Supplement to "Improving matching under hard distributional constraints"

(*Theoretical Economics*, Vol. 12, No. 2, May 2017, 863–908)

Daniel Fragiadakis
Department of Economics, Texas A&M University

Peter Troyan
Department of Economics, University of Virginia

In this supplementary appendix, we relax one of the key features of the DQDA mechanism: that the reduction sequence be exogenous to the submitted preferences. We define a new mechanism, the *endogenous-reduction DQDA* (EDQDA) mechanism that allows the reduction sequence to change depending on what preferences are submitted. Intuitively, this should allow the mechanism to respond more to changes in demand, and thus allocate seats even more flexibly than DQDA. There are two costs associated with this approach. First, EDQDA loses the important strategyproofness property satisfied by ACDA and DQDA (and standard DA) with no floor constraints; second, EDQDA will no longer Pareto dominate ACDA. However, we are able to show that EDQDA will be approximately strategyproof in large markets (in a formal sense defined below). In addition, we use simulations to study the magnitude of the welfare gains from our dynamic quotas mechanisms. While some students may be worse off under EDQDA, EDQDA tends to make students better off "on average," in the sense that in the simulations, the rank distribution of EDQDA first-order stochastically dominates that of ACDA.

To define EDQDA, we use a slightly different definition of a reduction sequence. A reduction sequence is now written as $\rho = \{(s^1, \theta^1), \dots, (s^K, \theta^K)\}$, where each $(s^k, \theta^k) \in S \times \Theta$. The ρ is a baseline order for reducing the ceilings, but, unlike for DQDA, an entry will be skipped if all floors for the corresponding type have already been met. In addition, we only reduce the type-specific ceilings, and not the capacities. The same entry may appear multiple times in ρ.[1]

Endogenous-reduction DQDA. Set $U^1 = U$.

Stage 1. Starting with the empty matching, run DA under (U^1, Q) and set $\mu^1 = \text{DA}^{(U^1,Q)}(P_I)$.

Daniel Fragiadakis: dfragiadakis@tamu.edu
Peter Troyan: troyan@virginia.edu

[1]Note that it is possible to describe any minimal reduction sequence η from the definition of DQDA by listing the school-type pair whose capacity and ceiling are reduced at each step.

DOI: 10.3982/TE2195

(a) If μ^1 is a feasible matching, end the algorithm and output this matching.

(b) Otherwise, let $\bar{\Theta} = \{\theta \in \Theta : |\mu^1_\theta(s)| < L_{s,\theta}$ for some $s \in S\}$ be the set of types for which at least one floor constraint is not yet satisfied, and let $Y = \{(s, \theta) \in S \times \bar{\Theta} : |\mu^1_\theta(s)| > L_{s,\theta}\}$ be the set of schools that have an excess of these types of students. Let (s, θ) be the element of Y that occurs earliest in ρ (if there is no such element, then choose the first element of ρ). Set $U^2_{s,\theta} = U^1_{s,\theta} - 1$ and $U^2_{s',\theta'} = U^1_{s',\theta'}$ for all other (s', θ'), and delete the earliest occurrence of (s, θ) from ρ. Proceed to stage 2.

In general, proceed as follows.

Stage k. Starting with the empty matching, run DA under (U^k, Q) and set $\mu^k = \text{DA}^{(U^k,Q)}(P_I)$.

(a) If μ^k is a feasible matching, end the algorithm and output this matching.

(b) Otherwise, let $\bar{\Theta} = \{\theta \in \Theta : |\mu^k_\theta(s)| < L_{s,\theta}$ for some $s \in S\}$ be the set of types for which at least one floor constraint is not yet satisfied, and let $Y = \{(s, \theta) \in S \times \bar{\Theta} : |\mu^k_\theta(s)| > L_{s,\theta}\}$ be the set of schools that have an excess of these types of students. Let (s, θ) be the element of Y that occurs earliest in ρ (if there is no such element, then choose the first element of ρ). Set $U^{k+1}_{s,\theta} = U^k_{s,\theta} - 1$ and $U^{k+1}_{s',\theta'} = U^k_{s',\theta'}$ for all other (s', θ'), and delete the earliest occurrence of (s, θ) from ρ. Proceed to stage $k+1$.

EDQDA functions similarly to DQDA, except instead of following the reduction sequence in order from start to finish, we find the earliest entry (s^k, θ^k) for which s^k has an excess of type θ^k students and not all type θ^k floors have been met. We then reduce the θ^k ceiling at s^k by 1 and leave everything else fixed. We thus may skip entries if the types corresponding to those entries have already met all floor constraints, which does not occur in DQDA. In addition, we only lower the type-specific ceilings, and not the capacities, so if a type θ ceiling is lowered so as to satisfy a floor elsewhere, that student's seat is not "wasted," and can be taken by a student of a different type. To ensure that EDQDA produces a feasible matching, ρ must be chosen in such a way that (U^K, Q) ensures a feasible match, where U^K is defined as $U^K_{s,\theta} = U_{s,\theta} - \sum_{k=1}^{K} 1\{(s^k, \theta^k) = (s, \theta)\}$, where $1\{\cdot\}$ is an indicator function that takes on a value of 1 if the kth entry of ρ is (s, θ).

These modifications intuitively make EDQDA a more efficient mechanism. While it turns out that EDQDA will not be more efficient in a Pareto sense (for reasons related to the discussion following Theorem 4, the simulations performed below show that, on average, the students will prefer EDQDA to both ACDA and DQDA. The cost of these welfare gains is that EDQDA is no longer strategyproof. However, EDQDA may still be a successful mechanism in practice, provided that the potential manipulations are not too easy to enact. This is formalized in the next section.

Remark S.1. The final matching produced by EDQDA is again equivalent to DA under some ceilings and capacities (U', Q'), and so EDQDA eliminates justified envy among same types (the argument is equivalent to the one used to prove the analogous statement in Theorem 4).

