}$,*

$$\Pr\{\exists i \text{ with } X_i \leq \delta\,|\hat{O}|\} \to 0$$

as $n \to \infty$.

Proof. X_i follows a binomial distribution $B(|\hat{O}|, \tilde{\epsilon})$ (recall that ξ_{io} follows a uniform distribution with support $[0,1]$). Hence,

$$
\begin{aligned}
\Pr\{\exists i \text{ with } X_i \leq \delta|\hat{O}|\} \leq &\quad \textstyle\sum_{i\in\hat{I}}\Pr\{X_i \leq \delta|\hat{O}|\} \\
= &\quad |\hat{I}|\Pr\{X_i \leq \delta|\hat{O}|\} \\
\leq &\quad |\hat{I}|\tfrac{1}{2}\exp\left(-2\tfrac{(|\hat{O}|\tilde{\epsilon}-\delta|\hat{O}|)^2}{|\hat{O}|}\right) \\
= &\quad \frac{|\hat{I}|}{2\exp\left(2(\tilde{\epsilon}-\delta)^2|\hat{O}|\right)} \to 0
\end{aligned}
$$

where the first inequality is by the union bound while the second equality is by Hoeffding's inequality. □

Proof of Lemma 2. Note that for $\epsilon > 0$ so small that for each $k = 1, \dots, K-1$, $U(u_k, 1) - \epsilon > U(u_{k+1}, 1)$, objects in $O_{\geq k}$ that yield a payoff greater than $U(u_k, 1) - \epsilon$ can only be in O_k. Hence, the first part of the Lemma implies the second part, so we prove the first part.

Let us fix $\epsilon > 0$. By the continuity of $U(u_k, \cdot)$, there exists $\tilde{\epsilon} > 0$ such that $U(u_k, 1-\tilde{\epsilon}) > U(u_k, 1) - \epsilon$. By the above claim, with $\hat{I} := I$ and $\hat{O} := O_k$, there exists $\delta < \tilde{\epsilon}$ such that with probability going to 1 as $n \to \infty$, all individuals in I have at least $\delta|O_k|$ objects o's in O_k for which $\xi_{io} > 1 - \tilde{\epsilon}$. By our choice of $\tilde{\epsilon}$, the payoffs that individuals enjoy for these objects must be higher than $U(u_k, 1) - \epsilon$. This implies that with probability going to 1, for every individual in I, his $\delta|O_k|$ most favorite objects in $O_{\geq k}$ yield a payoff greater than $U(u_k, 1) - \epsilon$, as claimed. $\square$

C Proof of Theorem 3

C.1 Preliminary Results

We first perform preliminary analysis on TTC, which will prove useful for the proof (which is contained in the next subsection). The important part of this analysis concerns the number of objects assigned via long cycles in TTC. This analysis requires delving deeply into stochastic (more precisely Markovian) structure of the number of the objects/agents assigned at any given round of TTC. And since this involves the setting up of the "random mapping" framework and is quite involved, we separate that result out as a new section in Appendix D. Here, we simply state the result developed from that section, that will be necessary for our proof.

To begin, define a random set:

$$\hat{O} := \{o \in O_1 | o \text{ is assigned in TTC via long cycles}\}.$$

Appendix D establishes the following result.

Lemma 4. *There exist* $\gamma > 0, \delta > 0, N > 0$ *s.t.*

$$\Pr\left\{\frac{|\hat{O}|}{n} > \delta\right\} > \gamma,$$

for all $n > N$.

PROOF. See Appendix D. □

For the next result, define

$$I_2 := \{i \in I | TTC(i) \in O_2\}$$

to be the (random) set of agents who are assigned under TTC to objects in O_2. We next establish that any randomly selected (unmatched) pair from $\hat{O}$ and I_2 forms an ϵ-block with positive probability for sufficiently small $\epsilon > 0$.

LEMMA 5. *There exist $\varepsilon > 0, \zeta > 0$ such that, for all $n > N$, for any $\epsilon \in [0, \varepsilon)$,*

$$\Pr\left[\eta_{jo} \geq \eta_{TTC(o)o} + \epsilon \,\middle|\, o \in \hat{O}, j \in I_2\right] > \zeta.$$

PROOF. Note first that since there are large common value differences, if $o \in \hat{O} \subset O_1$ and $j \in I_2$, it must be that o does not point to j in the cycle to which o belongs under TTC (otherwise, if j is part of the cycle in which o is cleared, since $o \in O_1$, this means that j must be pointing to an object in O_1 when she is cleared, which is a contradiction with $j \in I_2$). Note also that j is still in the market when o is cleared.

Define $E_1 := \{\eta_{jo} \geq \eta_{TTC(o)o}\} \wedge \{o \in \hat{O}\} \wedge \{j \in I_2\}$ and $E_2 := \{\eta_{jo} \leq \eta_{TTC(o)o}\} \wedge \{o \in \hat{O}\} \wedge \{j \in I_2\}$. We first show that $\Pr E_1 = \Pr E_2$.

Assume that under the realizations $\boldsymbol{\xi} := (\xi_{io})_{io}$ and $\boldsymbol{\eta} := (\eta_{io})_{io}$ event E_1 is true. Define $\hat{\boldsymbol{\eta}} := (\hat{\eta}_{io})_{io}$ where $\hat{\eta}_{jo} := \eta_{TTC(o)o}$ and $\hat{\eta}_{TTC(o)o} := \eta_{jo}$ – while $\hat{\boldsymbol{\eta}}$ and $\boldsymbol{\eta}$ coincide otherwise. It is easily checked that under the realizations $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\eta}}$, event E_2 is true. Indeed, that $\{\hat{\eta}_{jo} \leq \hat{\eta}_{TTC(o)o}\}$ holds true is trivial. Now, since, as we already claimed, under the realizations $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, j and $TTC(o)$ are never pointed by o, when j and $TTC(o)$ are switched in o's priorities, by definition of TTC, o still belongs to the same cycle and, hence, TTC runs exactly in the same way. This shows that $\{o \in \hat{O}\} \wedge \{j \in I_2\}$ also holds true under the realizations $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\eta}}$,

Given that $\Pr(\boldsymbol{\xi}, \boldsymbol{\eta}) = \Pr(\boldsymbol{\xi}, \hat{\boldsymbol{\eta}})$, we get that $\Pr E_1 = \Pr E_2$.

Next, let $E_\epsilon := \{\eta_{jo} \geq \eta_{TTC(o)o} + \epsilon\}$. Note that

$$\cup_{\epsilon > 0} E_\epsilon = \{\eta_{jo} > \eta_{TTC(o)o}\} =: E.$$

Since the distribution $\Pr[\cdot]$ of η_{io} has no atom, $\Pr\left[\cdot \,\middle|\, o \in \hat{O}, j \in I_2\right]$ has no atom as well ($\Pr(\eta_{jo} = \eta) = 0 \Rightarrow \Pr(\eta_{jo} = \eta \,\middle|\, o \in \hat{O}, j \in I_2) = 0$). Thus, we must have

$$\Pr\left[E \,\middle|\, o \in \hat{O}, j \in I_2\right] = \Pr\left[\{\eta_{jo} \geq \eta_{TTC(o)o}\} \,\middle|\, o \in \hat{O}, j \in I_2\right] = \frac{1}{2}.$$

Since E_ϵ is increasing when ϵ decreases, combining the above, we get[29]

$$\lim_{\epsilon\to 0}\Pr\left[E_\epsilon \middle| o\in\hat{O}, j\in I_2\right] = \Pr\left[\cup_{\epsilon>0}E_\epsilon \middle| o\in\hat{O}, j\in I_2\right] = \Pr\left[E \middle| o\in\hat{O}, j\in I_2\right] = \frac{1}{2}.$$

Thus, one can fix $\delta\in(0,1/2)$ arbitrarily close to 0 and find $\varepsilon>0$ so that for any $\epsilon\in(0,\varepsilon)$, $\Pr\left[E_\epsilon \middle| o\in\hat{O}, j\in I_2\right] \geq \frac{1}{2}-\delta>0$. □

Corollary 4. *For any $\epsilon>0$ sufficiently small, there exist $\zeta>0, N>0$ such that, for all $n>N$,*

$$\mathbb{E}\left[\frac{|\hat{I}_2(o)|}{n} \middle| o\in\hat{O}\right] \geq x_2\zeta$$

Proof. Then, for any ϵ sufficiently small, we have $\zeta>0$ and $N>0$ such that

$$\begin{aligned}
\mathbb{E}\left[|\hat{I}_2(o)| \middle| o\in\hat{O}\right] &= \mathbb{E}\left[\sum_{i\in I_2}\mathbf{1}_{\{\eta_{io}>\eta_{TTC(o)o}+\epsilon\}} \middle| o\in\hat{O}\right]\\
&= \mathbb{E}_{I_2}\left(\mathbb{E}\left[\sum_{i\in I_2}\mathbf{1}_{\{\eta_{io}>\eta_{TTC(o)o}+\epsilon\}} \middle| o\in\hat{O}, I_2\right]\right)\\
&= \mathbb{E}_{I_2}\left(\sum_{i\in I_2}\mathbb{E}\left[\mathbf{1}_{\{\eta_{io}>\eta_{TTC(o)o}+\epsilon\}} \middle| o\in\hat{O}, I_2, i\in I_2\right]\right)\\
&= \mathbb{E}_{I_2}\left(x_2 n\mathbb{E}\left[\mathbf{1}_{\{\eta_{io}>\eta_{TTC(o)o}+\epsilon\}} \middle| o\in\hat{O}, I_2, i\in I_2\right]\right)\\
&= x_2 n\left(\mathbb{E}\left[\mathbf{1}_{\{\eta_{io}>\eta_{TTC(o)o}+\epsilon\}} \middle| o\in\hat{O}, i\in I_2\right]\right)\\
&= x_2 n\Pr(\eta_{io}>\eta_{TTC(o)o}+\epsilon \,\middle|\, o\in\hat{O}, i\in I_2\,)\\
&= x_2 n\Pr(\eta_{io}\geq\eta_{TTC(o)o}+\epsilon \,\middle|\, o\in\hat{O}, i\in I_2\,)\\
&\geq x_2\zeta n,
\end{aligned}$$

for all $n>N$. □

C.2 Proof of Theorem 3

Proof. The theorem follows from Lemma 4 and Corollary 4. The former implies that as the economy grows, the number of objects assigned via long cycles remain significant. The

[29] Recall the following property. Let $\{E_n\}_n$ be an increasing sequence of events. Let $E:=\cup_n E_n$ be the limit of $\{E_n\}_n$. Then: $\Pr(E)=\lim_{n\to\infty}\Pr(E_n)$.

