Efficiency and Stability in Large Matching Markets *

Yeon-Koo Che[†] Olivier Tercieux[‡]

June 24, 2017

Abstract

We study efficient and stable mechanisms in matching markets when the number of agents is large and individuals' preferences and priorities are drawn randomly. When agents' preferences are uncorrelated, then both efficiency and stability can be achieved in an asymptotic sense via standard mechanisms such as deferred acceptance and top trading cycles. When agents' preferences are correlated over objects, however, these mechanisms are either inefficient or unstable even in an asymptotic sense. We propose a variant of deferred acceptance that is asymptotically efficient, asymptotically stable and asymptotically incentive compatible. This new mechanism performs well in a counterfactual calibration based on New York City school choice data.

JEL Classification Numbers: C70, D47, D61, D63.

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

---

*We are grateful to Ludovic Lelièvre, Charles Maurin and Xingye Wu for their excellent research assistance. We owe a special gratitude to Atila Abdulkadiroglu, Nikhil Agarwal and Parag Pathak for sharing programming codes for Gibbs sampling. We also thank Itai Ashlagi, Eduardo Azevedo, Eric Budish, Julien Combe, Olivier Compte, Tadashi Hashimoto, Yash Kanoria, Fuhito Kojima, Scott Kominers, SangMok Lee, Bobak Pakzad-Hurson, Debraj Ray, Al Roth, Rajiv Sethi, and seminar participants at Chicago, Columbia, Stanford, Maryland, NYU, Toronto, Wisconsin, UBC, UCL, Simon Fraser, Microsoft, KAEA Conference, NYC Market Design Workshop, PSE Market Design conference, UBC, Warwick Micro Theory conference and WCU Market Design conference for helpful comments. Both authors acknowledge financial support from Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF Project number 2016S1A2A2912564). Olivier Tercieux is grateful for the support from ANR grant SCHOOL_ CHOICE (ANR-12-JSH1-0004-01).

[†]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, is an 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 circumvent—a suggested assignment. Not only is stability crucial for the 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] For instance, in the school choice context, eliminating justified envy means that no student would lose a school seat to another student with a lower priority at that school.

Unfortunately, these two goals are incompatible (see Roth (1982)). Matching mechanisms such as serial dictatorship and top trading cycles (henceforth, TTC) attain efficiency but fail to be stable. Meanwhile, stable mechanisms such as Gale and Shapley's deferred acceptance algorithms (henceforth, DA) do not guarantee efficiency. In light of the impossibility of achieving both goals, the prevailing approach, particularly in the context of school choice, strives to attain one objective with the minimum possible sacrifice of the other goal. For instance, DA selects a stable matching that Pareto dominates all other stable matchings for the proposing side (Gale and Shapley, 1962). Similarly, there is a sense in which TTC, which allows agents to trade their priorities sequentially, satisfies efficiency at the minimal incidence of instabilities (Abdulkadiroglu, Che, Pathak, Roth, and Tercieux, 2017).[2]

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 at 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 either objective.

[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 is under public control, so strategic blocking is not a serious concern.

[2] This version of TTC was 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 to the agent with the highest priority. A cycle is then formed, the agents involved in that 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 exhausted. Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017) 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.

The purpose of the current paper is to answer these questions and, in the process, provide useful insights for practical market design. These questions remain outstanding because existing analytical frameworks are driven primarily by "qualitative" notions of the two goals. To make progress, we therefore need to relax them "quantitatively." Doing so requires imposing some structure on the model. First, we consider markets that are "large" in the number of participants and in the number of object types. Large markets are clearly relevant in many settings. For instance, in the US medical matching system, each year there are approximately 20,000 applicants for positions at 3,000 to 4,000 programs. In the New York City (NYC) school choice, approximately 80,000 students apply each year to over 700 high school programs. Second, we assume that agents' preferences are generated randomly according to some reasonable distributions. Specifically, we consider a model in which each agent's utility for an object depends on a common component (that does not vary across agents) and an idiosyncratic component that is independently drawn at random (and thus varies across the agents), and the agents' priorities over objects are drawn according to a distribution identically and independently.[3]

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 search 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 in a sufficiently large economy, with high probability, the proportion of agents who would have justified envy toward a significant number of agents vanishes.

Our first set of findings pertains to the tradeoff between DA and TTC. We find that, when agents' preferences for the objects are significantly correlated, the efficiency loss from DA remains significant even when the market grows large. Likewise, the instabilities in TTC do not disappear in large markets. The potential inefficiencies of DA and instabilities of TTC are well known from the existing literature; our novel finding here is that they remain "quantitatively" significant (even) in a large market.

These findings can be explained in intuitive terms. Suppose that the objects come in two tiers, high quality and low quality, and that every high-quality object is preferred to every low-quality object by 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 compete for a low-quality object. Such competition means that in a stable matching—

[3] We discuss in Sections 6 and 7 how our results carry over to richer environments in which agents' priorities are correlated.

including agent-optimal stable matching—the outcome is dictated largely by how the objects rank the agents and not by how the agents rank the objects. In other words, the competition among agents entails significant welfare loss in the presence of the stability requirement.

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. Typically, these agents have high priorities with the objects they are *trading off*, but they could well have very low priorities with the objects they are *trading in.*

Taken together, these two findings have an important practical market design implication, as they suggest that the standard approach of achieving one goal with a minimal sacrifice of the other may not be the best.[4]

Motivated by these results, we develop a new mechanism, called Deferred Acceptance with Circuit Breaker (DACB), that is both asymptotically efficient and asymptotically stable. This mechanism modifies DA to prevent participants from competing excessively. Specifically, all agents are ordered in some manner (for instance, at random), and following that order, each agent applies *one at a time* to the best object that has not yet rejected him,[5] and the proposed object accepts or rejects the applicant, much as in 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 continues until an agent makes a certain "threshold" number κ of offers for the first time. Then, the stage is terminated at that point, and all tentative assignments up to that point become final. The next stage then begins with the agent who was rejected at the end of the last stage applying to the best remaining object and the number of proposals for that agent being reset to zero. The stages proceed in this fashion until no rejection occurs.

This "staged" version of DA resembles 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 an agent reaching a threshold number of offers. Intuitively, the mechanism activates a "circuit breaker" whenever the competition "overheats" to an extent that places an agent at the risk of losing an object he ranks highly to an agent who ranks it relatively lowly (more precisely, above the threshold rank). This feature ensures that each object assigned at each stage goes to an agent who ranks it relatively highly

[4]In combinatorial assignment problems where transfers are not allowed, Budish (2011) makes a related point: he offers a market-like mechanism which makes relatively small compromises on efficiency and envy-freeness (while keeping desirable incentive properties) whereas known mechanisms satisfy one of these two objectives exactly.

[5]A version of DA in which offers are made according to a serial order was first introduced by McVitie and Wilson (1971).

among the objects available at that stage.

Given the independent drawing of idiosyncratic shocks, the "right" κ is shown to be sublinear in n no less than $\log^2(n)$ where n is the number of agents. Given the threshold, the DACB produces an assignment that is both asymptotically stable and asymptotically efficient. The analytical case for this mechanism rests on a limit analysis, but the mechanism performs well even away from the limit. Our simulation shows that, even for a moderately large market and 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 because the agent may then wish to apply to some safer object instead of a more popular one that has a high probability of rejecting him. The probability of this situation is one over the number of agents assigned in the current stage, which is on the order of n; hence, with a sufficient number of participants, the incentive issue is rather small.[6] Formally, we show that the mechanism induces truthful reporting as an ϵ-Bayes-Nash equilibrium.

Finally, another potential concern with this mechanism is the required bound on $\kappa \geq \log^2(n)$. In practice, applicants are often constrained to make a small number of applications, possibly below $\log^2(n)$ (a case in point is the high school assignment in NYC; see Section 6). To address such a situation, we generalize our mechanism so that for each κ, the termination of a stage is triggered *only* when at least $j \geq 1$ agents have each made more than κ offers. We provide a joint condition on (κ, j) that ensures that the generalized version of DACB is both asymptotically stable and asymptotically efficient. In particular, the required κ can be quite small for a sufficiently large j.

To study how our findings apply to a realistic market, we use the preference data supplied by the New York City Department of Education for public high school assignment during the 2009-2010 school year. Their main round employed a student-proposing DA in which each applicant submits a rank-order list (henceforth ROL) of up to 12 programs. Assuming the observed ROLs to prevail under alternative algorithms, we find a significant tradeoff between efficiency and stability. First, on average 5,189 students would be Pareto-improved if they were rematched efficiently starting from the DA. Meanwhile, TTC would entail 18,943 students with

[6] Unlike Azevedo and Budish (2015), the number of preference types grows without bound as the market grows large in the current model. Hence, their result on "strategy-proofness in the large" does not apply here. Nevertheless, it is simple to see from our arguments that DACB mechanism has a similar incentive property: truthful reporting is optimal in the limit economy against any *iid* distribution of reports provided that the distribution is one of those allowed in our paper. See also Remark 1.

justified envy.[7] This result is consistent with our theoretical finding that the tradeoffs do not disappear when the two prominent mechanisms are employed in a large market. Meanwhile, we show that DACB, with suitably-chosen parameters, would span a range of outcomes on the efficiency-stability frontier that are unattainable by the existing mechanisms. For reasons to be explained in detail, however, relying on the observed ROLs undersates the tradeoff between DA and TTC. We therefore performed structural estimation of students' preferences using the method developed by Abdulkadiroglu, Agarwal, and Pathak (2015), and simulated alternative algorithms based on these estimates. Under these counterfactual analyses, TTC and DA perform considerably worse; for instance, about 29,293 students can be made better off from a Pareto-improving reassignment under DA while 21,209 applicants would feel justified envy under TTC.[8] By contrast, DACB performs impressively. For instance, it can yield an outcome considerably more efficient than DA with very little sacrifice in stability.

The DACB mechanism bears some resemblance to features observed in popular real-world matching algorithms. The "staged termination" feature is present in the school assignment program in China (Chen and Kesten (2017)). More important, 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 most real-world implementations of DA. We provide a rationale for this practice that is common in the actual implementation of DA but has thus far been difficult to rationalize (see Haeringer and Klijn (2009), Calsamiglia, Haeringer, and Klijn (2010), Pathak and Sömez (2013) and Ashlagi, Nikzad, and Romm (2015)). Indeed, we show that DA with an appropriate limit on the ROLs can, to some extent, achieve an asymptotically efficient and stable *equilibrium* outcome.

The present paper is related to the growing literature that studies large matching markets, particularly those with a large number of object types and random preferences; see Immorlica and Mahdian (2005), Kojima and Pathak (2009), Lee (2017), Knuth (1997), Pittel (1989), Ashlagi, Braverman, and Hassidim (2014), Ashlagi, Kanoria, and Leshno (2017) and Lee and Yariv (2017). The first three papers are largely concerned with the incentive issues arising in DA. The last five papers are concerned with the ranks of the partners achieved by the agents on the two sides of the market under DA. In particular, the last three papers study the large market efficiency performance of DA, and their relationship with the current paper will be discussed more fully below. Unlike these papers, our paper considers not only DA

[7] These figures are broadly in line with Abdulkadiroglu, Pathak, and Roth (2009)'s analysis of the 2006-2007 choice data. Note that their efficient matching does not coincide with TTC. Instead, Abdulkadiroglu, Pathak, and Roth (2009) produce efficient matching by first running DA and then running a Shapley-Scarf TTC based on the DA assignment.

[8] As we explain in Section 6, the significant differences in the counterfactual analyses are attributed to the two methods which respectively provide lower- and upper-bound for the tradeoff between DA and TTC.

but also other mechanisms and adopts broader perspectives concerning both efficiency and stability.[9] Finally, Che and Tercieux (2015a) studies a large market properties of Pareto efficient mechanisms and provides some preliminary observation for the current paper.

2 Model

A finite set of agents are assigned a finite set of objects, at most one object for each agent. Because our analysis will involve studying the limit of a sequence of such finite economies as they become large, 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 **objects** O^n, where $|I^n| = |O^n| = n$. The assumption that these sets are of equal size is purely for convenience. Provided that they grow at the same rate, our results hold even if the sets are not of equal size. For much of the analysis, we suppress the superscript n for notational convenience.

2.1 Preliminaries

Throughout, we will 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.[10]

The common values take finite values $\{u_1, ..., u_K\}$ such that $u_1 > ... > u_K$. For each n-economy, the objects O^n are partitioned into tiers, $\{O_1^n, ..., O_K^n\}$, such that each object in tier O_k^n yields a common value of u_k to the agent who is assigned it. We assume that the proportion of tier-k objects, $|O_k^n|/n$, converges to $x_k > 0$ such that $\sum_{k \in K} x_k = 1$. We sometimes use the notation $O_{\geq k}$ to denote the set of objects in $\cup_{\ell \geq k} O_\ell$. Similarly, let $O_{\leq k} := \cup_{\ell \leq k} O_\ell$. One can imagine an alternative model in which the common value is drawn randomly from $\{u_1, ..., u_K\}$

[9] 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 (2015), Che and Kojima (2010), Kojima and Manea (2010), Azevedo and Leshno (2016) and Che, Kim, and Kojima (2013), Azevedo and Budish (2015), among others. The assumption of a finite number of object types enables one to use a continuum economy as a limit benchmark in these models. This feature makes substantial differences for both the analysis and the insights. The two strands of large matching market models capture issues that are relevant in different real-world settings and thus complement one another.

[10] This assumption is without loss of generality provided that the type distribution is atomless and has full and bounded support, as 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.

according to some distribution that converges to $\{x_1, ..., x_K\}$ as $n \to \infty$. Such a treatment will yield the same results as the current treatment, which can be regarded as considering each realization of such a random drawing.

We further assume that the function $U(\cdot, \cdot)$ takes values in $\mathbb{R}_+$, is strictly increasing in the common value and idiosyncratic shock and is continuous in the latter. The utility of remaining unmatched is assumed to be 0, which implies that each agent finds all objects acceptable.[11]

Next, the priorities agents have with different objects—or objects' "preferences" over agents—are drawn uniform randomly. Formally, we assume that individual i achieves a priority score:

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

at object $o \in O$, where *idiosyncratic shock* $\eta_{i,o}$ is a random variable drawn independently and identically from $[0, 1]$ according to the uniform distribution. This assumption simplifies the analysis. Although restrictive, this assumption captures a class of plausible circumstances under which a tradeoff between the two objectives persists and can be addressed more effectively by the novel mechanism we propose. Further, as explained in Section 5.1, our new mechanism can be easily modified to accommodate correlations in agents' priorities over objects. The function $V(\cdot)$ takes values in $\mathbb{R}_+$ and is strictly increasing and continuous. The utility of remaining unmatched is assumed to be 0, which implies that all objects find all individuals acceptable.

Fix an n-economy. We will consider a class of matching mechanisms that are Pareto efficient. A **matching** μ in an n-economy is a mapping $\mu : I \to O \cup \{\emptyset\}$ with the interpretation that agent i with $\mu(i) = \emptyset$ is unmatched. In addition, $\mu(i) \neq \mu(j)$ for any $j \neq i$, whenever $\mu(i) \neq \emptyset$ or $\mu(j) \neq \emptyset$. Let M_n denote the set of all matchings in n-economy. All of these objects depend on n, although their dependence is suppressed for notational convenience.