Large markets

In this section, we show formally that EDQDA has good incentive properties if the market is large. There are many ways to formalize the notion of a large market limit. We will use the concept of strategyproofness in the large (SPL) proposed by Azevedo and Budish (2013).[2] We choose this particular formalization because it is very broadly applicable (beyond matching algorithms), and whether a mechanism is SPL or not turns out to be a good predictor of whether it is a successful mechanism in practical applications.[3]

To show that our mechanism is SPL, we must expand the formal model. We consider a sequence of markets, indexed by $n \in \mathbb{N}$ (which grows large), where n is the number of agents, and $I^n = \{i_1, \ldots, i_n\}$ is the set of agents in market n. The set Θ is a finite set of *quota types*, and each student is of exactly one type in Θ. As previously, it may be useful to think of each $\theta \in \Theta$ as a different socioeconomic tier, though the practical meaning of this set may vary across applications. The set Θ is fixed for all n, but the number of students of each type θ, $|I^n_\theta|$, grows according to some fixed sequence. The set of schools $S = \{s_1, \ldots, s_m\}$ is also fixed for all n, but the capacities and quotas of the schools increase with n. Specifically, for each market $n \in \mathbb{N}$, school s has a capacity of Q^n_s, and type-specific floors and ceilings of $L^n_{s,\theta}$ and $U^n_{s,\theta}$. As in the original model, we collect these quotas into matrices L^n, U^n, and Q^n. We assume that the sequence $(L^n, U^n, Q^n)_{n\in\mathbb{N}}$ is such that at least one feasible matching exists for every market n.

Strategyproofness in the large is a cardinal concept. There is a finite set of *payoff (utility) types* T. Corresponding to each $t_i \in T$ is a von Neumann–Morgenstern expected utility function $u_{t_i} : \Delta S \to [0, 1]$, where ΔS is the set of lotteries over schools. Preferences are private, in that an agent's payoff depends only on her payoff type t_i and outcome (lottery). Each utility type t_i also has an associated ordinal preference relation over schools, which we denote P_{t_i}.

Each school has a finite number of *priority classes* $Z_s = \{1, \ldots, |Z_s|\}$. Each student i is assigned to one priority class at each school. Each s has a primitive (strict) ranking of priority classes $\hat{\succ}_s$, and ranks all students in a higher priority class above all students in a lower. Ties within a priority class are broken using a random lottery (see below). For each i, $z_i \in Z = \times_{s=1}^m Z_s$ is a list that denotes i's priority class at each school. One practical interpretation of priority classes is that each class corresponds to a certain zone, with students living within a certain radius of a school receiving higher priority for their neighborhood school than those living farther away. All students in each zone have

[2] For other large market incentive compatibility notions that are closely related to SPL, see Immorlica and Mahdian (2005), Kojima and Pathak (2009), and Kojima et al. (2013), who study the large market properties of DA, or Che and Kojima (2010) and Kojima and Manea (2010), who do the same for the probabilistic serial mechanism. One of the main advantages of SPL is that it is not tailored to a specific mechanism, and so can be applied more broadly.

[3] Azevedo and Budish (2013) show that non-SPL mechanisms (e.g., pay-as-bid auctions for Treasury bills, priority matching algorithms in hospital–residency markets) tend to perform poorly in the field and are eventually abandoned, while their SPL counterparts (uniform price auctions, DA) are successful and in continued use.

equal priority, up to the random lottery used to break ties. Consistent with this interpretation, we sometimes refer to Z as a set of *zones* for concreteness, but we emphasize that the priority classes can be based on factors other than geography.[4]

To summarize, let $\Lambda = \Theta \times Z \times T$. A student's overall *type* is then an element $\lambda_i \in \Lambda$. For each market n, define $n_{(\theta,z)}$ as the number of students of quota-zone type (θ, z). We assume that $n_{(\theta,z)} \to \infty$ for all (θ, z), so that the number of agents of each type (θ, z) grows large according to some fixed sequence.

Definition S.1. A *(direct) mechanism* $\{(\psi^n)_{n\in\mathbb{N}}, \Lambda\}$ is a sequence of allocation functions $\psi^n : \Lambda^n \to \Delta(S^n)$ such that every allocation in the support of $\psi^n(\lambda)$ is feasible for all n and all $\lambda \in \Lambda^n$.

Note that a mechanism as defined here produces a random allocation. For notational purposes, the inputs are vectors of types $\lambda \in \Lambda^n$, but each student i is restricted to reporting θ_i and z_i truthfully; the only private information is her payoff type t_i.

Given a distribution of payoff types $\pi \in \Delta T$, define for a student i of quota-zone type (θ_i, z_i) the quantity

$$\phi^n_{(\theta_i,z_i)}\big(t'_i, \pi\big) = \sum_{\lambda_{-i}\in\Lambda^{n-1}} \psi^n_i\big(\lambda'_i, \lambda_{-i}\big)\Pr\big(\lambda_{-i}|t_{-i} \sim \text{i.i.d.}(\pi)\big),$$

where $\lambda'_i = (\theta_i, z_i, t'_i)$ and $\Pr(\lambda_{-i}|t_{-i} \sim \text{i.i.d.}(\pi))$ gives the probability that the realized type profile of the $n-1$ other agents is $\lambda_{-i} = (\theta_{-i}, z_{-i}, t_{-i})$, given that the payoff types are drawn independently and identically distributed (i.i.d.) from some distribution π (recall that θ_{-i} and z_{-i} are fixed). In words, $\phi^n_{(\theta_i,z_i)}$ gives the outcome an agent of type (θ_i, z_i) receives when she reports her payoff type as t'_i and the payoff type reports of the other students are drawn according to π.