latter implies that each of such object finds many agents assigned by TTC to O_2 desirable for forming ϵ-blocks. More precisely, for any sufficiently small $\epsilon \in (0, U(u_1^0, 0) - U(u_2^0, 1))$, we get that, for any large n,

$$
\begin{aligned}
\mathbb{E}\left[\frac{|J_\epsilon(TTC)|}{n(n-1)}\right] &\geq \mathbb{E}\left[\sum_{o\in\hat{O}} \frac{|\hat{I}_2(o)|}{n(n-1)}\right] \\
&\geq \Pr\{|\hat{O}| \geq \delta n\}\mathbb{E}\left[\sum_{o\in\hat{O}} \frac{|\hat{I}_2(o)|}{n(n-1)} \Big| |\hat{O}| \geq \delta n\right] \\
&\geq \gamma\mathbb{E}\left(\mathbb{E}\left[\sum_{o\in\hat{O}} \frac{|\hat{I}_2(o)|}{n(n-1)} \Big| |\hat{O}| \geq \delta n, \hat{O}\right]\right) \\
&= \gamma\mathbb{E}\left(\sum_{o\in\hat{O}} \mathbb{E}\left[\frac{|\hat{I}_2(o)|}{n(n-1)} \Big| |\hat{O}| \geq \delta n, \hat{O}, o \in \hat{O}\right]\right) \\
&\geq \gamma\delta n\mathbb{E}\left[\frac{|\hat{I}_2(o)|}{n(n-1)} \Big| o \in \hat{O}\right] \\
&\geq \gamma\delta\mathbb{E}\left[\frac{|\hat{I}_2(o)|}{n} \Big| o \in \hat{O}\right]\gamma \geq \delta\zeta x_2 > 0.
\end{aligned}
$$

□

D Proof of Lemma 4: Random Structure of TTC

In this section, we provide an analysis of TTC in our random environment. Our ultimate goal is to prove Lemma 4. For our purpose, it is sufficient to consider the TTC assignment arising from the market consisting of the agents I and the objects O_1 in the top tier (recall that, irrespective of the realizations of the idiosyncratic values, all agents prefer every object in O_1 to any object in O_2). Hence, we shall simply consider an **unbalanced market** consisting of a set I of agents and a set O of objects such that (1) *the preferences of each side with respect to the other side are drawn i.i.d. uniformly,* and (2) *both* $|O|$ *and* $|I| - |O|$ *increase in the order of* $|O|$*, as the market size* $|O|$ *grows to infinity.* The analysis of this market requires a preliminary result on bipartite random mapping.

D.0.1 Preliminaries

Here, we develop a couple of preliminary results that we shall later invoke. Throughout, we shall consider two finite sets I and O, with cardinalities $|I| = n, |O| = o$.

Number of Spanning Rooted Forests. A **rooted tree** is a connected directed bipartite digraph where all vertices have out-degree 1 except the root which has out-degree 0.[30] A **rooted forest** is a bipartite graph which consists of a collection of disjoint rooted trees. A **spanning rooted forest over** $I \cup O$ is a forest comprising vertices $I \cup O$. From now on, a spanning forest will be understood as being over $I \cup O$. We will be using the following result.

Lemma 6 (Jin and Liu (2004)). *Let $V_1 \subset I$ and $V_2 \subset O$ where $|V_1| = \ell$ and $|V_2| = k$. The number of spanning rooted forests having k roots in V_1 and ℓ roots in V_2 is $f(n, o, k, \ell) := o^{n-k-1} n^{o-\ell-1}(\ell n + ko - k\ell)$.*

Random Bipartite Mapping. We now consider arbitrary mappings, $g : I \to O$ and $h : O \to I$, defined over our finite sets I and O. Note that such mappings naturally induces bipartite digraphs with vertices $I \cup O$ and directed edges with the number of outgoing edges equal to the number of vertices, one for each vertex. In this digraph, $i \in I$ points to $g(i) \in O$ while $o \in O$ points to $h(o) \in I$. A **random bipartite mapping** selects a composite map $h \circ g$ uniformly from a set $\mathcal{H} \times \mathcal{G} = I^O \times O^I$ of all bipartite mappings. Note that a random bipartite mapping induces a random bipartite digraph consisting of vertices $I \cup O$ and directed edges emanating from vertices, one for each vertex. We say that a vertex in a digraph is **cyclic** if it is in a cycle. The following lemma states the number of cyclic vertices in a random bipartite digraph induced by a random bipartite mapping.

Lemma 7 (Jaworski (1985), Corollary 3). *The number q of the cyclic vertices in a random bipartite digraph induced by a random bipartite mapping $g : I \to O$ and $h : O \to I$ has an expected value of*

$$\mathbb{E}[q] := 2 \sum_{i=1}^{o} \frac{(o)_i (n)_i}{o^i n^i},$$

[30] Sometimes, a tree is defined as an acyclic undirected connected graph. In such a case, a tree is rooted when we name one of its vertex a "root." Starting from such a rooted tree, if all edges now have a direction leading toward the root, then the out-degree of any vertex (except the root) is 1. So the two definitions are actually equivalent.

where $(x)_j := x(x-1)\cdots(x-j-1)$.

For the next result, consider agents I' and objects O' such that $|I'| = |O'| = m > 0$. We say a mapping $f = h \circ g$ is a **bipartite bijection**, if $g : I' \to O'$ and $h : O' \to I'$ are both bijections. Note that a bipartite bijection consists of disjoint cycles. A **random bipartite bijection** is a (uniform) random selection of a bipartite bijection from the set of all bipartite bijections. The following result will prove useful for a later analysis.

LEMMA 8. *Fix sets I' and O' with $|I'| = |O'| = m > 0$, and a subset $K \subset I' \cup O'$, containing $a \geq 0$ vertices in I' and $b \geq 0$ vertices in O'. The probability that each cycle in a random bipartite bijection contains at least one vertex from K is*

$$\frac{a+b}{m} - \frac{ab}{m^2}.$$

PROOF. We shall invoke Lovasz (1979) Exercise 3.6, which establishes that the probability that each cycle of a random permutation[31] of a finite set $|X|$ contains at least one element of a set $K \subset X$ is $|K|/|X|$.

To this end, observe first that a bipartite bijection $h \circ g$ induces a permutation of set I'. Thus, a random bipartite bijection defined over $I' \times O'$ induces a random permutation of I'. The probability that each cycle of the randomly selected bipartite bijection contains at least one vertex in K is identical to the probability that each cycle of the induced permutation of I' contains at least one of $a + Z$ vertices, where Z is the (random) number of vertices in $I' \setminus K$ that point to $K \cap O'$. For any $\max\{b-a, 0\} \leq z \leq \min\{m-a, b\}$,

$$\Pr\{Z = z\} = \frac{\binom{m-a}{z}\binom{a}{b-z}}{\binom{m}{b}}.$$

The above formula can be understood as follows. $\binom{m-a}{z}\binom{a}{b-z}$ is the number of ways one can choose z vertices from $I' \setminus K$ and $b - x$ vertices from $K \cap I'$. Thus, the total number of bipartite bijections having exactly z vertices in $I' \setminus K$ that point to $K \cap O'$ is $\binom{m-a}{z}\binom{a}{b-z}v$ where v is the total number of bipartite bijections in which the z vertices chosen from $I' \setminus K$ point to vertices in $K \cap O'$ and the $b - z$ vertices chosen from $K \cap I'$ point to vertices in $K \cap O'$. Note that v is precisely the number of bipartite bijections in which any b vertices arbitrarily chosen from I' point to vertices in $K \cap O'$. Hence, the total number of bipartite bijections having b vertices in I' pointing to $K \cap O'$ is $\binom{m}{b}v$. Thus, we get the above formula.

[31]Formally, a permutation of X is a bijection $f : X \to X$. A random permutation chooses randomly a permutation f from the set of all possible permutations.

Applying the earlier result, the desired probability is

$$
\begin{aligned}
&\sum_{x=\max\{b-a,0\}}^{\min\{m-a,b\}} \Pr\{Z=z\}\frac{a+z}{m} \\
=&\frac{a}{m}+\sum_{z=\max\{b-a,0\}}^{\min\{m-a,b\}} \Pr\{Z=z\}\frac{z}{m} \\
=&\frac{a}{m}+\sum_{x=\max\{b-a,0\}}^{\min\{m-a,b\}} \frac{\binom{m-a}{z}\binom{a}{b-z}}{\binom{m}{b}}\left(\frac{z}{m}\right) \\
=&\frac{a}{m}+\left(\frac{m-a}{m\binom{m}{b}}\right)\sum_{z=\max\{b-a,1\}}^{\min\{m-a,b\}} \binom{a}{b-z}\binom{m-a-1}{z-1} \\
=&\frac{a}{m}+\left(\frac{m-a}{m\binom{m}{b}}\right)\binom{m-1}{b-1} \\
=&\frac{a}{m}+\frac{b(m-a)}{m^2} \\
=&\frac{a+b}{m}-\frac{ab}{m^2},
\end{aligned}
$$

where the fourth equality follows from the Vandermonde's identity. □

D.1 Markov Chain Property of TTC

Again consider a TTC in an unbalanced market with agents I and objects O. As is well known, TTC assigns agents to objects via cycles formed recursively in multiple rounds. We shall call a cycle of length 2—an agent points to an object, which in turn points to the original agent—a **short-cycle**. Any cycles of length greater than 2 shall be called **long-cycles**. Our aim is to prove that the number of agents assigned via long-cycles in TTC grows in the same order n as the size of the market n grows. The difficulty with proving this result stems from the fact that the preferences of the agents and objects remaining after the first round of TTC need not be uniform, with their distributions affected nontrivially by the realized event of the first round TTC, and the nature of the conditioning is difficult to analyze in the large market. Our approach is to prove that, even though the exact composition of cycles are subject to the conditioning issue, the number of agents assigned in each round follows a Markov chain, and is thus free from the conditioning issue. We then combine this observation with the bound we shall establish on the number of agents assigned via short-cycles, to produce a desired result.

We shall begin with the Markov Chain result. This result parallels the corresponding result by Frieze and Pittel (1995) on the Shapley-Scarf version of TTC. The difference between the two versions of TTC is not trivial, so their proofs do not carry over.

Theorem 8. *Suppose any round of TTC begins with n agents and o objects remaining in the market. Then, the probability that there are $m \leq \min\{o, n\}$ agents assigned at the end of that round is*

$$p_{n,o;m} = \left(\frac{m}{(on)^{m+1}}\right)\left(\frac{n!}{(n-m)!}\right)\left(\frac{o!}{(o-m)!}\right)(o+n-m).$$

Thus, denoting n_i and o_i the number of individuals and objects remaining in the market at any round i, the random sequence (n_i, o_i) is a Markov chain.

D.2 Proof of Theorem 8

We begin by noting that TTC induces a random sequence of spanning rooted forests. Indeed, one could see the beginning of the first round of TTC as a situation where we have the trivial forest consisting of $|I| + |O|$ trees with isolated vertices. Within this step each vertex in I will randomly point to a vertex in O and each vertex in O will randomly point to a vertex in I. Note that once we delete the realized cycles, we again get a spanning rooted forest. So we can think again of the beginning of the second round of TTC as a situation where we start with a spanning rooted forest where the agents and objects remaining from the first round form this spanning rooted forest, where the roots consist of those agents and objects that had pointed to the entities that were cleared via cycles. Here again objects that are roots randomly point to a remaining individual and individuals that are roots randomly point to a remaining object. Once cycles are cleared we again obtain a forest and the process goes on like this.

