Supplementary Material for "Efficiency and Stability in Large Matching Markets"

(For Online Publication)

Yeon-Koo Che* Olivier Tercieux†

July 30, 2015

The supplementary material consists of eight sections. Section S.1 provides the proof of Lemma 4 on the probability that a significant fraction of top tier objects are assigned via long cycles. Section S.2 proves Theorem 3 on the asymptotic inefficiency of DA. Section S.3 gives the formal details on how Erdös-Renyi theorem can be used to achieve asymptotic stability and asymptotic efficiency and explain the drawbacks that the underlying mechanism would have. Section S.4 provides the details of the completion of the proof of Theorem 4. Section S.5 proves Theorem 5. In Section S.6, we analyze the generalized version of DACB. We provide a proof of Theorem 6. We also state and prove a result on incentives which is the analogous of Theorem 5 but for the extended DACB mechanism. Section S.7 contains simulations of alternative algorithms including DACB based on random preferences which confirm that our main results hold well beyond the setting we study, and particular for market sizes that are quite moderate. Finally, Section S.8 provide details of the NYC school choice procedures and the data, which would help the reader to understand the calibration performed in the main text.

S.1 Proof of Lemma 4

The proof of Lemma 4 requires a deeper understanding of the random structure of TTC. In the next section, we simply present the part of the results that are of direct use for the proof of Lemma 4. The result on the random structure of TTC is stated without proof, its proof can be found in Che and Tercieux (2015).

*Department of Economics, Columbia University, USA. Email: yeonkooche@gmail.com.

†Department of Economics, Paris School of Economics, France. Email: tercieux@pse.ens.fr.

We need to begin with some preliminary definitions and results from random graph/mapping theory.

S.1.1 Preliminaries

To begin, recall that it is sufficient to consider the TTC assignment arising from the market consisting of the agents I and the objects O_1 in the top tier (recall that, irrespective of the realizations of the idiosyncratic values, all agents prefer every object in O_1 to any object in O_2). Hence, we shall simply consider an **unbalanced market** consisting of a set I of agents and a set O of objects such that (1) *the preferences of each side with respect to the other side are drawn iid uniformly,* and (2) $n = |I| \geq |O|$.

Consider any two finite sets I and O, with cardinalities $|I| = n, |O| = o$. A **bipartite digragh** $G = (I \times O, E)$ consists of vertices I and O on two separate sides and directed edges $E \subset (I \times O) \cup (O \times I)$, comprising ordered pairs of the form (i, o) or (o, i) (corresponding to edge originating from i and pointing to o and an edge from o to i, respectively). A **rooted tree** is a bipartite digraph where all vertices have out-degree 1 except the root which has out-degree 0.[1] A **rooted forest** is a bipartite graph which consists of a collection of disjoint rooted trees. A **spanning rooted forest over** $I \cup O$ is a forest comprising vertices $I \cup O$. From now on, a spanning forest will be understood as being over $I \cup O$.

Consider an arbitrary mapping, $g : I \to O$ and $h : O \to I$, defined over our finite sets I and O. Note that such a mapping naturally induces a bipartite digraph with vertices $I \cup O$ and directed edges with the number of outgoing edges equal to the number of vertices, one for each vertex. In this digraph, $i \in I$ points to $g(i) \in O$ while $o \in O$ points to $h(o) \in I$. Such a mapping will be called a bipartite mapping. A **cycle** of a bipartite mapping is a cycle in the induced bipartite digraph, namely, distinct vertices $(i_1, o_1, ..., i_{k-1}, o_{k-1}, i_k)$ such that $g(i_j) = o_j, h(o_j) = i_{j+1}, j = 1, ..., k-1, i_k = i_1$. A **random bipartite mapping** selects a composite map $h \circ g$ uniformly from a set $\mathcal{H} \times \mathcal{G} = I^O \times O^I$ of all bipartite mappings. Note that a random bipartite mapping induces a random bipartite digraph consisting of vertices $I \cup O$ and directed edges emanating from vertices, one for each vertex. We say that a vertex in a digraph is **cyclic** if it is in a cycle of the digraph.

The following lemma states the number of cyclic vertices in a random bipartite digraph induced by a random bipartite mapping.

Lemma S1. *(Jaworski (1985), Corollary 3) The number q of the cyclic vertices in a*

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

random bipartite digraph induced by a random bipartite mapping $g : I \to O$ and $h : O \to I$ has an expected value of

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

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

The next theorem on the random structure of TTC proves crucial for the proof of Lemma 4. Its proof is contained in Che and Tercieux (2015).

Theorem S1. *Suppose any round of TTC begins with n agents and o objects remaining in the market. Then, the probability that there are $m \leq \min\{o, n\}$ agents assigned at the end of that round is*

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

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

This theorem shows that the numbers of agents and objects that are assigned in each round of TTC follow a simple Markov chain depending *only* on the numbers of agents and objects at the beginning of that round. It also characterizes the probability structure of the Markov chain. This theorem implies that there are no conditioning issues at least with respect to the *total* numbers of agents and objects that are assigned in each round of TTC. Namely, one does not need to keep track of the precise history leading up to a particular economy at the beginning of a round, as far as the numbers of objects assigned in that round is concerned. Obviously, this result is crucial in rendering Lemma 4 analytically provable. However, this result alone is not sufficient. We still need to understand the number of objects that are assigned via short versus long cycles in each round. Unfortunately, the *composition* of cycles—long versus short—cleared in each round depends on the precise history leading up to the economy at the beginning of that round. The next two sections deal with this issue.

S.1.2 The Number of Objects Assigned via Short Cycles

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

Formally, the random sequence of forests, $F_1, F_2, ...$ is defined as follows. First, we let F_1 be a trivial unique forest consisting of $|I| + |O|$ trees with isolated vertices, forming their own roots. For any $i = 2, ...$, we first create a random directed edge from each root of F_{i-1} to a vertex on the other side, and then delete the resulting cycles (these are the agents and objects assigned in round $i - 1$) and F_i is defined to be the resulting rooted forest.

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

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

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

F, the expectation of the number S_k of short-cycles is

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

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

Proposition S1. *If TTC round k begins with any forest F,*

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

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

Corollary S1. *For any round k of TTC, $\mathbb{E}\left[S_k\right] \leq 2$.*

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

S.1.3 The Number of Objects Assigned via Long Cycles

Again consider the unbalanced market in which $|I| \geq |O|$, and recall $n := |I|$ and $o := |O|$.

The Markov property established in Theorem S1 means that the number of agents and objects assigned in any TTC round depends only on the number of agents and objects that round begins with, regardless of how many rounds preceded that round and what happened in those rounds. Hence, the distribution of the (random) number M_k of objects that would be assigned in any round of TTC that begins with n_k agents and $o_k(\leq n_k)$ objects is the same as that in the first round of TTC when there are n_k agents and $o_k(\leq n_k)$. In particular, we can apply Lemma S1 to compute its expected value:[3]

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

We can make two observations: First, the expected number is increasing in o_k (and n_k) and goes to infinity as o_k (and n_k) increases. This can be seen easily by the fact that $\frac{k-l}{k}$ is increasing in k for any $k > l$. Second, given our assumption that $n_k \geq o_k$, there exists $\hat{o} \geq 1$[4] such that

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

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

Theorem S2. $\mathbb{E}\left[\frac{|\hat{O}|}{|O|}\right] \geq \frac{1}{3} - \frac{\hat{o}-2}{3|O|}$.

Proof. Consider the following sequence of random variables $\{\mathbb{E}(L_k \,|o_k)\}_{k=1}^{|O|}$ where o_k is the number of remaining objects at round k while L_k is the number of objects assigned at round k via long cycles. (Note both are random variables.) Thus, $o_1 = |O|$. Note that $\mathbb{E}(L_{|O|} \,|o_{|O|}) = 0$. By Theorem S1, we are defining here the process $\{\mathbb{E}(L_k \,|o_k)\}_{k=1}^{|O|}$ induced by the Markov chain $\{o_k\}$. Note also that $\mathbb{E}(L_k \,|o_k) = \mathbb{E}[M_k \,|o_k] - \mathbb{E}[S_k \,|o_k]$ where S_k is the number of objects assigned at round k via short cycles. By Proposition S1, $\mathbb{E}[S_k \,|F] \leq 2$ for any possible forest F, this implies that $\mathbb{E}[S_k \,|o_k] \leq 2$. Hence, we obtain that $\mathbb{E}(L_k \,|o_k) \geq 3 - 2 = 1$ if $o_k \geq \hat{o}$. (Recall that $\hat{o}$ is defined such that $o_k \geq \hat{o}$ implies $\mathbb{E}[M_k \,|o_k] \geq 3$.) Let T be first round at which the $\mathbb{E}(L_k \,|o_k)$ becomes smaller than 1:

[3] The number is half of that stated in Lemma S1 since the number of agents cleared in any round is precisely the half of the cyclic vertices in a random bipartite graph at the beginning of that round. Recall also that, by definition of TTC, together with our assumption that $o \leq n$, given the number of objects o_k, the number of individuals n_k is totally determined and is equal to $o_k + n - o$.

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

formally, $\mathbb{E}(L_k \,|o_k) \leq 1$ only if $k \geq T$ (this is well-defined since $\mathbb{E}(L_{|O|} \,\big|o_{|O|}) = 0$). Note that $o_T \leq \hat{o}$.

Now we obtain:

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

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

Once we have reached round T under TTC, at most $\hat{o}$ more short cycles can arise. Thus, the expected number of short cycles must be smaller than $2\mathbb{E}(T) + \hat{o}$. Indeed, the expected number of short cycles is smaller than 2 times the expected number of rounds for TTC to converge (recall that, by Corollary S1, the expected number of short cycles at each round is at most two) which itself is smaller than $2\mathbb{E}(T) + \hat{o}$. It follows that

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

Combining the above inequalities, we obtain that

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

from which the result follows. □

Lemma 4 is a direct corollary of Theorem S2.

S.2 Proof of Theorem 3

Since $U(u_1^0, 0) > U(u_2^0, 1))$, all objects in O_1 are assigned before any agent starts applying to objects in O_2. Hence, the assignment achieved by individuals assigned objects in O_1 is the same as the one obtained when we run DA in the submarket with individuals in I and objects in O_1. The following lemma shows that the agents assigned objects in O_1 suffer a significant number of rejections before getting assigned. This result is obtained by Ashlagi, Kanoria, and Leshno (2013) and by Ashlagi, Braverman, and Hassidim (2011). We provide a much simpler direct proof for this result here.[5]

Lemma S2 (Welfare Loss under Unbalanced Market). *Consider an unbalanced submarket consisting of agents I and objects O_1, where $|I| - |O_1| \to n(1 - x_1)$ as $n \to \infty$. Let I_1 be the (random) set of agents who are assigned objects in O_1, and let $I_1^\delta := \{i \in I_1 | i$ makes at least δn offers$\}$ be the subset of them who each suffer from more than δn rejections (before getting assigned objects in O_1). Then, there exist γ, δ, υ, all strictly positive, such that for all $n > N$ for some $N > 0$,*

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

Proof. Without loss, we work with the McVitie and Wilson's algorithm (which equivalently implement DA). Consider the individual $i = n$ at the last serial order, at the beginning of step n. By that step, each object in O_1 has surely received at least $|I| - |O_1|$ offers. This is because at least $|I| - |O_1| - 1$ preceding agents must be unassigned, so each of them must have been rejected by all objects in O_1 before the beginning of step n.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Lemma S2 implies that there exists $\epsilon' > 0, \upsilon' > 0, \gamma' > 0$ such that for all $n > N'$ for some $N' > 0$,

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

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

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

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

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

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

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

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

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

S.3 The Erdös-Renyi mechanim

We first briefly recall the Erdös-Renyi theorem mentioned in the paper. It is thus worth introducing the relevant model of random graph. A **bipartite graph** G consists in vertices, $V_1 \cup V_2$, and edges $E \subset V_1 \times V_2$ across V_1 and V_2 (with no possible edges within vertices in each side). A **perfect bipartite matching** is a bipartite graph in which each vertex is involved in exactly one edge. A random bipartite graph $B = (V_1 \cup V_2, p)$, $p \in (0, 1)$, is a bipartite graph with vertices $V_1 \cup V_2$ in which each pair $(v_1, v_2) \in V_1 \times V_2$ is linked by an edge with probability p independently (of edges created for all other pairs). In this context, the Erdös-Renyi theorem can be stated as follows.

Theorem S3. *Consider a random bipartite graph $B = (V_1 \cup V_2, p)$ where $0 < p < 1$ is a constant and for each $i \in \{1, 2\}$ and $|V_1| = |V_2| = n$. We have*

$$\Pr[\exists \textit{ a perfect bipartite matching }] \to 1 \textit{ as } n \to \infty.$$

Hence, in our environment where we draw randomly individuals' preferences and objects' priorities, one can build an associated random bipartite graph which consists of vertices $I \cup O$ and where an edge between (i, o) is added if and only if $\xi_{i,o} > 1 - \epsilon$ and $\eta_{i,o} > 1 - \epsilon$. Applying the Erdös-Renyi theorem we obtain that with probability approaching 1 as $|I| = |O| = n$ increases, a (perfect bipartite) matching exists where all objects and agents realize idiosyncratic payoffs greater than $1 - \epsilon$.[6] Hence, one could construct a mechanism in which (1) agents and objects (more precisely their suppliers) report their idiosyncratic shocks (2) given such a reports a bipartite graph is built where an edge between an agent and an object if and only if $\xi_{i,o} > 1 - \epsilon$ and $\eta_{i,o} > 1 - \epsilon$ and (3) a maximal bipartite matching is selected.[7] Such mechanism would select a perfect matching whenever it exists. Under truthful reports, this occurs with probability approaching 1 as the market size increases. Thus, by construction, this mechanism is asymptotically efficient and asymptotically stable. In addition, well-known polynomials algorithms such as the *augmenting path algorithm* would find a maximal matching.