We are now ready to formally define strategyproofness in the large. Given a finite set X, let $\bar{\Delta}X$ denote the set of full-support probability distributions over X.

Definition S.2 (Azevedo and Budish 2013). Mechanism $\{(\psi^n)_{n\in\mathbb{N}}, \Lambda\}$ is *strategyproof in the large* (SPL) if for any $\varepsilon > 0$ and any $\pi \in \bar{\Delta}T$, there exists an n_0 such that for all $n \geq n_0$, $(\theta, z) \in \Theta \times Z$, and all $t_i, t'_i \in T$,

$$u_{t_i}\big(\phi^n_{(\theta,z)}(t_i, \pi)\big) \geq u_{t_i}\big(\phi^n_{(\theta,z)}\big(t'_i, \pi\big)\big) - \varepsilon.$$

Last, we must define the EDQDA mechanism in this setting. Given some collection of reduction sequences $\{\rho^n\}_{n\in\mathbb{N}}$ defined as above, the mechanism proceeds as follows. First, draw a vector of lottery numbers $\ell \in [0,1]^n$ uniformly at random, where ℓ_i denotes the lottery number of student i. Then create a strict priority relation for each school s, $\succ^n_s$, as

$$i \succ^n_s j \quad \Longleftrightarrow \quad z_{i,s} \hat{\succ}_s z_{j,s} \text{ or } [z_{i,s} = z_{j,s} \text{ and } \ell_i > \ell_j],$$

[4]Formally, it is necessary to divide the students into a fixed number of priority classes to ensure semi-anonymity of the mechanisms, in the sense of Kalai (2004).

where $z_{i,s}$ is student i's priority class (zone) at s. Let $\mu^n(\lambda, \ell)$ be the matching produced by the EDQDA algorithm (as defined the previous section) using type-specific floors L^n, type-specific ceilings U^n, capacities Q^n, priorities $\succ_s^n$, reduction sequence ρ^n, and (ordinal) preferences $(P_{t_i})_{i \in I^n}$. Then define E^n as

$$E^n(\lambda) = \int_{\ell \in [0,1]^n} \mu^n(\lambda, \ell)\, d\ell.$$

THEOREM S.1. *The mechanism* $\{(E^n)_{n\in\mathbb{N}}, \Lambda\}$ *is strategyproof in the large.*

Intuitively, in a large enough market, it is unlikely that student i's report will have an effect on the final stage of the algorithm, and thus it is unlikely that she will be able to profitably manipulate. At a formal level, we prove the result by showing that EDQDA satisfies an envy-freeness condition identified by Azevedo and Budish (2013) as sufficient for strategyproofness in the large. The proof can be found at the end of this supplement.

SIMULATIONS

We have thus far argued from a theoretical perspective that our new mechanisms should lead to improved performance of matching markets with floor constraints, as they increase efficiency while still satisfying good incentive and fairness properties. However, the theoretical results do not say anything about the number of students who are made better off by the use of our mechanisms. To answer this question, we turn to simulations.

The purpose of these simulations is twofold: first, to approximate an actual school choice market and obtain a sense of the magnitude of the potential gains from our mechanisms; second, to conduct comparative statics with respect to important market parameters. With these dual goals in mind, we use the details of kindergarten assignment in Cambridge, MA (for which limited data are publicly available) as an anchor to set the number of students, schools, and student types, but also structure the simulations with enough flexibility to allow us to conduct comparative statics with respect to preference correlation, quotas, and capacities.[5]

Simulation parameters. There are $n = 750$ students, $m = 12$ schools, and two possible types $\Theta = \{\ell, h\}$. There are 250 students of type ℓ (low SES) and 500 students of type h (high SES).

Student preferences are determined as follows. Student i's utility for school s is $u_i(s) = \alpha v^c(s) + (1 - \alpha) v_i^p(s)$, where $v^c(s)$ is a common utility component that is the same across students, and $v_i^p(s)$ is i's private, idiosyncratic utility for school s. The common component $v^c(s)$ and all private components $v_i^p(s)$ are drawn i.i.d. uniformly from $[0, 1]$.[6] Ordinal preferences are then created from these cardinal preferences. Note that it is only the ordinal preferences that matter. The cardinal utility draws are used solely as

[5]The Cambridge data can be accessed at http://www3.cpsd.us/department/frc/FRC.

[6]This method of drawing preferences is common in the literature; see, for example, Hafalir et al. (2013) and Miralles (2009). Using other distributions (e.g., normal) leads to similar results.

		Floors	Ceilings	Capacities
Low flexibility	Primitive	(14, 36)	(39, 69)	90
	Artificial	–	(21, 42)	63
High flexibility	Primitive	(14, 14)	(76, 76)	90
	Artificial	–	(21, 44)	65

TABLE S.1. The floors, ceilings, and capacities at each school for the low and high flexibility cases. For entries (x, y), x corresponds to the low type ℓ and y corresponds to the high type h.

a simple way to vary the correlation in the ordinal preferences. The measures of welfare we use to compare the mechanisms are ordinal measures.

By varying α, we can study how mechanism performance varies as a function of preference correlation. A value of $\alpha = 0$ corresponds to uncorrelated preferences, while $\alpha = 1$ corresponds to common preferences. To get a sense of the degree of preference correlation in a real market, we use the Cambridge data, which list for each school the number of students who rank it as their first choice. The value of α corresponding to the data is $\alpha = 0.13$.[7]

School priority relations are drawn uniformly at random, independently across schools. For the school floors, ceilings, and capacities, we consider two cases: low flexibility and high flexibility. Table S.1 gives the type-specific floors, ceilings, and capacities for the two cases. For simplicity, we treat all schools symmetrically, and so the numbers in the table correspond to the floors, ceilings, and capacities for each school. In the high flexibility case, the floors are lower and the ceilings are higher (compared to the low flexibility case), meaning there is a wider range of possible final assignments in the high flexibility case. Given the primitive floors, ceilings, and capacities (L, U, Q), the artificial capacities $(\bar{U}, \bar{Q})$ are chosen as the highest symmetric values that ensure a feasible matching.