Formally, the random sequence of forests, $F_1, F_2, ...$ is defined as follows. First, we let F_1 be a trivial unique forest consisting of $|I| + |O|$ trees with isolated vertices, forming their own roots. For any $i = 2, ...$, we first create a random directed edge from each root of F_{i-1} to a vertex on the other side, and then delete the resulting cycles (these are the agents and objects assigned in around $i-1$) and F_i is defined to be the resulting rooted forest. For any rooted forest F_i, let $N_i = I_i \cup O_i$ be its vertex set and $k_i = (k_i^I, k_i^O)$ be the vector denoting the numbers of roots on both sides, and use (N_i, k_i) to summarize this information. And let $\mathcal{F}_{N_i,k_i}$ denote the set of all rooted forests having N_i as the vertex set and k_i as the vector of its root numbers.

LEMMA 9. *Given $(N_j, k_j), j = 1, ..., i$, every (rooted) forest of $\mathcal{F}_{N_i,k_i}$ is equally likely.*

PROOF. We prove this result by induction on i. Since for $i = 1$, by construction, the trivial forest is the unique forest which can occur, this is trivially true for $i = 1$. Fix $i \geq 2$, and assume our statement is true for $i - 1$.

Fix $N_i = I_i \cup O_i \subset N_{i+1} = I_{i+1} \cup O_{i+1}$, and k_i and k_{i+1}. For each forest $F \in \mathcal{F}_{N_{i+1},k_{i+1}}$, we consider a possible pair (F', ϕ) that could have given rise to F, where $F' \in \mathcal{F}_{N_i,k_i}$ and ϕ maps the roots of F' in I_i to its vertices in O_i as well as the roots of F' in O_i to its vertices in I_i. In words, such a pair (F', ϕ) corresponds to a set N_i of agents and objects remaining at the end of TTC round i, of which k_i^I agents of I_i and k_i^O objects have lost their favorite parties (and thus must their repoint to new partners in N_i at TTC round $i + 1$), and the way in which they repoint to the new partners at TTC round $i+1$ causes a new forest F to emerge at the end of TTC round $i + 1$. There are typically multiple such pairs that could give rise to F.

We start by showing that each forest $F \in \mathcal{F}_{N_{i+1},k_{i+1}}$ arises from the same number of such pairs—i.e., that the number of pairs (F', ϕ), $F' \in \mathcal{F}_{N_i,k_i}$, causing F to arise does not depend on the particular $F \in \mathcal{F}_{N_{i+1},k_{i+1}}$. To this end, for any given $F \in \mathcal{F}_{N_{i+1},k_{i+1}}$, we construct all such pairs by choosing a quadruplet (a, b, c, d) of four non-negative integers with $a + c = k_i^I$ and $c + d = k_i^O$,

(i) choosing c old roots from I_{i+1}, and similarly, d old roots from O_{i+1},

(ii) choosing a old roots from $I_i \backslash I_{i+1}$ and similarly, b old roots from $O_i \backslash O_{i+1}$,

(iii) choosing a partition into cycles of $N_i \backslash N_{i+1}$, each cycle of which contains at least one old root from (ii),

(iv) choosing a mapping of the λ new roots to $N_i \backslash N_{i+1}$ satisfying the bipartite graph constraint.

Clearly, the number of pairs (F', ϕ), $F' \in \mathcal{F}_{N_i,k_i}$, satisfying the above restrictions depends only on $|I_i|$, $|O_i|$, k_i, k_{i+1}, and $|N_{i+1}| - |N_i|$. We denote the number of such pairs by $\beta(|I_i|, |O_i|, k_i; |I_{i+1}| - |I_i|, k_{i+1})$. Let $\phi_i = (\phi_i^I, \phi_i^O)$ where ϕ_i^I is the random mapping from the roots of F_i in I_i to O_i and ϕ_i^O is the random mapping from the roots of F_i in O_i to I_i. Let $\phi = (\phi^I, \phi^O)$ be a generic mapping of that sort. Since, conditioned on $F_i = F'$, the mappings ϕ_i^I and ϕ_i^O are uniform, we get (where k_i^I and k_i^O denote the number of roots of F_i in I_i and O_i respectively)

$$\Pr(F_{i+1} = F | F_i = F') = \frac{1}{|O_i|^{k_i^I}} \frac{1}{|I_i|^{k_i^O}} \sum_{\phi} \Pr(F_{i+1} = F | F_i = F', \phi_i = \phi), \tag{7}$$

where we used the fact that the conditional probability in the sum above is 1 or 0, depending upon whether the forest F arises from the pair $(F^{'}, \phi)$ or not.

Therefore, we obtain

$$\begin{aligned}
&\Pr(F_{i+1} = F | (N_1, k_1), ..., (N_i, k_i)) \\
=& \sum_{F' \in \mathcal{F}_{N_i,k_i}} \Pr(F_{i+1} = F, F_i = F' | (N_1, k_1), ..., (N_i, k_i)) \\
=& \sum_{F' \in \mathcal{F}_{N_i,k_i}} \Pr(F_{i+1} = F | (N_1, k_1), ..., (N_i, k_i), F_i = F') \Pr(F_i = F' | (N_1, k_1), ..., (N_i, k_i)) \\
=& \frac{1}{|\mathcal{F}_{N_i,k_i}|} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \Pr(F_{i+1} = F | F_i = F') \\
=& \frac{1}{|\mathcal{F}_{N_i,k_i}|} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \frac{1}{|O_i|^{k_i^I}} \frac{1}{|I_i|^{k_i^O}} \sum_{\phi} \Pr(F_{i+1} = F | F_i = F', \phi_i = \phi) \\
=& \frac{1}{|\mathcal{F}_{N_i,k_i}|} \frac{1}{|O_i|^{k_i^I}} \frac{1}{|I_i|^{k_i^O}} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \sum_{\phi} \Pr(F_{i+1} = F | F_i = F', \phi_i = \phi) \\
=& \frac{1}{|\mathcal{F}_{N_i,k_i}|} \frac{1}{|O_i|^{k_i^I}} \frac{1}{|I_i|^{k_i^O}} \beta(|I_i|, |O_i|, k_i; |I_{i+1}| - |I_i|, k_{i+1}),
\end{aligned} \tag{8}$$

where the third equality follows from the induction hypothesis and the Markov property of $\{F_j\}$, the fourth follows from (7), and the last follows from the definition of β. Note this probability is independent of $F \in \mathcal{F}_{N_{i+1},k_{i+1}}$. Hence,

$$\begin{aligned}
&\Pr(F_{i+1} = F | (N_1, k_1), ..., (N_i, k_i), (N_{i+1}, k_{i+1})) \\
=& \frac{\Pr(F_{i+1} = F | (N_1, k_1), ..., (N_i, k_i))}{\Pr(F_{i+1} \in \mathcal{F}_{N_{i+1},k_{i+1}} | (N_1, k_1), ..., (N_i, k_i))} \\
=& \frac{\Pr(F_{i+1} = F | (N_1, k_1), ..., (N_i, k_i))}{\sum_{\tilde{F} \in \mathcal{F}_{N_{i+1},k_{i+1}}} \Pr(F_{i+1} = \tilde{F} | (N_1, k_1), ..., (N_i, k_i))} = \frac{1}{|\mathcal{F}_{N_{i+1},k_{i+1}}|},
\end{aligned} \tag{9}$$

which proves that, given $(N_j, k_j), j = 1, ..., i$, every rooted forest of $\mathcal{F}_{N_i,k_i}$ is equally likely. $\square$

The next lemma then follows easily.

LEMMA 10. *Random sequence* (N_i, k_i) *forms a Markov chain.*

PROOF.

$$\Pr((N_{i+1}, k_{i+1})|(N_1, k_1), ..., (N_i, k_i)) = \sum_{F \in \mathcal{F}_{N_{i+1},k_{i+1}}} \Pr(F_{i+1} = F|(N_1, k_1), ..., (N_i, k_i))$$

$$= \sum_{F \in \mathcal{F}_{N_{i+1},k_{i+1}}} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \Pr(F_{i+1} = F, F_i = F'|(N_1, k_1), ..., (N_i, k_i))$$

$$= \sum_{F \in \mathcal{F}_{N_{i+1},k_{i+1}}} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \Pr(F_{i+1} = F|(N_1, k_1), ..., (N_i, k_i), F_i = F') \Pr(F_i = F'|(N_1, k_1), ..., (N_i, k_i))$$

$$= \sum_{F \in \mathcal{F}_{N_{i+1},k_{i+1}}} \sum_{F' \in \mathcal{F}_{N_i,k_i}} \frac{1}{|\mathcal{F}_{N_i,k_i}|} \Pr(F_{i+1} = F|F_i = F')$$

$$= \sum_{F \in \mathcal{F}_{N_{i+1},k_{i+1}}} \frac{1}{|\mathcal{F}_{N_i,k_i}|} \frac{1}{|O_i|^{k_i^I}} \frac{1}{|I_i|^{k_i^O}} \beta(|I_i|, |O_i|, k_i; |I_{i+1}| - |I_i|, k_{i+1}),$$

where third equality follows from Lemma 9, specifically (9), and the Markov property of $\{F_i\}$, and the last equality follows from (8). Observing that the conditional probability depends only on (N_{i+1}, k_{i+1}) and (N_i, k_i), the Markov chain property is established. □

The proof of Lemma 10 reveals in fact that the conditional probability of (N_{i+1}, k_{i+1}) depends on (N_i, k_i) only through its cardinalities $(|I_i|, |O_i|)$, leading to the following conclusion.