A **matching mechanism** is a function that maps states to matchings, where a state $\omega = (\{\xi_{i,o}, \eta_{i,o}\}_{i \in I, o \in O})$ consists of the realized profile $\{\xi_{i,o}\}_{i \in I, o \in O}$ of the idiosyncratic component of agents' payoffs and the realized profile $\{\eta_{i,o}\}_{i \in I, o \in O}$ of agents' priorities with the objects.[12] With a slight abuse of notation, we will 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. The set of all states is denoted by Ω. Let $\mathcal{M}_n$ denote the set of all matching mechanisms in n-economy. For convenience, we

[11] This feature does not play a crucial role in our results, which hold provided that a linear fraction of objects are acceptable to all agents.

[12] Note that matching mechanisms depend on cardinal preferences/priorities in our model, whereas standard mechanisms such as top trading cycles and deferred acceptance depend only on ordinal preferences and priorities. Obviously, cardinal preferences/priorities induce ordinal preferences/priorities, and the current treatment clearly encompasses these mechanisms. We focus on cardinal utilities and priorities to operationalize the asymptotic notions of efficiency and stability. Our results do not depend on the particular cardinalization of utilities.

will often suppress the dependence of the matching mechanism on ω and on n.

For a limit analysis, we are interested in a sequence $\{\mu_n\}$ of matching mechanisms for the corresponding n-economies. We call such a sequence a **matching outcome**. In what follows, when no confusion arises, we suppress the dependence of matching mechanisms on the index n when we refer to a matching outcome.

2.2 Welfare and Fairness Concepts in Large Markets

A matching $\mu \in M_n$ is Pareto efficient if there is no other matching $\mu' \in M_n$ 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 $\mu \in \mathcal{M}_n$ 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. A matching μ at a given state is stable if there is no pair (i, o) such that $U_i(o) > U_i(\mu(i))$ and $V_o(i) > V_o(\mu(o))$—i.e., a pair wish to match with each other rather than their partners in matching μ. A matching mechanism $\mu \in \mathcal{M}_n$ is stable if, for each state $\omega \in \Omega$, the matching it induces, i.e., $\mu_\omega(\cdot)$, is stable.

Throughout, we will invoke the following implication of Pareto efficiency.

LEMMA 1 (Che and Tercieux (2015c)). *For any Pareto-efficient matching outcome* $\{\mu_n\}$,

$$\frac{\sum_{i\in I} U_i(\mu_n(i))}{n} \xrightarrow{p} \sum_{k=1}^{K} x_k U(u_k, 1).$$

Notice that the right-hand side gives the (normalized) total utility that would be obtained if all agents attained the highest possible idiosyncratic value; hence, it is the utilitarian upper bound. The lemma states that the aggregate utilities agents enjoy in any Pareto efficient mechanism approach that bound in probability as $n \to \infty$. Recall that our model allows for agents' preferences to be correlated; in particular agents tend to prefer objects with higher common value than lower common value. The striking implication of Lemma 1 is that this conflict of interests does not cause a significant welfare loss if the allocation is Pareto efficient. As will be seen, the same will not be the case with a stable matching.

We next discuss how efficiency and stability can be weakened in the large market setting. We say that a matching outcome $\{\mu_n\}$ is **asymptotically efficient** if, for any $\epsilon > 0$ and for any mechanism $\{\mu'_n\}$ that Pareto dominates $\{\mu_n\}$:

$$\frac{|I_\epsilon({\mu'}_n|\mu_n)|}{n} \xrightarrow{p} 0 \text{ as } n \to \infty,$$

where

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

is the set of agents who would benefit more than ϵ by switching from μ_n to μ'_n. In words, a matching outcome is asymptotically efficient if the fraction of 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 that a matching outcome $\{\mu_n\}_n$ is **asymptotically stable** if, for any $\epsilon > 0$:

$$\frac{|J_\epsilon(\mu_n)|}{n(n-1)} \xrightarrow{p} 0 \text{ as } n \to \infty,$$

where

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

is the set of ϵ**-block**s—namely, the set of pairs of an unmatched agent and an object who would each gain ϵ or more from matching with one another rather than matching according to μ_n. Asymptotic stability requires that for any $\epsilon > 0$, the fraction of these ϵ-blocks as a share 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 ϵ-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 $\hat{O}^i_\epsilon(\mu_n) := \{o \in O | (i, o) \in J_\epsilon(\mu_n)\}$ be the set of objects agent i can form an ϵ-block with against μ_n. 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 in probability, i.e., for any $\epsilon, \delta > 0$:

$$\frac{|I_{\epsilon,\delta}(\mu_n)|}{n} \xrightarrow{p} 0 \text{ as } n \to \infty,$$

where

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

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 becomes aggrieved from justifiably envying, for example, someone from a random sample of finite agents (e.g., friends or neighbors), then the property will guarantee that only a vanishing fraction of individuals will suffer significant aggrievement as the economy grows large.

2.3 Two Prominent Mechanisms

As mentioned above, the existing literature and school choice programs in practice center on the following two mechanisms, and the tradeoff between the two will be an important part of our inquiry.

☐ **Top Trading Cycles (TTC) Mechanism:**

The Top Trading Cycles algorithm, originally introduced by Shapley and Scarf (1974) and later adapted by Abdulkadiroglu and Sonmez (2003) to the school choice context, has been an influential method for achieving efficiency.[13] The mechanism has some notable applications. For instance, the TTC mechanism was used until recently to assign students to public high schools in the New Orleans school system. A version of TTC is also used for kidney exchange among donor-patient pairs with incompatible donor kidneys (see Sonmez and Unver (2011)).

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 who has the highest priority with that object. Because the numbers of individuals and objects are finite, the directed graph thus obtained has at least one cycle. Every individual who belongs to a cycle is assigned the object at which he is pointing. The assigned individuals and objects 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 least one individual is removed at the end of each round. The TTC mechanism selects a matching via this algorithm for each realization of individuals' preferences and objects' priorities.

As is well known, the TTC mechanism is Pareto efficient and strategy proof (i.e., it is a dominant strategy for agents to report their preferences truthfully). As mentioned, TTC is unstable but it is constrained-stable in the sense that it is not envy-dominated by any other Pareto efficient and strategy proof mechanisms (see Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017)).

☐ **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 matching system in the US and other countries adopt DA for assigning 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)).

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

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

Step 0: Linearly order individuals in I.

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

Step $i \geq 2$**:** Let individual i make an offer to his favorite object o in O from among the objects to which he has not yet made an offer. If o does not have a tentatively accepted agent, then o tentatively accepts i. If $i = n$, end the algorithm; otherwise, iterate to Step $i + 1$. If, however, o has a tentatively accepted agent—call him i^*—object o chooses between i and i^* and tentatively accepts the one with the higher priority (or who is more preferred by o) and rejects the other. The rejected agent is named i, and we return to the beginning of Step i.

Note that the algorithm iterates to Step $i + 1$ *only* after all offers made in Step i are processed and there are no more rejections. The algorithm terminates in n Steps, with finite offers having been made. The DA mechanism selects a matching via this process for each realization of individuals' preferences and 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)). However, the DA matching is not Pareto efficient, meaning that the agents may all be better off under another matching (which is not stable).

3 Efficiency and Stability with 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. This case has been considered extensively in the computer science literature (Wilson (1972), Pittel (1989), Pittel (1992), Frieze and Pittel (1995), Knuth (1997)). In particular, those papers characterize the asymptotics of the ranks of individuals and objects under DA. Specifically, let R_i^{DA} be the rank of individual i under DA, i.e., $R_i^{DA} = \ell$ if i obtains his ℓth most favorite object under DA. Similarly, we define R_o^{DA} to be the rank of object o under DA. We will repeatedly utilize the following results.

Lemma 2 (Pittel (1989, 1992)). *Assume $K = 1$. Then,*

$$\Pr\left\{\max_{i\in I} R_i^{DA} \leq \log^2(n)\right\} \to 1 \text{ as } n \to \infty.$$

In addition, for any $\delta > 0$,

$$\Pr\left\{\frac{1}{n}\sum_{o\in O} R_o^{DA} \leq (1+\delta)\frac{n}{\log(n)}\right\} \to 1 \text{ as } n \to \infty.$$

Since both $\log^2(n)$ and $n/\log(n)$ are small relative to n as $n \to \infty$, this lemma implies that the agents and objects attain both very high payoffs in a large market. In fact, both DA and TTC involve little tradeoff when preferences are uncorrelated:

THEOREM 1. *If $K = 1$ (i.e., agents' preferences are uncorrelated), then any outcome of Pareto-efficient mechanism, and hence that of TTC, is asymptotically stable, and the outcome of DA is asymptotically efficient.*[14]

PROOF. The asymptotic stability of a Pareto-efficient mechanism follows from Lemma 1, which implies that for any $\epsilon > 0$, the proportion of the set $I_\epsilon(\tilde{\mu})$ of agents who realize payoffs less than $U(u_1, 1) - \epsilon$ in any Pareto-efficient matching mechanism $\tilde{\mu} \in \mathcal{M}_n^*$ vanishes in probability as $n \to \infty$. Because $I_{\epsilon,\delta}(\tilde{\mu}) \subset I_\epsilon(\tilde{\mu})$, asymptotic stability then follows.

The asymptotic efficiency of DA is shown as follows. Let E_1 be the event that all agents are assigned objects in DA that they rank within $\log^2(n)$. By Lemma 2, the probability of that event goes to 1 as $n \to \infty$. Now, fix any small $\epsilon > 0$ and let E_2 be the event that all agents would receive a payoff greater than $U(u_1, 1) - \epsilon$ from each of their top $\log^2(n)$ objects. Because for any $\delta > 0$, $\log^2(n) \leq \delta|O_1| = \delta n$ for a sufficiently large n, by Lemma 3-(i) in Appendix A, the probability of that event goes to 1 as $n \to \infty$. Clearly, whenever both events occur, all agents will receive a payoff greater than $U(u_1, 1) - \epsilon$ under DA. As the probability of both events occurring goes to 1, the DA mechanism is asymptotically efficient.[15] □

It is worth noting that the tradeoffs of the two mechanisms do not disappear *qualitatively* even in 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 required for the efficiency of DA and stability of TTC according to Ergin (2002) and Kesten (2006), respectively, fail almost surely as the market grows large. What Theorem 1 suggests is that the tradeoff disappears *quantitatively*, provided that the agents have uncorrelated preferences.

Uncorrelated preferences mean that the conflicts that agents may have over the goods disappear as the economy grows large, as each agent is increasingly able to find an object that he likes that others do not like. This, in turn, implies that the agents can attain high payoffs, in fact, arbitrarily close to their payoff upper bound as $n \to \infty$ under DA. This eliminates (probabilistically) the possibility that a significant fraction of agents can be made discretely better off from rematching, thus explaining the asymptotic efficiency of DA. Similarly, under

[14]Using Wilson (1972), one can show that the result of the theorem holds regardless of the objects' priorities. Hence, there is no need here to draw these randomly.

[15]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, asymptotic efficiency would still follow from the second part of Lemma 2 despite our use of a DA in which individuals are the proposers (see the proof of Proposition 1 for a formal argument). This also implies that under a DA in which schools are the proposers, in our environment (in which priorities are drawn randomly), asymptotic efficiency on the individual side can be achieved. Thus, any stable mechanism is asymptotically efficient in this context.

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 Efficiency and Stability under General Preferences

We now consider our main model in which agents' preferences are correlated. In particular, we assume that some objects are regarded by "all" agents as better than the other objects. This situation is common in many contexts such as school assignment, as students and parents tend to value similar qualities about schools (teacher and peer qualities, safety, etc.).

To consider such an environment in a simple way, we suppose that the objects are divided into two tiers O_1 and O_2 such that $|I| = |O_1| + |O_2| = n$. As assumed above, $\lim_{n\to\infty} \frac{|O_k|}{n} = x_k > 0$ for $k = 1, 2$. Our arguments in this section generalize in an obvious way to a case with more than two tiers. In addition, we assume that every agent considers each object in O_1 to be better than each object in O_2: $U(u_1, 0) > U(u_2, 1)$. In the school choice context, for instance, this feature corresponds to a situation in which students agree on the preference rankings over schools across different districts but may disagree on the rankings of schools within each district. Agents' priorities with objects are given by idiosyncratic random shocks, as assumed above.

In this environment, we will show that the standard tradeoff between DA and TTC extends to large markets even in the asymptotic sense—namely, DA is not asymptotically efficient and TTC is not asymptotically stable.

4.1 Asymptotic Instability of TTC

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

Theorem 2. *In our model with two tiers, the matching outcome of TTC is not asymptotically stable. More precisely, there exists $\epsilon > 0$ such that:*

$$\frac{|J_\epsilon(TTC)|}{n(n-1)} \stackrel{p}{\not\to} 0.$$

We provide the main idea of the proof here; the full proof is in Appendix B. In essence, the asymptotic instability of TTC arises from the key feature of this mechanism. In TTC, agents attain efficiency by "trading" among themselves the objects at which they have high priorities. This process entails instabilities because an agent could have a very low priority with an object and yet could obtain it if he has a high priority with an object that is demanded by another agent who has a high priority with the former object. This insight is well known

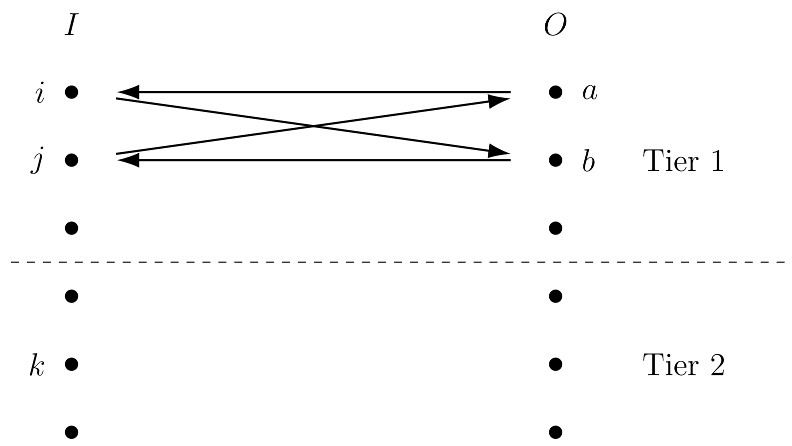


Figure 1: Possible justified envy by k toward j

but silent on the magnitude of the instabilities for a large economy. Recall, for instance, that instabilities are not significant in a large economy when agents' preferences are uncorrelated. In that case, the agents' preferences do not conflict with one another, 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 case is different, however, when their preferences are correlated. 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 obtain 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 lower priorities (with the objects they obtain) 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 points to an object, which in turn points back to that agent—a **short cycle** and any cycle of length greater than two a **long cycle**. Intuitively, the agents who are assigned via short cycles are likely to have high priorities with their assigned objects.[16] 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* (because the objects must have pointed to them), but they could have very low priorities with the objects they *trade in*. For instance, in Figure 1, agent i need not have a high priority with

[16]This is obvious for the agents assigned in the first round, as they have the highest priorities. However, even those assigned in later rounds are likely to have high priorities provided that they are assigned via short cycles: Lemma 1 implies that almost all agents are assigned within a 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.

b, although agent j does. In fact, their priorities with the objects they are assigned play no (contributory) role in the formation of such a cycle.[17] Hence, their priorities with the objects they are assigned (in O_1) are at best simple *iid* draws, and hence each of them has at most a one-half probability of having a higher priority than an agent assigned an object in O_2. This suggests that any agent assigned an object 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. In Figure 1, agent k (who is assigned an object in O_2) has probability $1/2$ of having a higher priority with b than agent i.