However, as stated in the paper, this mechanism would not be desirable for several reasons. First, it would not work if the agents cannot tell apart common values from idiosyncratic values. More importantly, the mechanism would not have a good incentive property. An agent will be reluctant to report the objects in lower tiers even though they have high idiosyncratic preferences. Indeed, if he expects that with significant probability, he will not get any object in the highest tier, he will have incentives to claim that he enjoys high idiosyncratic payoffs with a large number of high tier objects and that all his idiosyncratic payoffs for the other tiers are low. It is very likely that there is a perfect matching even under this misreport and this will ensure him to get matched with a high tier object.

S.4 Completion of the Proof of Theorem 4

We start by showing that DACB with $\kappa \geq \log^2(n)$ and $\kappa = o(n)$ is asymptotically efficient. Let $\mu :=$ DACB, we have to show that for any $\epsilon > 0$, and any mechanism μ' that weakly Pareto-dominates μ

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

where we recall that $I_\epsilon(\mu' \,|\mu) := \{i \in I \,|U(\mu(i)) < U(\mu'(i)) - \epsilon\}$.

We first introduce some notations: In the sequel, for an arbitrary collection of sets $\{X_k\}_{k=1}^K$, we let $X_{<k}$ (resp. $X_{\leq k}$) be the set $\cup_{\ell<k} X_k$ (resp. $\cup_{\ell \leq k} X_k$). Note that $X_{<1} = \emptyset$.

[6] The unbalanced case can be treated with almost no modifications.

[7] A maximal bipartite matching is a matching that contains the largest possible number of edges.

We fix any mechanism μ' that weakly Pareto-dominates μ and any $\epsilon > 0$ small enough so that $U(u_k, 1) - \epsilon > U(u_{k+1}, 1)$ for all $k = 1, ..., K-1$.

Now, fix $k \geq 1$ and let $I'_k := \{i \in I | \mu(i) \in O_k\}$ be the set of individuals matched to objects in O_k under matching μ' and recall that I_k is the set of individuals matched to objects in O_k under matching μ. By definition, $|I'_{<k}| = |I_{<k}|$. Let

$$\bar{I}_k := \{i \in I_k \,|U(\mu(i)) \geq U(u_k, 1) - \epsilon\} \text{ and } \hat{I}_{\epsilon,k}(\mu' \,|\mu) := \{i \in I_k \,|U(\mu(i)) \geq U(\mu'(i)) - \epsilon\}.$$

Then, for any $i \in \bar{I}_k \backslash I'_{<k}$, $U(\mu(i)) \geq U(u_k, 1) - \epsilon \geq U(\mu'(i)) - \epsilon$. Hence, we must have $\bar{I}_k \backslash I'_{<k} \subset \hat{I}_{\epsilon,k}(\mu' \,|\mu)$. Observe next $\bar{I}_{<k} \subset I'_{<k}$. This follows since, if $i \in \bar{I}_{<k}$ (note that this implies $k \geq 2$), then by definition $U(\mu'(i)) \geq U(\mu(i)) \geq U(u_{k-1}, 1) - \epsilon > U(u_k, 1)$ and so $i \in I'_{<k}$.

We fix $\gamma' < 1$ and show that, as $n \to \infty$,

$$\Pr\left\{ \frac{\left|\hat{I}_{\epsilon,k}(\mu' \,|\mu)\right|}{|I_k|} \geq \gamma' \right\} \to 1.$$

For $k = 1$, as $n \to \infty$,

$$\Pr\left\{ \frac{\left|\hat{I}_{\epsilon,1}(\mu' \,|\mu)\right|}{|I_1|} \geq \gamma' \right\} \geq \Pr\left\{ \frac{\left|\bar{I}_1 \backslash I'_{<1}\right|}{|I_1|} \geq \gamma' \right\} = \Pr\left\{ \frac{\left|\bar{I}_1\right|}{|I_1|} \geq \gamma' \right\} \to 1,$$

by Proposition 1. Next consider $k \geq 2$ and fix any $\delta > 0$ such that $\gamma' + \delta < 1$. Then, as

$n \to \infty$,

$$\begin{aligned}
\Pr\left\{\frac{\left|\hat{I}_{\epsilon,k}(\mu'\,|\mu)\right|}{|I_k|} \geq \gamma'\right\} &\geq \Pr\left\{\frac{\left|\bar{I}_k \backslash I'_{<k}\right|}{|I_k|} \geq \gamma'\right\} \\
&= \Pr\left\{\frac{\left|\bar{I}_k \backslash \left((I'_{<k} \cap \bar{I}_{<k}) \cup (I'_{<k} \backslash \bar{I}_{<k})\right)\right|}{|I_k|} \geq \gamma'\right\} \\
&\geq \Pr\left\{\frac{\left|\bar{I}_k \backslash (I'_{<k} \cap \bar{I}_{<k})\right| - \left|I'_{<k} \backslash \bar{I}_{<k}\right|}{|I_k|} \geq \gamma'\right\} \\
&= \Pr\left\{\frac{\left|\bar{I}_k\right|}{|I_k|} - \frac{\left|I'_{<k} \backslash \bar{I}_{<k}\right|}{|I_k|} \geq \gamma'\right\} \\
&= \Pr\left\{\frac{\left|\bar{I}_k\right|}{|I_k|} - \left(\frac{|I'_{<k}|}{|I_k|} - \frac{\left|\bar{I}_{<k}\right|}{|I_k|}\right) \geq \gamma'\right\} \\
&\geq \Pr\left\{\frac{\left|\bar{I}_k\right|}{|I_k|} \geq \gamma' + \delta \text{ and } \left(\frac{|I'_{<k}|}{|I_k|} - \frac{\left|\bar{I}_{<k}\right|}{|I_k|}\right) \leq \delta\right\} \\
&= \Pr\left\{\frac{\left|\bar{I}_k\right|}{|I_k|} \geq \gamma' + \delta \text{ and } \frac{\left|\bar{I}_{<k}\right|}{|I_{<k}|} \geq 1 - \delta\frac{|I_k|}{|I_{<k}|}\right\} \to 1,
\end{aligned}$$

where the second equality holds since $\bar{I}_k \cap \bar{I}_{<k} = \emptyset$, which implies $\bar{I}_k \backslash (I'_{<k} \cap \bar{I}_{<k}) = \bar{I}_k$; the third equality holds since $\bar{I}_{<k} \subset I'_{<k}$; the last equality uses $|I'_{<k}| = |I_{<k}|$; and the convergence result follows from Proposition 1.

To complete the argument, fix any $\gamma > 0$. The, the desired result holds since

$$\begin{aligned}
\Pr\left\{\frac{|I_\epsilon(\mu'\,|\mu)|}{n} \geq \gamma\right\} &= \Pr\left\{\sum_{k=1}^{K}\frac{\left|\hat{I}_{\epsilon,k}(\mu'\,|\mu)\right|}{n} < 1 - \gamma\right\} \\
&\leq \Pr\left\{\frac{\left|\hat{I}_{\epsilon,k}(\mu'\,|\mu)\right|}{n} < \frac{1-\gamma}{K} \text{ for some } k\right\} \\
&\leq \sum_{k=1}^{K}\Pr\left\{\frac{\left|\hat{I}_{\epsilon,k}(\mu'\,|\mu)\right|}{n} < \frac{1-\gamma}{K}\right\} \\
&= \sum_{k=1}^{K}\left(1 - \Pr\left\{\frac{\left|\hat{I}_{\epsilon,k}(\mu'\,|\mu)\right|}{n} \geq \frac{1-\gamma}{K}\right\}\right) \to 0
\end{aligned}$$

as $n \to \infty$, where the second inequality uses the union bound, and the convergence follows from the above arguments (with $\gamma' := \frac{1-\gamma}{K}$).

We now prove that DACB with $\kappa \geq \log^2(n)$ and $\kappa = o(n)$ is asymptotically stable. We must show that

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

where we recall that $J_\epsilon(\mu) := \{(i,o) \in I \times O \,|U_i(o) > U_i(\mu(i)) + \epsilon \text{ and } V_o(i) > V_o(\mu(o)) + \epsilon\}$. Fix any $\gamma > 0$, we have

$$\begin{aligned}
\Pr\left\{\frac{|J_\epsilon(\mu)|}{n(n-1)} \geq \gamma\right\} &\leq \Pr\left\{\frac{|I \times \{o \in O\,|V(1) > V_o(\mu(o)) + \epsilon\}|}{n(n-1)} \geq \gamma\right\} \\
&= \Pr\left\{\frac{|\{o \in O\,|V(1) > V_o(\mu(o)) + \epsilon\}|}{n-1} \geq \gamma\right\} \\
&= \Pr\left\{\frac{1}{n-1}\sum_{k=1}^{K} |\{o \in O_k\,|V(1) > V_o(\mu(o)) + \epsilon\}| \geq \gamma\right\} \\
&\leq \Pr\left\{\frac{1}{n-1}\,|\{o \in O_k\,|V(1) > V_o(\mu(o)) + \epsilon\}| \geq \frac{\gamma}{K} \text{ for some } k\right\} \\
&\leq \sum_{k=1}^{K} \Pr\left\{\frac{1}{n-1}\,|\{o \in O_k\,|V(1) > V_o(\mu(o)) + \epsilon\}| \geq \frac{\gamma}{K}\right\} \\
&= \sum_{k=1}^{K} \Pr\left\{\frac{|\{o \in O_k\,|V(1) > V_o(\mu(o)) + \epsilon\}|}{|O_k|} \geq \frac{\gamma}{K}\frac{n-1}{|O_k|}\right\} \to 0
\end{aligned}$$

as $n \to \infty$, where the last inequality is by the union bound while the convergence result holds by Proposition 1.

S.5 Proof of Theorem 5

Let $f : I \to \{1, \ldots, n\}$ denote the serial orders for the agents. Recall by the basic uncertainty assumption, the distribution of the serial orders is such that for each agent i and any $\ell = 1, \ldots, n$, $\Pr\{f(i) = \ell\}$ goes to 0 as n goes to infinity.

Let us fix $\varepsilon > 0$ and $k = 1, ..., K$. Assume that f gives to agent i a serial order in $\{|O_{\leq k-1}|+2, ..., |O_{\leq k}|\}$ with the convention that $|O_{\leq 0}|+2 = 1$. We show that there is $N \geq 1$ such that for any $n \geq N$, for any vector of cardinal utilities $(\hat{u}_o)_{o\in O} := (U_i(u_o, \xi_{io}))_{o\in O}$, i cannot gain more than ε by deviating given that everyone else reports truthfully. As will be clear, the argument does not depend on the specific serial order of i within $\{|O_{\leq k-1}| + 2, ..., |O_{\leq k}|\}$ and since there are finitely many tiers, N can be taken to be uniform across all individuals with serial order in $\cup_{k=1}^{K}\{|O_{\leq k-1}| + 2, ..., |O_{\leq k}|\}$. Hence, conditional on the event that i's serial order is in $\{|O_{\leq k-1}| + 2, ..., |O_{\leq k}|\}$ for some $k = 1, ..., K$, it will follow that for any $n \geq N$, for any vector of cardinal utilities, i cannot gain more than ε by deviating given that everyone else reports truthfully. Now, by assumption, the probability of the event that i's serial order is in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$ for some $k=1,...,K$ tends to 1 as n goes to infinity and so even without conditioning, i cannot gain more than ε by deviating given that everyone else reports truthfully.

Before starting the proof of Theorem 5, we state the following lemma.

Lemma S3. *Let us assume that $\kappa(n) \geq \log^2(n)$ and $\kappa(n) = o(n)$ and consider DACB mechanism where all agents report truthfully. Fix any $k=1,\ldots,K$ and any agent i with a serial order in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$. Assuming all agents report truthfully, the probability that i is matched at Stage k converges to 1 as n goes to infinity.*

Proof. By the argument in the proof of Proposition 1, we know that with probability approaching 1 as n goes to infinity, Step $|O_{\leq k}|$ ends under DACB and, for any agent i with a serial order in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$, the outcome of DACB at the end of that step coincides with that of DA in the submarket composed only of individuals with serial orders in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$ together with the individual who was rejected at the last step of Stage $k-1$ (if $k \geq 2$) and only of objects in O_k. Since the outcome of DA does not depend on the specific serial order used, under the event that the outcome of DACB at the end of Step $|O_{\leq k}|$ coincide with that of DA in that submarket, if we permute the ordering of agents with serial orders in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$ the outcome of DACB at Step $|O_{\leq k}|$ remains the same and so the final outcome of DACB remains the same. Thus, conditional on this event, for each agent with a serial order in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$, the probability of not being matched by the end of Stage k is the same. Hence, since the number of such agents goes to infinity as n grows, this probability must go to 0 as n goes to infinity. Since the conditional event has a probability converging to 1 as n goes to infinity, (unconditionally) the probability of not being matched by the end of Stage k must go to 0 as n goes to infinity. □

In the sequel, we assume that all individuals other than i report truthfully their preferences. We partition the set of possible reports into two sets $\mathcal{T}_1$ and $\mathcal{T}_2$ as follows. $\mathcal{T}_1$ consists of the set of reports that, when restricted to objects in $O_{\geq k}$, only contain objects in O_k within the κ first objects. $\mathcal{T}_2$ consists of the set of reports which, when restricted to objects in $O_{\geq k}$ contain some object outside O_k within the κ first objects.