COROLLARY 5. *Random sequence* $\{(|I_i|, |O_i|, k_i^I, k_i^O)\}$ *forms a Markov chain.*

PROOF. Let $n_i := |I_i|$ and $o_i := |O_i|$. By symmetry, given $(n_1, o_1, k_1^I, k_1^O), ..., (n_i, o_i, k_i^I, k_i^O)$, the set (I_i, O_i) is chosen uniformly at random among all the $\binom{n}{n_i}\binom{o}{o_i}$ possible sets. Hence,

$$
\begin{aligned}
&\Pr((n_{i+1}, o_{i+1}, k_{i+1}^I, k_{i+1}^O)|(n_1, o_1, k_1^I, k_1^O), ..., (n_i, o_i, k_i^I, k_i^O))\\
=&\sum_{(I_i,O_i):|I_i|=n_i,|O_i|=o_i} \Pr\{(n_{i+1}, o_{i+1}, k_{i+1}^I, k_{i+1}^O)|(n_1, o_1, k_1^I, k_1^O), ..., (n_i, o_i, k_i^I, k_i^O), (I_i, O_i)\}\\
&\qquad\qquad \times \Pr\left\{(I_i, O_i) \mid (n_1, o_1, k_1^I, k_1^O), ..., (n_i, o_i, k_i^I, k_i^O)\right\}\\
=&\left(\sum_{(I_i,O_i):|I_i|=n_i,|O_i|=o_i} \Pr\{(n_{i+1}, o_{i+1}, k_{i+1}^I, k_{i+1}^O)|(n_1, o_1, k_1^I, k_1^O), ..., (I_i, O_i, k_i^I, k_i^O)\}\right) \frac{1}{\binom{n}{n_i}\binom{o}{o_i}}\\
=&\left(\sum_{\substack{(I_i,O_i):|I_i|=n_i,|O_i|=o_i\\(I_{i+1},O_{i+1}):|I_{i+1}|=n_{i+1},|O_{i+1}|=o_{i+1}}} \Pr\{(I_{i+1}, O_{i+1}, k_{i+1}^I, k_{i+1}^O)|(n_1, o_1, k_1^I, k_1^O), ..., (I_i, O_i, k_i^I, k_i^O)\}\right)\\
&\qquad \times \frac{1}{\binom{n}{n_i}\binom{o}{o_i}}\\
=&\frac{1}{\binom{n}{n_i}\binom{o}{o_i}} \sum_{\substack{(I_i,O_i):|I_i|=n_i,|O_i|=o_i\\(I_{i+1},O_{i+1}):|I_{i+1}|=n_{i+1},|O_{i+1}|=o_{i+1}}} \Pr\{(I_{i+1}, O_{i+1}, k_{i+1}^I, k_{i+1}^O)|(I_i, O_i, k_i^I, k_i^O)\},
\end{aligned}
$$

where the second equality follows from the above reasoning and the last equality follows from the Markov property of $\{(I_i, O_i, k_i^I, k_i^O)\}$. The proof is complete by the fact that the last line, as shown in the proof of Lemma 10, depends only on $(n_{i+1}, o_{i+1}, k_{i+1}^I, k_{i+1}^O), (n_i, o_i, k_i^I, k_i^O))$. $\square$

We are now in a position to obtain our main result:

LEMMA 11. *The random sequence* (n_i, o_i) *is a Markov chain, with transition probability given by*

$$
\begin{aligned}
p_{n,o;m} &:= \Pr\{n_i - n_{i+1} = o_i - o_{i+1} = m | n_i = n, o_i = o\}\\
&= \left(\frac{m}{(on)^{m+1}}\right)\left(\frac{n!}{(n-m)!}\right)\left(\frac{o!}{(o-m)!}\right)(o+n-m).
\end{aligned}
$$

PROOF. We first compute the probability of transition from (n_i, o_i, k_i^I, k_i^O) such that $k_i^I + k_i^O = \kappa$ to $(n_{i+1}, o_{i+1}, k_{i+1}^I, k_{i+1}^O)$ such that $k_{i+1}^I = \lambda^I$ and $k_{i+1}^O = \lambda^O$:

$$
\begin{aligned}
&\mathbf{P}(n, o, \kappa; m, \lambda^I, \lambda^O)\\
:=&\Pr\left\{n_i - n_{i+1} = o_i - o_{i+1} = m, k_{i+1}^I = \lambda^I, k_{i+1}^O = \lambda^O \mid n_i = n, o_i = o, k_i^I + k_i^O = \kappa\right\}.
\end{aligned}
$$

This will be computed as a fraction $\frac{\Theta}{\Upsilon}$. The denominator Υ counts the number of rooted forests in the bipartite digraph with k_i^I roots in I_i and k_i^O roots in O_i, multiplied by the ways in which k_i^I roots of I_i could point to O_i and k_i^O roots of O_i could point to I_i.[32] Hence, letting $f(n,o,k^I,k^O)$ denote the number of rooted forests in a bipartite graph (with n and o vertices on both sides) containing k^I and k^O roots on both sides.

$$\begin{aligned}\Upsilon =& \sum_{(k^I,k^O):k^I+k^O=\kappa} o^{k^I} n^{k^O} f(n,o,k^I,k^O)\\ =& \sum_{k^I+k^O=\kappa} o^{k^I} n^{k^O} \binom{n}{k^I}\binom{o}{k^O} o^{n-k^I-1} n^{o-k^O-1}(nk^O+ok^I-k^Ik^O)\\ =& \sum_{k^I+k^O=\kappa} \binom{n}{k^I}\binom{o}{k^O} o^{n-1} n^{o-1}(nk^O+ok^I-k^Ik^O)\\ =& o^n n^o \left(2\binom{n+o-1}{\kappa-1}-\binom{n+o-2}{\kappa-2}\right).\end{aligned}$$

The first equality follows from the fact that there are $o^{k^I} n^{k^O}$ ways in which k^I roots in I_i point to O_i and k^O roots in O_i could point to I_i. The second equality follows from Lemma 6. The last uses Vandermonde's identity.

The numerator Θ counts the number of ways in which m agents are chosen from I_i and m objects are chosen from O_i to form a bipartite bijection each cycle of which contains at least one of κ old roots, and for each such choice, the number of ways in which the remaining vertices form a spanning rooted forest and the λ^I roots in I_{i+1} point to objects in $O_i \setminus O_{i+1}$ and λ^O roots in O_{i+1} point to agents in $O_i \setminus O_{i+1}$. To compute this, we first compute

$$\alpha(n,o,\kappa;m,\lambda^I,\lambda^O) = \sum_{(k^I,k^O):k^I+k^O=\kappa} \beta(n,o,k^I,k^O;m,\lambda^I,\lambda^O),$$

where β is defined in Lemma 9. In words, α counts, for any F with $n-m$ agents and $o-m$ objects and roots λ^I and λ^O on both sides, the total number of pairs (F',ϕ) that could have given rise to F, where F' has n agents and o objects with κ roots and ϕ maps the roots to the remaining vertices. Following the construction in the beginning of Lemma

[32] Given that we have $n_i = n$ individuals, $o_i = o$ objects and $k_i^I + k_i^O = \kappa$ roots at the beginning of step i under TTC, one may think of this as the total number of possible directed bipartite digraph one may obtain via TTC at the end of step i when we let k_i^I roots in I_i point to their remaining most favorite object and k_i^O roots in O_i point to their remaining most favorite individual.

9, the number of such pairs is computed as

$$
\begin{aligned}
&\alpha(n, o, \kappa; m, \lambda^I, \lambda^O) \\
:=& \sum_{a+b+c+d=\kappa} \binom{n-m}{c}\binom{o-m}{d}\binom{m}{a}\binom{m}{b}\left(\frac{a+b}{m}-\frac{ab}{m^2}\right)(m!)^2 m^{\lambda^I+\lambda^O} \\
=&(m!)^2 m^{\lambda^I+\lambda^O} \times \left(\sum_{a+b+c+d=\kappa} \binom{n-m}{c}\binom{o-m}{d}\binom{m-1}{a-1}\binom{m}{b}\right. \\
&\left. + \sum_{a+b+c+d=\kappa} \binom{n-m}{c}\binom{o-m}{d}\binom{m}{a}\binom{m-1}{b-1} - \sum_{a+b+c+d=\kappa} \binom{n-m}{c}\binom{o-m}{d}\binom{m-1}{a-1}\binom{m-1}{b-1}\right) \\
=&(m!)^2 m^{\lambda^I+\lambda^O}\left(2\binom{n+o-1}{\kappa-1}-\binom{n+o-2}{\kappa-2}\right).
\end{aligned}
$$

The first equality follows from Lemma 8, along with the fact that there are $(m!)^2$ possible bipartite bijections between $n-m$ agents and $o-m$ objects, and the fact that there are $m^{\lambda^I} m^{\lambda^O}$ ways in which new roots λ^I agents and λ^O objects) could have pointed to $2m$ cyclic vertices (m on the individuals' side and m on the objects' side), and the last equality follows from Vandermonde's identity.

The numerator Θ is now computed as:

$$
\begin{aligned}
\Theta &= \binom{n}{m}\binom{o}{m} f(n-m, o-m, \lambda^I, \lambda^O)\alpha(n, o, \kappa; m, \lambda^I, \lambda^O) \\
&= \binom{n}{m}\binom{o}{m} f(n-m, o-m, \lambda^I, \lambda^O)(m!)^2 m^{\lambda^I+\lambda^O}\left(2\binom{n+o-1}{\kappa-1}-\binom{n+o-2}{\kappa-2}\right). \\
&= \left(\frac{n!}{(n-m)!}\right)\left(\frac{o!}{(o-m)!}\right) m^{\lambda^I+\lambda^O} f(n-m, o-m, \lambda^I, \lambda^O)\left(2\binom{n+o-1}{\kappa-1}-\binom{n+o-2}{\kappa-2}\right).
\end{aligned}
$$

Collecting terms, let us compute

$$
\mathbf{P}(n, o, \kappa; m, \lambda^I, \lambda^O) = \frac{1}{o^n n^o}\left(\frac{n!}{(n-m)!}\right)\left(\frac{o!}{(o-m)!}\right) m^{\lambda^I+\lambda^O} f(n-m, o-m, \lambda^I, \lambda^O).
$$

A key observation is that this expression does not depend on κ, which implies that (n_i, o_i) forms a Markov chain.

Its transition probability can be derived by summing the expression over all possible (λ^I, λ^O)'s:

$$
p_{n,o;m} := \sum_{0 \le \lambda^I \le n-m, 0 \le \lambda^O \le o-m} P(n, o, \kappa; m, \lambda^I, \lambda^O).
$$

To this end, we obtain:

$$
\begin{aligned}
&\sum_{0\le\lambda^I\le n-m}\sum_{0\le\lambda^O\le o-m} m^{\lambda^I} m^{\lambda^O} f(n-m, o-m, \lambda^I, \lambda^O)\\
=&\sum_{0\le\lambda^I\le n-m}\sum_{0\le\lambda^O\le o-m} m^{\lambda^I} m^{\lambda^O} \binom{n-m}{\lambda^I}\binom{o-m}{\lambda^O}\times\\
&\quad (o-m)^{n-m-\lambda^I-1}(n-m)^{o-m-\lambda^O-1}((n-m)\lambda^O+(o-m)\lambda^I-\lambda^I\lambda^O)\\
=&m\left(\sum_{0\le\lambda^I\le n-m}\binom{n-m}{\lambda^I}m^{\lambda^I}(o-m)^{n-m-\lambda^I}\right)\left(\sum_{1\le\lambda^O\le o-m}\binom{o-m-1}{\lambda^O-1}m^{\lambda^O-1}(n-m)^{o-m-\lambda^O}\right)\\
&+m\left(\sum_{1\le\lambda^I\le n-m}\binom{n-m-1}{\lambda^I-1}m^{\lambda^I-1}(o-m)^{n-m-\lambda^I}\right)\left(\sum_{0\le\lambda^O\le o-m}\binom{o-m}{\lambda^O}m^{\lambda^O}(n-m)^{o-m-\lambda^O}\right)\\
&-m^2\left(\sum_{1\le\lambda^I\le n-m}\binom{n-m-1}{\lambda^I-1}m^{\lambda^I-1}(o-m)^{n-m-\lambda^I}\right)\left(\sum_{1\le\lambda^O\le o-m}\binom{o-m-1}{\lambda^O-1}m^{\lambda^O-1}(n-m)^{o-m-\lambda^O}\right)\\
=&mo^{n-m}n^{o-m-1}+mo^{n-m-1}n^{o-m}-m^2o^{n-m-1}n^{o-m-1}\\
=&mo^{n-m-1}n^{o-m-1}(n+o-m),
\end{aligned}
$$

where the first equality follows from Lemma 6, and the third follows from the Binomial Theorem.