The crucial part of the proof of Theorem 2 is to show that the number of agents assigned an O_1 object 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" (see Bollobas (2001, Chapter 14)), we can compute the expected number of objects in O_1 that are assigned via long cycles in the first round of TTC. However, this is insufficient for our purpose because the number of objects that are assigned in the first round of TTC (which is on the order of $\sqrt{n}$) comprises a vanishing proportion of n as the market becomes large. However, extending the random mapping analysis to the subsequent rounds of TTC is difficult because the distribution of the preferences and priorities of the agents remaining after the first round depends on the specific realization of the first round of TTC. In particular, their preferences for the remaining objects in O_1 are no longer *iid*. This conditioning issue requires a deeper understanding of the precise random structure through which the algorithm evolves over rounds. We do this in Che and Tercieux (2015a). In particular, we establish that the *number* of objects (and thus of 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.[18] However, the *composition* of the cycles, in particular short versus long cycles, is subject to the conditioning issue. Nevertheless, in the Supplementary Material S.1, we are able to bound the number of short cycles formed in each round of TTC, and this bound, combined with the Markov property of the number of objects assigned in each round, produces the result.

[17] If anything, the role of their priorities 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 the object she receives in that round.

[18] The Markov Chain result is of independent interest and likely to be useful beyond the current paper. For instance, given the number of agents and objects remaining at a given stage of TTC, we explicitly derived the formula for the distribution of the number of agents matched at that step. This formula can be used to analyze the welfare of agents under TTC even for a finite economy.

4.2 Asymptotic Inefficiency of DA

Given correlated preferences, we also find that the inefficiency of DA is significant in the large market:

THEOREM 3. *In our two-tier model, the matching outcome of DA is not asymptotically efficient. More precisely, there exist $\epsilon > 0$ and a matching μ that Pareto dominates DA in each n-economy such that:*

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

as $n \to \infty$.[19]

COROLLARY 1. *Any stable matching outcome fails to be asymptotically efficient in our two-tier model.*

PROOF. The DA matching Pareto dominates all other stable matchings, as shown by Gale and Shapley (1962). Hence, any matching μ that Pareto dominates DA and satisfies the property stated in Theorem 3 will Pareto dominate any stable matching and satisfy the same property. □

The proof of Theorem 3 is in the Supplementary Material S.2; we explain its intuition here. When the agents' preferences are correlated, agents 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 apply for objects in O_1 before they ever apply for any object in O_2. The first phase of the DA is then effectively a sub-market consisting of I agents and O_1 objects with random preferences and priorities. As there are excess agents of size $|I| - |O_1|$, which grows linearly in n, even those agents who are fortunate enough to be assigned objects in O_1 must have competed to the extent that they have suffered a significant welfare loss.[20]

Indeed, note that each of the agents who is eventually assigned an object in O_2 must have made $|O_1|$ offers to the objects in O_1 before he/she is rejected by all of them. This means that each object in O_1 must receive at least $|I| - |O_1|$ offers. Then, from an agent's perspective, to be assigned an object in O_1, he must survive competition from at least $|I| - |O_1|$ other agents. The odds of this equal $\frac{1}{|I|-|O_1|}$, as the agents are all ex ante symmetric. Hence, the odds that

[19] Recall that the dependence of the matching μ that Pareto dominates DA on the n-economy is suppressed for notational simplicity.

[20] This result is obtained by Ashlagi, Kanoria, and Leshno (2017) and Ashlagi, Braverman, and Hassidim (2014) building on the algorithm originally developed by Knuth, Motwani, and Pittel (1990) and Immorlica and Mahdian (2005). Here, we provide a direct proof that is much simpler. This proof is sketched here and detailed in Appendix S.2.

an agent is rejected by his top δn choices, for small enough $\delta > 0$, is at least

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

because $\frac{|I|-|O_1|}{n} \to (1-x_1)$ as $n \to \infty$. Note that this probability approaches one as δ becomes sufficiently small. This probability is not conditional on whether an agent is assigned an object in O_1, and clearly that probability is one, conditional on the agent not being assigned any object in O_1. However, each agent *is* assigned an object in O_1 with positive probability (i.e., approaching $x_1 > 0$), and hence 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 precisely in Appendix S.2, therefore, even those agents who are fortunate 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 Lemma 1, with probability going to 1, a fraction arbitrarily close to 1 of agents assigned to O_1 objects 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.[21]

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

5 Deferred Acceptance with Circuit Breaker

As we just saw, two of the most prominent mechanisms fail to find matchings that are asymptotically efficient and asymptotically stable. Is there a mechanism that satisfies both properties? In the sequel, we propose a new mechanism that satisfies both desiderata.[22] To be more precise, we define a class of mechanisms indexed by some integer κ (allowed to be ∞). For

[21]This result is related to Ashlagi and Nikzad (2015) who show that many pairs of students would benefit from directly exchanging assignments ex post when there is a shortage of seats. Besides the "types" of reassignment, the notion of welfare gain is different, however: we focus on the agents who would benefit "discretely" from reassignment.

[22]The feasibility of attaining both asymptotic efficiency and asymptotic stability can be seen directly by appealing to the Erdös-Renyi theorem. Exploiting this theorem, one can construct a mechanism that is

a given κ, the new mechanism then modifies (the McVitie-Wilson version of) DA to finalize tentative assignments whenever an agent has made κ offers for the first time. We will show how κ can be chosen to achieve our goal. For further applicability, we next generalize the mechanism such that the assignment is triggered when a certain number $j \geq 1$ of agents have each made κ offers.[23] For expositional clarity, we begin with the simplest version and present the generalized version in a second step.

5.1 Basic Algorithm

Given a value κ, the DA with Circuit Breaker (DACB) begins by collecting agents' preference rankings of objects. Next, the agents are given serial orders: $1, \dots n$. We do not specify how the **serial orders** of the agents are determined, except to assume that they admit **basic uncertainty** from the agents' perspective: *for each* $k = 1, ..., n$, *the probability that any agent* i *receives the serial order* k *goes to zero as* $n \to \infty$. This property holds trivially if the agents' serial orders are chosen randomly according to the uniform distribution but holds much more generally, for instance, even when an agent could anticipate this distribution to some extent based on his priorities.

Given the agents' preference rankings and serial orders, DACB with index κ 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 offers an agent makes. We first initialize $\hat{I} = I$ and $\hat{O} = O$ and set the counter for each agent to zero.

Step $i \geq 1$: The agent with index i (i.e., i-th lowest serial order) in $\hat{I}$ makes an offer to his 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 agent, then o tentatively holds the agent that made the offer, and the algorithm iterates to Step $i + 1$. (The algorithm is terminated if no more students are left). If o is already holding an agent tentatively, it tentatively accepts the agent who is higher on 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 tentatively assigned 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 agents.

asymptotically efficient and asymptotically stable. However, this mechanism would not be desirable for several reasons. In particular, as we discuss in the Supplementary Material S.3, it would not have good incentive properties. By contrast, the mechanism that is proposed here does have a good incentive property, as we show below.

[23] The extended version of DACB has the additional benefit of making the mechanism robust when there are (a small number of) agents who may act irrationally and trigger the circuit-breaker prematurely. Clearly, the extended version of the DACB algorithm is not significantly affected by the presence of such agents.

Reset the counter for the agent rejected at step i to zero. If $\hat{I}$ is non-empty, return to Step 1; otherwise, terminate the algorithm.

2. If the counter for the agent who has made an offer is strictly below κ, then if he made an offer to an unmatched object, we move to Step $i+1$; otherwise, we return to the beginning of Step i where—instead of the agent with serial order i—the agent rejected by o makes an offer.

The Steps $1, ..., i$, taken until a threshold κ is reached, are called a **Stage**. Specifically, a Stage begins whenever $\hat{O}$ is reset, and the Stages are numbered $1, 2, ...$ serially. Each Stage has finite Steps, and there will be finite Stages. This algorithm modifies the McVitie and Wilson (1971) version of DA such that tentative assignments are periodically finalized.

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 predetermined serial order. A serial dictatorship is efficient but obviously fails to satisfy (even asymptotic) stability because it completely ignores the agents' priorities with the objects. By contrast, if $\kappa = +\infty$, then the DACB mechanism coincides with the DA mechanism. As demonstrated above, DA is stable but fails to be asymptotically efficient. Thus, intuitively, κ should be sufficiently large to allow agents to make enough offers (otherwise, we would not achieve asymptotic stability) but sufficiently small to avoid excessive competition by the agents (otherwise, 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 4. *If $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$, then the matching outcome of DACB is asymptotically efficient and asymptotically stable.*[24]

PROOF. See Appendix C.□

Theorem 4 shows that DACB is superior to DA and TTC in large markets when the designer cares about both asymptotic efficiency and asymptotic stability.

Roughly speaking, the idea of DACB is to endogenously segment the market into "balanced" submarkets. To appreciate this idea, consider a thought experiment wherein the designer partitions agents (for example, randomly) into K groups with the number of agents I_k in group $k = 1, ..., K$ set equal to $|O_k|$; the designer then runs DA separately for each submarket consisting of I_k and O_k. Lemma 2 then implies that, with high probability,[25] all

[24] Recall that $\kappa(n) = o(n)$ means that $\lim_{n\to\infty} \frac{\kappa(n)}{n} = 0$

[25] For a sequence of events E_n, we say that this sequence occurs *with high probability* if $\Pr(E_n)$ converges to 1 as n goes to infinity.

except for a vanishing fraction of agents would enjoy idiosyncratic payoffs and priorities arbitrarily close to the upper bounds in each submarket.[26] Asymptotic efficiency and asymptotic stability would thus follow. In particular, the segmentation avoids the significant welfare loss that would result from excessive competition for top-tier objects under DA (without segmentation). In practice, however, such a precise segmentation would be difficult to achieve because the designer would not know the exact preference structure of the agents; for instance, the designer would not know exactly which set of objects belongs to the top tier, which set belongs to the second tier, and so forth. Moreover, such an exogenous segmentation could be highly susceptible to possible misspecification of segments by the designer. DACB, with its periodic clearing of markets, achieves the necessary segmentation of the market *endogenously*, without exact knowledge on the part of the designer.

How the segmentation works under DACB—namely the proof of Theorem 4—is explained as follows. First, as $\kappa(n)$ is sublinear in n (i.e., $\kappa(n) = o(n)$), with probability approaching one as $n \to \infty$, *all* agents find their $\kappa(n)$ most preferred objects to be in O_1 (see Lemma 3 in Appendix A). Therefore, all first $|O_1|$ agents in terms of serial order would compete for objects in O_1. Because $\kappa(n) \geq \log^2(n)$, Lemma 2 implies that, with high probability, the first $|O_1|$ Steps of DACB would proceed without the threshold $\kappa(n)$ being reached by any agent, meaning that with high probability, the first $|O_1|$ Steps would proceed precisely the same as if DA were run on the "hypothetical" submarket consisting of the first $|O_1|$ agents and the objects O_1. It then follows that with high probability, the entire O_1 would be assigned without triggering the termination of the first Stage.

Next comes Step $|O_1|+1$. By then, with high probability, all objects in O_1 are assigned, and hence given the first observation (i.e., that all agents find their $\kappa(n)$ most preferred objects to be in O_1), some agent must be rejected at least $\kappa(n)$ times before the Step $|O_1|+1$ concludes, and thus the end of Stage 1 must be triggered at that Step. Since, with high probability, all objects' payoffs are arbitrarily close to the upper bound by the end of Step $|O_1|$ (by the second part of Lemma 2), this must also be the case by the end of Step $|O_1|+1$ because these objects will have received even more offers. Further, by definition, all the $|O_1|+1$ agents (except for one) participating in this stage will be matched to one of their $\kappa(n)$ top choices. Because $\kappa(n)$ is sublinear in n, by the end of Stage 1, these agents will still receive payoffs arbitrarily close to the maximum achievable when matched to O_1 objects (this is proven in Appendix A). Although the first stage is likely to end at Step $|O_1|+1$ and thus involves one more agent than the number of objects, the resulting market is "approximately" balanced, and

[26] Recall from Lemma 2 that all individuals and objects except for a fraction vanishing in probability enjoy ranks that are sublinear in n. This in turn implies that all but a vanishing fraction of these agents attain idiosyncratic payoffs arbitrarily close to the upper bounds (see Lemma 3 in Appendix A and Lemma 6 in Appendix C).

the competition among agents is still moderate because of the offer bound $\kappa(n)$.[27] The same observation applies to the subsequent Stages, suggesting that a segmentation of the market into balanced submarkets would emerge endogenously under DACB.

Several remarks are in order on the parameter $\kappa(n)$. Unlike the exogenous segmentation, the threshold $\kappa(n)$ does not depend on the precise tier structure of the objects and thus can be implemented without knowing it. Second, there is a fairly broad range of $\kappa(n)$ that produces asymptotic efficiency and asymptotic stability. This means that the performance of DACB is robust to the possible misspecification of $\kappa(n)$ on the part of the designer. Third, the precise range of $\kappa(n)$ will certainly depend on the preference structure, which may depart from that assumed in our model, but, as we illustrate in Section 6, it can be fine-tuned to a specific market based on a careful study of its data. Fourth, as proven in Supplementary Appendix S.5, the convergence is pretty fast. More specifically, the probability that the DACB with a suitably chosen $\kappa(n)$ achieves any desired degrees of efficiency and stability converges at rates faster than $1/n$.

This result is further reinforced by simulations we performed. Figure 2 shows the utilitarian welfare—more precisely the average idiosyncratic utility enjoyed by the agents—achieved by alternative algorithms, including DACB with $\kappa(n) = \log^2(n)$, under the assumption that $U(u_o, \xi_{i,o}) = u_o + \xi_{i,o}$ and $V(\eta_{i,o}) = \eta_{i,o}$, and that each of $u_o, \xi_{i,o}$ and $\eta_{i,o}$ are distributed uniformly from $[0, 1]$. Importantly, the welfare is measured under the varying market size ranging from $n = 10$ to $10,000$.[28] As expected, the TTC achieves higher utilitarian welfare, followed by DACB, and DA, and they all increase with the market size. But the levels of the utilitarian welfare as well as the rates at which the welfare increases with the market size differ across different mechanisms in a significant way. The efficiency of DACB rises quickly with market size, reaching 90% for $n = 1,000$, and above 96% for $n = 10,000$. By contrast, the efficiency ratio is fairly steady around 80-81% for DA, regardless of the size. As n rises in the range $1,000 \leq n \leq 10,000$, the DACB's gap relative to TTC narrows to 3%, while its gap relative to DA widens to 15%. This result shows that DACB performs well in efficiency even for relatively small markets. Figure 3 shows the fraction of 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

[27] Ashlagi, Kanoria, and Leshno (2017) show that a small imbalance of only one agent is enough to increase the average rank enjoyed by the agent from the order of $\log n$ to $n/\log(n)$. While even the latter rank will give rise to a high payoff in our setup, the first stage of DACB differs from the DA with a small imbalance. Due to the bound on the offer, the maximal rank to be enjoyed by the (matched) agents is $\kappa(n)$, which differs from $n/\log(n)$.