We will be using the following terminology: Fix a set of possible reports $\mathcal{T}$. Given an event E_{P_i} which may depend on i's report P_i, we will say that the probability of E_{P_i} converges to 1 *uniformly across all reports in $\mathcal{T}$* if for any $\epsilon > 0$, there is N such that for any $n \geq N$, $\Pr(E_{P_i}) \geq 1-\epsilon$ for any report P_i in $\mathcal{T}$.

Recall that i's serial order is in $\{|O_{\leq k-1}|+2,...,|O_{\leq k}|\}$. In the sequel, given agent i's report, we define $p(r)$ to be the probability of obtaining the r-th ranked object within the $O_{\geq k}$ objects (we abuse notations and forget about the dependence of $p(r)$ on i's report).

Lemma S4. *If i's report is of type $\mathcal{T}_1$, then $\sum_{r=1}^{\kappa} p(r)$ converges to 1. In addition, the convergence is uniform across all possible reports in $\mathcal{T}_1$. If i's report is of type $\mathcal{T}_2$, then $\sum_{r=1}^{\ell} p(r)$ converges to 1 where $\ell \leq \kappa$ is the rank (within the $O_{\geq k}$ objects) of the first object outside O_k. In addition, the convergence is uniform across all possible reports in $\mathcal{T}_2$.*[8]

Proof. Consider $\mathcal{E}$ the event under which, independently of i's reported preferences, provided that all individuals from 1 to $|O_{\leq k-1}| + 1$ report truthfully their preferences over objects in $O_{\leq k-1}$, for each $k' = 1, ..., k-1$, the objects assigned in stage k' are exactly those in $O_{k'}$. By our argument in the proof of Proposition 1, the probability of that event tends to 1. By construction, the convergence is uniform over all of i's possible reports. From now on, let us condition on the realization of event $\mathcal{E}$. By Lemma S3, we know that with (conditional) probability going to 1, i is matched within Stage k.[9] In addition, by ex ante symmetry of objects within a given tier (given our conditioning event $\mathcal{E}$, the way i ranks objects in $O_{\leq k-1}$ does not matter), the rate at which the conditional probability goes to 1 is the same for each report in $\mathcal{T}_1$. But given that $\Pr(\mathcal{E})$ goes to 1 uniformly across all possible i's reports, the unconditional probability that i is matched within Stage k also converges to 1 uniformly across these reports. This proves the first part of the lemma.

Now, we move to the proof of the second part of the lemma. Let us consider the event that for each $k' = 1, ..., K$, all individuals other than i only rank objects in $O_{k'}$ within their κ most favorite objects in $O_{\geq k'}$. Consider as well event $\mathcal{E}$ as defined above and let $\mathcal{F}$ be the intersection of these two events. By Lemma 3-(ii) as well as Proposition 1, we know that $\Pr(\mathcal{F})$ goes to 1 as n goes to infinity. By construction, the convergence is uniform over all of i's possible reports. Note that under event $\mathcal{F}$ no individual other than i will make an offer to an object outside O_k within Stage k. In addition, under $\mathcal{F}$, the probability that i is matched in Stage k goes to 1 as n goes to infinity uniformly across any report of i in $\mathcal{T}_2$.[10] Combining these observations, it must be the case that the probability that "i

[8]More precisely, we mean that for any $\epsilon > 0$, there is N such that for any $n \geq N$, $\sum_{r=1}^{\ell} p(r) \geq 1 - \epsilon$ for any $\ell \leq \kappa$ and for any report of i which, restricted to objects in $O_{\geq k}$ contains an object outside O_k at rank ℓ (where the rank is within objects in $O_{\geq k}$).

[9]The only difference with Lemma S3 is that i's report on his preferences may not be truthful and so cannot be considered to be drawn randomly. However, it should be clear that the argument goes through as long as i's report is independent of his opponents' preferences which must be true in the environment we are considering where types are drawn independently and so where players play independently. That the conditional probability of the event "i is matched within Stage k" tends to 1 comes from the facts that the unconditional probability of the event tends to 1 and that the conditioning event $\mathcal{E}$ has a probability which tends to 1.

[10]Indeed, the probability that i gets matched in Stage k under a report in $\mathcal{T}_2$ is larger than the probability that i gets matched in Stage k under the report where within the κ first objects in $O_{\geq k}$, any object outside O_k is replaced by an object in O_k. To see this, let ℓ be the rank (within $O_{\geq k}$) of the first object outside O_k under the original report. Observe that under $\mathcal{F}$, i cannot get matched to an object in a tier $k' < k$.

is matched to the object outside O_k with rank ℓ or to a better-ranked object" converges to 1 uniformly across all possible i' reports in $\mathcal{T}_2$. In addition, we know that, under $\mathcal{F}$, i can only be matched to an object in $O_{\geq k}$, hence, we get that conditional on $\mathcal{F}$, $\sum_{r=1}^{\ell} p(r)$ converges to 1 uniformly across i's possible reports. Given that $\Pr(\mathcal{F})$ goes to 1 uniformly across all possible i's reports, this statement holds for unconditional probabilities as well, as was to be shown. □

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

Corollary S2. *For any i's report, we have that $\sum_{r=1}^{\kappa} p(r)$ converges to 1. Convergence is uniform across all of i's reports.*

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

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

Let us consider a lie of individual i. Given our conditioning event, what matters are the reports within objects in $O_{\geq k}$. In addition, given the symmetry of objects within each tier, it is optimal, and thus we assume, that agent i orders the objects within each tier truthfully among them.[11] Thus, we can restrict attention to a lie in which an agent lists truthfully objects in O_k for ranks 1 to $\ell - 1$ and lists an object in $O_>k$ for rank ℓ, for some ℓ (the ranks here are that within $O_{\geq k}$ objects). Clearly, if $\ell > \kappa$ then, by the previous lemma, irrespective of the exact form of the lie, for n large enough, the lie cannot benefit agent i by more than ε. Thus, we assume without loss that $\ell \leq \kappa$. By definition of DACB,

In addition, if under the modified report, i gets matched to an object with a rank (within $O_{\geq k}$) strictly smaller than ℓ then, by definition of DACB, i will obtain the same match under the original report. Now, if i gets matched to an object with a rank (within $O_{\geq k}$) larger than ℓ then, under the original report, i applies to an object in $O_{>k}$ and so, under $\mathcal{F}$, by definition of DACB, i gets matched to that object. Since the modified report is in $\mathcal{T}_1$, and, as we already showed, the probability that i is matched in Stage k goes to 1 as n goes to infinity uniformly across any report of i in $\mathcal{T}_1$, this completes our argument.

[11] That is, for any object o and o' which both belong to the same tier k, if i prefers o to o' then i ranks o ahead of o'.

for each $r = 1, ..., \ell - 1$ individual i still has probability $p_T(r)$ to get the object with rank r under the false report. But now he has probability $p_L(\ell)$ to get matched to the object in $O_{>k}$. Now, recall that for all $n \geq N_1$, $\sum_{r=1}^{\kappa} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq 1$. This implies that for all $n \geq N_1$, $\sum_{r=1}^{\kappa} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{r=1}^{\ell-1} p_T(r) + p_L(\ell)$ and so $\sum_{r=\ell}^{\kappa} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq p_L(\ell)$. In addition, we know by the above lemma that $\sum_{r=1}^{\ell-1} p_T(r) + p_L(\ell)$ converges to 1 uniformly across all possible deviations of individual i. Since, given our conditioning event, i has zero probability to get an object in $O_{\leq k-1}$, there must exist some $N_2 \geq 1$ so that for all $n \geq N_2$, i's expected payoff when he lies is smaller than $\sum_{r=1}^{\ell-1} p_T(r)\hat{u}_r + p_L(\ell)\hat{u}_* + \frac{\varepsilon}{2}$, where $\hat{u}_*$ is the utility for agent i of the best object in $O_{>k}$. Clearly, $\hat{u}_* < \hat{u}_r$ for each $r = 1, ..., \kappa$ (recall that i's type falls into $\mathcal{T}_1$). In the sequel, we fix $n \geq \max\{N_1, N_2\}$. We obtain that, conditional on $\mathcal{E}$, the expected payoff when lying is smaller than

$$\begin{aligned}
\sum_{r=1}^{\ell-1} p_T(r)\hat{u}_r + p_L(\ell)\hat{u}_* + \frac{\varepsilon}{2} &\leq \sum_{r=1}^{\ell-1} p_T(r)\hat{u}_r + \sum_{r=\ell}^{\kappa} p_T(r)\hat{u}_* + \hat{u}_* \frac{\varepsilon}{2U(1,1)} + \frac{\varepsilon}{2} \\
&\leq \sum_{r=1}^{\ell-1} p_T(r)\hat{u}_r + \sum_{r=\ell}^{\kappa} p_T(r)\hat{u}_* + \varepsilon \\
&\leq \sum_{r=1}^{\ell-1} p_T(r)\hat{u}_r + \sum_{r=\ell}^{\kappa} p_T(r)\hat{u}_r + \varepsilon
\end{aligned}$$

where the first inequality uses the fact that $n \geq N_1$ and so that $\sum_{r=\ell}^{\kappa} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq p_L(\ell)$. The second inequality holds since $\hat{u}_* \leq U(1,1)$. The last inequality holds because $\hat{u}_* < \hat{u}_r$ for each $r = 1, ..., \kappa$. Since the expected payoff of the truth is larger than $\sum_{r=1}^{\kappa} p_T(r)\hat{u}_r$, we conclude that, conditional on $\mathcal{E}$, lying cannot make i gain more than ε whenever $n \geq \max\{N_1, N_2\}$.[12] Since $\mathcal{E}$ has a probability going to 1 uniformly across all possible deviations of individual i, a same result holds for unconditional expected payoffs.

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

Let us consider a lie by individual i. Given our conditioning event, what matters are the reports within objects in $O_{\geq k}$. In addition, given the symmetry of objects within each tier, we can assume without loss of generality that agent i orders the objects within

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

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

$$\begin{aligned}
\sum_{r=1}^{R'-1} p_T(r)\hat{u}_r + p_L(R')\hat{u}_* + \frac{\varepsilon}{2} &\leq \sum_{r=1}^{R'-1} p_T(r)\hat{u}_r + \sum_{r=R'}^{R} p_T(r)\hat{u}_* + \hat{u}_* \frac{\varepsilon}{2U(1,1)} + \frac{\varepsilon}{2} \\
&\leq \sum_{r=1}^{R'-1} p_T(r)\hat{u}_r + \sum_{r=R'}^{R} p_T(r)\hat{u}_* + \varepsilon \\
&\leq \sum_{r=1}^{R'-1} p_T(r)\hat{u}_r + \sum_{r=R'}^{R} p_T(r)\hat{u}_r + \varepsilon
\end{aligned}$$

where the first inequality uses the fact that $n \geq N_1$ which implies $\sum_{r=R'}^{R} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq p_L(R')$. The second inequality holds since $\hat{u}_* \leq U(1,1)$. The last inequality holds because $\hat{u}_* \leq \hat{u}_r$ for each $r = 1, ..., R$. Since the expected payoff of the truth is larger than $\sum_{r=1}^{R} p_T(r)\hat{u}_r$, we conclude that, conditional on $\mathcal{E}$, lying cannot benefit agent i by more than ε whenever $n \geq \max\{N_1, N_2\}$. Since, as n increases, the probability of $\mathcal{E}$ converges to 1 uniformly across all possible deviations of individual i, the same result holds for unconditional expected payoffs.