Multiplying the term $\frac{1}{o^n n^o}\left(\frac{n!}{(n-m)!}\right)\left(\frac{o!}{(o-m)!}\right)$, we get the formula stated in the Lemma.
□

This last lemma concludes the proof of Theorem 8.

D.3 The Number of Objects Assigned via Short Cycles

We begin with the following question: If round k of TTC begins with a rooted forest F, what is the expected number of short-cycles that will form at the end of that round? We will show that, irrespective of F, this expectation is bounded by 2. To show this, we will make a couple of observations.

To begin, let n_k be the cardinality of the set I_k of individuals in our forest F and let o_k be the cardinality of O_k, the set of F's objects. And, let $A \subset I_k$ be the set of roots on the individuals side of our given forest F and let $B \subset O_k$ be the set of its roots on the objects side. Their cardinalities are a and b, respectively.

Now, observe that for any $(i,o) \in A \times B$, the probability that (i,o) forms a short-cycle is $\frac{1}{n_k}\frac{1}{o_k}$. For any $(i,o) \in (I_k \backslash A) \times B$, the probability that (i,o) forms a short-cycle is $\frac{1}{n_k}$ if i points to o and 0 otherwise. Similarly, for $(i,o) \in A \times (O_k \backslash B)$, the probability that (i,o) forms a short-cycle is $\frac{1}{o_k}$ if o points to i and 0 otherwise. Finally, for any $(i,o) \in (I_k \backslash A) \times (O_k \backslash B)$, the probability that (i,o) forms a short-cycle is 0 (by definition of a forest, i and o cannot be pointing to each other in the forest F). So, given the forest F, the expectation of the number S_k of short-cycles is

$$\begin{aligned}
\mathbb{E}\left[S_k | F_k = F\right] &= \mathbb{E}\left[\sum_{(i,o)\in I_k \times O_k} \mathbf{1}_{\{(i,o) \text{ is a short-cycle}\}} \middle| F_k = F\right] \\
&= \sum_{(i,o)\in I_k \times O_k} \mathbb{E}\left[\mathbf{1}_{\{(i,o) \text{ is a short-cycle}\}} | F_k = F\right] \\
&= \sum_{(i,o)\in A \times B} \mathbb{E}\left[\mathbf{1}_{\{(i,o) \text{ is a short-cycle}\}} | F_k = F\right] \\
&\quad + \sum_{(i,o)\in (I_k \backslash A) \times B} \mathbb{E}\left[\mathbf{1}_{\{(i,o) \text{ is a short-cycle}\}} | F_k = F\right] \\
&\quad + \sum_{(i,o)\in A \times (O_k \backslash B)} \mathbb{E}\left[\mathbf{1}_{\{(i,o) \text{ is a short-cycle}\}} | F_k = F\right] \\
&= \sum_{(i,o)\in A \times B} \Pr\{(i,o) \text{ is a short-cycle} | F_k = F\} \\
&\quad + \sum_{(i,o)\in (I_k \backslash A) \times B} \Pr\{(i,o) \text{ is a short-cycle} | F_k = F\} \\
&\quad + \sum_{(i,o)\in I_k \times (O_k \backslash B)} \Pr\{(i,o) \text{ is a short-cycle} | F_k = F\} \\
&\leq \frac{ab}{n_k o_k} + \frac{n_k - a}{n_k} + \frac{o_k - b}{o_k} \\
&= 2 - \frac{a o_k + b n_k - ab}{n_k o_k} \leq 2.
\end{aligned}$$

Observe that since $o_k \geq b$, the above term is smaller than 2. Thus, as claimed, we obtain the following result[33]

[33]Note that the bound is pretty tight: if the forest F has one root on each side and each node which is not a root points to the (unique) root on the opposite side, the expected number of short-cycles given F is $\frac{1}{n_k o_k} + \frac{n_k - 1}{n_k} + \frac{o_k - 1}{o_k} \to 2$ as $n_k, o_k \to \infty$. Thus, the conditional expectation of s_k is bounded by 2 and, asymptotically, this bound is tight. However, we can show, using a more involved computation, that the unconditional expectation of s_k is bounded by 1. The details of the computation are available upon request.

PROPOSITION 2. *If TTC round k begins with any forest F,*

$$\mathbb{E}\left[S_k \,|F_k + F\right] \leq 2.$$

Given that our upper bound holds for any forest F, we get the following corollary.

COROLLARY 6. *For any round k of TTC,* $\mathbb{E}\left[S_k\right] \leq 2$.

D.4 The Number of Objects Assigned via Long Cycles

Again consider the unbalanced market in which $|I| - |O|$ is in the same order of magnitude as $|O|$, and recall $n := |I|$ and $o := |O|$.

The Markov property established in Theorem 8 means that the number of agents and objects assigned in any TTC round depends only on the number of agents and objects that round begins with, regardless of how many rounds preceded that round and what happened in those rounds. Hence, the distribution of the (random) number M_k of objects that would be assigned in any round of TTC that begins with n_k agents and $o_k(\leq n_k)$ objects is the same as that in the first round of TTC when there are n_k agents and $o_k(\leq n_k)$. In particular, we can apply Lemma 7 to compute its expected value:[34]

$$\mathbb{E}[M_k \mid |O_k| = o_k] = \sum_{i=1}^{o} \frac{(o_k)_i (n_k)_i}{o_k^i n_k^i}.$$

We can make two observations: First, the expected number is increasing in o_k (and n_k) and goes to infinity as o_k (and n_k) increases. This can be seen easily by the fact that $\frac{k-l}{k}$ is increasing in k for any $k > l$. Second, given our assumption that $n_k \geq o_k$, there exists $\hat{o} \geq 1$[35] such that

$$\mathbb{E}[M_k \,|o_k] \geq 3 \text{ if } o_k \geq \hat{o}.$$

We are now ready to present the main result. Recall that $\hat{O}$ is the (random) set of objects that are assigned via long cycles in TTC.

THEOREM 9. $\mathbb{E}\left[\frac{|\hat{O}|}{|O|}\right] \geq \frac{1}{3} - \frac{\hat{o}+2}{3|O|}$.

[34]The number is half of that stated in Lemma 7 since the number of agents cleared in any round is precisely the half of the cyclic vertices in a random bipartite graph at the beginning of that round. Recall also that, by definition of TTC, together with our assumption that $o \leq n$, given the number of objects o_k, the number of individuals n_k is totally determined and is equal to $o_k + n - o$.

[35]One can check that $\hat{o} = 13$ works. In particular, if $n_k = o_k$, $\mathbb{E}[m_k \,|o_k] \geq 3$ if and only if $o_k \geq 13$.

PROOF. Consider the following sequence of random variables $\{\mathbb{E}(L_k\,|o_k\,)\}_{k=1}^{|O|}$ where o_k is the number of remaining objects at round k while L_k is the number of objects assigned at round k via long cycles. (Note both are random variables.) Thus, $o_1 = |O|$. Note that $\mathbb{E}(L_{|O|}\,\big|o_{|O|}) = 0$. By Theorem 8, we are defining here the process $\{\mathbb{E}(L_k\,|o_k\,)\}_{k=1}^{|O|}$ induced by the Markov chain $\{o_k\}$. Note also that $\mathbb{E}(L_k\,|o_k\,) = \mathbb{E}[M_k\,|o_k] - \mathbb{E}[S_k\,|o_k]$ where S_k is the number of objects assigned at round k via short cycles. By Proposition 2, $\mathbb{E}[S_k\,|F] \leq 2$ for any possible forest F, this implies that $\mathbb{E}[S_k\,|o_k] \leq 2$. Hence, we obtain that $\mathbb{E}(L_k\,|o_k\,) \geq 3-2=1$ if $o_k \geq \hat{o}$. (Recall that $\hat{o}$ is defined such that $o_k \geq \hat{o}$ implies $\mathbb{E}[M_k\,|o_k] \geq 3$.) Let T be first round at which the $\mathbb{E}(L_k\,|o_k\,)$ becomes smaller than 1: formally, $\mathbb{E}(L_k\,|o_k) \leq 1$ only if $k \geq T$ (this is well-defined since $\mathbb{E}(L_{|O|}\,\big|o_{|O|}) = 0$). Note that $o_T \leq \hat{o}$.

Now we obtain:

$$\begin{aligned}
\mathbb{E}[|\hat{O}|] &= \mathbb{E}(\sum_{k=1}^{|O|} L_k) \\
&= \sum_{o\in O} \Pr\{\tilde{o}_k = o\} \sum_{k=1}^{|O|} \mathbb{E}\left[L_k \middle| \tilde{o}_k = o\right] \\
&= \sum_{t} \Pr\{T=t\} \sum_{o\in O} \Pr\{\tilde{o}_k = o \big| T=t\} \sum_{k=1}^{|O|} \mathbb{E}\left[L_k \middle| \tilde{o}_k = o\right] \\
&\geq \sum_{t} \Pr\{T=t\} \sum_{k=1}^{t-1} \sum_{o\in O} \Pr\{\tilde{o}_k = o \big| T=t\} \mathbb{E}\left[L_k \middle| \tilde{o}_k = o\right] \\
&\geq \sum_{t} \Pr\{T=t\}(t-1) \\
&= \mathbb{E}[T] - 1
\end{aligned}$$

where the last inequality holds by definition of the random variable T. Indeed, whenever $\Pr\{\tilde{o}_k = o \big| T=t\} > 0$ (recall that $k < t$), $\mathbb{E}\left[L_k \middle| \tilde{o}_k = o\right] \geq 1$ must hold.

Once we have reached round T under TTC, at most $\hat{o}$ more short cycles can arise. Thus, the expected number of short cycles must be smaller than $2\mathbb{E}(T) + \hat{o}$. Indeed, the expected number of short cycles is smaller than 2 times the expected number of rounds for TTC to converge (recall that, by Corollary 6, the expected number of short cycles at each round is at most two) which itself is smaller than $2\mathbb{E}(T) + \hat{o}$. It follows that

$$2\mathbb{E}[T] + \hat{o} \geq \mathbb{E}[|O| - |\hat{O}|].$$

Combining the above inequalities, we obtain that

$$\mathbb{E}[|\hat{O}|] \geq \frac{1}{3}\left(|O| - \hat{o} + 2\right),$$

from which the result follows. □

Corollary 7. *There exists* $\gamma > 0, \delta > 0, N > 0$

$$\Pr\left\{\frac{|\hat{O}|}{|O|} > \delta\right\} > \gamma,$$

for all $|O| > N$.