We last must discuss how to construct η (for DQDA) and ρ (for EDQDA). As noted before, there are many possible ways to do this. Since it is not obvious ex ante which should be chosen, we do so randomly and symmetrically. That is, at each stage we randomly choose which quota to reduce, subject to the constraint that all (s, θ) must be reduced once before any quota is reduced for a second time (and all must be reduced twice before any quota is reduced a third time, etc.). For η, we also reduce the capacity of school s and each stage, while for ρ, we do not. For more on the specification of η and ρ, see the additional simulation details at the end of this supplement.

[7] This value was obtained by finding the α that minimized the distance between 1000 simulated distributions of first choices and the empirical distribution from the Cambridge data. While indicative of student preferences, care should be taken in interpreting this number, because Cambridge uses a non-strategyproof mechanism (immediate acceptance), and so it is unknown whether the reported preferences are truthful (which is one of the common arguments made in support of switching to a strategyproof mechanism). However, because our proposed mechanisms are strategyproof, by varying α we are able to get a sense of how our mechanisms *would* perform if implemented, even if the true α were different from that implied by the current data.

Simulation results. We run 150 iterations for each set of parameters. The outcome metric is the average rank distribution over these 150 iterations. The rank distribution plots, for each x, the number of students who receive their xth or better choice. It is one of the common metrics publicly reported by school districts when evaluating mechanisms (for example, on its website, the Cambridge school district says that "85% of students receive one of their top 3 choices," which is a simplified reporting of the rank distribution).[8]

In Figure S.1, we plot the rank distributions for our mechanisms under different choices of parameters. To read the figures, take the top left panel as an example: it says that under ACDA, about 620 students on average get their first choice, about 725 get their first or second choice, etc., while under EDQDA, about 700 students get their first choice, 745 get their first or second choice, etc. For clarity, we only plot the beginning of the rank distributions, because at higher ranks x, essentially all students are getting their xth or better choice under all mechanisms.

The figure includes plots for six parameterizations: three values of α ($\alpha = 0, 0.13, 0.26$) and two levels of flexibility. At the end of this supplement, we report different results for other parameter values, but the three main takeaways can be obtained from Figure S.1.

(i) Comparing ACDA, DQDA, and EDQDA: In every case, the average rank distribution for ACDA is stochastically dominated by DQDA, which in turn is itself stochastically dominated by EDQDA. This confirms the Pareto dominance of DQDA over ACDA (Theorem 4). While there is no analogous Pareto dominance of EDQDA over DQDA, the simulations show that it is the case that the students will on average be better off under EDQDA compared to both DQDA and ACDA. The results suggest that there are significant gains to be had from our mechanisms, with on average over 20% more students receiving their first choice under EDQDA compared to ACDA for some parameterizations (see the Appendix).

(ii) Comparative statics with respect to preference correlation: Within each column, the flexibility is fixed, but the preference correlation increases as we move from the top row to the bottom. As can be seen in the figure, the gains from our mechanism are larger when the preference correlation is smaller. Intuitively, this is because when preferences are uncorrelated, the floors are more likely to be filled in the early stages of the dynamic quotas mechanisms. As the correlation increases, this becomes less likely, and the dynamic quotas mechanism is more likely to end in later stages, making it closer to artificial caps. However, even at high correlations, there are still gains to be had from our mechanisms.

(iii) Comparative statics with respect to flexibility: Within each row, the preference correlation α is fixed, but the flexibility increases as we move from the left column to the right. As can be seen in the figure, the gains from our mechanism are

[8] Motivated by the fact that authorities are often concerned with rank distributions, Featherstone (2011) investigates mechanisms that optimize the rank distribution in the context of object allocation without priorities.