Consider next a lie which lists the best object in $O_{>k}$ for rank $R' \geq R$ (recall that the rankings are those within $O_{\geq k}$). Here again, without loss of generality, one can think of such a lie as a report in which for ranks 1 to $R-1$ (the rank here is that within $O_{\geq k}$ objects) the agent lists truthfully among objects in O_k. In that case, by definition of DACB, for each $r = 1, ..., R-1$ individual i still has probability $p_T(r)$ to get the object with rank r. But now he has probability $p_L(r)$ to get matched to the object with rank

r for each $r = R, ..., R'$. Now, recall that for all $n \geq N_1$, $\sum_{r=1}^{R} p_T(r) + \frac{\varepsilon}{2U(1,1)} \geq 1$. This implies that for all $n \geq N_1$, $\sum_{r=1}^{R-1} p_T(r) + p_T(R) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{r=1}^{R-1} p_T(r) + \sum_{r=R}^{R'} p_L(r)$ and so $p_T(R) + \frac{\varepsilon}{2U(1,1)} \geq \sum_{r=R}^{R'} p_L(r)$. In addition, we know by the above lemma that $\sum_{r=1}^{R-1} p_T(r) + \sum_{r=R}^{R'} p_L(r)$ converges to 1 uniformly across all possible deviations of individual i. Since, given our conditioning event, i has zero probability to get an object in $O_{\leq k-1}$, there must exist some $N_2 \geq 1$ so that for all $n \geq N_2$, i's expected payoff when he lies is less than $\sum_{r=1}^{R-1} p_T(r)\hat{u}_r + \sum_{r=R}^{R'} p_L(r)\hat{v}_r + \frac{\varepsilon}{2}$, where $\hat{v}_r \leq \hat{u}_R$ for each $r = R, ..., R'$. In the sequel, we fix $n \geq \max\{N_1, N_2\}$. Conditional on $\mathcal{E}$, the expected payoff when lying is no greater than

$$\begin{aligned}
\sum_{r=1}^{R-1} p_T(r)\hat{u}_r + \sum_{r=R}^{R'} p_L(r)\hat{v}_r + \frac{\varepsilon}{2} &\leq \sum_{r=1}^{R-1} p_T(r)\hat{u}_r + \sum_{r=R}^{R'} p_L(r)\hat{u}_R + \frac{\varepsilon}{2} \\
&\leq \sum_{r=1}^{R-1} p_T(r)\hat{u}_r + p_T(R)\hat{u}_R + \frac{\varepsilon}{2U(1,1)}\hat{u}_R + \frac{\varepsilon}{2} \\
&\leq \sum_{r=1}^{R-1} p_T(r)\hat{u}_r + p_T(R)\hat{u}_R + \varepsilon
\end{aligned}$$

where the first inequality uses the fact that $\hat{u}_R \geq \hat{v}_r$ for all $r = R, ..., R'$. The second inequality uses the fact that $n \geq N_1$ which implies $\sum_{r=R}^{R'} p_L(r) \leq p_T(R) + \frac{\varepsilon}{2U(1,1)}$. The last inequality follows from the fact that $\hat{u}_R \leq U(1,1)$. Since the expected payoff of the truth is larger than $\sum_{r=1}^{R} p_T(r)\hat{u}_r$, we conclude that, conditional on $\mathcal{E}$, lying cannot make i gain more than ε whenever $n \geq \max\{N_1, N_2\}$. Since, as n increases, the probability of $\mathcal{E}$ converges to 1 uniformly across all possible deviations of individual i, a same result holds for unconditional expected payoffs.

S.6 Analysis of the Extended DACB Algorithm

S.6.1 Proof of Theorem 6

The following lemma will be instrumental for the proof and is inspired from an observation by Wilson (1972) that in the environment where preferences are uncorrelated, priorities are arbitrary and the market is balanced, a modification of the DA algorithm can be studied as a standard urn model.[13] Using this analogy Wilson (1972) shows that the expected total

[13] Wilson (1972) shows that in this environment, assuming individuals are memoryless, the number of offers needed for all objects to be matched is equivalent to the number of trials needed to collect all n coupons in the Coupon Collector Problem (in this problem, coupons are being collected, equally likely, within an urn of n different coupons).

number of offers made by individuals under DA is smaller $n \log(n)$. We strengthen this result by proving a concentration result.

Lemma S5. *Consider any arbitrary objects' priorities. In the case individuals' preferences are uncorrelated (i.e., where $U_i(o) = \xi_{i,o}$ for all i,o), for any $\gamma > 1$, with probability going to 1 as n goes to infinity, the total number of offers made by individuals under DA is smaller than $\gamma n \log(n)$.*

PROOF. We start from the version of the DA mechanism as defined in McVitie and Wilson (1971). By the principle of deferred decisions, we assume that whenever an individual has an opportunity to make an offer, he makes his offer randomly to an object to which he has not yet made any offer. Now, as proposed in Wilson (1972), we modify this mechanism assuming that agents are memoryless: whenever an agents has an opportunity to make an offer he makes this offer randomly to an object in O, including those to which he has already made offers to. For each realization of individuals' preferences, the outcome is the same as with memory. Indeed, if an agent makes an offer to an object which already rejected him, he will continue to be rejected and the final outcome remains unchanged. The total number of offers when individuals have no memory must be larger than in the original case where agents have memory. We let X be the total number of offers needed for all objects in O to be matched under the mechanism where agents are memoryless. Given that $\gamma > 1$, it is enough to show that $\Pr(X \geq \gamma n \log(n)) \leq \frac{1}{n^{\gamma-1}}$. For any particular object $o \in O$, the probability that o receives no offer by time $\gamma n \log(n)$ offers are made in McVitie and Wilson's algorithm[14] is

$$(1 - \frac{1}{n})^{\gamma n \log(n)} = \left((1 - \frac{1}{n})^n\right)^{\gamma \log(n)} \leq e^{-(\gamma \log(n))} = \frac{1}{n^\gamma} \tag{S0}$$

where the first term is the probability that all the first $\gamma n \log(n)$ offers have been directed to objects other than o.

Next observe that the algorithm ends (and the assignment is complete) once every object receives at least one offer. Hence, the probability that $X \geq \gamma n \log(n)$ is the probability that the algorithm (with memorlyess agents) is not complete after $n \log(n)$ offers have been made, which in turn equals the probability that at least one object has not received an offer by the time $n \log(n)$ total offers have been made. By the union bound and using Equation (S0), this latter probability is no greater $n \frac{1}{n^\gamma} = \frac{1}{n^{\gamma-1}}$. $\square$

[14] At each stage of the algorithm, there is an individual who is rejected (the identity of the individual may depend on objects' priorities but his identity does not matter for our computations) and, using the principle of deferred decisions and the fact that individuals are memoryless, this individual makes an offer randomly to an object in O.

In the sequel, for each size of the market n, we let $j(n)$ and $\kappa(n)$ be the two parameters of DACB. The following result is a straightforward implication of the above lemma.

Corollary S3. *Consider any arbitrary objects' priorities. In the case individuals' preferences are uncorrelated, if* $\liminf_{n\to\infty} \frac{j(n)\kappa(n)}{n\log(n)} > 1$ *then under DA we must have*

$$\Pr\{\textit{fewer than } j(n) \textit{ agents make more than } \kappa(n) \textit{ offers}\} \to 1$$

as $n \to \infty$.

Proof. Proceed by contradiction and assume that there is $\delta > 0$ and a sequence $n_k \to \infty$ as $k \to \infty$ such that along that sequence

$$\Pr\{\text{more than } j(n_k) \text{ individuals make more than } \kappa(n_k) \text{ offers}\} > \delta.$$

This implies that along the sequence $\{n_k\}$, there is a probability greater than $\delta > 0$ that the total number of offers made under DA is greater than $j(n_k)\kappa(n_k)$. Now, note that since $\liminf_{n\to\infty} \frac{j(n)\kappa(n)}{n\log(n)} > 1$, we must have that for some $\gamma > 1$ and for n_k large enough, $j(n_k)\kappa(n_k) > \gamma n_k \log(n_k)$. Hence, we obtain that along the sequence $\{n_k\}$, there is a probability greater than $\delta > 0$ that the total number of offers made under DA is strictly greater than $\gamma n_k \log(n_k)$ which yields a contradiction. □

In the sequel we fix the two parameters of the DACB mechanism to be $j(n)$ and $\kappa(n)$. Theorem 6 directly follows from the proposition below.

Proposition S2. *Fix any* $k \geq 1$*. As* $n \to \infty$*, with probability approaching one, at the end of Stage* k *of DACB, all objects in* O_k *are assigned and at most* $j(n)$ *objects outside* O_k *are assigned. In addition, for any* $\epsilon > 0$

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

where $\hat{I}_k$ *is the set of individuals matched at Stage* k *of DACB. Similarly,*

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

where $\hat{O}_k$ *is the set of objects matched at Stage* k *of DACB.*

Proof of Proposition S2. We focus on Stage $k = 1$, as will become clear, the other cases can be treated in a similar 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 objects. If we were to run standard DA just for this submarket, then because preferences are drawn *iid*, by Lemma 2 and Corollary S3, with probability approaching 1 as $n \to \infty$, at the end of (standard) DA, (a) *all* agents have made fewer than $\log^2(|O_1|) \le \log^2(n)$ offers and (b) fewer than $j(|O_1|) \le j(n)$ individuals have made more than $\kappa(|O_1|) \le \kappa(n)$ offers.

Consider now the original market. For any $\delta > 0$, since $k(n) = o(n)$, we must have that $k(n) \le \delta |O_1|$ for any n large enough. Hence, by Lemma 3-(ii) the event that for each agent's $\max\{\kappa(n), \log^2(n)\}$ favorite objects are in O_1 has probability approaching 1 as $n \to \infty$. Let us condition on this event, labeled $\mathcal{E}$.

We now show that, conditional on $\mathcal{E}$, with probability approaching 1 as $n \to \infty$, *all* objects in O_1 and no more than $j(n)$ objects outside O_1 are assigned by the end of Stage 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 $\mathcal{E}$, as long as each agent has made fewer than $\log^2(n)$ offers (which ensures that offers are only made to O_1 objects) and fewer than $j(n)$ individuals have made more than $\kappa(n)$ offers (which ensures that the end of Stage 1 is not triggered), the $|O_1|$ first steps of DACB proceed exactly in the same way as DA in the submarket composed of the $|O_1|$ first agents (according to the ordering used in DACB) and of all objects in O_1. Hence, as mentioned above, with probability going to 1 as $n \to \infty$, we then reach the end of Step $|O_1|$ of DACB before Stage 1 ends (i.e., before more than $j(n)$ individuals applied to their $\kappa(n)$ most favorite object). Hence, conditional on $\mathcal{E}$, with probability going to 1, all objects in O_1 are assigned before Stage 1 ends. In addition, under event $\mathcal{E}$, if more than $j(n)$ individuals make offers to objects outside O_1 before Stage 1 ends, then more than $j(n)$ individuals make more than $\kappa(n) = o(n)$ offers which is not possible by definition of a stage in DACB. Hence, under event $\mathcal{E}$, Stage 1 must end before more than $j(n)$ objects outside O_1 are assigned. Now, since $\Pr(\mathcal{E}) \to 1$ as $n \to \infty$, we obtain that at the end of stage 1 of DACB, all objects in O_1 are assigned and at most $j(n)$ objects outside O_1 are assigned with probability going to 1. This completes the proof of the first part of Proposition S2.

Now, we move to the proof of the second part of Proposition S2. We fix any $\epsilon > 0$ and $\gamma < 1$ and want to show that as $n \to \infty$,

$$\Pr\left\{\frac{|\{i \in \hat{I}_1 | U_i(DACB(i)) \ge U(u_1, 1) - \epsilon\}|}{|\hat{I}_1|} > \gamma\right\} \longrightarrow 1$$

and

$$\Pr\left\{\frac{|\{o \in \hat{O}_1 | V_o(DACB(o)) \ge V(1) - \epsilon\}|}{|\hat{O}_1|} > \gamma\right\} \longrightarrow 1.$$

In the sequel, we condition on event $\mathcal{E}$. Recall that, with probability going to 1, the number of individuals matched in Stage 1 is between $|O_1|+1$ and $|O_1|+j(n)$ and all these individuals except possibly for $j(n)$ of them obtain an object within their $\kappa(n)$ most favorite objects. By Lemma 3-(i) this implies that, with probability going to 1, all these individuals but potentially $j(n) = o(n)$ of them enjoy a payoff above $U(u_1, 1) - \epsilon$. As we have shown, with probability going to 1, the first $|O_1|$ Steps of Stage 1 of DACB proceed exactly in the same way as DA in the submarket that consists of the $|O_1|$ first agents (according to the ordering used in DACB) and of all objects in O_1. We first note that, by Lemma 2, under DA in this submarket, with probability going to 1 the proportion of objects in O_1 with a rank smaller than $\frac{2}{1-\gamma}|O_1|/\log(|O_1|)$ is larger than γ. To see this, suppose to the contrary that with probability bounded away from 0, as the market grows, the proportion of objects with a rank above $\frac{2}{1-\gamma}|O_1|/\log(|O_1|)$ is more than $1-\gamma$. Then, with probability bounded away from 0, as the market grows: $\frac{1}{|O_1|}\sum_{o \in O_1} R_o^{DA} > \frac{1}{|O_1|}(1-\gamma)|O_1|\frac{2}{1-\gamma}(|O_1|/\log(|O_1|)) = 2|O_1|/\log(|O_1|)$ which yields a contradiction to Lemma 2. Hence, we obtain that with probability going to 1 by the end of Step $|O_1|$ of Stage 1 of DACB, the proportion of objects in O_1 with a rank smaller than $\frac{2}{1-\gamma}|O_1|/\log(|O_1|)$ is larger than γ. Given that for any $\delta > 0$, for n large enough, $|O_1|/\log(|O_1|) \leq \delta\,|I|$, by Lemma 3-(iii), we must also have that, with probability going to 1, the proportion of objects o in O_1 with $V(DACB(o)) \geq 1 - \epsilon$ is above γ. Since objects in O_1 will have received even more offers at the end of Stage 1, it must still be that, with probability going to 1, the proportion of objects in O_1 for which $V(DACB(o)) \geq 1-\epsilon$ is above γ when n is large enough. We ignore the remaining objects matched in Stage 1 since there are fewer than $j(n) = o(n)$ such objects. Thus, for $k = 1$, the second statement in Proposition S2 is proved provided that our conditioning event $\mathcal{E}$ holds. Since, this event has probability going to 1 as $n \to \infty$, the result must hold even without the conditioning. Thus, we have proved Proposition S2 for the case $k = 1$.