E Proof of Theorem 4

Since $U(u_1^0, 0) > U(u_2^0, 1))$, all objects in O_1 are assigned before any agent starts applying to objects in O_2. Hence, the assignment achieved by individuals assigned objects in O_1 is the same as the one obtained when we run DA in the submarket with individuals in I and objects in O_1. The following lemma shows that the agents assigned objects in O_1 suffer a significant number of rejections before getting assigned. This result is obtained by Ashlagi, Kanoria, and Leshno (2013) and by Ashlagi, Braverman, and Hassidim (2011). We provide a much simpler direct proof for this result here.[36]

Lemma 12 (Welfare Loss under Unbalanced Market). *Consider an unbalanced submarket consisting of agents* I *and objects* O_1, *where* $|I| - |O_1| \to n(1 - x_1)$ *as* $n \to \infty$. *Let* I_1 *be the (random) set of agents who are assigned objects in* O_1, *and let* $I_1^\delta := \{i \in I_1 | i$ *makes at least* δn *offers*$\}$ *be the subset of them who each suffer from more than* δn *rejections (before getting assigned objects in* O_1*). Then, there exist* γ, δ, υ, *all strictly positive, such that for all* $n > N$ *for some* $N > 0$,

$$\Pr\left\{\frac{|I_1^\delta|}{|I|} > \gamma\right\} > \upsilon.$$

Proof. Without loss, we work with the McVitie and Wilson's algorithm (which equivalently implement DA). Consider the individual $i = n$ at the last serial order, at the beginning of step n. By that step, each object in O_1 has surely received at least $|I| - |O_1|$

[36] The main case studied by Ashlagi, Kanoria, and Leshno (2013) deals with the situation in which the degree of unbalancedness is small; i.e., $|I| - |O_1|$ is sublinear in n. Our proof does not apply to that case.

offers. This is because at least $|I| - |O_1| - 1$ preceding agents must be unassigned, so each of them must have been rejected by all objects in O_1 before the beginning of step n.

Each object receives offers randomly and selects its most preferred individual among those who have made offers to that object. Since each object will have received at least $|I| - |O_1|$ offers, its payoff must be at least $\max\{\eta_{1,o}, ...\eta_{|I|-|O_1|,o}\}$, i.e., the maximum of $|I| - |O_1|$ random draws of its idiosyncratic payoffs. At the beginning of step n, agent n makes an offer to an object o (i.e., his most favorite object which is drawn iid). Then, for n to be accepted by o, it must be the case that $\eta_{i,o} \geq \max\{\eta_{1,o}, ...\eta_{|I|-|O_1|,o}\}$. This occurs with probability $\frac{1}{|I|-|O_1|}$. Thus, the probability that n is assigned o is at most $\frac{1}{|I|-|O_1|}$.

Hence, for any $\delta \in (0, x_1)$, the probability that agent n is rejected δn times in a row is at least

$$\left(1 - \frac{1}{|I| - |O_1|}\right)^{\delta n} \to \left(\frac{1}{e}\right)^{\frac{\delta}{1-x_1}}.$$

Since agent n is ex ante symmetric with all other agents, for any agent $i \in I$,

$$\liminf \Pr\left\{\mathcal{E}_i^\delta\right\} \geq \left(\frac{1}{e}\right)^{\frac{\delta}{1-x_1}},$$

for any $\delta \in (0, x_1)$, where $\mathcal{E}_i^\delta$ denotes the event that i makes at least δn offers.

Let $\mathcal{F}_i := \{i \in I_1\}$ denote the event that agent i is assigned an object in O_1, and let $\mathcal{F}_i^c := \{i \notin I_1\}$ be its complementary event. Then, by ex ante symmetry of all agents, $\Pr\{\mathcal{F}_i\} = |O_1|/n \to x_1$ as $n \to \infty$. For $\delta \in (0, x_1)$, we obtain

$$\begin{aligned}
\Pr\left\{\mathcal{E}_i^\delta\right\} &= \Pr\{\mathcal{F}_i\} \Pr\left\{\mathcal{E}_i^\delta \mid \mathcal{F}_i\right\} + \Pr\{\mathcal{F}_i^c\} \Pr\left\{\mathcal{E}_i^\delta \mid \mathcal{F}_i^c\right\} \\
&\to x_1 \Pr\left\{\mathcal{E}_i^\delta \mid \mathcal{F}_i\right\} + (1 - x_1) \cdot 1 \text{ as } n \to \infty,
\end{aligned}$$

where the last line obtains since, with probability going to one as $n \to 1$, an agent who is not assigned an object in O_1 must make at least $\delta n < x_1 n$ offers. Combining the two facts, we have

$$\liminf \Pr\left\{\mathcal{E}_i^\delta \mid \mathcal{F}_i\right\} \geq \frac{1}{x_1}\left(\left(\frac{1}{e}\right)^{\frac{\delta}{1-x_1}} - (1 - x_1)\right).$$

Observe that the RHS tends to a strictly positive number as $\delta \to 0$. Thus, for $\delta > 0$ small enough (smaller than $(1 - x_1)\log(\frac{1}{x_1})$), $\Pr\left\{\mathcal{E}_i^\delta \mid \mathcal{F}_i\right\}$ is bounded below by some positive constant for all n large enough.

It thus follows that there exist $\delta \in (0, x_1)$, $\gamma > 0$ such that

$$\begin{aligned}
\mathbb{E}\left[\frac{|I_1^\delta|}{|I|}\right] &= \frac{1}{|I|}\mathbb{E}\left[\sum_{i\in I_1}\mathbf{1}_{\mathcal{E}_i^\delta}\right] \\
&= \frac{1}{|I|}\mathbb{E}_{I_1}\mathbb{E}\left[\sum_{i\in I_1}\mathbf{1}_{\mathcal{E}_i^\delta}\middle| I_1\right] \\
&= \frac{|I_1|}{|I|}\mathbb{E}_{I_1}\mathbb{E}\left[\mathbf{1}_{\mathcal{E}_i^\delta}|i \in I_1\right] \\
&= \frac{|I_1|}{|I|}\mathbb{E}\left[\mathbf{1}_{\mathcal{E}_i^\delta}|i \in I_1\right] \\
&= \frac{|I_1|}{|I|}\Pr\left\{\mathcal{E}_i^\delta|\mathcal{F}_i\right\} \\
&> \gamma
\end{aligned}$$

for all $n > N$ for some $N > 0$, since $\frac{|I_1|}{|I|} \to x_1$ as $n \to \infty$. The claimed result then follows. □

Lemma 12 implies that there exists $\epsilon' > 0, \upsilon' > 0, \gamma' > 0$ such that for all $n > N'$ for some $N' > 0$,

$$\Pr\left\{\frac{|\tilde{I}_\epsilon|}{|I|} \geq \gamma'\right\} \geq \upsilon',$$

where $\tilde{I}_{\epsilon'} := \{i \in I | DA(i) \in O_1, U_i(DA(i)) \leq U(u_1, 1-\epsilon')\}$ is the set of agents assigned to objects in O_1 but receive payoffs bounded above by $U(u_1, 1-\epsilon')$.

Now consider a matching mechanism μ that first runs DA and then runs a Shapley-Scarf TTC afterwards, namely the TTC with the DA assignments serving as the initial endowments for the agents. This mechanism μ clearly Pareto dominates DA. In particular, if $DA(i) \in O_1$, then $\mu(i) \in O_1$. For any ϵ'', let

$$\check{I}_{\epsilon''} := \{i \in I | \mu(i) \in O_1, U_i(DA(i)) \geq U(u_1, 1-\epsilon'')\},$$

be those agents who attain at least $U(u_1, 1-\epsilon'')$ from μ. By Theorem 1, we have for any ϵ'', γ'' and υ'', such that

$$\Pr\left\{\frac{|\check{I}_{\epsilon''}|}{|I|} \leq \gamma''\right\} < \upsilon'',$$

for all $n > N''$ for some $N'' > 0$.

Now set ϵ', ϵ'' such that $\epsilon = \epsilon' - \epsilon'' > 0$, γ', γ'' such that $\gamma := \gamma' - \gamma'' > 0$, and υ', υ'' such that $\upsilon := \upsilon' - \upsilon'' > 0$. Observe that $I_\epsilon(\mu|DA) \supset \tilde{I}_{\epsilon'} \setminus \check{I}_{\epsilon''}$, so $|I_\epsilon(\mu|DA)| \geq |\tilde{I}_{\epsilon'}| - |\check{I}_{\epsilon''}|$.

It then follows that for all $n > N := \max\{N', N''\}$,

$$\begin{aligned}
\Pr\left\{\frac{|I_\epsilon(\mu|DA)|}{|I|} \geq \gamma\right\} &\geq \Pr\left\{\frac{|\tilde{I}_{\epsilon'}|}{|I|} - \frac{|\check{I}_{\epsilon''}|}{|I|} \geq \gamma\right\} \\
&\geq \Pr\left\{\frac{|\tilde{I}_{\epsilon'}|}{|I|} \geq \gamma' \text{ and } \frac{|\check{I}_{\epsilon''}|}{|I|} \leq \gamma''\right\} \\
&\geq \Pr\left\{\frac{|\tilde{I}_{\epsilon'}|}{|I|} \geq \gamma'\right\} - \Pr\left\{\frac{|\check{I}_{\epsilon''}|}{|I|} > \gamma''\right\} \\
&\geq \upsilon' - \upsilon = \upsilon.
\end{aligned}$$

F Proof of Theorem 6

Let us fix $\varepsilon > 0$. Fix $k = 1, ..., K$ and $i \in \{|O_{\leq k-1}| + 2, ..., |O_{\leq k}| + 1\}$ with the convention that $|O_{\leq 0}| + 2 = 1$ and $|O_{\leq K}| + 1 = n$. We show that there is $N \geq 1$ s.t. for any $n \geq N$, for any vector of cardinal utilities $(u_o)_{o\in O} := (U_i(u_o, \xi_{io}))_{o\in O}$, i cannot gain more than ε by deviating given that everyone else reports truthfully. Given the symmetry of the problem among individuals in $\{|O_{\leq k-1}|+2, ..., |O_{\leq k}|+1\}$ and given that there are finitely many tiers, N can be taken to be uniform across all individuals which will show the desired result.

In the sequel, we assume that all individuals other than i report truthfully their preferences. We partition the set of i's possible reports into two sets $\mathcal{T}_1 \cup \mathcal{T}_2$ as follows. $\mathcal{T}_1$ consists of the set of i's reports that, restricted to objects in $O_{\geq k}$, only contain objects in O_k within the κ first objects. $\mathcal{T}_2$ consists of the set of i's reports which, restricted to objects in $O_{\geq k}$ contain some object outside O_k within the κ first objects. We will be using the following lemma.