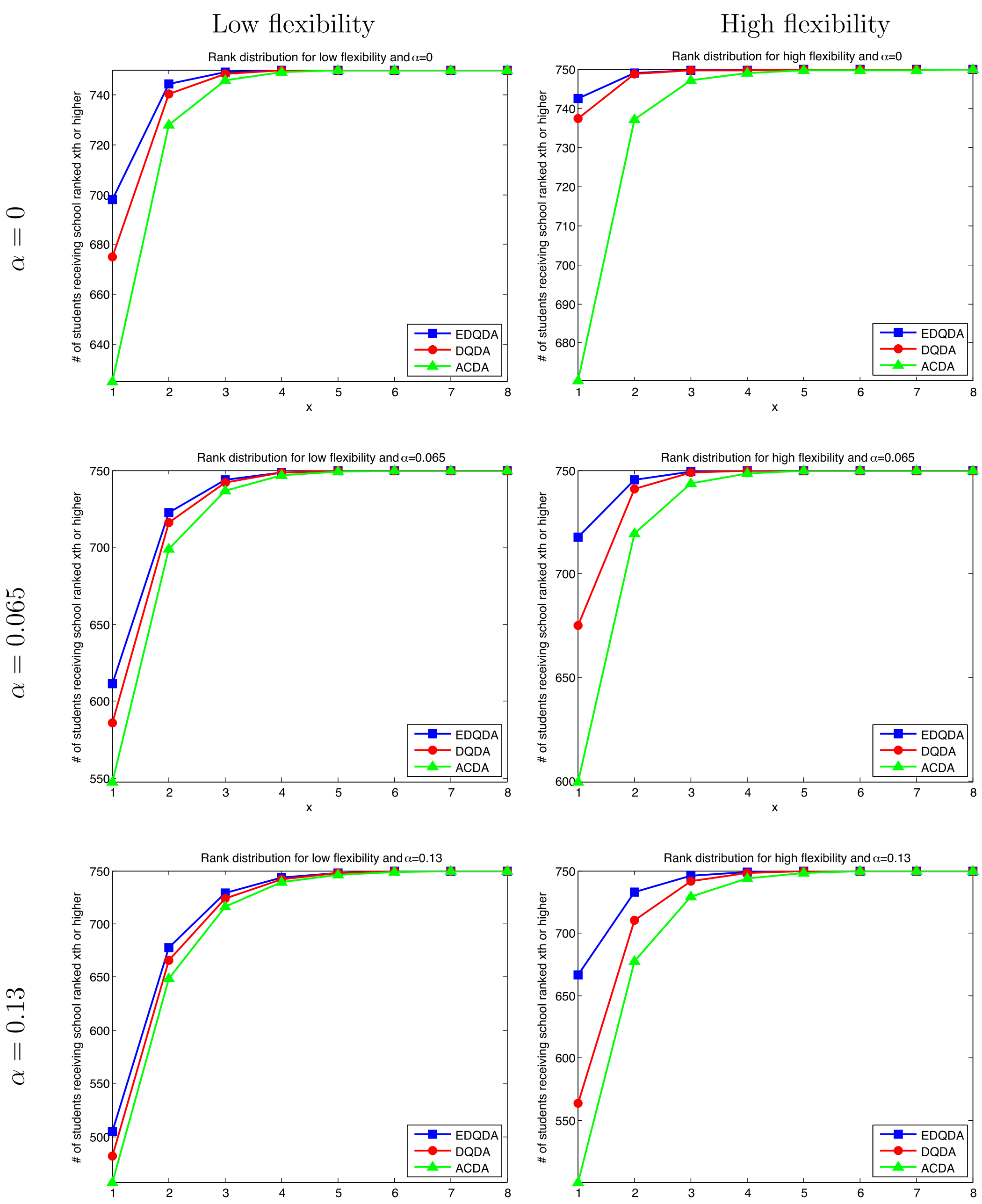


FIGURE S.1. The number of students assigned to their xth or better choice for various parameter values, averaged over 150 iterations. Higher plots correspond to better mechanisms (on average) for the students. For clarity, we only plot the beginning of the distributions, but the first-order stochastic dominance (FOSD) relation EDQDA $>_{\text{FOSD}}$ DQDA $>_{\text{FOSD}}$ ACDA holds for all values of x.

larger when there is more flexibility. Intuitively, when there is less flexibility (i.e., the floors and ceilings are "close"), dynamic quotas becomes more similar to artificial caps. In the extreme case when all ceilings are equal to all floors and there is no flexibility, DQDA is equivalent to ACDA (and to EDQDA). As the flexibility increases, there is more room for our dynamic quotas mechanisms to respond to the submitted preferences of the agents, and hence the gains obtained from using a dynamic quotas mechanism increase.

In summary, while our mechanisms will produce efficiency gains for any parameterization, the potential gains are increasing in flexibility and decreasing in preference correlation. However, we still recommend the use of DQDA or EDQDA even in markets with low flexibility or high preference correlation, because they will make the students better off on average, without sacrificing fairness or incentive properties.

Proofs and additional simulation results

Proof of Theorem S.1

To simplify notation, for a student of type $\lambda_i = (\theta_i, z_i, t_i)$, we define $g_i = (\theta_i, z_i)$ and refer to g_i as student i's *group*. Recall that students cannot misreport their group and, as described in the text, the number of students in each group, n_g, grows according to some fixed sequence such that $n_g \to \infty$ for all groups g.

We now state a no-envy condition that will be sufficient for a mechanism to be SPL, and then show that EDQDA does indeed satisfy this condition.

Definition S.3 (Azevedo and Budish 2013). A mechanism $\{(\psi^n)_{n\in\mathbb{N}}, \Lambda\}$ is *envy-free but for tie-breaking* (EFTB) if for each n there exists a function $x^n : (\Lambda \times [0,1])^n \to \Delta(S^n)$ that is symmetric over its coordinates and such that

$$\psi^n(\lambda) = \int_{\ell\in[0,1]^n} x^n(\lambda, \ell)\, d\ell,$$

and if $\ell_i \geq \ell_j$ and i and j belong to the same group g, then $u_{t_i}(x_i^n(\lambda, \ell)) \geq u_{t_i}(x_j^n(\lambda, \ell))$.

Lemma S.1. *The mechanism $\{(E^n)_{n\in\mathbb{N}}, \Lambda\}$ is envy-free but for tie-breaking.*

Proof. To show this, we must exhibit a function x^n that satisfies the properties of Definition S.3. Such a function is given immediately by $\mu^n(\lambda, \ell)$, as defined above. Recall that by definition we have

$$E^n(\lambda) = \int_{\ell\in[0,1]^n} \mu^n(\lambda, \ell)\, d\ell$$

for all $n \in \mathbb{N}$. The function μ^n is clearly symmetric over its coordinates by construction. Thus, the last thing we need to show is that an agent i in group g never envies another agent j in group g with a lower lottery number. Assume that i and j belong to the same group, but $\ell_i > \ell_j$. Recall that $\succ_s^n$ is the (post-lottery) priority relation for school s. The fact that i and j belong to the same group and $\ell_i > \ell_j$ imply that the priority relations are such that $i \succ_s^n j$ for all $s \in S$. Now, $\mu^n(\lambda, \ell)$ is equivalent to the matching produced by standard DA under some quotas $(U^{n,k}, Q^n)$, priorities $(\succ_s^n)_{s\in S}$, and (ordinal) preferences $(P_{t_i})_{i\in I^n}$. This matching eliminates all justified envy for same types with respect to the strict priorities $(\succ_s^n)_{s\in S}$, which implies that $\mu_i^n(\lambda, \ell)\ R_{t_i}\ \mu_j^n(\lambda, \ell)$, and so $u_{t_i}(\mu_i^n(\lambda, \ell)) \geq u_{t_i}(\mu_j^n(\lambda, \ell))$, which proves the lemma. □