[28] 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 $n = 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000$.

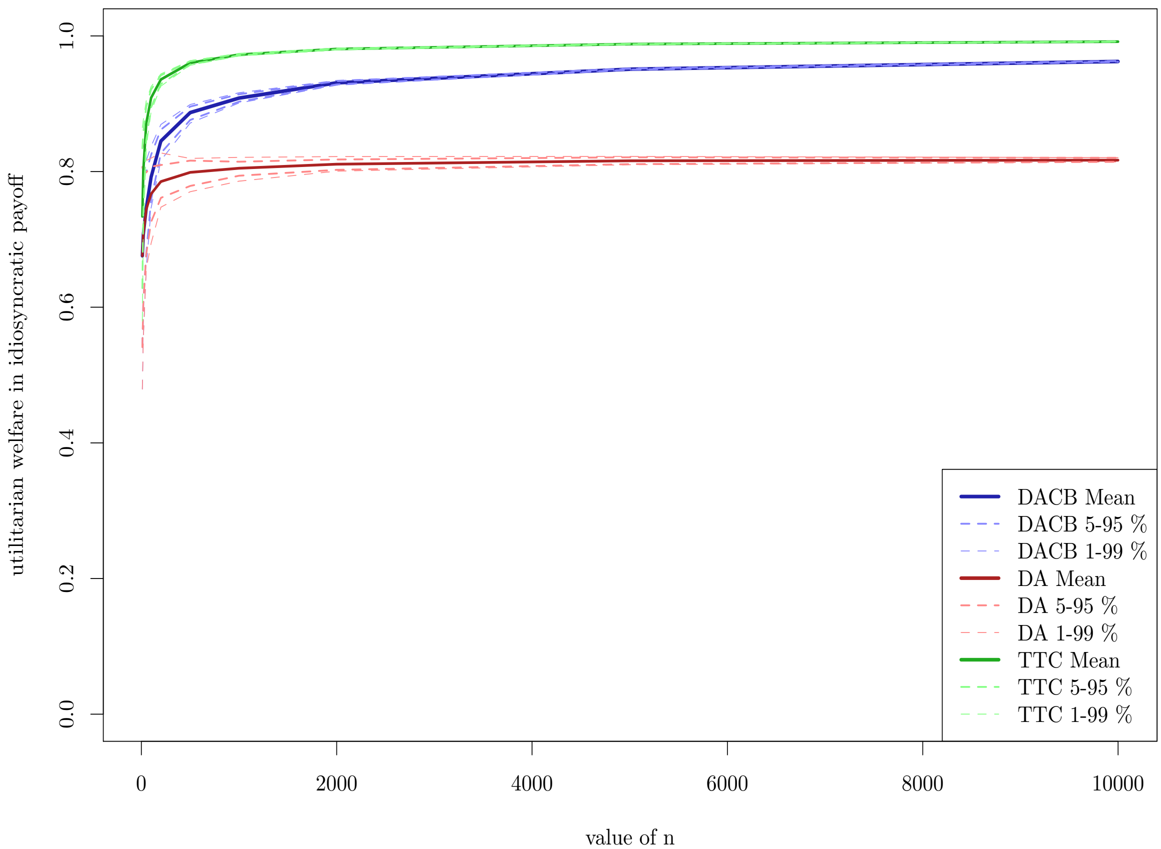


Figure 2: Utilitarian welfare under alternative mechanisms

possible pairs, and these proportions do not vary much with the market size. Our convergence rates (computed in Supplementary Appendix S.5) as well as these simulation results suggest that the DACB performs reasonably well even for moderate-size markets.

While our analysis has assumed that agents' priorities over objects are uncorrelated, one can handle correlations in priorities by approriately selecting the serial order of agents in DACB. Suppose for instance that agents' priorities—more precisely objects' utilities over agents—consist of common values in finite tiers and randomly-drawn idiosyncratic components, just like agents' utilities over objects. Then, the designer can run DACB with a serial order reflecting their priorities; namely, agents with higher common values (high tier) are ordered ahead of agents with lower common values (low tier). The asymptotic efficiency and the asymptotic stability of DACB are then preserved. Of course, this approach requires knowl-

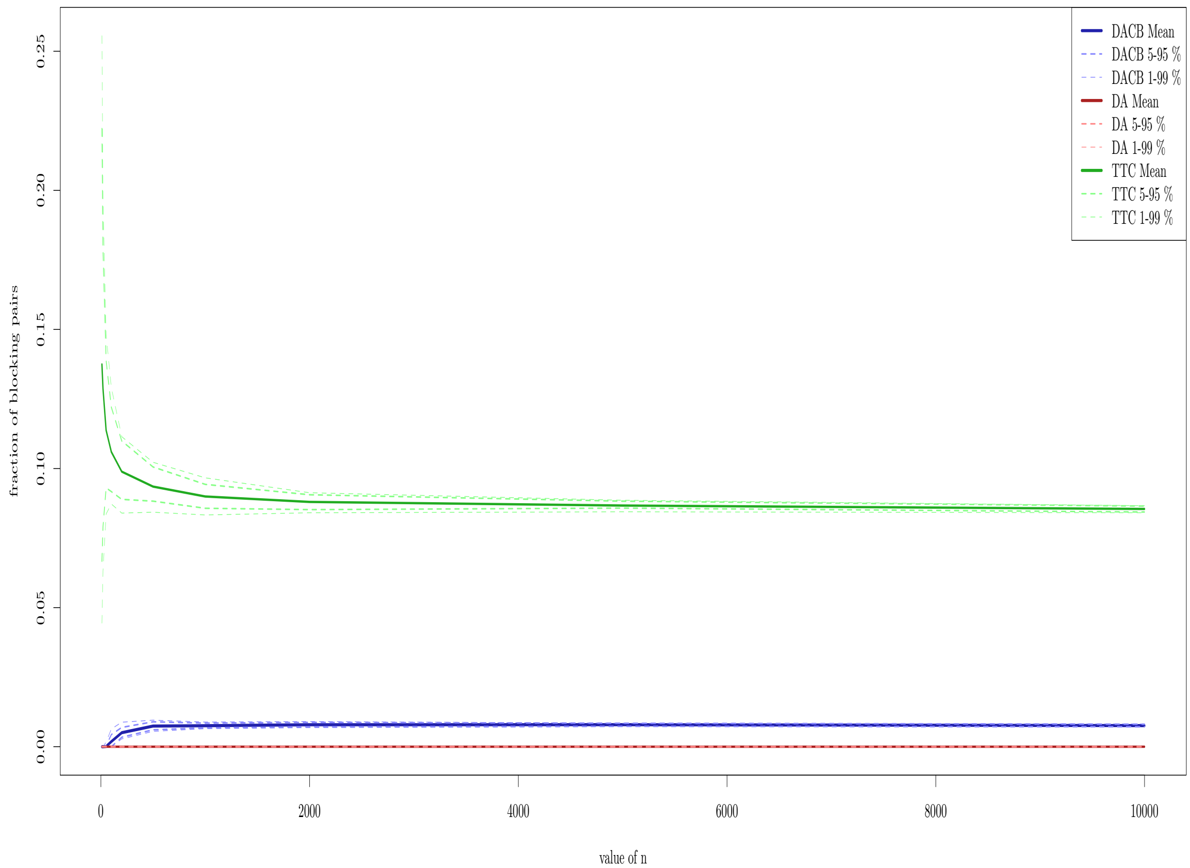


Figure 3: The fraction of blocking pairs under alternative mechanisms.

edge about agents' priorities by the designer. Such a knowledge is often available. In school choice, students' priorities are public information and known to the school system (designer). Further, our simulations with randomly generated data provided in Supplementary Material S.8 show that our results, again with the serial order chosen similarly, are robust to more general forms of priorities (and preferences).[29]

One potential drawback of DACB is that it is not strategy proof. In particular, the agent who is eventually unassigned at each Stage may wish to misreport his preferences by including

[29]When agents' priorities over objects are perfectly correlated, DA and TTC are both equivalent to serial dictatorship where the ordering is given by the common ranking of agents by the objects. Obviously, the outcome is efficient and stable. Interestingly, it is implemented by DACB for any possible κ where the serial order is also given by the common ranking.

among his κ-best ranked objects a "safe" item that is outside his top κ favorite objects but is unlikely to be popular with other agents. Such misreporting could benefit the agent because the safe item may not have received any other offer and thus would accept him, whereas truthful reporting leaves him unassigned at that stage and results in the agent receiving a worse object.[30]

However, the odds of becoming such an agent is roughly one over the number of agents assigned in the Stage; hence for an appropriate choice of κ and given the basic uncertainty over one's serial order, the odds are very small from the perspective of each agent 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 agent 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 thus far. The underlying informational environment is Bayesian: each agent only knows his own preferences, labeled his "type," and knows the distribution of others' preferences and the distribution of priorities (including his own).

DACB is an ordinal mechanism, i.e., it maps profiles of ordinal preferences reported by the agents and agents' priorities with objects into matchings. In the game induced by DACB, the set of actions by agent 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 that we denote $P_i = P(\xi_i)$. This is interpreted as the truthful report of agent i of type ξ_i. Given any $\epsilon > 0$, truth-telling is an interim ϵ-Bayes-Nash equilibrium if, for each agent i, each type ξ_i and any possible report of ordinal preferences P_i', we have

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

where $U_i(DACB(P_i,.))$ denotes the utility that i receives when he reports P_i.

THEOREM 5. *Fix any $\epsilon > 0$. Then, under DACB with $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$, there exists $N > 0$ such that for all $n > N$, truth-telling is an interim ϵ-Bayes-Nash equilibrium.*

[30] This observation can be made precise. Suppose that there are four agents and four objects. Agent 1 prefers o_1 most and o_2 second most, but he has the lowest priority with each of these two objects. Agent 1's third most preferred object is o_3, but he enjoys the highest priority with that object. Agents 2 and 3 rank o_2 and o_3, respectively, at the top of their preference lists, while agent 4 ranks o_1 first. Consider DACB with $\kappa = 2$ for this economy. Suppose first all agents report truthfully, including agent 1. One can verify that agent 1 triggers the end of Stage 1 and is assigned o_4. Specifically, in the first three Steps, agents 1, 2, and 3 apply to o_1, o_2 and o_3, respectively, and are tentatively accepted by them. In Step 4, agent 4 applies to o_1, which keeps him and rejects agent 1. Agent 1 then applies to o_2 and is rejected, at which point Stage 1 ends. In Stage 2, agent 1 is assigned object o_4. Suppose next agent 1 misreports by ranking o_3 among his two most favorite objects. Then, he can guarantee himself o_3. In sum, agent 1 benefits from misreporting his preference, suggesting that truthful reporting is not a Bayes-Nash equilibrium behavior. Nevertheless, we argue below that in the large economy, truthful reporting is a ϵ-Bayes-Nash equilibrium.

PROOF. See Supplementary Material S.6. □

REMARK 1. [Virtual asymptotic strategy-proofness] It is easy to see from our argument that a stronger incentive property can be obtained. Suppose that agents draw their reports according to any arbitrary *iid* distribution which lies in the class of distributions allowed by the current model, hence not necessarily truthfully.[31] Fix any $\epsilon > 0$. Then, under DACB with $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$, there exists $N > 0$ such that for all $n > N$, truth-telling is an (interim) ϵ-best response against any such strategy (which is not necessarily truthful). This is reminiscent of Azevedo and Budish (2015)'s notion of "strategy-proofness in the large."

Thus 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 information such as his priorities. In such a case, agent i's type would be a pair $(\xi_i, \eta_i) := (\{\xi_{i,o}\}_{o\in O}, \{\eta_{i,o}\}_{o\in O})$. Note that DACB still has good incentive properties even in this richer context. Indeed, given any $\kappa(n)$ (i.e., $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$), for any $\epsilon > 0$, it is an ex ante ϵ-Bayes-Nash equilibrium to report truthfully when the number of agents is large enough.[32]

To see this, fix $k = 1, ..., K$ and agent with serial order $i \in \{|O_{\leq k-1}| + 2, ..., |O_{\leq k}| + 1\}$ (with the convention that $|O_{\leq 0}| + 2 = 1$ and $|O_{\leq K}| + 1 = n$). As shown in Theorem 4, given truthful reporting by all agents, the agent is assigned one of his $\kappa(n)$ most preferred objects in O_k—and hence enjoys a payoff arbitrarily close to $U(u_k, 1)$—with high probability. Further, given truthful reporting by the other agents, with high probability, Stage $k' < k$ ends before agent i takes his turn, irrespective of his behavior. These two facts imply that a deviation from truthful behavior cannot make the deviating agent ϵ-better off in ex ante terms for sufficiently large n. Hence, truthful reporting is an ex ante ϵ-Bayes-Nash equilibrium. This does not imply that *all* types (ξ_i, η_i) have incentives for reporting truthfully, but it does imply that *almost all* types of agents will have incentives for truth-telling.

5.2 Extended Algorithm

In many real-world matching mechanisms, applicants are allowed or willing to list only a small number of objects. A case in point is the NYC school choice, wherein an applicant can rank only 12 choices in his/her application. In such cases, our lower bound on $\kappa(n)$ stated in

[31] More precisely, let $\{\tilde{\xi}_{io}\}_o$ be a collection of iid random variables drawn from an arbitrary distribution in $[0, 1]$. For any k and $o \in O_k$, draw $U(u_o, \tilde{\xi}_{io})$. Then, the admissable strategy is to report the ordering induced by the realized cardinal utilities. Note that $\{\tilde{\xi}_{io}\}_o$ and $\{\xi_{io}\}_o$ are all independent.

[32] Truthful reporting means reporting one's true preferences *irrespective of one's priorities.* Such a behavior is an ex ante ϵ-Bayes-Nash equilibrium if for any $\epsilon > 0$, the gain from deviating from that behavior is less than ϵ ex ante (i.e., prior to the realization of preferences and priorities) for an n that is sufficiently large.

Theorem 4 may in some instances be too large. Hence, we consider a generalization of our mechanism under which a significantly smaller lower bound can be achieved.

The new version of DACB collects preference rankings from agents and assigns them serial orders in the same manner as before. However, it is indexed by two integers j and κ. Termination of a Stage is now triggered whenever there are j individuals, each having made at least κ offers. In other words, we allow up to $j-1$ individuals to make more than κ offers before the circuit breaker is activated. Obviously, when $j = 1$, we return to our original version of DACB. Under this version of DACB indexed by $j(n)$ and $\kappa(n)$ (where n is the size of the market), we obtain the following result.