Lemma 13. *If i's report is of type $\mathcal{T}_1$, then $\sum_{k=1}^{\kappa} p(k)$ converges to 1 (where $p(k)$ stands for the probability of getting object with rank k within the $O_{\geq k}$ objects). In addition, the convergence is uniform across all possible reports in $\mathcal{T}_1$. If i's report is of type $\mathcal{T}_2$, then $\sum_{k=1}^{\ell} p(k)$ converges to 1 where ℓ is the rank (within the $O_{\geq k}$ objects) of the first object outside O_k. In addition, the convergence is uniform across all possible reports in $\mathcal{T}_2$.*

Proof. Consider $\mathcal{E}$ the event under which, independently of i's reported preferences, provided that all individuals from 1 to $|O_{\leq k-1}| + 1$ report truthfully their preferences over objects in $O_{\leq k-1}$, for each $k' = 1, ..., k-1$, the objects assigned in stage k' are exactly those in $O_{k'}$. By our argument in the proof of Proposition 1, the probability of that event tends to 1. By construction, the convergence is uniform over all of i's possible reports. From now on, let us condition w.r.t. the realization of event $\mathcal{E}$. By Proposition 1, we know that with (conditional) probability going to 1, i is matched within stage k.[37] In addition, by ex ante symmetry of objects within a given tier (given our conditioning event $\mathcal{E}$, the way i ranks objects in $O_{\leq k-1}$ does not matter), the rate at which the conditional probability goes to 1 is the same for each report in $\mathcal{T}_1$. But given that $\Pr(\mathcal{E})$ goes to 1 uniformly across all possible i's reports, the unconditional probability that i is matched within stage k also converges to 1 uniformly across these reports. This proves the first part of the lemma.

Now, we move to the proof of the second part of the lemma. Let us consider the event that for each $k' = 1, ..., K$, all individuals other than i only rank objects in $O_{k'}$ within their κ most favorite objects in $O_{\geq k}$. Consider as well event $\mathcal{E}$ as defined above and let $\mathcal{F}$ be the intersection of these two events. By Lemma 2 as well as Proposition 1, we know that $\Pr(\mathcal{F})$ goes to 1 as n goes to infinity. By construction, the convergence is uniform over all of i's possible reports. Note that under event $\mathcal{F}$ no individual other than i will make an offer to an object outside O_k within stage k. In addition, under $\mathcal{F}$, the probability that i is matched in stage k goes to 1 as n goes to infinity and the convergence rate does not depend on i's specific report.[38] Combining these observations, it must be the case that the probability that "i is matched to the object outside O_k with rank ℓ or to a better object" converges to 1 at a rate independent of i' reports. In addition, we know that with probability going to 1, i can only be matched to an object in $O_{\geq k}$, thus we get that $\sum_{k=1}^{\ell} p(k)$ converges to 1 uniformly across i's possible reports. $\square$

Since $\mathcal{T}_1$ and $\mathcal{T}_2$ cover the set of all possible reports of individual i, we get

[37]That the unconditional probability of the event "i is matched within stage k" tends to 1 comes from Proposition 1. The only difference with the setting of Proposition 1 is that i's report on his preferences is not drawn randomly. However, it should be clear that the argument in the proof of Proposition 1 goes through as long as i's report is independent of his opponents' preferences which must be true in the environment we are considering where types are drawn independently and so where players play independently.
Now, that the conditional probability of our event goes to 1 comes from the facts that the unconditional probability of the event tends to 1 and that the conditioning event $\mathcal{E}$ has a probability which tends to 1.

[38]Indeed, if i makes an offer to the object that is not in O_k then we know that under $\mathcal{F}$, i will eventually be matched to that object at the end of stage k. If i only makes offers to objects in O_k then, by symmetry, we know that the probability that the agent who triggers the end of stage k (who cannot be agent i) makes an offer to the very same object is $|\frac{1}{O_k}|$. Hence, overall, if i has an opportunity to make an offer, with probability going to 1 (uniformly across reports in $\mathcal{T}_2$), agent i is assigned in stage k. Now, even if we consider a worst case scenario where $i = |O_{\leq k}| + 1$, given our assumption that all i's opponents are telling the truth, our argument in the proof of Proposition 1 implies that, with probability going to 1 (and irrespective of i's report), i will have an opportunity to make an offer in stage k.

Corollary 8. *For any i's report, we have that $\sum_{k=1}^{\kappa} p(k)$ converges to 1. Convergence is uniform across all of i's reports.*

In the sequel, we condition w.r.t. event $\mathcal{E}$ defined in the proof of the above lemma, i.e., the event that, irrespective of i's reported preferences, all objects in $O_{\leq k-1}$ are gone when stage k starts. As we already said, the probability of $\mathcal{E}$ converges to 1 uniformly across all possible i's reports. Now, we fix a type $(u_o)_{o\in O} := (U_i(u_o, \xi_{io}))_{o\in O}$ of individual i and consider two cases depending on whether this type falls into $\mathcal{T}_1$ or $\mathcal{T}_2$.

Case 1: Assume that individual i's type falls into $\mathcal{T}_1$. Clearly, the expected utility of telling the truth is higher than $\sum_{k=1}^{\kappa} p_T(k)u_k$ where $p_T(k)$ is the probability of getting object with rank k within the $O_{\geq k}$ objects (and hence within the O_k objects) when reporting truthfully. By the above lemma, if i reports truthfully, then with probability going to 1, he gets one of his κ most favorite object within $O_{\geq k}$, thus, for some $N_1 \geq 1$, and for all $n \geq N_1$, $\sum_{k=1}^{\kappa} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq 1$.

Let us consider a lie of individual i. Given our conditioning event, what matters are the reports within objects in $O_{\geq k}$. In addition, given the symmetry of objects within each tier, agent i will order the objects within each tier truthfully among them. Thus, one can think of a lie there as a report where from rank 1 to some rank $\ell - 1$ (the rank here is that within $O_{\geq k}$ objects) agent i is sincere and ranks truthfully the objects that are all in O_k but then at rank ℓ, i lists an object in $O_{>k}$. In that case, by definition of DACB, for each $k = 1, ..., \ell - 1$ individual i still has probability $p_T(k)$ to get the object with rank k. But now he has probability $p_L(\ell)$ to get matched to the object in $O_{>k}$. Now, recall that for all $n \geq N_1$, $\sum_{k=1}^{\kappa} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq 1$. This implies that for all $n \geq N_1$, $\sum_{k=1}^{\kappa} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{k=1}^{\ell-1} p_T(k) + p_L(\ell)$ and so $\sum_{k=\ell}^{\kappa} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq p_L(\ell)$. In addition, we know by the above lemma that $\sum_{k=1}^{\ell-1} p_T(k) + p_L(\ell)$ converges to 1 uniformly across all possible deviations of individual i. Since, given our conditioning event, i has zero probability to get an object in $O_{\leq k-1}$, there must exist some $N_2 \geq 1$ so that for all $n \geq N_2$, i's expected payoff when he lies is smaller than $\sum_{k=1}^{\ell-1} p_T(k)u_k + p_L(\ell)u_* + \frac{\varepsilon}{2}$ where u_* is the utility of the best object in $O_{>k}$ and so must satisfy $u_* < u_k$ for each $k = 1, ..., \kappa$. In the sequel, we fix $n \geq \max\{N_1, N_2\}$. We obtain that, conditional on $\mathcal{E}$, the expected payoff when lying is smaller than

$$\begin{aligned}
\sum_{k=1}^{\ell-1} p_T(k)u_k + p_L(\ell)u_* + \frac{\varepsilon}{2} &\leq \sum_{k=1}^{\ell-1} p_T(k)u_k + \sum_{k=\ell}^{\kappa} p_T(k)u_* + u_*\frac{\varepsilon}{2U(1,1)} + \frac{\varepsilon}{2} \\
&\leq \sum_{k=1}^{\ell-1} p_T(k)u_k + \sum_{k=\ell}^{\kappa} p_T(k)u_* + \varepsilon \\
&\leq \sum_{k=1}^{\ell-1} p_T(k)u_k + \sum_{k=\ell}^{\kappa} p_T(k)u_k + \varepsilon
\end{aligned}$$

The first inequality uses the fact that $n \geq N_1$ and so that $\sum_{k=\ell}^{\kappa} p_T(k)+\frac{\varepsilon}{2U(1,1)} \geq p_L(\ell)$. The second inequality comes from the simple fact that $u_* \leq U(1,1)$. The last inequality holds because $u_* < u_k$ for each $k = 1, ..., \kappa$. Since the expected payoff of the truth is larger than $\sum_{k=1}^{\kappa} p_T(k)u_k$, we conclude that, conditional on $\mathcal{E}$, lying cannot make i gain more than ε whenever $n \geq \max\{N_1, N_2\}$.[39] Since $\mathcal{E}$, has a probability going to 1 uniformly across all possible deviations of individual i, a same result holds for unconditional expected payoffs.

Case 2: Assume that individual i has a type which falls into $\mathcal{T}_2$. Consider the κ first objects in $O_{\geq k}$ and let R be the rank (here again, the rank is taken among $O_{\geq k}$ objects) of the first object in $O_{>k}$ appearing there. Clearly, the expected utility of truth-telling is higher than $\sum_{k=1}^{R} p_T(k)u_k$ where $p_T(k)$ is the probability of getting object with rank k within the $O_{\geq k}$ objects when reporting truthfully. By the above lemma, if i reports truthfully, then with probability going to 1, i gets one of his R most favorite object within $O_{\geq k}$, thus, for some $N_1 \geq 1$, and for all $n \geq N_1$, $\sum_{k=1}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq 1$.

Let us consider a lie of individual i. Given our conditioning event, what matters are the reports within objects in $O_{\geq k}$. In addition, given the symmetry of objects within each tier, agent i will order the objects within each tier truthfully among them. Let us first consider a lie where the first object in $O_{>k}$ is ranked at $R' < R$ (here again, the rankings are those within $O_{\geq k}$). Thus, one can think of such a lie there as a report where from rank 1 to some rank $R' - 1$ (the rank here is that within $O_{\geq k}$ objects) agent i is sincere and ranks truthfully the objects that are all in O_k but then at rank R', i lists an object in $O_{>k}$. In that case, by definition of DACB, for each $k = 1, ..., R' - 1$ individual i still has probability $p_T(k)$ to get the object with rank k. But now he has probability $p_L(R')$ to get matched to the object in $O_{>k}$ listed at rank R'. Now, recall that for all $n \geq N_1$, $\sum_{k=1}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq 1$. This implies that for all $n \geq N_1$, $\sum_{k=1}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq$

[39] Notice that by the uniform convergence result in the above lemma, N_1 and N_2 are independent on i's specific report.