Given this lemma, the theorem follows from Proposition 1 of Azevedo and Budish (2013), which states that if a mechanism satisfies EFTB, it is SPL.[9]

Constructing η and ρ

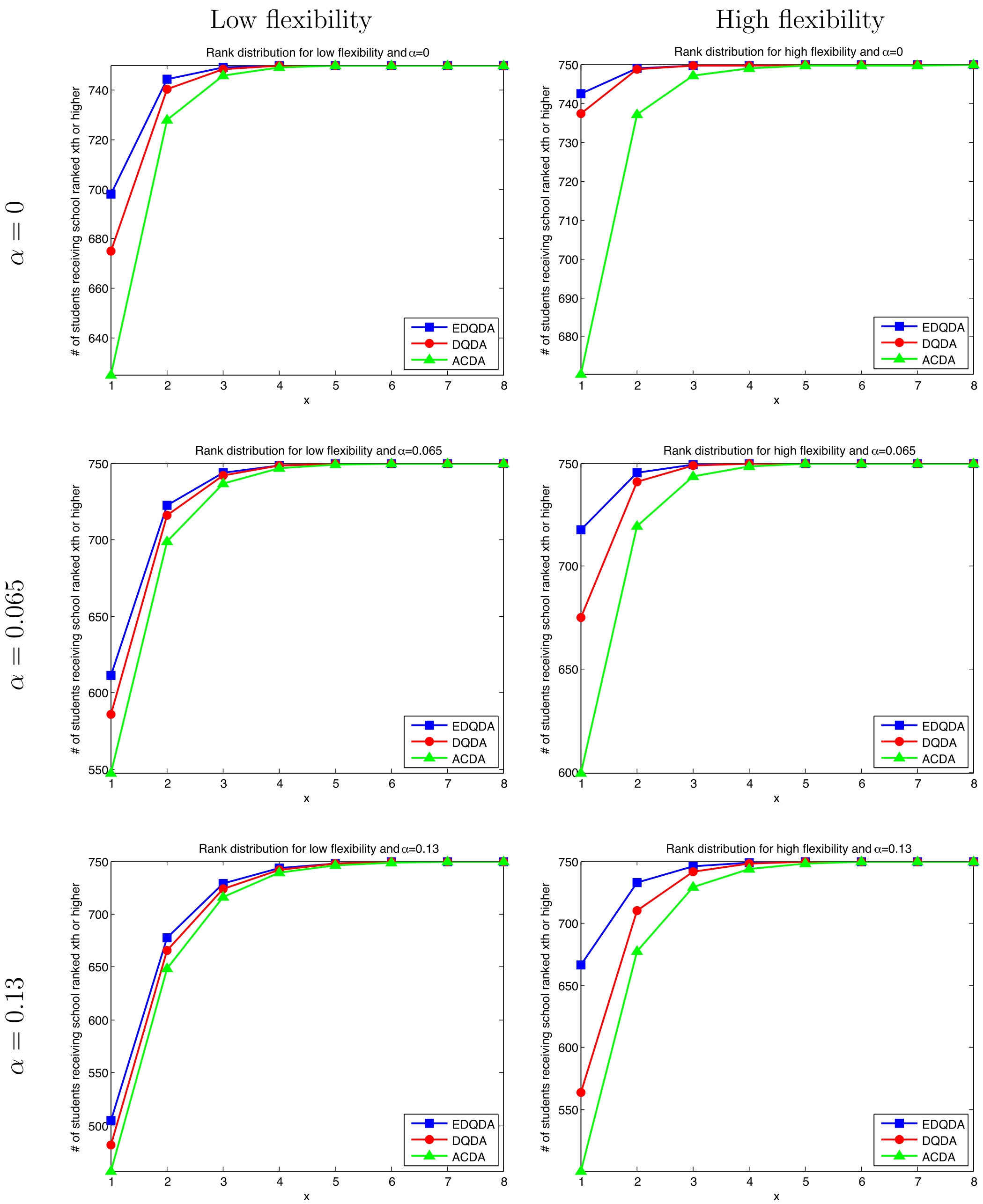
In this section, we describe in detail how we choose η and ρ to run the simulations. Essentially, at each stage, we randomly choose one school-type pair (s, θ) for which the ceiling/capacity will be lowered, subject to feasibility constraints. More specifically, to construct η, we start by setting $(U^K, Q^K) = (\bar{U}, \bar{Q})$. Then we randomly choose some pair (s, θ) such that $U_{s,\theta}^K < U_{s,\theta}$ and $Q_s^K < Q_s$, and set $U_{s,\theta}^{K-1} = U_{s,\theta}^K + 1$ and $Q_s^{K-1} = Q^K + 1$. For the remaining (s', θ'), $U_{s',\theta'}^{K-1} = U_{s,',\theta'}^{K-1}$ and $Q_{s'}^K = Q_{s'}^{K-1}$. For (U^{K-2}, Q^{K-2}), we again randomly select another school-type pair *different* from (s, θ), and raise its ceiling and capacity by 1 (again, subject to the constraint that doing so does not violate the true ceilings and capacities). We continue in this manner until it is impossible to raise capacities any further without violating the true (U, Q). This produces a sequence $\eta = \{(U^1, Q^1), \ldots, \{(U^K, Q^K)\}$ that can then be used to run DQDA. We construct η "backward," starting from (U^K, Q^K), only to simplify the coding. It can be done "forward" as well.