Theorem 6. *If* $\liminf_{n\to\infty} \frac{j(n)\kappa(n)}{n\log(n)} > 1$ *and* $j(n)$ *and* $\kappa(n)$ *are* $o(n)$, *then the matching outcome of DACB is asymptotically efficient and asymptotically stable.*

Proof. See the Supplementary Material S.7. □

Remark 2. [Relationship with Theorem 6] Theorem 4 is not a special case of Theorem 6. Indeed, for $j(n) = 1$, the above theorem gives $n\log(n)$ as a lower bound on $\kappa(n)$, which is obviously much greater than $\log^2(n)$ and, more generally, has no bite because, trivially, agents rank at most n objects. Theorem 6 is therefore useful only for a sufficiently large j. Finally, the arguments in the proof of each of these two results are distinct.

The new feature of DACB with (κ, j), $j \gg 1$, is that among those taking turns in Stage k, up to j agents will likely fail to receive objects in O_k. However, as the market grows large, j becomes very small relative to the number of agents assigned during that Stage. Hence, given (a suitable generalization of) basic uncertainty regarding the serial order, each agent finds the odds of being one of such agents or unmatched negligible in the large economy. This feature ensures that the extended algorithm retains the same desirable incentive properties as the basic algorithm. Specifically, in the Supplementary Material S.7, we show that Theorem 5 extends to DACB with (κ, j), satisfying the condition of Theorem 6 (see Theorem S4).

Remark 3. [Effects of preference correlation] Theorems 2 and 3, together with Theorem 1, suggest that increased correlation in agents' preferences over objects exacerbates the tradeoff between efficiency and stability, worsening welfare under DA and reducing stability under TTC. Meanwhile, Theorems 4 and 6 show that DACB is asymptotically efficient and stable, even when common value differences among objects are large. This latter result is a limit result, however, so it does not speak to the effect of preference correlation *for a given finite economy*. To gain more insight on this latter issue, we performed simulations (see Supplementary Material S.8 for details). These simulations reveal that, as preferences become more correlated, DACB performs worse in efficiency and stability in *absolute terms*, but its performance improves in *relative terms* compared with DA or TTC. These simulations also suggest how the designer may adjust (κ, j) with the change in preference correlations.

5.3 DA with Constrained ROLs

The main feature of DACB—namely, limiting participants' choices—is reminiscent of mechanisms employed in some centralized matching procedure. A prominent example is the DA with constrained ROLs (henceforth DAC), a variant of DA in which applicants' ROLs cannot exceed a fixed length. While DAC is very common in practice,[33] authors have found it difficult to rationalize the constraint on choices.[34] Our perpective recognizes one redeeming quality of this practice in limiting the harmful effect of competition. Indeed, we can show at least in a two-tier model that the DAC admits an ε-Bayesian Nash equilibrium whose outcome is asymptotically efficient and asymptotically stable, *provided that the constraint is chosen appropriately.* This offers some justification for the use of DAC.

To begin, consider DAC with length chosen at $\kappa(n)$ for each n-economy. Then, for each n-economy, DAC induces a Bayesian game in which each agent i's type comprises a vector $\xi_i := (\xi_{io})_o$, and his (pure) strategy maps his type to a ROL of $\kappa(n)$ objects. We are interested in the matching outcome induced by the sequence of strategy profiles adopted by the agents for all n-economies.

Theorem 7. *Assume that $K = 2$. Consider the DAC with length $\kappa(n)$. Let $\kappa(n) = o(n)$ and satisfy $\kappa(n)/\log^2(n) \to \infty$ as $n \to \infty$. There is a sequence of strategy profiles satisfying the following two properties. (1) For any $\varepsilon > 0$, and n large enough, the strategy profile for each n-economy is an ex-ante ε-Bayes Nash equilibrium, and (2) the induced matching outcome is asymptotically efficient and asymptotically stable.*

The proof of Thoerem 7, which is available in Appendix D, is sketched as follows. For the argument, we consider a sequence of strategies $\sigma_{x(n)}$, indexed by $x(n) \in (0, 1)$, wherein each agent lists his $x(n)\kappa(n)$ most favorite objects in O_1 first and then $(1-x(n))\kappa(n)$ most favorite objects in O_2 next, truthfully ordered within each block. The argument then mimics that of Theorem 1. If $x(n)$ converges sufficiently slowly to 1 as $n \to \infty$, then, for the chosen $\kappa(n)$, as the market grows large both $x(n)\kappa(n)$ and $(1 - x(n))\kappa(n)$ become large enough so that with high probability all objects are assigned and "enjoy" payoffs arbitrarily close to $V(1)$. At the same time, both $x(n)\kappa(n)$ and $(1-x(n))\kappa(n)$ are sublinear in n, which ensures that all agents enjoy idiosyncratic payoffs arbitrarily close to the upper bound. Finally, the incentive to deviate unilaterally from such a strategy vanishes as the market grows large. First of all, given ex ante symmetry, there is no incentive to manipulate the relative rankings of objects within a tier. Further, deviating by lowering x below $x(n)$ does not significantly raise the

[33] For example, Chicago, New York City, Ghana, Spain, Turkey adopt DAC in school assignment. See Fack, Grenet, and He (2015) for references.

[34] Calsamiglia, Haeringer, and Klijn (2010) find in their laboratory experiment that constrained choices lead to preference manipulations and unstable matchings. In light of these problems, the practice is viewed as puzzling (see Pathak (2016) for instance).

probability of assignment (in light of the near full assignment result above), and raising x above $x(n)$ does not significantly raise the probability of getting a first-tier object when $x(n)$ is already close to 1.

While Theorem 7 suggests that DAC may have a similar benefit as DACB, we view the former as inferior to the latter in several respects. First, the DAC requires agents to behave strategically in the right way to achieve the desirable outcome.[35] By contrast, the DACB is virtually strategyproof in the sense discussed in Remark 1, and hence requires little strategic coordination among agents. Second, a mistake in strategizing under DAC may be severe since an agent may simply be unassigned. In DACB, such a risk is mitigated by the feature that those unassigned in any given stage can still participate in the subsequent stage with an additional "budget" of κ offers. Finally, Theorem 7 requires the length of ROLs to be at least $\log^2(n)$, which is often impractically large.[36] We expect the mechanism to perform much worse when the constraint is more realistic. By contrast, DACB can be adapted to work with a "small" number of choices, as suggested in Section 5.2.

Remark 4 (Chinese parallel mechanism). As we already pointed out, a drawback of the DAC is that once a student exhausts her (constrained) ROL he gets unassigned. DACB mitigates this risk by "replenishing" the choice quota for students unassigned after each stage. Alternatively, one can repeat the DAC in multiple stages: in each stage all remaining agents participate in a DAC with some length κ. This is precisely what Chen and Kesten (2017) calls (symmetric) Chinese-parallel mechanism.[37] A result similar to Theorem 7 would apply to this mechanism for an appropriately chosen κ. While this mechanism mitigates the strategic risks facing participants better than DAC, it will not eliminate them, particularly for a realistic level of κ. By comparison, the (extended) DACB virtually eliminates the risk and thus incentivizes the agents to report truthfully.[38]

[35] Note further that notion of equilibrium here is "ex ante" ε-Bayesian Nash equilibrium, which is not as satisfactory as "interim" ε-Bayesian Nash equilibrium used for DACB.

[36] For instance, $\log^2(n)$ is 73 if $n = 5,000$, 85 if $n = 10,000$, and 133 if $n = 100,000$.

[37] As noted in Chen and Kesten (2017), the mechanisms used for assigning students to Chinese universities are slightly different. College programs are partitioned into several tiers differing in "prestige." And, the Chinese parallel mechanism is run first for the first-tier (most prestigious) colleges. Next, those unassigned move to the second round, where the same mechanism is run for the second-tier colleges. The same process is repeated for each subsequent tier. See also Wu and Zhong (2014) for additional details.

[38] More precisely, under the Chinese-parallel mechanism, if all agents were to report truthfully, a significant number of agents (i.e., linear in n) would be active (i.e., make offers) but end up unassigned in each round (except for the final) for any $\kappa(n)$ sublinear in n. This means that each participant faces a significant incentive to manipulate preferences by moving up safe items in his ROL. By contrast, in DACB, only a small (sublinear) number of active agents get unassigned in each stage when reporting truthfully.

6 Field Application: NYC School Choice

Thus far, we have considered a large one-to-one matching market (in the limiting sense) with a class of random preferences and priorities. Real-world matching markets often depart from this model; for instance, school choice involves many-to-one matching. Further, the size of the market may not be very large. Our theory may not apply exactly in these settings. Nevertheless, it is interesting to assess whether DACB could offer a more desirable compromise on the tradeoff between efficiency and stability in such a case. We thus study the school choice in NYC.

In New York City, each year approximately 80,000 middle school students (mostly 8th graders) are assigned to over 700 public high school programs through a centralized matching process. The process involves multiple rounds, but the main round, Round 2, employs the deferred acceptance algorithm to assign participants to programs in several categories: *screened, limited unscreened, unscreened, ed-op, zoned* and *audition*.[39] Each participant in the main round may submit a rank-order list (ROL) of up to 12 programs, and each program ranks applicants—who listed the program in their ROLs—according to its priority criteria, which depend on the type of the program. The priorities are coarse for many programs, and ties are broken by a single (uniform) lottery for all programs.

Our analysis focuses on the 2009-2010 assignment. We calibrate the performances of DACB with several (κ, j)'s against DA and TTC as benchmarks. In so doing, we take two different approaches.

- **Counterfactual based on observed ROLs**: This approach postulates that applicants would submit the same ROLs under counterfactual scenarios as they submitted under the NYC matching.
- **Counterfacutal based on structual estimates of preferences**: This approach structurally estimates the preferences of applicants and simulates their ROLs under counterfactual scenarios, assuming that all programs are acceptable.

The first approach essentially rests on the following two assumptions: (i) that the programs in ROL are truthfully ranked and dominate all other unranked ones; and (ii) that the unranked programs are not acceptable for the applicants. Assumption (i), often called *weak truth-telling*, is a standard assumption made when dealing with a strategyproof mechanism such as DA.[40]

[39] Assignment to the so-called specialized "exam" schools is processed through the first round, which takes place before the main round. Since 2010, the first round and the main round have been merged into a single round, but the process for the main round remains unchanged.

[40] This assumption is not entirely innocuous though since the strategy-proofness of DA does not apply when the applicants' ROLs are truncated (see Haeringer and Klijn (2009)). Nevertheless, about 80% of participants

Assumption (ii) is more controversial since the presence of a supplementary round means that some applicants will choose not to list all acceptable programs in the main round. In fact, as many as 5,241 students out of 5,611 unassigned by the end of the main round listed new additional programs in a supplementary round. For this reason, the observed ROLs do not typically include all acceptable programs. In other words, this method postulates too "short" an ROL for an applicant. As will be seen, the short ROLs will *understate* the effect of competition, and therefore the tradeoff between efficiency and stability and the performance of DACB.

This problem motivates the second approach—the structural estimation method based on Abdulkadiroglu, Agarwal, and Pathak (2015) (henceforth AAP)—which invokes only the weak truth-telling assumption (i). Under this approach, we first estimate (random) utilities as functions of student and program characteristics, and use the estimates to "complete" the applicants' ROLs. One advantage of this approach is that we can actually represent the true DA, without any constraints on ROLs or a follow-up supplementary round, just as considered in the theory. At the same time, since the AAP method does not consider outside options, the simulated/predicted ROLs would include *all* programs, including those that applicants may find unacceptable. Hence, this method posits too "long" an ROL for an applicant. As will be seen, this feature will tend to *overstate* the effect of competition, and therefore the tradeoff between efficiency and stability and the performance of DACB.

In sum, the two approaches allow us to provide (lower and upper) bounds on the relative performance of DACB, and the tradeoff between DA and TTC. We thus view them as mutually complementary. We now present the results for each of the two approaches described above.[41]

6.1 Comparison of Mechanisms Based on Observed ROLs

Table 1 describes average performances of alternative algorithms according to various measures.[42]

do not fill up their ROLs, suggesting that the tuncation is not binding. A similar approach is followed by Abdulkadiroglu, Agarwal, and Pathak (2015), Abdulkadiroglu, Pathak, and Roth (2009), Abdulkadiroglu, Che, Pathak, Roth, and Tercieux (2017).

[41] A student's priorities at alternative NYC schools are likely to be correlated. As we already suggested, in such an environment, one can improve the performance of DACB by adjusting the agents' serial orders to reflect their average priorities. In the subsequent analysis, we ignore this possibility and simply employ a random serial ordering. In a previous version of this work (Che and Tercieux (2015b)), we also measured the performance of DACB when the serial order reflects the agents' priorities. As expected, the performance is significantly better. In this sense, the subsequent results understate the potential benefit of the DACB if one can "optimize" serial orders.

[42] The average here is taken over 100 independent draws of a single lottery used to break ties in schools' priorities. In particular, this means that the values reported for DA does not coincide with the realized outcome in 2009-10 assignment.

Table 1: The efficiency and stability of alternative mechanisms

	DA	DACB-$(2, 20000)$	DACB-$(4, 2000)$	DACB-$(6, 1)$	TTC
# Pareto Improvable	5189.89 (241.98)	3654.19 (80.07)	2409.60 (65.56)	449.43 (64.89)	0
# students with envy	0	2041.05 (179.83)	4620.43 (251.46)	13268.26 (502.03)	18943.21 (324.67)
# assigned to top choice	40370.88 (389.10)	41696.43 (239.16)	42966.22 (208.03)	45098.69 (190.77)	45109.77 (200.80)
# unassigned	4362.98 (166.72)	4645.11 (175.51)	4978.56 (176.47)	5601.75 (161.91)	5624.98 (158.68)

Note: We ran 100 iterations of each algorithm with independent draws of lotteries, and focus on the average performance of each algorithm, including DA. The numbers inside the parentheses are standard errors.