Consider next Stage $k > 1$. The objects remaining in Stage k have received no offers in Stages $1, ..., k-1$ (or else the objects would have been assigned in those stages). Hence, by the principle of deferred decisions, we can assume that the individuals' preferences over those objects are yet to be drawn in the beginning of Stage k. Similarly, we can assume that priorities of those objects are also yet to be drawn. Put in another way, conditional on Stage $k-1$ being over, we can assume without loss that the distribution of preferences and priorities is the same as the unconditional one. Thus, we can consider the market composed of the individuals and objects not matched in previous stages. We can set O_1 to be equal to the set of remaining objects in O_k, O_2 to be equal to the set of remaining objects in O_{k+1}, etc... (with high probability, the cardinality of each tier defined in this way is linear in n, i.e., between $|O_\ell| - (\ell-1)j(n)$ and $|O_\ell|$ for each ℓ) so the exact same reasoning as above completes the argument. □

S.6.2 Incentives under the Extended Algorithm

In the sequel, we consider the extended version of DACB with parameters $j(n)$ and $\kappa(n)$. We first slightly strengthen our assumption that the serial orders admit some *basic uncertainty* from the agents' perspective: for each sequence of sets $E^n \subset \{1,\ldots,n\}$ such that $\lim |E^n|/n$ goes to 0 as n goes to infinity, we assume that the probability that any agent i receives a serial order in E^n goes to zero as $n \to \infty$.

We show that the following result.

Theorem S4. *Consider a sequence of economies satisfying* $\liminf_{n\to\infty} \frac{j(n)\kappa(n)}{n\log(n)} > 1$ *and* $j(n)$ *and* $\kappa(n)$ *are* $o(n)$*. Fix any* $\epsilon > 0$*. Under DACB, there exists* $N > 0$ *such that for all* $n > N$*, truthtelling is an interim* ϵ*-Bayes-Nash equilibrium.*

The proof is rather similar to that of Theorem 5. The two main lemmas (Lemmas S4 and S3) from the proof of Theorem 5 have to be adapted. The rest of the proof is in essence the same and is thus omitted.

Let us fix $\varepsilon > 0$ and $k = 1,...,K$. Assume that the ordering of DACB gives to agent i a serial order in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$ with the convention that $|O_{\leq 0}|+1+j(n) = 1$. We show that there is $N \geq 1$ such that for any $n \geq N$, for any vector of cardinal utilities $(\hat{u}_o)_{o\in O} := (U_i(u_o,\xi_{io}))_{o\in O}$, i cannot gain more than ε by deviating given that everyone else reports truthfully. As will be clear, the argument does not depend on the specific serial order of i within $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$ and so given that there are finitely many tiers, N can be taken to be uniform across all individuals with serial order in $\cup_{k=1}^{K}\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$. Hence, conditional on the event that i's serial order is in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$ for some $k = 1,...,K$, it will follow that for any $n \geq N$, for any vector of cardinal utilities, i cannot gain more than ε by deviating given that everyone else reports truthfully. Now, given our assumption on the distribution from which the ordering of DACB is drawn, the probability of the conditioning event goes to 1. Hence, for any $n \geq N$, for any vector of cardinal utilities $(\hat{u}_o)_{o\in O} := (U_i(u_o,\xi_{io}))_{o\in O}$, i cannot gain more than ε by deviating given that everyone else reports truthfully – which shows the desired result.

Lemma S6. *Let us assume that* $\liminf_{n\to\infty} \frac{j(n)\kappa(n)}{n\log(n)} > 1$*. Consider the DACB mechanism. Fix any* $k = 1,\ldots,K$ *and any agent* i *with a serial order in* $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$*. Assuming all agents report truthfully, the probability that* i *is matched at Stage* k *to one of his* κ *most favorite choices within remaining objects at that stage converges to 1 as* $n \to \infty$*.*

Proof. By the argument in the proof of Proposition S2, we know that with probability approaching 1 as n goes to infinity, Step $|O_{\leq k}|$ ends under DACB[15] and, for any agent i

[15] Here, when counting the number of steps which occurred by the end of Stage k, we consider the total

with a serial order in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$, the outcome of DACB at the end of that step coincides with that of DA in the submarket composed only of remaining individuals at that stage with a serial order below $|O_{\leq k}|$ (which contains all individuals with serial order in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$) and the remaining objects in O_k. Since the outcome of DA does not depend on the specific linear order used, under the event that the outcome of DACB at the end of Step $|O_{\leq k}|$ coincide with that of DA in that submarket, if we switch the ordering (of the linear order of DACB) of agents with serial order in $\{|O_{\leq k-1}|+j(n)+1,...,|O_{\leq k}|\}$ the outcome of DACB at Step $|O_{\leq k}|$ remains the same and so the final outcome of DACB remains the same. Thus, conditional on this event, for each agent with a serial order in $\{|O_{\leq k-1}|+j(n)+1,...,|O_{\leq k}|\}$, the probability of either not being matched by the end of Stage k or of being one of the $j(n)$ individuals making more than κ offers is the same. Hence, since the number of such agents goes to infinity as n grows, this probability must go to 0 as n goes to infinity. Since the conditional event has a probability converging to 1 as n goes to infinity, (unconditionally) this probability must go to 0 as n goes to infinity. □

In the sequel, we let $\hat{O}_{\geq k}$ be the set of available objects in Stage k. Note that this does not depend on i's reports (given that i's serial order is in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$). We assume that all individuals other than i report truthfully their preferences. We partition the set of i's possible reports into two sets $\mathcal{T}_1$ and $\mathcal{T}_2$ as follows. $\mathcal{T}_1$ consists of the set of i's reports that, when restricted to objects in $\hat{O}_{\geq k}$, only contain objects in O_k within the κ first objects. $\mathcal{T}_2$ consists of the set of i's reports which, when restricted to objects in $\hat{O}_{\geq k}$, contain some object outside O_k within the κ first objects.

We will again be using the following terminology: Fix a set of possible reports $\mathcal{T}$. Given an event E_{P_i} which may depend on i's report P_i, we will say that the probability of E_{P_i} converges to 1 *uniformly across all reports in* $\mathcal{T}$ if for any $\epsilon > 0$, there is N such that for any $n \geq N$, $\Pr(E_{P_i}) \geq 1-\epsilon$ for any report P_i in $\mathcal{T}$.

Recall that i's serial order is in $\{|O_{\leq k-1}|+1+j(n),...,|O_{\leq k}|\}$. In the sequel, given agent i's report, we define $p(r)$ to be the probability of obtaining the r-th ranked object within the $\hat{O}_{\geq k}$ objects (we abuse notations and forget about the dependence of $p(r)$ on i's report).

Lemma S7. *If i's report is of type $\mathcal{T}_1$, then $\sum_{k=1}^{\kappa} p(k)$ converges to 1. In addition, the convergence is uniform across all possible reports in $\mathcal{T}_1$. If i's report is of type $\mathcal{T}_2$, then $\sum_{k=1}^{\ell} p(k)$ converges to 1 where $\ell \leq \kappa$ is the rank (within the $\hat{O}_{\geq k}$ objects) of the first object outside O_k. In addition, the convergence is uniform across all possible reports in $\mathcal{T}_2$.*[16]

---

number of steps from the beginning of Stage 1. Hence, we say that Step $|O_{\leq k}|$ of Stage k ends if, from the beginning of Stage 1, $|O_{\leq k}|$ steps have occurred.

[16]More precisely, we mean that for any $\epsilon > 0$, there is N such that for any $n \geq N$, $\sum_{k=1}^{\ell} p(k) \geq 1-\epsilon$

PROOF. Consider $\mathcal{E}$ the event under which, independently of i's reported preferences, provided that all individuals from 1 to $|O_{\leq k-1}|+j(n)$ report truthfully their preferences, for each $k' = 1, ..., k-1$, the objects assigned in Stage k' contain all those in $O_{k'}$ and contain no more than $j(n)$ objects from $O_{>k'}$. By our argument in the proof of Proposition 1, the probability of this event tends to 1. By construction, the convergence is uniform over all of i's possible reports. From now on, let us condition on the realization of event $\mathcal{E}$. By Lemma S6, we know that with (conditional) probability going to 1 as $n \to \infty$, i makes fewer than κ offers and is matched within Stage k. In addition, by ex ante symmetry of objects within a given tier (given our conditioning event $\mathcal{E}$, the way i ranks objects in $O_{\leq k-1}$ does not matter), the rate at which the conditional probability goes to 1 is the same for each report in $\mathcal{T}_1$. But given that $\Pr(\mathcal{E})$ goes to 1 uniformly across all possible i's reports, the unconditional probability that i is matched within Stage k to an object within his κ most favorite in $\hat{O}_{\geq k}$ also converges to 1 uniformly across these reports. This proves the first part of the lemma.

We next move to the proof of the second part of the lemma. Let us consider the event that for each $k' = 1, ..., K$, all individuals other than i only rank objects in $O_{k'}$ within their κ most favorite objects in $\hat{O}_{\geq k'}$ (recall that with probability approaching 1 as n goes to infinity, the size of $\hat{O}_{\geq k'} \cap O_{k'}$ is linear in n). Consider as well event $\mathcal{E}$ as defined above and let $\mathcal{F}$ be the intersection of these two events. By Lemma 3-(ii) as well as Proposition 1, we know that $\Pr(\mathcal{F})$ goes to 1 as n goes to infinity. By construction, the convergence is uniform over all of i's possible reports. Note that under event $\mathcal{F}$ no individual other than i will make an offer to an object outside O_k within Stage k. In addition, under $\mathcal{F}$, the probability that i is matched in Stage k and obtains an object with his κ most favorite with remaining objects goes to 1 as n goes to infinity uniformly across any report of i in $\mathcal{T}_2$.[17] Combining these observations, it must be the case that the probability that "i is matched to the object outside O_k with rank ℓ or to a better object" converges to 1 uniformly across all

for any $\ell \leq \kappa$ and for any report of i which, restricted to objects in $O_{\geq k}$ contains an object outside O_k at rank ℓ (where the rank is within objects in $O_{\geq k}$).

[17] Indeed, the probability that i gets matched in Stage k to one of his κ most favorite object in $\hat{O}_{\geq k}$ under a report in $\mathcal{T}_2$ is larger than the probability that i gets matched in Stage k to one of his κ most favorite object in $\hat{O}_{\geq k}$ under the report where within the κ first objects in $\hat{O}_{\geq k}$, any object outside O_k is replaced by an object in O_k. To see this, let ℓ be the rank (within $\hat{O}_{\geq k}$) of the first object outside O_k under the original report. Observe that under $\mathcal{F}$, i cannot get matched to an object in a tier $k' < k$. In addition, if under the modified report, i gets matched to an object with a rank (within $\hat{O}_{\geq k}$) strictly smaller than ℓ then, by definition of DACB, i will obtain the same match under the original report. Now, if i gets matched to an object with a rank (within $\hat{O}_{\geq k}$) in $\{\ell, \ldots, \kappa\}$ then, under the original report, i applies to an object in $O_{>k}$ and so, under $\mathcal{F}$, by definition of DACB, i gets matched to that object. Since the modified report is in $\mathcal{T}_1$, and, as we already showed, the probability that i is matched in Stage k to one of his κ most favorite object in $\hat{O}_{\geq k}$ goes to 1 as n goes to infinity uniformly across any report of i in $\mathcal{T}_1$, this completes our argument.

possible i' reports in $\mathcal{T}_2$. In addition, we know that under $\mathcal{F}$, agent i can only be matched to an object in $\hat{O}_{\geq k}$, hence, $\sum_{k=1}^{\ell} p(k)$ converges to 1 uniformly across i's possible reports. Given that $\Pr(\mathcal{F})$ goes to 1 uniformly across all possible i's reports, this statement holds for unconditional probabilities as well, as was to be shown. □

S.7 Simulation Based on Randomly Generated Data

.

Our analytical results in the paper are obtained in a limit market as $n \to \infty$, and one concern could be their validity for realistic markets, with large but finite participants. Our calibration work based on the NYC School Choice data already suggests that the DACB would work well in the field, such as in NYC school choice, offering a range of compromises between the two objectives superior to the simple convexification of DA and TTC. But what about smaller markets? To test the robustness of our results to market size, we performed a number of simulations to see how alternative mechanisms would work under market with different number of participants.

S.7.1 Varying Market Sizes

Assume 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 $[0,1]$.[18] We later change the support of u_o to $[0,3]$ to investigate the impact of increased correlation in agents' preferences. For the case of DACB, we set $\kappa(n) = \log^2(n)$ throughout.

Figure 1 shows the utilitarian welfare—more precisely the average idiosyncratic utility enjoyed by the agents—under the alternative algorithms for the varying market size ranging from $n = 10$ to $10,000$.[19]

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 increase 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

[18] We also performed several simulations in which the common values are distributed over two values, as would be in the two tier case. The simulations outcomes in this case are largely similar to the ones reported here.