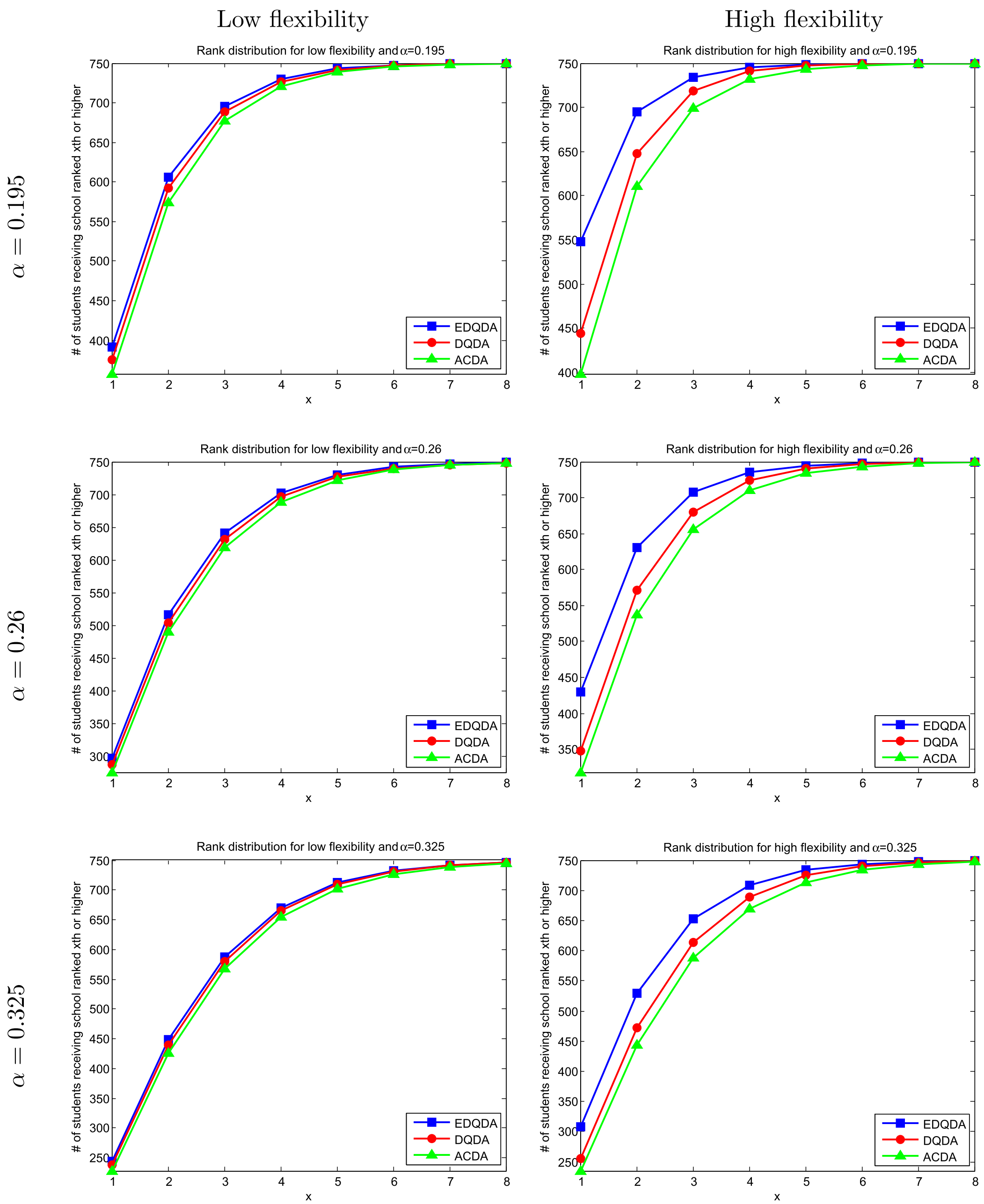
To construct ρ for use in EDQDA, we start with the η just constructed (we begin with this η so as to make a fair comparison with DQDA). Starting from where we left off in the previous paragraph, we choose a pair (s, θ) randomly from those that have not yet reached their true type-specific ceiling, and raise this type-specific ceiling by 1 (note that just the ceilings are considered, not the capacities, since they are fixed in EDQDA). We continue to do so until we can no longer raise any (s, θ) ceiling further without violating the true $U_{s,\theta}$. We then convert this sequence of ceiling-capacity vectors into the corresponding sequence of school-type pairs ρ, as described above, which is then used to run EDQDA.

Additional simulation results

Below we present simulation results for additional parameter values. The first column corresponds to the low flexibility case, while the second column corresponds to the high flexibility case. Each row corresponds to different values of the correlation parameter α. Note that for image clarity, the vertical axes differ across figures.

[9] Our model is slightly different, since the number of students in each group g grows deterministically, rather than stochastically. In Appendix C of their paper, Azevedo and Budish (2013) note that their results continue to hold when the number of students in each group grows deterministically and the utility types are drawn i.i.d. within each group, which is the case here.


Low flexibility
High flexibility
$\alpha = 0$
$\alpha = 0.065$
$\alpha = 0.13$
Rank distribution for low flexibility and α=0
Rank distribution for high flexibility and α=0
Rank distribution for low flexibility and α=0.065
Rank distribution for high flexibility and α=0.065
Rank distribution for low flexibility and α=0.13
Rank distribution for high flexibility and α=0.13
# of students receiving school ranked xth or higher
x
EDQDA
DQDA
ACDA