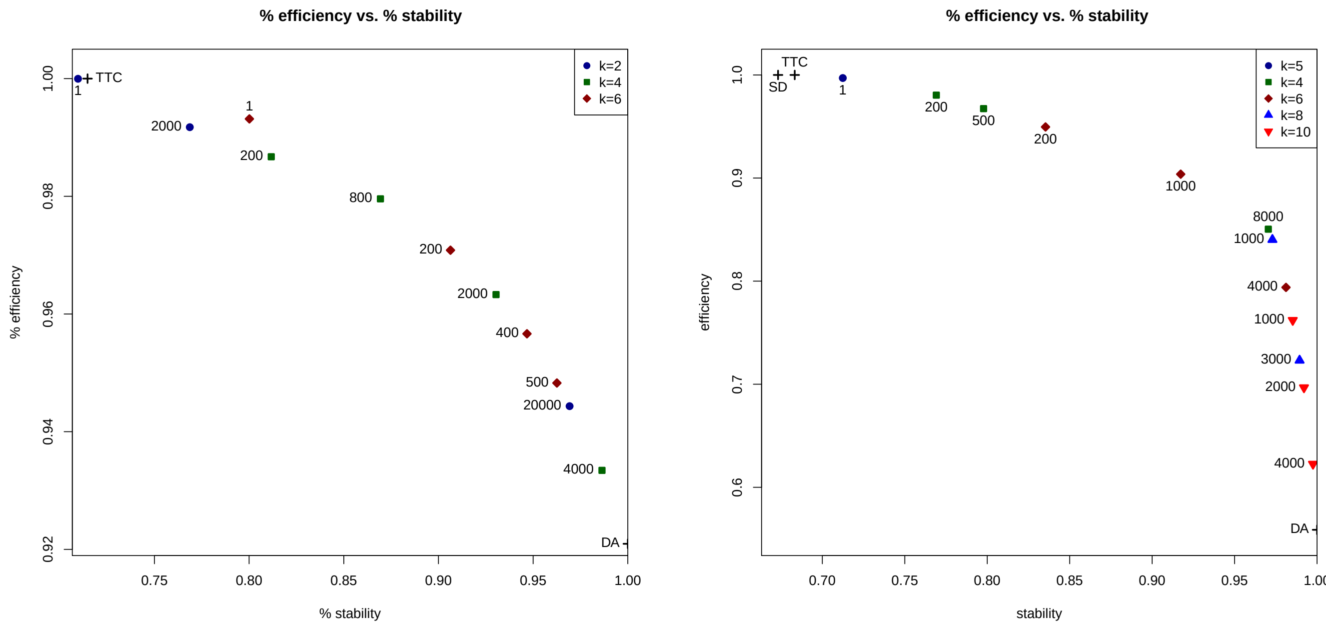
The first row describes, for each mechanism, the average number of students who can be made better off from a Pareto-improving reassignment of the original outcome using the Shapley-Scarf TTC. This number is zero for an efficient matching algorithm such as TTC. Arguably, the larger this number is the more inefficient a matching is. Hence, this number can be interpreted as a measure of inefficiency. The second row counts the number of agents with justified envy (i.e., who are part of at least one blocking pair) in each mechanism. Obviously, DA admits no such agent. As expected, TTC may admit a large number of agents with justified envy.[43] DACB provides a compromise between DA and TTC, yielding higher efficiency than DA and lower instability (i.e., a lower number of applicants with justified envy) than TTC.

As noted, the DACB is flexible enough to admit many new options to the designer's arsenal of policy tools. Figure 4-(a) depicts the range of ways in which the tradeoff between efficiency and stability is resolved via DACB with various (κ, j)'s. In these figures, efficiency (the vertical axis) is measured as the percentage of agents who cannot be made better off through Pareto-improving reallocation, while stability (the horizontal axis) is measured as the percentage of students who do not have any justified envy.[44]

Not surprisingly, DA and TTC occupy the southeast and northwest extreme corners of the figure. In between the two, DACB with various (κ, j) values spans a rich array of compromises between the objectives. As expected, the efficiency of DACB increases as κ falls, while its stability increases as κ rises. The last row presents the number of applicants assigned to their top choice. The "frontier" is outside the linear segment between DA and TTC, suggesting

[43]Serial dictatorship which corresponds to DACB with $(\kappa, j) = (1, 1)$ totally ignores the school priorities and admits slightly more agents with envy than TTC.

[44]Specifically, efficiency is defined as $1 - \frac{\text{\# of Pareto-improvable agents}}{\text{\# of total agents}}$ where the # of Pareto-improvable agents corresponds to the number of agents who are better off when running Shapley-Scarf TTC on top of the mechanism. As for stability, our percentage is defined by $1 - \frac{\text{\# of agents with envy}}{\text{\# of total agents}}$ where the # of agents with envy corresponds to the # of agents who are part of at least one blocking pair.



(a) Based on Observed ROLs

(b) Based on Structural Estimates

Figure 4: Efficiency vs. Stability: based on two methods (in ordinal measures).

Note: The shape of each coordinate corresponds to κ, while the associated integer refers to parameter j.

that the outcomes of DACB are superior to a simple convexification of DA and TTC.

As we already argued, the short ROLs observed in the data means that the exercise understates the true competition that applicants will have under unrestricted DA. Thus, our calibration potentially overstates the efficiency performance of DA. Likewise, the short ROLs also mean that our calibration is likely to "miss" incidences of justified envy that will arise under TTC, meaning it overstates the stability of TTC. Overall, the calibration may understate the magnitude of the tradeoff between efficiency and stability and therefore understate the relative performance of DACB.

In sum, the above results can be interpreted as conservative estimates of the tradeoff between DA and TTC and the benefits achievable from DACB.

6.2 Comparison of Mechanisms Based on Structural Estimates

In this section, we use demand estimation for school programs to complete applicants' ROLs. To this end, we estimate a random utility model. In this model, the utility of student i for school program j is given by

$$u_{ij} = \delta_j + \sum_{\ell} \alpha^{\ell} z_i^{\ell} x_j^{\ell} + \sum_k \gamma_i^k x_j^k - d_{ij} + \epsilon_{ij}, \tag{2}$$

where $\delta_j = \mathbf{x_j}\beta + \xi_j$, $\mathbf{x_j}$ is a vector of program j's observed characteristics, $\mathbf{z_i}$ is a vector of observed students' characteristics, ξ_j is a program-specific unobserved vertical characteristic, γ_i captures idiosyncratic tastes for program characteristics and ϵ_{ij} captures idiosyncratic tastes for programs. Finally, d_{ij} is the distance measured in miles between student i and program j's geographic locations. We further assume that $\gamma_i \sim \mathcal{N}(0, \Sigma_\gamma)$, $\xi_j \sim \mathcal{N}(0, \sigma_\xi^2)$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The vector of parameters we estimate is $(\alpha, \beta, \Sigma_\gamma, \sigma_\xi^2, \sigma_\epsilon^2)$. As noted by AAP, this model is an ordered choice version of the model used by Rossi, McCulloch, and Allenby (1996), who show that these distributional assumptions allow for estimation via Gibb's sampling.[45] The specification treats distance as a numeraire: the coefficient -1 on distance is a scale normalization (assuming students dislike to travel) which allows us to measure utility in distance units.

School program characteristics include the math achievement of the student body, the percentage of students receiving subsidized lunch, the percentage of white students, the size of 9th grade, a dummy indicating a language-focused program (coded as Asian, Spanish or others) and a dummy on the program type.[46] Student characteristics include academic performance (in math and English), ethnicity, a subsidized lunch status, neighborhood income, proficiency in English and a special ed status.

Our estimates are reported in Supplementary Material S.10. They are largely consistent with those found by AAP based on 2003-2004 data. More important, the sign and magnitude of parameters are reasonable and intuitive. We use these estimates to draw each student's ROLs according to (2) as well as his/her priorities (which includes single tie-breaking lottery).[47] With these as inputs, we simulate counterfactual algorithms, including DA (without truncation and without the supplementary round), TTC, and DACB with various (κ, j)'s. We compute the average performance of each mechanism over 100 realizations of these random draws.

Table 2 reports the analogues of Table 1. The first three rows are the same as the corresponding rows in Table 1, measuring the number of students who can be made better-off from

[45] See the online appendix of AAP for details.

[46] The NYC high school directory describes numerous program types. As in AAP, these program types were aggregated into different categories: Arts, Humanities/Interdisciplinary, Business / Accounting, Math / Science, Career, Vocational, Government/law, Zoned and Others.

[47] There are two reasons why students' priorities are random. First, priorities at school programs can be coarse and we use a single tie-breaking rule to break ties. Second, as explained in the supplementary material (see Section S.9), we had to estimate the distribution of priorities of students (as a function of students' observables) at schools which are not in their observed ROLs.

Table 2: The efficiency and stability of alternative mechanisms

	DA	DACB-$(8, 3000)$	DACB-$(4, 8000)$	DACB-$(4, 200)$	TTC
# Pareto Improvable	29293.28 (339.54)	18382.59 (141.92)	9931.34 (101.85)	1299.98 (62.09)	0
# students with envy	0	703.73 (20.33)	1963.13 (32.89)	15329.54 (162.49)	21029.31 (250.66)
# assigned to top choice	23823.26 (215.65)	30671.82 (102.43)	36502.27 (71.70)	43674.35 (64.08)	44060.72 (48.77)
average welfare	-2.63 (0.031)	-1.93 (0.006)	-1.60 (0.006)	-1.38 (0.008)	-1.31 (0.005)
average justified envy	0	0.11 (0.004)	0.25 (0.006)	0.99 (0.011)	1.28 (0.018)

Note: We ran 100 iterations of each algorithm with independent draws of lotteries, and focus on the average performance of each algorithm, including DA. The numbers inside the parentheses are standard errors.

a Pareto-improving reassignment, the number of students with justified envy and the number of students getting their top choice, all averaged over 100 iterations.[48]

The "competition" effect of long lists is apparent: TTC performs considerably worse in stability than in Table 1 (2,000 more applicants would suffer justified envy), and DA performs considerably worse than in Table 1 (about 25,000 more students would benefit from a Pareto-improvement). As expected, DACB performs impressively. For instance, DACB with $(\kappa, j) = (8, 3000)$ yields considerably higher welfare with very little sacrifice in stability, compared with DA. In general, Figure 4-(b) promises a much more significant improvement to be achievable by DACB, compared with Figure 4-(a).

So far, we have used our cardinal utility estimates only to predict ordinal performances of alternative mechanisms. We can also measure their performances in cardinal utilities. The penultimate row of Table 2 measures the average (utilitarian) welfare of students for each mechanism, and the last row measures the average justified envy in terms of utility gain from fufilling one's justified envy.[49] Figure 5 replicates Figure 4-(b) using cardinal measures of efficiency and stability. Here again, DACB with various (κ, j)'s spans a rich array of compromises between efficiency and stability. As with Figure 4-(b), the "frontier" is significantly curved, suggesting the potential for DACB to act as desirable compromise on efficiency and stability.

To conclude, these outcomes provide a rich set of new choices from which a policy maker can choose. A careful study of data, as illustrated here, could help a policy maker to tailor the design of DACB to fit his/her sense of the social weighting of the two objectives.

[48]As explained in the Supplementary Material (Section S.9), the number of seats equals the number of students. Since we assume students rank all programs in our counterfactual analysis, there are no unassigned students under the mechanisms we study.

[49]For each student, we compute the utility difference between his assignment and his most preferred school program with which he could block. We then average this number over all students.

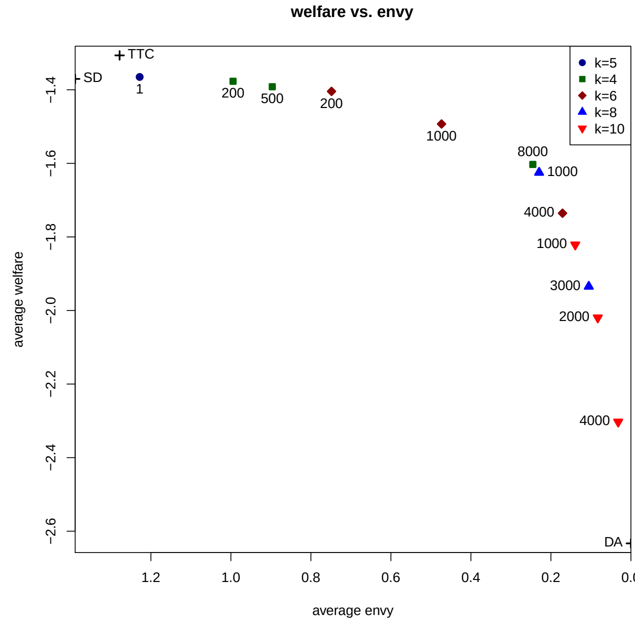


Figure 5: Efficiency vs. Stability: based on structural estimates (in cardinal measures).

Note: The shape of each coordinate corresponds to κ, while the associated integer refers to parameter j.

7 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 Shapley's deferred acceptance (DA) and top trading cycles (TTC), each satisfy one property but fail to satisfy the other. Considering a plausible class of situations in which individual agents have preferences drawn randomly according to common and idiosyncratic shocks and priorities drawn in an *iid* fashion, we show that these failures—the inefficiency of DA and instability of TTC—remain significant even in large markets.

We have therefore proposed a new mechanism, deferred acceptance with a circuit breaker (DACB), which modifies DA to keep agents from competing excessively for over-demanded objects—a root cause of DA's significant efficiency loss in a large market. Specifically, the proposed mechanism builds on McVitie and Wilson's version of DA in which agents make offers one at a time along a predetermined serial order. However, during the process, whenever an agent makes a certain threshold number of offers for the first time, the process is stopped, and what had been a tentative assignment up to that point is finalized. Thereafter, a new stage of the serialized process is begun with the remaining agents and objects, again with the same circuit-breaker feature, and this process is repeated until all agents are processed.

We have shown that DACB with suitably chosen parameters κ 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.

Although our analytical model is not without restriction, our analysis of the NYC school choice data validates our overall findings. Specifically, we have found that the inefficiencies of DA and instabilities of TTC are significant and that DACB offers viable compromises on the tradeoff between efficiency and stability. In addition, the numerous simulations we performed confirm that the main results hold well beyond the setting we study and in particular for market sizes that are quite moderate. In that respect, it is interesting to compare our results with those obtained by Lee and Yariv (2017). They show that *stable* mechanisms are asymptotically efficient in a balanced market if the agents' preferences and priorities have common shocks distributed continuously over an interval. By contrast, Ashlagi, Kanoria, and Leshno (2017) and the current paper note that DA is likely to be asymptotically inefficient when there is competition among agents for desirable objects—either because of a scarcity of objects (when there is imbalance) or because of a positive correlation in agents' preferences. Indeed, Supplementary Material S.8 shows that the inefficiency of DA vanishes very slowly even in the environment of Lee and Yariv (2017) and that the magnitude of the difference between DACB and DA can be considerable for realistic market sizes. Recall also our analysis of NYC school choice, which shows that DA entails a significant efficiency loss compared with DACB.[50] Finally, and potentially more important, our results regarding the asymptotic efficiency and asymptotic stability of DACB are robust to the introduction of market imbalances, which is not the case for DA.

Another important design parameter of DACB is the threshold number of offers that triggers assignment. Here, again, it can be optimized relative to the detailed features of the market in question. While theoretical results show that DACB achieves an asymptotically efficient and stable outcome when the market grows arbitrarily large, for a finite market, there will always remain some (potentially small) trade-off between the two objectives. Our calibration work on NYC as well as our simulations reported the in Supplementary Material S.8 show that DACB offers a range of possible compromises between efficiency and stability depending on the specific value of the trigger chosen by the designer. Thus, the serial order of agents and the condition that triggers the circuit breaker can be fine-tuned toward the specifics of a given market.

Finally, our proposed mechanism shares several features of mechanisms that are already in use. As discussed, the truncation of the ROLs is another common feature employed in many centralized matching procedures (see Haeringer and Klijn (2009), Calsamiglia, Haeringer, and

[50]See also Che and Tercieux (2015c) for a discussion of the inefficiencies of DA in the Lee and Yariv (2017) environment, particularly compared with standard efficient mechanisms.

Klijn (2010) and Ashlagi, Nikzad, and Romm (2015)). The "staged" clearing of markets is observed in matching markets such as college admissions in China. The current paper sheds some light on the roles that these features may play, particularly in mitigating the harmful effect of excessive competition among participants, and suggests a method for harnessing these features without jeopardizing participants' incentives.

References

Abdulkadiroglu, A., N. Agarwal, and P. A. Pathak (2015): "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match," *American Economic Review*, forthcoming.

Abdulkadiroglu, A., Y.-K. Che, P. Pathak, A. Roth, and O. Tercieux (2017): "Minimizing Justified Envy in School Choice : The Design of New Orleans OneApp," NBER Working Paper No. 23265.