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

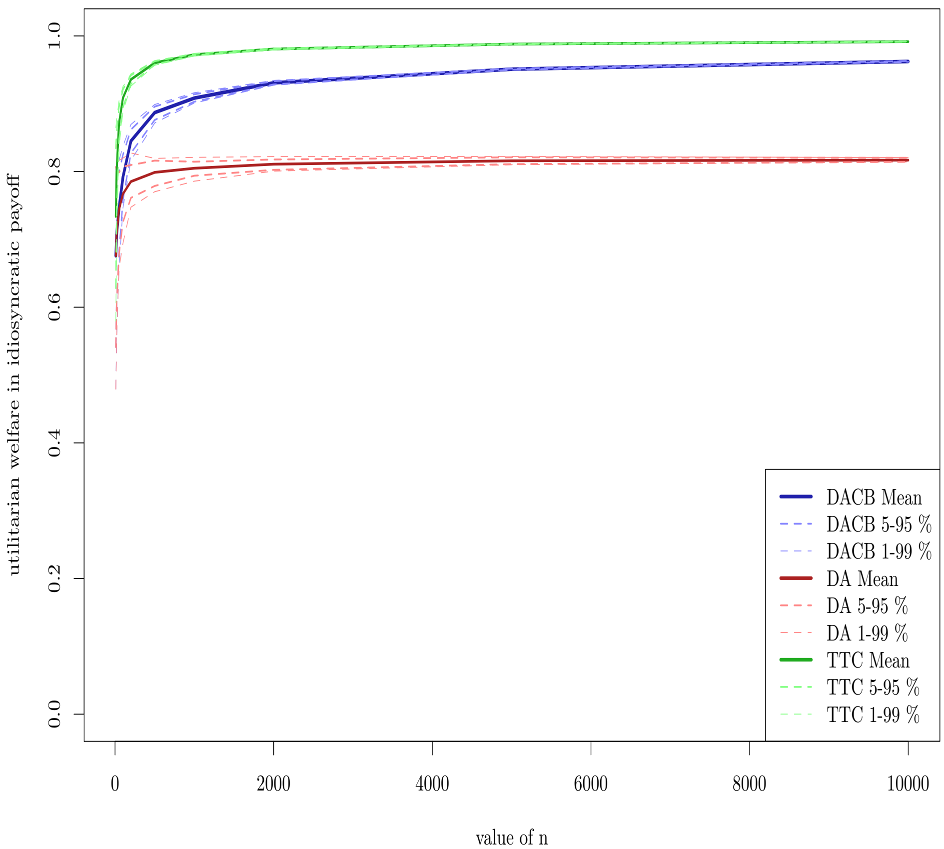

Figure 1: Utilitarian welfare under alternative mechanisms

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 2 shows the fraction of ϵ-blocking pairs under DA, DACB and TTC. Clearly, DA admits no blocking pairs, so the fraction is always zero. Between TTC and DACB, there is a substantial difference. Blocking pairs admitted by TTC comprise almost 9% of all possible pairs, whereas DACB admits blocking pairs that are less than 1% of all possible pairs, and these proportions do not vary much with the market size. This result is consistent with our calibration result showing that the number of blocking pairs shrinks dramatically when the mechanism is shifted from TTC to DACB. When we weaken the notion of stability to $\epsilon-$stability, focusing only on the pairs of the

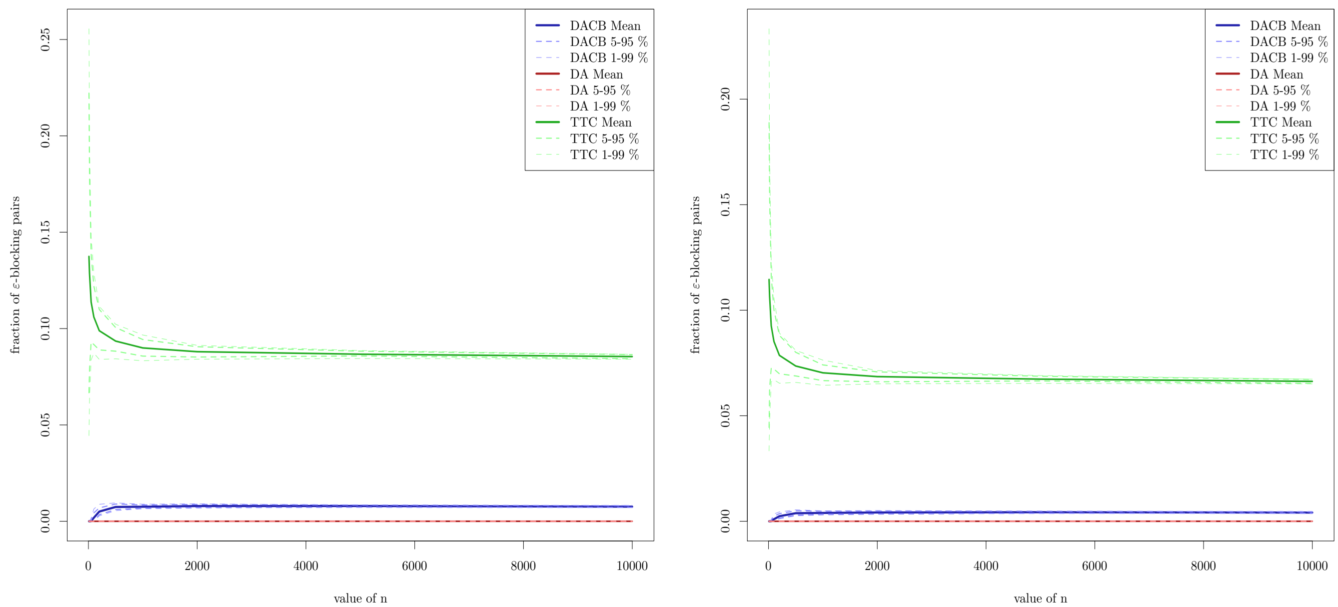

Figure 2: The fraction of ϵ-blocking pairs under alternative mechanisms with $\epsilon = 0$ (left panel) and $\epsilon = 0.05$ (right panel)

agent and object that would benefit from blocking more than ϵ percentile improvement in rankings, the fractions decline (not surprisingly), but the differences between the two mechanisms remain significant. Under $\epsilon = 0.05$, the fractions of $\epsilon-$blocking pairs are around 7% for TTC, but the fraction is close to zero under DACB, again largely irrespective of the market size. These results are consistent with the earlier calibration results showing that the number of blocking pairs decrease dramatically when we shift from TTC to DACB.

S.7.2 Many-to-one matching with varying κ

Recall our analysis focuses on one-to-one matching. In order to test the robustness of our results to a realistic many-to-one matching setting, we consider a model in which 100,000 students are to be assigned to 500 schools each with 200 seats. We assume that common and idiosyncratic values are both drawn uniformly from $[0, 1]$. We simulate DACB under a variety of values of κ's. Figure 3 shows the sense in which DACB with different κ's span different ways to compromise on efficiency (left panel) and stability (right panel). For a low value of κ, DACB performs similarly to TTC, attaining about 97.5% of the utilitarian welfare upper bound and admitting about 18% of blocking pairs, whereas DA attains close

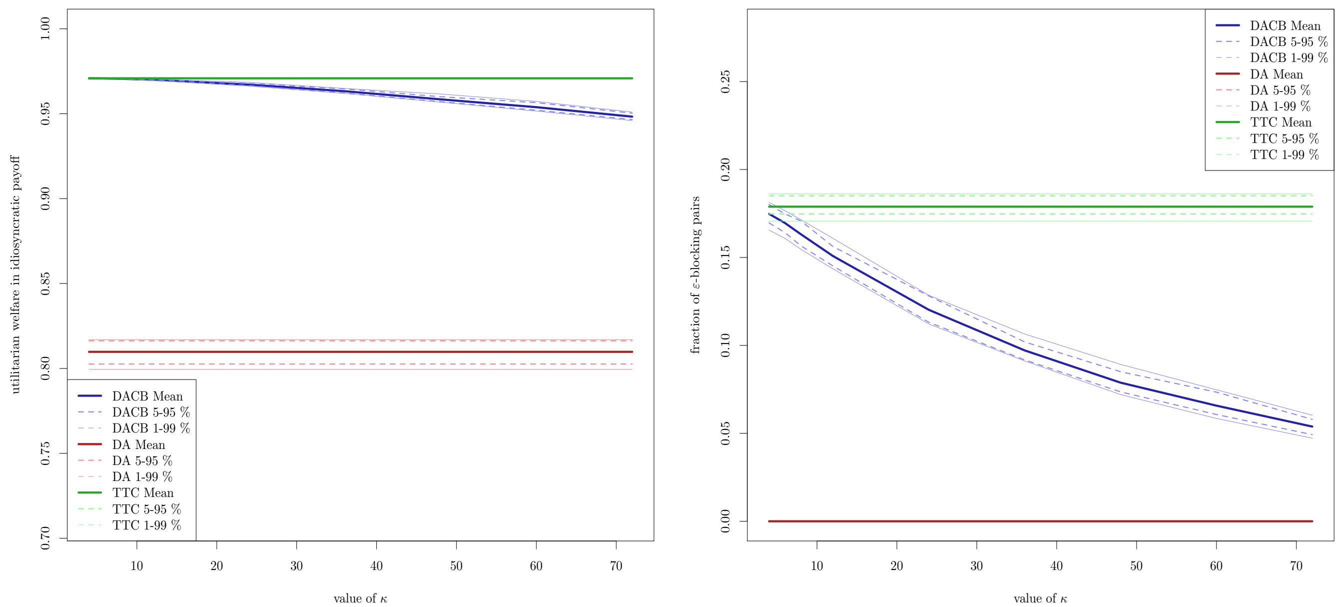

Figure 3: Utilitarian welfare (left panel) and the fraction of 0-blocking pairs (right panel) under DACB with varying κ's

to 80% of the utilitarian welfare upper bound. For a low κ, DACB resembles TTC with high welfare and high incidences of blocking pairs. As κ rises, the utilitarian welfare under DACB falls, but the number of blocking pairs declines as well.

In practice, the value of κ can be chosen to attain a good compromise between utilitarian efficiency and stability. For instance, DACB with $\kappa = 35$ keeps the utilitarian welfare still very high at 97% but the proportion of blocking pairs below 10%. For DACB with $\kappa = 70$, the utilitarian welfare remains still high at 95% but the fraction of blocking pairs drops down to almost 5%.

The "right" value of κ will of course depend on the precise preference structure as well as the allowable margin for violations of the exact objectives—i.e., the values of ϵ's in ϵ-efficiency and ϵ-stability. To illustrate, Figure 4 below shows again the performances of DACB with different κ's, when the common values are drawn uniformly from $[0, 1/2]$ (idiosyncratic values are still drawn uniformly from $[0, 1]$). DACB with $\kappa = 15$ attains 97% of the utilitarian welfare upper bound at admits less than 5% of ϵ-blocking pairs for $\epsilon = 0.05$.

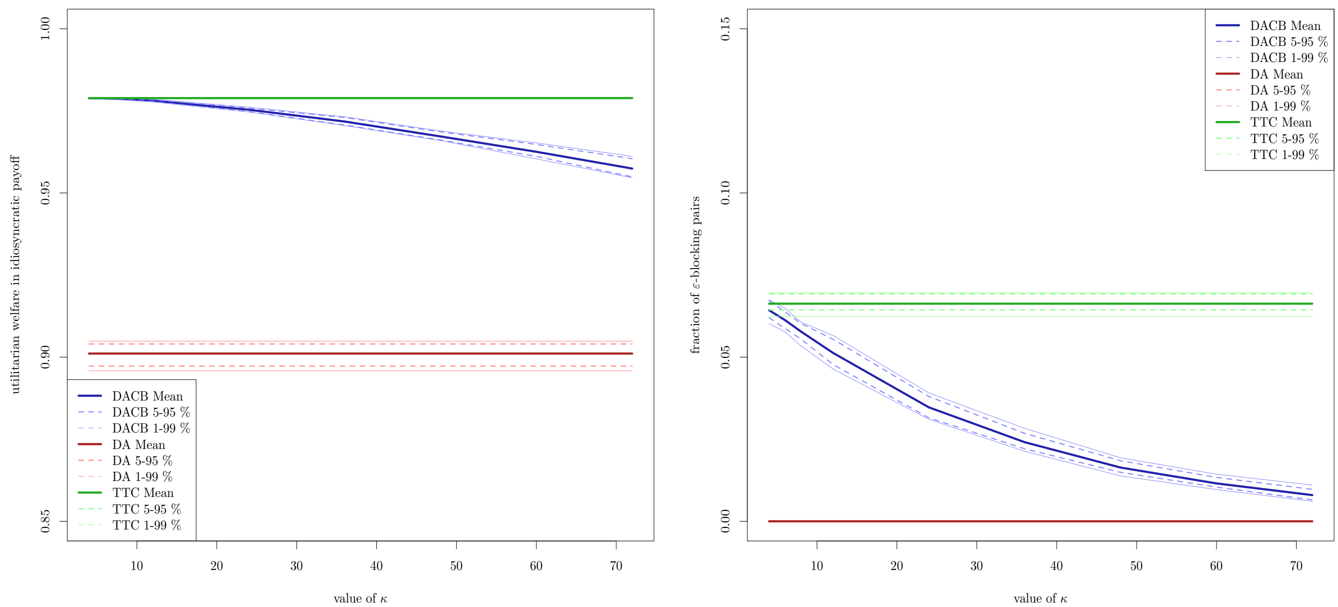

Figure 4: Utilitarian welfare (left panel) and the fraction of 0.05-blocking pairs (right panel) under DACB with varying κ's

S.7.3 Correlated Priorities

Lee and Yariv (2014) show that any stable matching (and thus DA) is asymptotically efficient if agents' priorities have common component that is drawn from a non-degenerate interval according to an absolutely continuous distribution function. To see how DACB compares with DA under this environment, we performed several simulations under the assumption that agent i's priority at object o is given by $V(v_i, \eta_{i,o}) = v_i + \eta_{i,o}$, where both v_i and $\eta_{i,o}$ are each uniformly distributed from $[0, 1]$. As before, we assume that agent's preference is given by $U(u_o, \xi_{i,o}) = u_o + \xi_{i,o}$where $\xi_{i,o}$ is distributed uniformly from $[0, 1]$, and u_o is uniform on $[0, 1]$ or $[0, 3]$. In keeping with Sections 6 and 7 in the paper, when we run DACB in this environment, the serial orders of the agents are determined based on their average priorities, i.e., the agent with the best average priority gets the first position in the serial order, the one with the second average priority gets the second, and so on. Hence, the name priority-based (PB) DACB.