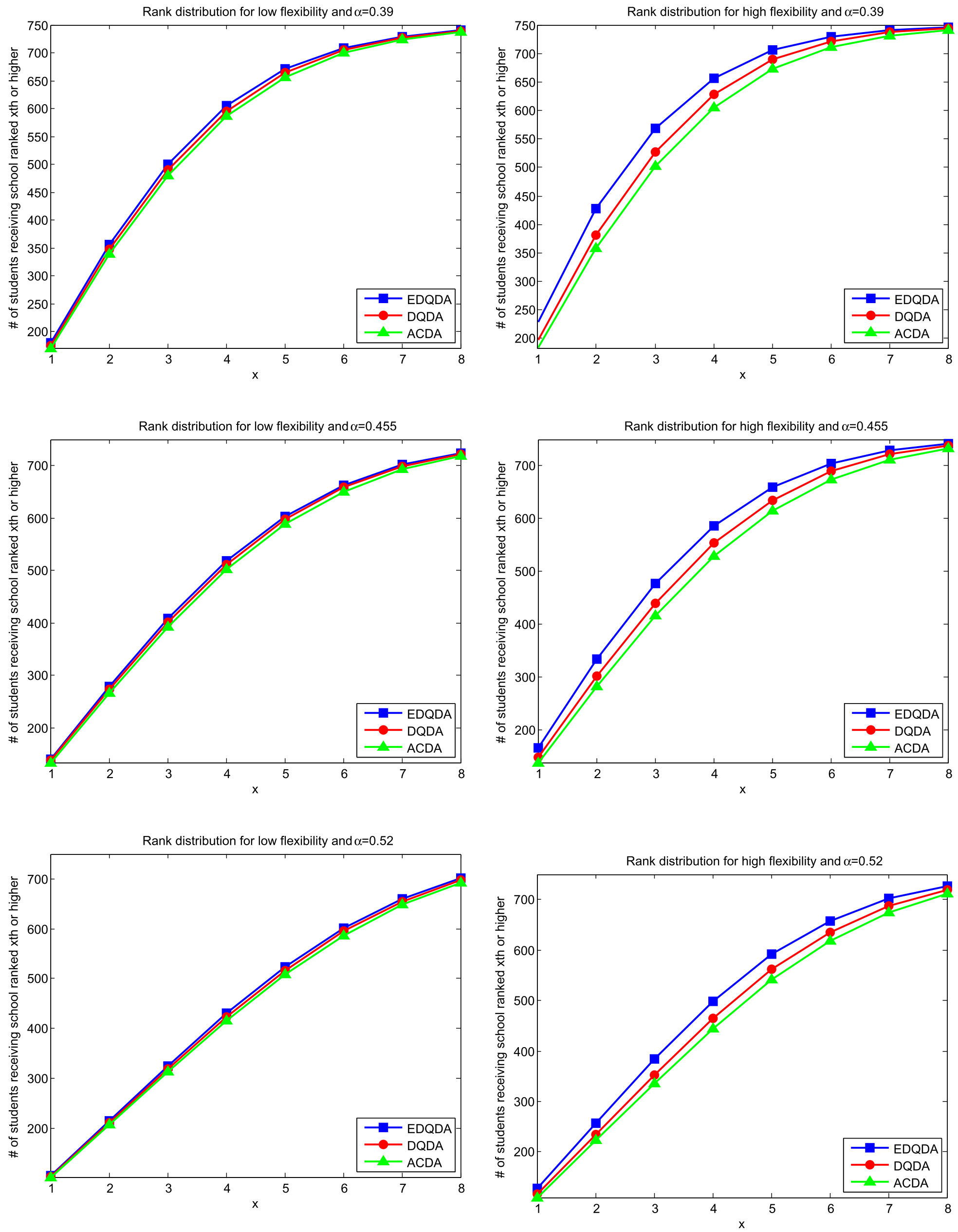

Low flexibility
High flexibility
α = 0.195
α = 0.26
α = 0.325
Rank distribution for low flexibility and α=0.195
Rank distribution for high flexibility and α=0.195
Rank distribution for low flexibility and α=0.26
Rank distribution for high flexibility and α=0.26
Rank distribution for low flexibility and α=0.325
Rank distribution for high flexibility and α=0.325
# of students receiving school ranked xth or higher
x
EDQDA
DQDA
ACDA

Low flexibility High flexibility

$\alpha = 0.39$

$\alpha = 0.455$

$\alpha = 0.52$



References

Azevedo, Eduardo M. and Eric Budish (2013), "Strategyproofness in the large." Working paper, University of Chicago. [3, 4, 5, 9, 10]

Che, Yeon-Koo and Fuhito Kojima (2010), "Asymptotic equivalence of probabilistic serial and random priority mechanisms." *Econometrica*, 78, 1625–1672. [3]

Featherstone, Clayton R. (2011), "A rank-based refinement of ordinal efficiency and a new (but familiar) class of ordinal assignment mechanisms." Working paper, The Wharton School, University of Pennsylvania. [7]

Hafalir, Isa E., M. Bumin Yenmez, and Muhammed A. Yildirim (2013), "Effective affirmative action in school choice." *Theoretical Economics*, 8, 325–363. [5]

Immorlica, Nicole and Mohammad Mahdian (2005), "Marriage, honesty, and stability." In *Proceedings of the 16th ACM–SIAM Symposium on Discrete Algorithms*, 53–62. [3]

Kalai, Ehud (2004), "Large robust games." *Econometrica*, 72, 1631–1665. [4]

Kojima, Fuhito and Mihai Manea (2010), "Incentives in the probabilistic serial mechanism." *Journal of Economic Theory*, 145, 106–123. [3]

Kojima, Fuhito and Parag A. Pathak (2009), "Incentives and stability in large two-sided matching markets." *American Economic Review*, 99, 608–627. [3]

Kojima, Fuhito, Parag A. Pathak, and Alvin E. Roth (2013), "Matching with couples: Stability and incentives." *Quarterly Journal of Economics*, 128, 1585–1632. [3]

Miralles, Antonio (2009), "School choice: The case for the Boston mechanism." In *Auctions, Market Mechanisms and Their Applications*, 58–60, Springer, Berlin. [5]

Co-editor Faruk Gul handled this manuscript.

Manuscript received 29 May, 2015; final version accepted 8 April, 2016; available online 20 April, 2016.