$\sum_{k=1}^{R'-1} p_T(k) + p_L(R')$ and so $\sum_{k=R'}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq p_L(R')$. In addition, we know by the above lemma that $\sum_{k=1}^{R'-1} p_T(k) + p_L(R')$ converges to 1 uniformly across all possible deviations of individual i. Since, given our conditioning event, i has zero probability to get an object in $O_{\leq k-1}$, there must exist some $N_2 \geq 1$ so that for all $n \geq N_2$, i's expected payoff when he lies is smaller than $\sum_{k=1}^{R'-1} p_T(k)u_k + p_L(R')u_* + \frac{\varepsilon}{2}$ where u_* is the utility of the object in $O_{>k}$ listed at rank R' and so must satisfy $u_* \leq u_k$ for each $k = 1, ..., R$. In the sequel, we fix $n \geq \max\{N_1, N_2\}$. We obtain that, conditional on $\mathcal{E}$, the expected payoff when lying is smaller than

$$\begin{aligned}
\sum_{k=1}^{R'-1} p_T(k)u_k + p_L(R')u_* + \frac{\varepsilon}{2} &\leq \sum_{k=1}^{R'-1} p_T(k)u_k + \sum_{k=R'}^{R} p_T(k)u_* + u_*\frac{\varepsilon}{2U(1,1)} + \frac{\varepsilon}{2} \\
&\leq \sum_{k=1}^{R'-1} p_T(k)u_k + \sum_{k=R'}^{R} p_T(k)u_* + \varepsilon \\
&\leq \sum_{k=1}^{R'-1} p_T(k)u_k + \sum_{k=R'}^{R} p_T(k)u_k + \varepsilon
\end{aligned}$$

The first inequality uses the fact that $n \geq N_1$ and so that $\sum_{k=R'}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq p_L(R')$. The second inequality comes from the simple fact that $u_* \leq 2U(1,1)$. The last inequality holds because $u_* \leq u_k$ for each $k = 1, ..., R$. Since the expected payoff of the truth is larger than $\sum_{k=1}^{R} p_T(k)u_k$, we conclude that, conditional on $\mathcal{E}$, lying cannot make i gain more than ε whenever $n \geq \max\{N_1, N_2\}$. Since $\mathcal{E}$, has a probability going to 1 uniformly across all possible deviations of individual i, a same result holds for unconditional expected payoffs.

Similarly, let us consider a lie so that the first object in $O_{>k}$ is ranked at $R' \geq R$ (recall that the rankings are those within $O_{\geq k}$). Thus, one can think of such a lie there as a report where from rank 1 to some rank $R-1$ (the rank here is that within $O_{\geq k}$ objects) agent i is sincere and ranks truthfully the objects that are all in O_k. In that case, by definition of DACB, for each $k = 1, ..., R-1$ individual i still has probability $p_T(k)$ to get the object with rank k. But now he has probability $p_L(k)$ to get matched to the object with rank k for each $k = R, ..., R'$. Now, recall that for all $n \geq N_1$, $\sum_{k=1}^{R} p_T(k) + \frac{\varepsilon}{2U(1,1)} \geq 1$. This implies that for all $n \geq N_1$, $\sum_{k=1}^{R-1} p_T(k) + p_T(R) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{k=1}^{R-1} p_T(k) + \sum_{k=R}^{R'} p_L(k)$ and so $p_T(R) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{k=R}^{R'} p_L(k)$. In addition, we know by the above lemma that $\sum_{k=1}^{R-1} p_T(k) + \sum_{k=R}^{R'} p_L(k)$ converges to 1 uniformly across all possible deviations of individual i. Since, given our conditioning event, i has zero probability to get an object in $O_{\leq k-1}$, there must exist some $N_2 \geq 1$ so that for all $n \geq N_2$, i's expected payoff when he lies is smaller than

$\sum_{k=1}^{R-1} p_T(k)u_k + \sum_{k=R}^{R'} p_L(k)v_k + \frac{\varepsilon}{2}$ where $v_k \leq u_R$ for each $k = R, ..., R'$. In the sequel, we fix $n \geq \max\{N_1, N_2\}$. We obtain that, conditional on $\mathcal{E}$, the expected payoff when lying is smaller than

$$\begin{aligned}
\sum_{k=1}^{R-1} p_T(k)u_k + \sum_{k=R}^{R'} p_L(k)v_k + \frac{\varepsilon}{2} &\leq \sum_{k=1}^{R-1} p_T(k)u_k + \sum_{k=R}^{R'} p_L(k)u_R + \frac{\varepsilon}{2} \\
&\leq \sum_{k=1}^{R-1} p_T(k)u_k + p_T(R)u_R + \frac{\varepsilon}{2U(1,1)}u_R + \frac{\varepsilon}{2} \\
&\leq \sum_{k=1}^{R-1} p_T(k)u_k + p_T(R)u_R + \varepsilon
\end{aligned}$$

The first inequality uses the fact that $u_R \geq v_k$ for all $k = R, ..., R'$. The second inequality uses the fact that $n \geq N_1$ and so that $\sum_{k=R}^{R'} p_L(k) \leq p_T(R) + \frac{\varepsilon}{2U(1,1)}$. The last inequality follows from the simple fact that $u_R \leq U(1,1)$. Since the expected payoff of the truth is larger than $\sum_{k=1}^{R} p_T(k)u_k$, we conclude that, conditional on $\mathcal{E}$, lying cannot make i gain more than ε whenever $n \geq \max\{N_1, N_2\}$. Since $\mathcal{E}$, has a probability going to 1 uniformly across all possible deviations of individual i, a same result holds for unconditional expected payoffs.

References

Abdulkadiroglu, A., Y.-K. Che, and O. Tercieux (2013): "The Role of Priorities in Top Trading Cycles," Columbia University, Unpublished mimeo.

Abdulkadiroglu, A., Y.-K. Che, and Y. Yasuda (2008): "Expanding 'Choice' in School Choice," mimeo.

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

——— (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99(5), 1954–1978.

Abdulkadiroglu, A., P. A. Pathak, A. E. Roth, and T. Sonmez (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 368–372.

Abdulkadiroglu, A., and T. Sönmez (1998): “Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems,” *Econometrica*, 66, 689–701.

Abdulkadiroglu, A., and T. Sonmez (2003): “School Choice: A Mechanism Design Approach,” *American Economic Review*, 93, 729–747.

Ashlagi, I., M. Braverman, and A. Hassidim (2011): “Stability in large matching markets with complementarities,” Discussion paper, Discussion paper, Mimeo.

Ashlagi, I., Y. Kanoria, and J. D. Leshno (2013): “Unbalanced Random Matching Markets,” MIT, Unpublished mimeo.

Azevedo, E. M., and J. W. Hatfield (2012): “Complementarity and Multidimensional Heterogeneity in Matching Markets,” mimeo.

Azevedo, E. M., and J. D. Leshno (2011): “A supply and demand framework for two-sided matching markets,” *Unpublished mimeo, Harvard Business School*.

Balinski, M., and T. Sönmez (1999): “A tale of two mechanisms: student placement,” *Journal of Economic Theory*, 84, 73–94.

Bollobas, B. (2001): *Random Graphs*. Cambridge University Press, Cambridge, United Kingdom.

Calsamiglia, C., G. Haeringer, and F. Klijn (2010): “Constrained School Choice: An Experimental Study,” *American Economic Review*, 100, 1860–1874.

Che, Y.-K., J. Kim, and F. Kojima (2013): “Stable Matching in Large Economies,” mimeo.

Che, Y.-K., and F. Kojima (2010): “Asymptotic equivalence of probabilistic serial and random priority mechanisms,” *Econometrica*, 78(5), 1625–1672.

Chen, Y., and O. Kesten (2013): “From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms,” Unpublished mimeo.

Dawande, M., P. Keskinocak, J. Swaminathan, and S. Tayur (2001): “On Bipartite and Multipartite Clique Problems,” *Journal of Algorithms*, 41, 388403.

Dubins, L. E., and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *American Mathematical Monthly*, 88, 485–494.

Ergin, H. (2002): "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 70, 2489–2498.

Frieze, A., and B. Pittel (1995): "Probabilistic Analysis of an Algorithm in the Theory of Markets in Indivisible Goods," *The Annals of Applied Probability*, 5, 768–808.

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

Haeringer, G., and F. Klijn (2009): "Constrained School Choice," *Journal of Economic Theory*, 144, 1921–1947.

Hylland, A., and R. Zeckhauser (1979): "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 87, 293–314.

Immorlica, N., and M. Mahdian (2005): "Marriage, Honesty, and Stability," *SODA 2005*, pp. 53–62.

Jaworski, J. (1985): "A Random Bipartite Mapping," *The Annals of Discrete Mathematics*, 28, 137–158.

Jin, Y., and C. Liu (2004): "Enumeration for spanning forests of complete bipartite graphs," *Ars Combinatoria - ARSCOM*, 70, 135–138.

Kesten, O. (2006): "On Two Competing Mechanisms for Priority Based Allocation Problems," *Journal of Economic Theory*, 127, 155–171.

Knuth, D. E. (1996): *Stable marriage and its relation to other combinatorial problems.* American Mathematical Society, Providence.

Knuth, D. E., R. Motwani, and B. Pittel (1990): "Stable Husbands," *Random Structures and Algorithms*, 1, 1–14.

Kojima, F., and M. Manea (2010): "Incentives in the Probabilistic Serial Mechanism," *Journal of Economic Theory*, 145, 106–123.

Kojima, F., and P. A. Pathak (2008): "Incentives and Stability in Large Two-Sided Matching Markets," forthcoming, *American Economic Review*.

Lee, S. (2012): "Incentive compatibility of large centralized matching markets," URL http://www. hss. caltech. edu/˜ sangmok/research. html.

Lovasz, L. (1979): *Combinatorial Problems and Exercises*. North Holland, Amsterdam.

McVitie, D. G., and L. Wilson (1971): "The stable marriage problem," *ACM*, 14, 486–490.

Pathak, P., and J. Sethuraman (2011): "Lotteries in Student Assignment: An Equivalence Result," *Theortical Economics*, 6, 1–17.

Pathak, P. A., and T. Sömez (2013): "School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation," *American Economic Review*, 103, 80–106.

Pittel, B. (1989): "The Average Number of Stable Matchings," *SIAM Journal on Discrete Mathematics*, 2, 530–549.

——— (1992): "On Likely Solutions of a Stable Marriage Problem," *The Annals of Applied Probability*, 2, 358–401.

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

Roth, A. E., and M. A. O. Sotomayor (1990): *Two-sided matching: a study in game-theoretic modeling and analysis*. Cambridge University Press, Cambridge.

Shapley, L., and H. Scarf (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 22–37.

Sonmez, T., and U. Unver (2011): "Market Design for Kidney Exchange," in *Oxford Handbook of Market Design*, ed. by M. N. Z. Neeman, and N. Vulkan. forthcoming.