Figure 5 compares utilitarian welfare under DA, DACB and TTC when u_o is uniform on $[0, 1]$.

As can be seen, the TTC performs best, followed by DACB, which performs in turn

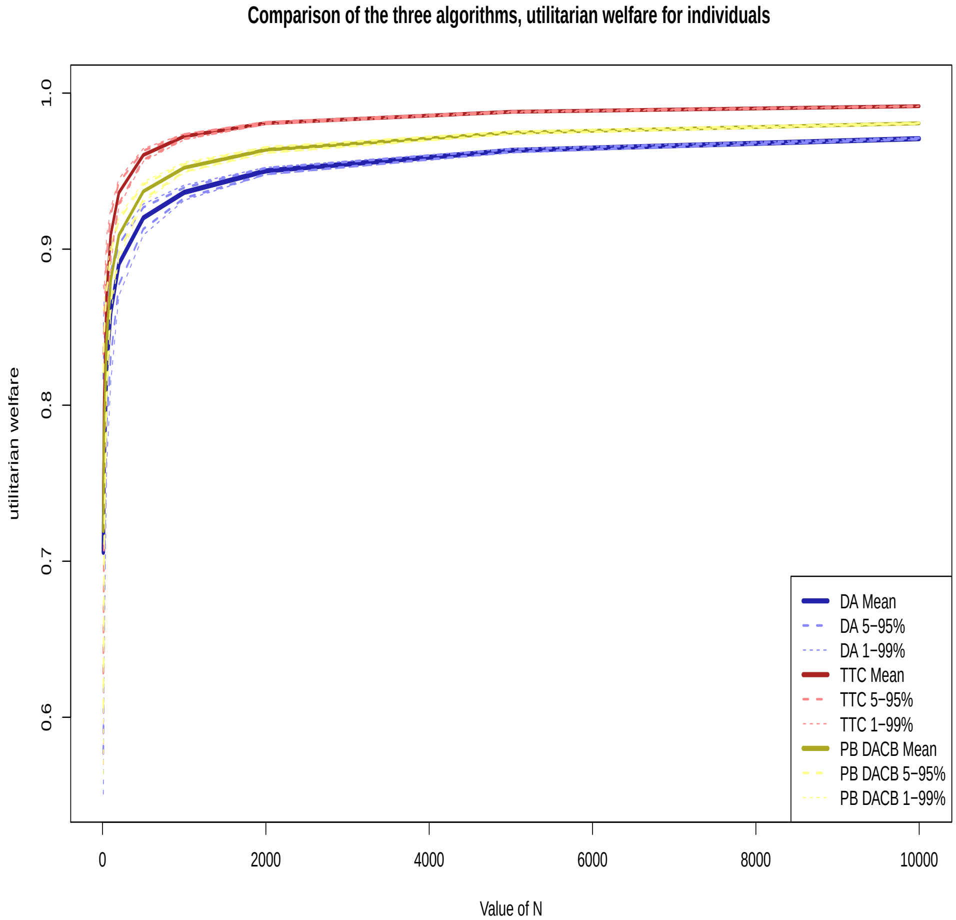

Figure 5: Utilitarian welfare under correlated priorities; u_o uniform on $[0, 1]$

better than DA. While both DA and DACB attain higher welfare as the size of the market increases, one can see that the difference between the two is significant for a reasonable market size.

Figure 6 shows the same comparison when u_o is uniform on $[0, 3]$. This change simply means that agents' preferences exhibit higher correlation than before. Hence, as discussed in the text, this case implicitly involves market imbalance and excessive competition toward high quality objects. In light of our result in Section 4.0.2 and Ashlagi, Kanoria, and Leshno (2013), one would expect the gap between DACB and DA to widen in this case. This is indeed what we see here.

Last, Figures 7 and 8 compare the number of blocking pairs under alternative mecha-

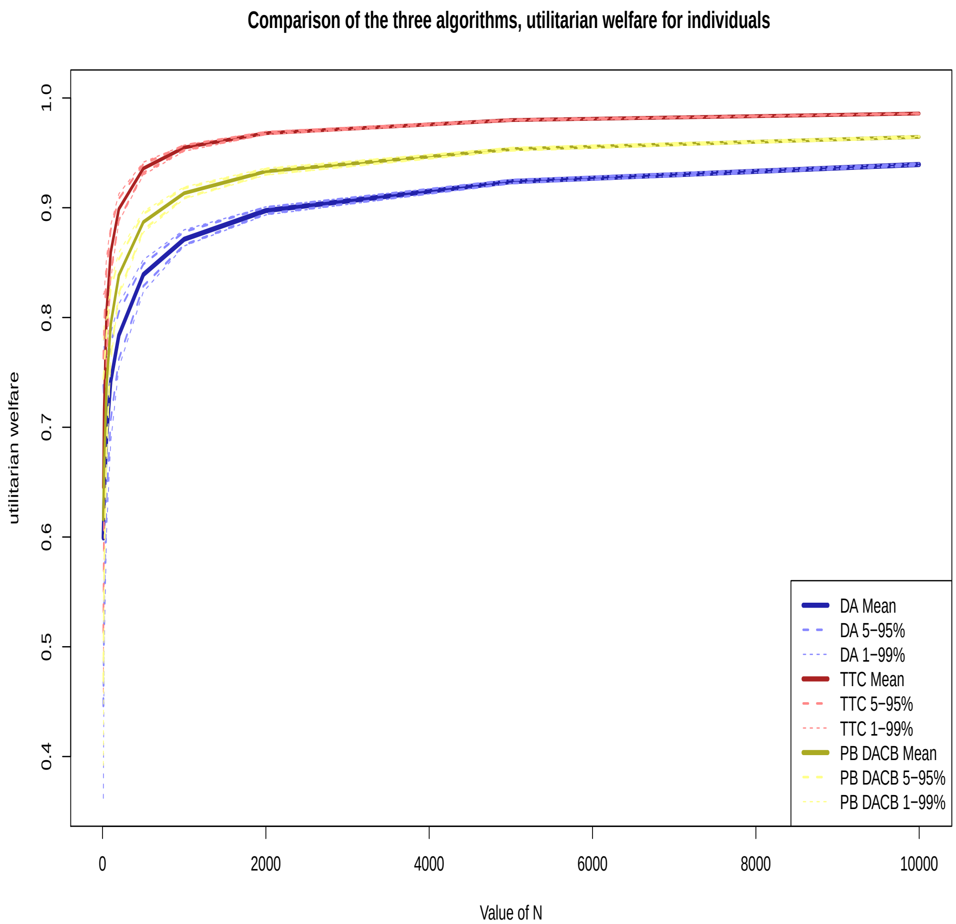

Figure 6: Utilitarian welfare under correlated priorities; u_o uniform on $[0, 3]$

nisms in this environment.

Compared with the case without correlation (see Figure 2), the fraction of blocking pairs under TTC is significantly lower (close to 3-6% as opposed to 10% in Figure 2). The reason for this is that the agents assigned in early rounds of TTC tend to have high priorities even when they are assigned via long cycles, unlike the case of uncorrelated priorities. Nevertheless, the fraction of blocking pairs does not fall as the market grows large. This fact suggests that the asymptotic instability we find in Section 4.0.1 is robust to the introduction of correlated priorities. By contrast, the PB DACB eliminates almost all blocking pairs; the fraction of blocking pairs hovers around 1%.

The simulations show that the results obtained in the paper hold broadly in terms of

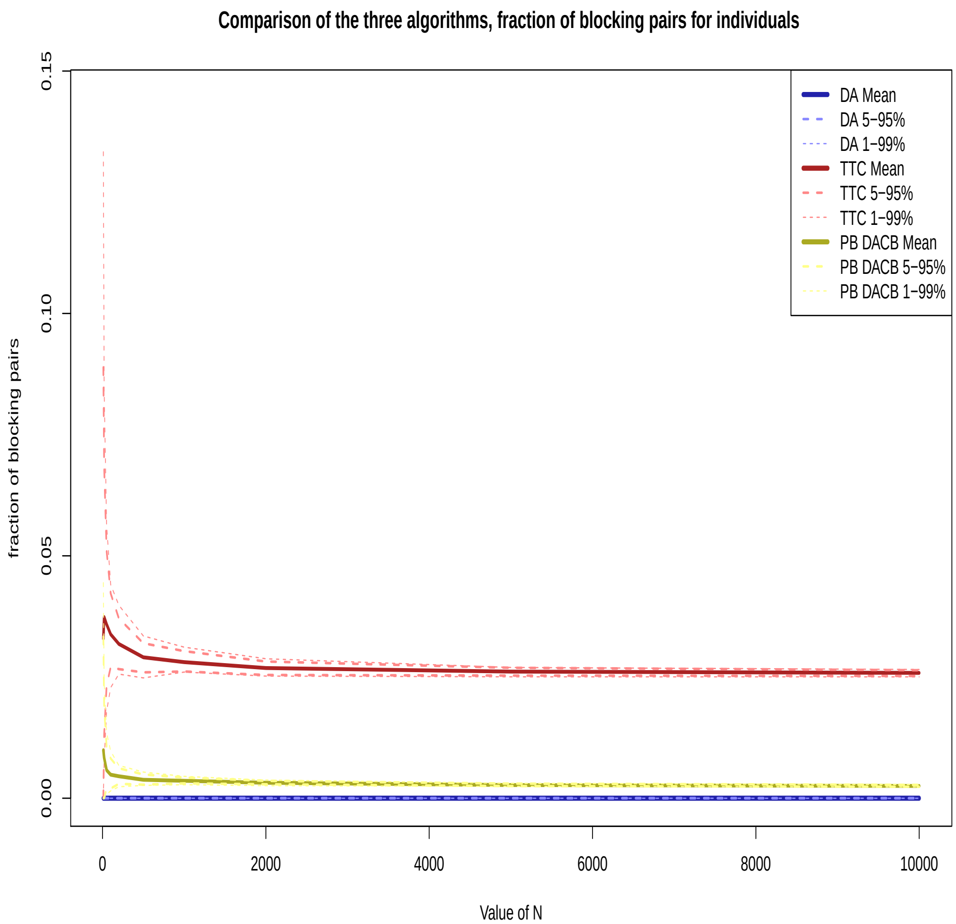

Figure 7: Number of blocking pairs under correlated priorities; u_o uniform on $[0, 1]$

the market sizes and in terms of the distribution of the common values.