Abdulkadiroglu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding 'Choice' in School Choice," *American Economic Journal: Microeconomics*, 7, 1–42.

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. Sonmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Ashlagi, I., M. Braverman, and A. Hassidim (2014): "Stability in large matching markets with complementarities," *Operations Research*, 62, 713–732.

Ashlagi, I., Y. Kanoria, and J. D. Leshno (2017): "Unbalanced Random Matching Markets," *Journal of Political Economy*, 125, 69–98.

Ashlagi, I., and A. Nikzad (2015): "What Matters in School Choice Tie-Breaking? How Competition Guides Design," MIT, Unpublished mimeo; abridged version published in ACM16.

Ashlagi, I., A. Nikzad, and A. Romm (2015): “Assigning More Students to Their Top Choices: A Tie-Breaking Rule Comparison,” MIT, Unpublished mimeo.

Azevedo, E. M., and E. Budish (2015): “Strategy-proofness in the Large,” *Unpublished mimeo, University of Chicago and University of Pennsylvania.*

Azevedo, E. M., and J. D. Leshno (2016): “A supply and demand framework for two-sided matching markets,” *Journal of Political Economy*, 124, 1235–1268.

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.

Budish, E. (2011): “The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes,” *Journal of Political Economy*, 119, 1061–1103.

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.

Che, Y.-K., and O. Tercieux (2015a): “An Analysis of Top Trading Cycles in Two-sided Matching Markets,” Columbia University and PSE, Unpublished mimeo.

——— (2015b): “Efficiency and Stability in Large Matching Markets,” Columbia University and PSE, Unpublished mimeo.

——— (2015c): “Payoff Equivalence of Efficient Mechanisms in Large Markets,” forthcoming, *Theoretical Economics*.

Chen, Y., and O. Kesten (2017): “Chinese College Admissions and School Choice Reforms: A Theoretical Analysis,” *Journal of Political Economy*, 125, 99–139.

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.

Fack, G., J. Grenet, and Y. He (2015): "Beyond Truth-Telling: Preference Estimation with Centralized School Choice," Paris School of Economics, Unpublished mimeo.

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.

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

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

Knuth, D. E. (1997): *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 (2009): "Incentives and Stability in Large Two-Sided Matching Markets," *American Economic Review*, 99, 608–627.

Lee, S. (2017): "Incentive Compatibility of Large Centralized Matching Markets," *Review of Economic Studies*, 84, 444–463.

Lee, S., and L. Yariv (2017): "On the Efficiency of Stable Matchings in Large Markets," University of Pennsylvania, Unpublished mimeo.

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

Pathak, P. (2016): "What Really Matters in Designing School Choice Mechanisms," forthcoming,Advances in Economics and Econometrics, 11th World Congress of the Econometric Society eds. Larry Samuelson.

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.

Rossi, P., R. McCulloch, and G. Allenby (1996): "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15, 321–340.

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.

Wilson, L. (1972): "An analysis of the stable marriage assignment algorithm," *BIT*, 12, 569–575.

Wu, B., and X. Zhong (2014): "Matching mechanisms and matching quality: Evidence from a top university in China," *Games and Economic Behavior*, 84, 196 – 215.

A Preliminary Lemma

The following lemma will be used several times in what follows.

Lemma 3. *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 each agent in I, (i) each of his top $\delta|O_k|$ favorite objects in $O_{\geq k} := \cup_{\ell \geq k} O_\ell$ yields a payoff greater than $U(u_k, 1) - \epsilon$* and *(ii) every such object belongs to O_k; moreover, for each object in O, (iii) each of its top $\delta|I|$ individuals in I have priority scores of at least $V(1) - \epsilon$.*

For the proof, 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 $\xi_{io} \geq 1 - \tilde{\epsilon}$. Then, for any $\delta < \tilde{\epsilon}$, $\Pr\{\exists i$ 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:

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

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

Proof of Lemma 3. It should be clear that the third part of the statement can be proven using the same argument used for the second part. In addition, for a $\epsilon > 0$ sufficiently small that for each $k = 1, \ldots, K-1$, $U(u_k, 1) - \epsilon > U(u_{k+1}, 1)$, 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. Thus, in the sequel, we only prove the first part of the statement.

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 high probability, all individuals in I have at least $\delta|O_k|$ objects os 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. □

B Proof of Theorem 2

To begin, define a random set:

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

Lemma 4. *There exist $\gamma > 0, \delta > 0, N > 0$ such that*

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

for all $n > N$.

PROOF. See the Supplementary Material S.1. □

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 objects in O_2 under TTC. 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 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 because the common value difference between O_1 and O_2 objects is large, if $o \in \hat{O} \subset O_1$ and $j \in I_2$, it must be the case 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, as $o \in O_1$, this means that j must be pointing to an object in O_1 when she is cleared, which contradicts $j \in I_2$). 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 given realizations $\boldsymbol{\xi} := (\xi_{io})_{io}$ and $\boldsymbol{\eta} := (\eta_{io})_{io}$, event E_1 occurs. Define $\hat{\boldsymbol{\eta}} := (\hat{\eta}_{io})_{io}$, where $\hat{\eta}_{jo} := \eta_{TTC(o)o}$ and $\hat{\eta}_{TTC(o)o} := \eta_{jo}$. $\hat{\boldsymbol{\eta}}$ and $\boldsymbol{\eta}$ coincide otherwise. It is easily verified that under the realizations $\boldsymbol{\xi}$ and $\hat{\boldsymbol{\eta}}$, event E_2 occurs. Indeed, that $\{\hat{\eta}_{jo} \leq \hat{\eta}_{TTC(o)o}\}$ holds true is trivial. Now, because, as we already claimed, under the realizations $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, j and $TTC(o)$ are never pointed to 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 $\hat{\boldsymbol{\eta}}$,

Since $\Pr(\boldsymbol{\xi}, \boldsymbol{\eta}) = \Pr(\boldsymbol{\xi}, \hat{\boldsymbol{\eta}})$, we can easily conclude that $\Pr E_1 = \Pr E_2$.

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

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

Because the distribution of η_{jo} has no atom, $\Pr\left[\cdot \,\middle|\, o \in \hat{O}, j \in I_2\right]$, the distribution of η_{jo} conditional on $o \in \hat{O}$ and $j \in I_2$ also has no atom ($\Pr(\eta_{jo} = \eta) = 0 \Rightarrow \Pr(\eta_{jo} = \eta \,\middle|\, o \in \hat{O}, j \in I_2) = 0$). Thus, we must have:

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

where the last equality holds because $\Pr E_1 = \Pr E_2$.

As E_ϵ increases when ϵ decreases, combining the above, we obtain:[51]

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

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

Corollary 2. *For any $\epsilon>0$ sufficiently small, there exist $\zeta>0, N>0$ such that, for all $n>N$:*

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

where $\hat{I}_2^\epsilon(o) := \{i\in I_2 | \eta_{io} > \eta_{TTC(o)o}+\epsilon\}$.

Proof. For any ϵ sufficiently small, we have $\zeta>0$ and $N>0$ such that for all $n>N$:

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

where the inequality follows from Lemma 5 . □

We are now ready to prove Theorem 2. The proof follows from Lemma 4 and Corollary 2. The former implies that as the economy grows, the expected number of objects in tier 1 assigned via long cycles remains significant. The latter implies that each of such objects finds many agents assigned by TTC to O_2 desirable for forming ϵ-blocks with. Specifically, for any

[51] 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 $\Pr(E) = \lim_{n\to\infty}\Pr(E_n)$.

sufficiently small $\epsilon \in (0, U(u_1, 0) - U(u_2, 1))$, we obtain that, for any $n > N$:

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

where the first inequality follows from the observation that if $i \in \hat{I}_2^\epsilon(o)$ and $o \in \hat{O}$, then (i, o) ϵ-blocks TTC, while the penultimate inequality follows from Corollary 2.

C Proof of Theorem 4

Fix any $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$. The following proposition is crucial for the proof.

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 $\epsilon > 0$:*

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

where $I_k := \{i \in I | DACB(i) \in O_k\}$. Similarly:

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

Recall that Lemma 2 gives a sense in which, in the uncorrelated case, the average rank achieved by objects under DA is small relative to the size of the market. In the sequel, we will need the following lemma which shows the implications of this for the values of $V_o(DA(o))$ in such an environment.

LEMMA 6. *Fix any $\epsilon > 0$. In the uncorrelated case,*

$$\frac{|\{o \in O | V_o(DA(o)) \geq V(1) - \epsilon\}|}{|O|} \xrightarrow{p} 1.$$

Proof. Fix $\epsilon, \gamma > 0$. We first claim that, with probability going to 1, the proportion of objects in O that achieve a rank below $\frac{2}{1-\gamma}|O|/\log(|O|)$ is greater than γ. To prove this, suppose to the contrary that with probability bounded away from 0, as the market grows, the proportion of objects enjoying ranks above $\frac{2}{1-\gamma}|O|/\log(|O|)$ is greater than $1-\gamma$. Then, with probability bounded away from 0, as the market grows,

$$\frac{1}{|O|}\sum_{o\in O_1} R_o^{DA} > \frac{1}{|O|}(1-\gamma)|O|\frac{2}{1-\gamma}(|O|/\log(|O|)) = 2|O|/\log(|O|),$$

which yields a contradiction of Lemma 2. Hence, with probability going to 1, the proportion of objects in O enjoying ranks below $\frac{2}{1-\gamma}|O|/\log(|O|)$ is larger than γ. Given that for any $\delta > 0$, for a sufficiently large n, $|O|/\log(|O|) \leq \delta\,|I|$, by Lemma 3-(iii) we must also have that, with probability going to 1, the proportion of objects o in O with $V_o(DA(o)) \geq 1-\epsilon$ is above γ. □

Proof of Proposition 1. We focus on $k = 1$; the other cases can be treated in exactly the same way.

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. If we were to run standard DA just for this submarket, then because preferences are drawn *iid*, by Lemma 2, with probability approaching 1 as $n \to \infty$, *all* agents would have made fewer than $\log^2(n)$ offers at the end of (standard) DA.

Consider now the original market. For any $\delta > 0$, as $\kappa(n) = o(n)$, we must have $\kappa(n) \leq \delta\,|O_1|$ for any sufficiently large n. Hence, by Lemma 3-(ii), 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 someone reaches his $\kappa(n)$-th offer. Moreover, because not all of the first $|O_1|+1$ agents can receive 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 until then. In sum, conditional on $\mathcal{E}$, Stage 1 will end at Step $|O_1|+1$ or before, and only objects in O_1 will have been assigned by the end of Stage 1.

We now show that, conditional on $\mathcal{E}$ with probability approaching 1 as $n \to \infty$, *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 (and the same is true for the distribution of objects' priorities over individuals). Given event $\mathcal{E}$, provided that all agents have made fewer than $\kappa(n)$ offers, the $|O_1|$ first steps of DACB proceed exactly in the same way as in 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. Because $\kappa(n) \geq \log^2(n)$, by Lemma 2, with probability going to 1 as $n \to \infty$, we reach the end of Step $|O_1|$ of DACB before Stage 1 ends (i.e., before any agent has applied to his $\log^2(n) \leq \kappa(n)$ most favorite object). Thus, with probability going to 1, the outcome thus far coincides with that attained in DA in the submarket composed of the $|O_1|$ first agents and of all objects in O_1. 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 fact, because $\Pr(\mathcal{E}) \to 1$ as $n \to \infty$, Step $|O_1|+1$ will be triggered with probability going to 1. This completes the proof of the first part of Proposition 1.

Now, we turn to the second part of the proof of Proposition 1. (Recall, we are still considering $k=1$.) We fix any ϵ and $\gamma < 1$ and wish to show that as $n \to \infty$:

$$\Pr\left\{\frac{|\{i \in I_1 | U_i(DACB(i)) \geq U(u_1, 1) - \epsilon\}|}{|I_1|} > \gamma\right\} \longrightarrow 1$$

and

$$\Pr\left\{\frac{|\{o \in O_1 | V_o(DACB(o)) \geq V(1) - \epsilon\}|}{|O_1|} > \gamma\right\} \longrightarrow 1.$$

In the sequel, we condition on event $\mathcal{E}$. First, by construction, every matched individual obtains an object within his/her $\kappa(n)$ most favorite objects which by Lemma 3-(i) implies that, with probability going to 1, they *all* enjoy payoffs above $U(u_1, 1) - \epsilon$.[52] This proves the first statement.

We next prove the second statement again for $k=1$. As we have shown, with high probability, the first $|O_1|$ Steps (i.e., Stage 1) of DACB proceed exactly the same way as in DA in the submarket that consists of the $|O_1|$ first agents and of all objects in O_1, where individuals' preferences and objects' priorities are drawn according to the unconditional distribution (which in this submarket is uncorrelated). According to Lemma 6, under DA in this submarket, with probability going to 1, the proportion of objects o in O_1 with $V_o(DA(o)) \geq 1 - \epsilon$ is above γ. Because objects in O_1 will have received even more offers at the end of Stage 1 of DACB than under the DA in the corresponding subeconomy, it must still be the case that, at the end of that stage, with probability going to 1, the proportion of objects in O_1 for which $V(DACB(o)) \geq 1 - \epsilon$ is above γ when n is large enough. Thus, for $k=1$, the second statement in Proposition 1 is proven provided that our conditioning event $\mathcal{E}$ holds. Because this event has probability going to 1 as $n \to \infty$, the result must hold even without the conditioning. Thus, we have proven Proposition 1 for the case $k=1$.

Consider next Stage $k > 1$. The objects remaining in Stage k have received no offers in Stages $j = 1, ..., k-1$ (otherwise, the objects would have been assigned during those stages). Hence, by the principle of deferred decisions, we can assume that the individuals' preferences over those objects are yet to be drawn at the beginning of Stage k. Similarly, we can assume that the priorities of those objects are also yet to be drawn. In other words, conditional on Stage $k-1$ being complete, we can assume without loss of generality that the distribution

[52] Note that this implies $\Pr\left[\{i \in I_1 | U_i(DACB(i)) \geq U(u_1, 1) - \epsilon\} = I_1\right] \to 1$ as $n \to \infty$. Hence, part of the statement of Proposition 1 can be strengthened.

of preferences and priorities is the same as the unconditional one. Thus, we can proceed inductively to complete the proof. □

Given Proposition 1, Theorem 4 follows straightforwardly. The first statement means that with high probability, all but a vanishing fraction of agents realize arbitrarily close to the highest idiosyncratic payoff. This implies that the proportion of the agents who would benefit discretely from a Pareto-dominating reassignment of DACB must vanish in probability. This observation, together with the second statement of Proposition 1, implies that the fraction of agent-object pairs that would gain discretely from blocking the DACB matching also vanishes in probability. The formal proof is in the Supplementary Material S.4.

D Proof of Theorem 7

Recall that in the two tier environment under consideration (i.e., $K = 2$), given $x \in (0, 1)$, we denote σ_x for the symmetric profile of strategies where each agent of each type lists first his $x\kappa(n)$ most favorite objects in O_1 and then his $(1 - x)\kappa(n)$ most favorite objects in O_2.