S.8 NYC Public High School Choice and the Associated Data Set

S.8.1 Institutional elements on how NYC actual system operates

The possibility to choose schools has a long history in New-York City. For instance, selective high schools appeared in the early 20th schools. The current system established by the Bloomberg administration started in 2004 and is the most highly centralized that the city

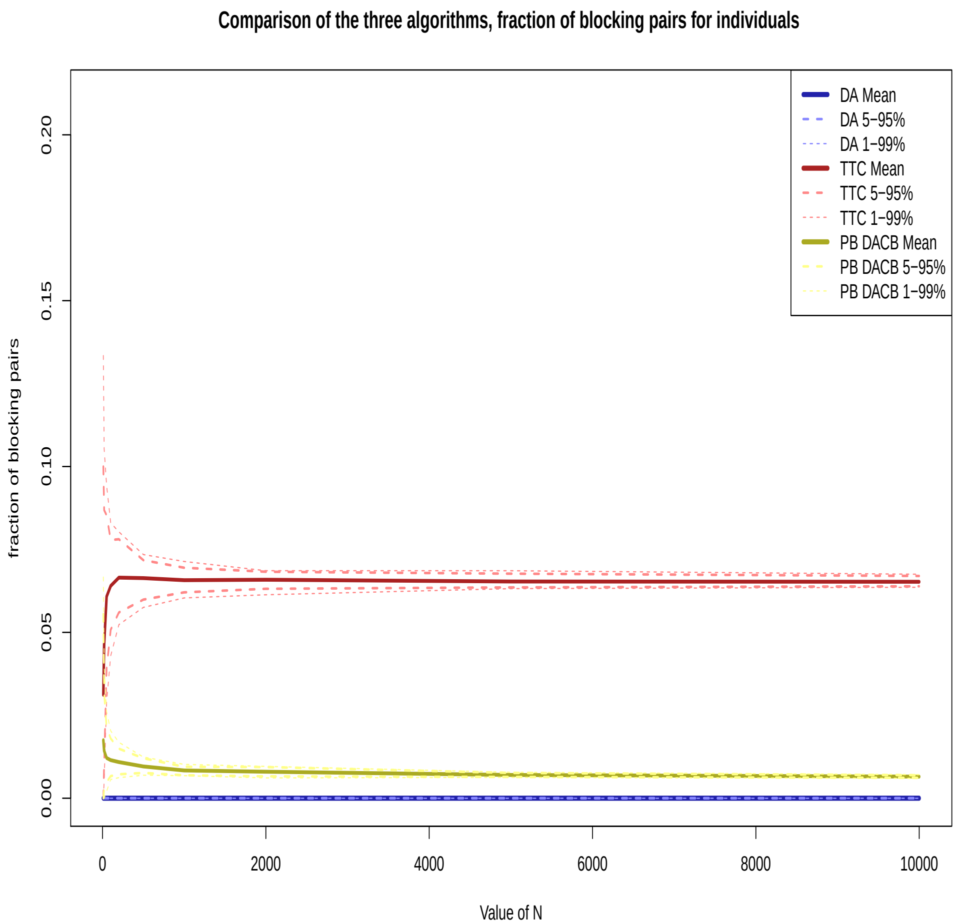

Figure 8: Number of blocking pairs under correlated priorities; u_o uniform on $[0, 3]$

has ever known.[20] In this section, we only provide a brief description of the procedure and refer to Corcoran and Levin (2011) or Abdulkadiroglu, Pathak, and Roth (2009) for further descriptions. Each fall, about 90,000 students apply to enter a high school in New-York City in the following year. Most of them are 8th graders.[21] NYC has over 700 school

[20]It replaced a system where students could submit at most five choices outside of their zone or attend their zoned schools if any. Under this old system, schools offers were uncoordinated so that some students received multiple offers while a large fraction of students received no offers. After the main round of assignment more than a third of students were left with no assignment and ultimately got administratively assigned. See Abdulkadiroglu, Agarwal, and Pathak (2015).

[21]A small fraction of the students are 9th graders. Those are mainly coming from junior high schools including 9th grade.

programs.[22] There are seven types of "admission methods" which determine how a program orders students by priorities. As we will see, these priorities are often weak, causing a set of students to fall in the same priority class. In order to resolve these indeterminacies, a single tie-breaking rule is used: at the beginning of the process, a single random number is assigned to each student. Whenever necessary this number is used to break an indeterminacy in the priorities of a program. We now describe these seven admission methods.

Unscreened programs do not rank students at all, i.e., all students are in the same priority class. *Zoned* school programs give priority to students who live within a pre-defined zone area (some students can have no zoned program). For these two programs, there is no evaluation by schools of students. However, the other five programs, to some extent, evaluate students' abilities in some way or condition their priorities to the interest students show in attending the school. For instance, for *limited unscreened* programs, all else being equal, students' who attended an information session or visit the school's exhibit at high school fairs or open houses will have a higher priority. Similarly, *Screened* programs (all else equal) assign higher priority to students based on several criteria such as their 7th grade report card, reading and math standardized scores, attendance and punctuality, interview, essay or additional diagnostic tests. *Audition* programs partly base their priorities on auditions aimed at evaluating their proficiency in specific performing of visual arts, music, or dance. *Educational option* programs target a distribution in terms of reading ability measured by their score on the 7th grade standardized reading test. If possible, 16% of seats are assigned high performing, 68% middle performing students and the 16% remaining are assigned low performing students. Moreover, for half of their seats, educational programs can actively rank students (while still respecting the target distribution) as screened schools do. For the other half, the random tie-breaking rule is used subject to the distributional constraint.[23] In addition, the system ensures admission to an educational option program for students within the top 2 percent on the seventh grade

[22]One special feature of NYC is that one school can offer several school programs. Indeed, there are about 400 high schools housing these programs. Students apply to programs, and the programs within the same high school can be very distinct. Hence, for our purpose, the relevant unit to focus on is the program.

[23] More precisely, each educational option program is splitted into six separate programs: LR, LS, MR, MS, HR, HS where L, M and H stand for low, middle or high achievement (in the reading test) while R and H stand for random and select. For a given educational option program, 50%×16% of the seats are for LR, 50%×68% are for MR and 50%×16% HR. Similarly, 50%×16% of the seats are for LS, 50%×68% are for MS and 50%×16% HS. For these "virtual subprograms", high-level students are ranked above middle- and low-level students at HS/HR, middle-level students are ranked above high- and low-level students at MS/MR, and low-level students are ranked above high- and middle-level students at LS/LR. Within each class the selection process depends on whether the virtual subprogram is select (S) or random (R). For select subprograms, the selection within a category is similar to that of a screened programs. For random subprograms, the single random number is used to break ties.

standardized reading test provided that this program is top ranked. Finally, *specialized high school* programs have a special status. Students willing to apply to these schools have to pass the Specialized High School Admission Test (SHSAT). Students' priorities in these schools are purely determined by their SHSAT score.[24]

In the summer following seventh grade, families are encouraged to review the Directory of the NYC Public High Schools published by the Department of Education. This document provides information on schools' locations, contact information, enrollment, academic performances along with details on how priorities are set. In the fall, students who are interested in entering a specialized high school program have to pass the SHSAT test. Programs requiring auditions and interviews conduct them during the fall semester as well. Early December, students have to participate in the *first round*. In that round, students have to submit a list of up to twelve school programs by order of preferences. Students who passed the SHSAT test are allowed to submit an additional list of specialized school programs ordered by preferences.[25] The DA (student proposing) mechanism is used to produce the assignment.[26] In March, students receiving an offer from a specialized program (hence, who passed the SHSAT test) are informed of that offer together with the additional offer (if any) they may have received from a non-specialized program. They are asked to pick one of the two offers. Students who did not get any offer in round 1 or did not participate in round 1 go through a *second round* (sometimes referred to as the main round). These students submit a rank ordered list of up to twelve schools.[27] Capacities of schools are adjusted based on the decisions made by the students assigned in the first round. The same algorithm is used to assign students. The vast majority of students is assigned in that round. However, if a student is unassigned, there is a third round in April where this student can again submit a new set of up to twelve choices among remaining school programs. In this third round a random serial dictatorship is used to assign students. In case a student is still unmatched he will be assigned a school as close as possible to his residence. In addition, if there are sufficient grounds, a student may appeal in May. Finally, students who were not present for this high school admission process have to meet with an admissions counselor at the enrollment office in order to get assigned a high school. This last round is usually referred to as over-the-counter.[28]

[24]Only a small fraction of seats at these schools are opened to disadvantaged students who performed well at the SHSAT but who were below the cut-off score for acceptance.

[25]There are nine specialized high school programs in NYC.

[26]To handle the target distribution of Educational Option programs (see footnote 23), the rank order list of a student who applies to such a program is modified to rank the six "virtual" subprograms according to the order HR, HS, MR, MS, LR, LS.

[27]See for details on the second round application process.

[28]The number of such students is surprisingly large (around 36,000 every year). These are among the school system's highest-needs students. See Arvidsson, Fruchter, and Mokhtar (2013)

S.8.2 Data sets

The NYC Department of Education (DOE) provided us with several data sets. We used a file for academic year 2009-2010 which for each round (round 1, 2 and 3), contains the rank order list of students who participated in that round as well as the assignment achieved at the end of the round. In addition to this, for each school program that a student ranks, the file provides information on the priority of the student at that school. Thus, we can reconstruct each school program's priorities (at least over students who ranked that program) – see next section for additional details. We focused on the main round (i.e., second round) of the admission process. In this round we have 78,112 students (out of 83,127 total participants) and 652 school programs.

The information about each program's capacity is not available. In order to estimate the capacities of a school, we used the enrollments through Rounds 2 and 3 at that school. We also observed that some seats are released in the third round: some schools rejecting students in the second round have some additional available seats in the third round. Hence, we also tested the robustness of our results using the following alternative strategy to estimate capacities: if in Round 2 a school program rejects a student then this must mean that the enrollment at Round 2 at that school is equal to its capacity in that round. In such a case, the alternative approach sets the capacity to be equal to the enrollment in Round 2. For other school programs, capacities were still set to their enrollments through Rounds 2 and 3.[29] It turns out that this different way of estimating does not result in a significant difference in analysis.[30]

S.8.3 Assumptions on how the NYC system operates and strategies to complete missing data

Priorities. For each school program that a student ranks, the student is assigned a *priority group* and a *priority rank*. These two numbers are meant to allow us to reconstruct each school program's priorities. Priority group specifies coarse equivalent classes. Students with small priority groups numbers always have higher priority than students with larger ones.

[29] There are only small number of programs in the latter category.

[30] Abdulkadiroglu, Pathak, and Roth (2009) use the final fall enrollments as estimates of capacities. However, these figures include a sizable number of students who are admitted through the over-the-counter round. Evidently, the seats for these applicants are created during this administrative assignment process and they are not available during Round 2 assignment; many programs who reject some applicants in Round 2 end up admitting students in administrative assignment. Given this, the fall enrollments would over-estimate the capacities used for the Round 2 assignment. We believe that the qualitative nature of the comparison of alternative matching algorithms performed here and also in Abdulkadiroglu, Pathak, and Roth (2009) is robust to alternative methods of estimating capacities.

Within each coarse priority class so defined, whenever available, the priority rank number is used to further discriminate between students' priorities. This lexicographic order between priority group and priority rank is (based on the explanation by Department of Education staff) the alleged procedure.[31]

As mentioned in the paper, an applicant's priority is observed in the current data set only for the programs he lists in his rank order list (ROL). The missing priority information for programs not listed in ROL does not cause any issue for the implementation of DA or DACB (the way the priorities of students are set at schools they find unacceptable has no impact on the final assignment). However, it presents some issues with calibrating TTC. Indeed, under standard TTC, a student may be able to trade a seat of a school at which he has a high priority – even though he may find that school unacceptable – with a seat at a school that he likes (but where his priority may be low). Here we assume that programs not listed in an applicant's ROL assign a lower priority to that student than those who listed them in their ROL's. This treatment potentially understates both the efficiency benefit as well as the incidence of justified envy associated with TTC. Nevertheless, the outcome is broadly in line with alternative efficient mechanisms, for which this issue does not arise, such as that of TTC that allow the agents to trade their DA assignments (i.e., Gale's original top trading cycle with an initial assignment given by DA). In particular, this does not appear to be a significant issue for the use of TTC as a benchmark for evaluating DACB's.

Reported Preferences. Following the existing literature, we assume that the observed ROLs of the applicants are their truthful preference ranking of top programs. This assumption is not entirely innocuous since the strategyproofness of DA does not apply when the applicants' ROLs are truncated (see Haeringer and Klijn (2009)). Nevertheless, about 80% of the participants did not fill up their ROLs, suggesting that the truncation was not a binding constraint. However, interpreting the programs not ranked in ROL of an applicant as unacceptable is not well justified since many applicants unassigned in Round 2 listed additional programs (not in their original ROLs) in the third round.[32] This issue results in overstatement of the efficiency performance of DA in our calibration. At the same time, that the same assumption is adopted by Abdulkadiroglu, Pathak, and Roth

[31] We do note, however, that there are incidences of these lexicographic rule being violated. The incidences comprise a rather small fraction. For instance, there are 2645 students who see their priority group number violated, i.e. who are rejected from a school program while others with higher priority group number are accepted; and 4051 students who see their priority group number violated or are rejected from a school program while others with a same priority group number and a higher priority rank are accepted. In any case, the priority rule serves our purpose, which is to consider a realistic market setting.

[32] About 5000 of them listed additional programs in Round 3 which they did not list in Round 2 ROL, contradicting the hypothesis that applicants find programs they did not rank in ROL unacceptable.

(2009) makes our result more easily comparable to theirs. Finally, as pointed out by Abdulkadiroglu, Pathak, and Roth (2009), the system does not give incentives to students who scored within the top 2 percent on the seventh grade standardized reading test to report truthfully their preferences. Since these only represent a small fraction of our set of students (a lot of them are actually matched in the first round: only 400 remain in the second round), following Abdulkadiroglu, Pathak, and Roth (2009) and Abdulkadiroglu, Agarwal, and Pathak (2015), we excluded all such students (i.e., who both scored within the top 2 percent of the reading test and ranked an educational option school at the top of their list) who remain in the main round when we ran our algorithms.

References

Abdulkadiroglu, A., N. Agarwal, and P. A. Pathak (2015): "The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match," MIT, unpublished mimeo.

Abdulkadiroglu, A., P. A. Pathak, and A. E. Roth (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99(5), 1954–1978.

Arvidsson, T. S., N. Fruchter, and C. Mokhtar (2013): "Over the Counter, Under the Radar," Discussion paper, Annenberg Institute for School Reform, Brown University.

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

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

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

Corcoran, S. P., and H. M. Levin (2011): "School Choice and Competition in the New York City Schools," in *Education Reform in New York City: Ambitious Change in the Nation's Most Complex School System*, ed. by J. O'Day, C. Bitter, and L. M. Gomez. Harvard Education Press.

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

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

Lee, S., and L. Yariv (2014): "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.

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