For each integer n, define $x(n) := 1 - \sqrt{\log^2(x_2 n)/\kappa(n)}$. Observe that, since we assumed that $\kappa(n)/\log^2(n) \to \infty$ as $n \to \infty$, $x(n)$ converges to 1 as n grows large. We show that the sequence of symmetric strategies $\{\sigma_{x(n)}\}_n$ satisfies the following two properties. (1) For any $\varepsilon > 0$, and any n large enough, $\sigma_{x(n)}$ is an ex-ante ε-Bayes Nash equilibrium; (2) the induced matching outcome is asymptotically efficient and asymptotically stable.

For the proof, we first establish the following claim:

Claim: *Given the sequence of profiles of strategies $\{\sigma_{x(n)}\}_n$, the probability that all agents are matched goes to 1 as n goes to infinity.*

PROOF. Given an n-economy, recall that under the profile of strategies $\sigma_{x(n)}$, the assignment of DA (with truncation $\kappa(n)$) can be obtained in two steps. In the first step, we run DA with only first tier objects and the agents ranking only $x(n)\kappa(n)$ objects. To complete the matching, in the second step, we run DA with only second tier objects and all unmatched individuals ranking only $(1 - x(n))\kappa(n)$ objects.[53] Now, in Step 1, there are more individuals than objects and ordinal preferences are drawn iid uniformly. We will show that with probability going to 1 in this unbalanced market all objects are matched. Pick randomly $x_1 n$ individuals. We obtain a new market composed of these $x_1 n$ individuals together with the $x_1 n$ objects in O_1. In this balanced market, the number of offers received by each object is weakly smaller than in the market with all n individuals and the $x_1 n$ objects in O_1. Thus, it is

[53] There may exist O_1 objects which are not assigned by the end of the first step. We do not have to consider them in the second step because they would in any case receive no offer in that stage (since all agents remaining in the second step have exhausted all their offers to objects in O_1).

enough for our purpose to show that, with probability going to 1, in this balanced market all objects are matched. In order to show this, consider DA without any constraint on offers. We know, by Pittel (1992), that in this balanced market, with probability going to 1 as n goes to infinity, all objects in O_1 are matched before agents make more than $\log^2(x_1 n)$ offers. In Step 1, our mechanism is DAC with the length of ROL at $x(n)\kappa(n)$. Because $\kappa(n)/\log^2(n) \to \infty$ as $n \to \infty$ and $x(n)$ goes to 1 as $n \to \infty$, there is $N > 0$ such that $x(n)\kappa(n) \geq \log^2(x_1 n)$ for any $n > N$ (recall that x does not depend on n). Combining the two previous observations, it follows that, in the balanced market, with probability going to 1 as n goes to infinity, all objects in O_1 are matched, as claimed.

Now, in Step 2 we have all remaining individuals and all objects in O_2. Since we only have objects in O_1 in Step 1, the number of individuals who remain unmatched at the end of Step 1 must be (weakly) greater than $|O| - |O_1| = |O_2|$. Hence, Step 2 is a (weakly) unbalanced market where there are more individuals than objects and where the length of each agent's ROL is $(1 - x(n))\kappa(n)$. In addition, by the principle of deferred decisions, we can simply assume that preferences/priorities are yet to be drawn. Finally, we observe that $(1 - x(n))\kappa(n)/\log^2(x_2 n) = \sqrt{\kappa(n)/\log^2(x_2 n)}$ goes to ∞ as n grows large. So for n large enough, $(1 - x(n))\kappa(n)$ is greater than $\log^2(x_2 n)$. Thus, we can mimic the reasoning we just made for Step 1 to show that with probability going to 1 as n grows, all objects in O_2 are matched.

By combining the two above results (for Step 1 and 2), we conclude that the probability that all objects (and so that all agents) are matched goes to 1 as n goes to infinity, as claimed. □

Because agents are symmetric, by the previous result, for each individual, the probability of being matched an O_1 object and that of being matched an O_2 object converges to x_1 and x_2 respectively. We state this in the following corollary.

Corollary 3. *Given the sequence of profiles of strategies $\{\sigma_{x(n)}\}_n$, for each $k = 1, 2$, the probability that an agent is matched to an object in O_k goes to x_k as n goes to infinity.*

The above corollary can be used to prove that the strategy profiles in the sequence have a desirable incentive property.

Proposition 2. *For any $\varepsilon > 0$, for any n large enough, $\sigma_{x(n)}$ is an ex-ante ε-Bayes Nash equilibrium in the n-economy.*

Proof. Fix any $\varepsilon > 0$. Without loss of generality, assume that $\varepsilon > 0$ is small enough so that $U(u_1, 1) - \varepsilon > U(u_2, 1)$. We claim that for n large enough $\sigma_{x(n)}$ is an ex-ante ε-Bayes Nash equilibrium. Let n be large enough so that $(1 - x(n))U(u_1, 1) \leq \varepsilon$.

Let E_i be the event that agent i realizes utility of at least $U(u_k, 1) - \varepsilon$ from each of her $\kappa(n)$ most favorite objects in O_k for $k = 1, 2$. This last condition ensures that under event E_i, individual i prefers each object in O_1 that he ranks to any object in O_2 (since $U(u_1, 1) - \varepsilon > U(u_2, 1)$). Because $\kappa(n) = o(n)$, by Lemma 3, the probability of E_i goes to 1 as n increases. Thus, the expected payoffs conditional on E_i of agent i converge to the (unconditional) ex-ante payoffs.

Note that, by symmetry, the probability of getting matched an O_1 object for individual i does not depend on ξ.[54] Hence, it must be the same as the ex-ante probability of being matched to an O_1 object which, by Corollary 3, converges to x_1 under the profile of strategies $\sigma_{x(n)}$. Similarly, the probability that ξ gets matched to an O_2 object converges to x_2 and the probability that he gets unmatched converges to 0. Fix n large enough so that the probability of being unmatched given an arbitrary type ξ is bounded from above by $\varepsilon / U(u_2, 1)$.

In the sequel, we fix an agent and consider a deviation from the strategy prescribed by $\sigma_{x(n)}$. We fix a realization of type $\xi \in E_i$ and assume without loss of generality that, under the deviating strategy, ξ lists first his $z(n)$ most favorite objects in O_1 and then his $\kappa(n) - z(n)$ most favorite objects in O_2.[55] There are two cases to consider.

Case 1. $z(n) \leq x(n)\kappa(n)$. One of the following events must be true.

1a. Under the profile of strategies $\sigma_{x(n)}$, ξ gets matched to an O_1 object within his $z(n)$ most favorite objects. Here, the deviation has no impact on ξ's assignment.

1b. Under the profile of strategies $\sigma_{x(n)}$, ξ gets matched to an O_1 object between his $z(n) + 1$ and $x\kappa(n)$ most favorite objects. Here, the deviation entails a payoff loss strictly greater than $U(u_1, 1) - \varepsilon - U(u_2, 1) > 0$ (recall that each object in O_1 listed under the profile of strategies σ_x all yields payoff of least $U(u_1, 1) - \varepsilon > U(u_2, 1)$).

1c. Under the profile of strategies σ_x, ξ gets matched to an O_2 object. Under the deviating strategy, ξ cannot be assigned an object in O_1. While the exact impact of the deviating strategy on the size of the set of participants in Step 2 is not easy to analyze, the gain (if any) from the deviating strategy is bounded by $U(u_2, 1) - (U(u_2, 1) - \varepsilon) = \varepsilon$ since $\xi \in E_i$.

1d. Under the profile of strategies $\sigma_{x(n)}$, ξ gets unmatched. Under the deviating strategy, ξ cannot be assigned an object in O_1. The gain of the deviating strategy is at most $U(u_2, 1)$.

Hence, the expected gain of the deviating strategy is bounded from above by $\Pr\{1c\}\varepsilon + \Pr\{1d\}U(u_2, 1)$. By construction, $\Pr\{1d\}U(u_2, 1) \leq \varepsilon$ and so the expected gain of the deviating strategy is at most 2ε.

Case 2. $z(n) \geq x(n)\kappa(n)$. One of the following events must occur.

2a. Under the deviating strategy, ξ gets matched to an O_1 object within his $x(n)\kappa(n)$ most

[54] Agents with types ξ and ξ' rank the same number of objects in O_1 and the same number of objects in O_2. In addition, objects within each tier are ex-ante symmetric.

[55] This follows from the ex ante symmetry of the objects within each tier and the use of symmetric strategies adopted by the opponents. Listing objects within a tier untruthfully or dropping a more preferred object within a tier can only do worse.

favorite objects. Here, the deviation has no impact on ξ's assignment.

2b. Under the deviating strategy, ξ gets matched to an O_1 object between his $x(n)\kappa(n)+1$ and $z(n)$ most favorite objects. Here, the deviation yields a payoff gain of at most $U(u_1,1)>0$ (the worst case scenario being that ξ gets unassigned under $\sigma_{x(n)}$).

2c. Under the deviating strategy, ξ gets matched to an O_2 object while under the strategy associated with $\sigma_{x(n)}$, ξ gets assigned an object in O_2 (he cannot be assigned an object in O_1). The gain (if any) obtained using the deviating strategy is at most $U(u_2,1)-(U(u_2,1)-\varepsilon)=\varepsilon$ since $\xi\in E_i$.

2d. Under the deviating strategy, ξ gets matched to an O_2 object while under $\sigma_{x(n)}$, ξ gets unassigned. The gain obtained using the deviating strategy is at most $U(u_2,1)$.

2e. Under the deviating strategy, ξ gets unmatched. Trivially, there can be no gain from using the deviation.

Hence, the expected gain of the deviating strategy is bounded from above by $\Pr\{2b\}U(u_1,1)+\Pr\{2c\}\varepsilon+\Pr\{2d\}U(u_2,1)$. By construction, $\Pr\{2d\}U(u_2,1)\leq\varepsilon$. In addition, we show below that $\Pr\{2b\}U(u_1,1)\leq\varepsilon$ which implies that the expected gain of the deviating strategy is bounded from above by 3ε.

Thus, it remains to show that $\Pr\{2b\}U(u_1,1)\leq\varepsilon$. Recall that $\{2b\}$ is defined as the event where, under the deviating strategy, ξ gets matched to an O_1 object between his $x(n)\kappa(n)+1$ and $z(n)$ most favorite objects. Recall the basic property that if o and o' are in O_1 and o is listed ahead of o' in i's list then the probability of getting assigned object o is higher than that of being assigned o'. Hence, the probability that under the deviating strategy, ξ is assigned an object between his $x(n)\kappa(n)+1$ and $z(n)$ most favorite objects is bounded from above by $1-x(n)$.[56] Thus, $\Pr\{2b\}\leq 1-x(n)$. Our assumption that $(1-x(n))U(u_1,1)\leq\varepsilon$ completes the argument.

To recap, there is n large enough so that, given any realization $\xi\in E_i$, the expected gain from deviating is at most 3ε. Since this is true for any $\xi\in E_i$, the expected gains from deviating conditional on E_i must be bounded by 3ε. Since payoffs are bounded and $\Pr(E_i)$ goes to 1 as n grows, we are done. $\square$

[56] To see this, let us denote by p_ℓ the probability of being matched an object listed in the ℓth position for $\ell=1,...,z(n)$. We claim that $\sum_{\ell=x(n)\kappa(n)+1}^{z(n)} p_\ell\leq 1-x(n)$. Suppose to the contrary that $\sum_{\ell=x(n)\kappa(n)+1}^{z(n)} p_\ell>1-x(n)$. Then because $\{p_\ell\}$ is a decreasing sequence, we would have $p_{x(n)\kappa(n)+1}>\frac{1-x(n)}{z(n)-x\kappa(n)}$. Again because $\{p_\ell\}$ is a decreasing sequence, we obtain that $\sum_{\ell=1}^{x(n)\kappa(n)} p_\ell>x(n)\frac{1-x(n)}{z(n)/\kappa(n)-x(n)}$ where the term $\frac{1-x(n)}{z(n)/\kappa(n)-x(n)}$ is greater than 1 because $z(n)/\kappa(n)\leq 1$. Thus, $\sum_{\ell=1}^{x(n)\kappa(n)} p_\ell>x(n)$ and so eventually we get $\sum_{\ell=1}^{z(n)} p_\ell>x(n)+1-x(n)=1$, a contradiction.

Proof of Theorem 7. Fix any $\varepsilon > 0$ and let n be large enough so that $\sigma_{x(n)}$ is an ex ante ε-equilibrium (this is well-defined by Proposition 2). We claim that the resulting outcome is asymptotically efficient and asymptotically stable.

Asymptotic efficiency. We know by Claim 1 that the probability that all agents are matched goes to 1 as n goes to infinity. In addition, for any $\delta > 0$, as we showed in Lemma 3, the probability that all agents list only objects with which they enjoy idiosyncratic payoffs greater than $1 - \delta$ goes to 1 as grows. Taken together, these two statements imply that, for any $\delta > 0$, with probability going to 1 as n grows, all agents get matched to object with which they enjoy idiosyncratic payoffs greater than $1 - \delta$. Hence, the outcome is asymptotically efficient.

Asymptotic stability. To show the asymptotic stability of the induced outcome, we come back to the proof of the last claim. Recall that the outcome of DA (with truncation $\kappa(n)$) was obtained in two steps. In Step 1, the submarket is composed of all n individuals and only objects in the first tier O_1. We are in an environment where there are more individuals than objects. Let us pick $x_1 n$ individuals randomly and consider the balanced market composed of these $x_1 n$ individuals and the $x_1 n$ objects in the first tier O_1. Clearly, under our mechanism, each object is weakly worse-off in this new market. If we run standard DA in this market, we know from Pittel (1992) that with probability going to 1, the algorithm will end before agents make more than $\log^2(x_1 n)$ offers. In Step 1, our mechanism is DA with truncation $x(n)\kappa(n)$. Because $\kappa(n)/\log^2(n) \to \infty$ as $n \to \infty$ and $x(n)$ goes to 1 as n grows large, there exists $N > 0$ such that $x(n)\kappa(n) \geq \log^2(n)$ for any $n > N$. Combining the two previous observations, in the balanced submarket, with probability going to 1 as n goes to infinity, the matching given by DA and the matching given by DA with truncation $x(n)\kappa(n)$ are the same. By Lemma 6, in the balanced submarket, for any $\delta > 0$, the proportion of objects in O_1 receiving a (idiosyncratic) payoff greater than $1 - \delta$ converges in probability to 1. Thus, in the whole market where there are (weakly) more individuals than objects this must remain true. Now, to treat objects in O_2, we must consider the second step. Again, there are (weakly) more individuals than objects and where preferences/priorities over O_2 objects (i.e. remaining ones) are yet to be drawn. In addition, as we already claimed, $(1 - x(n))\kappa(n)/\log^2(x_2 n) = \sqrt{\kappa(n)/\log^2(x_2 n)}$ goes to ∞ as n grows large, and so, for n large enough, $(1 - x(n))\kappa(n)$ is greater than $\log^2(x_2 n)$. Hence, the same argument as for O_1 objects can be applied to O_2 objects To recap, for any $\delta > 0$, the proportion of objects in O receiving a (idiosyncratic) payoff greater than $1 - \delta$ converges in probability to 1. This completes the proof of the asymptotic stability of the induced outcome. □