# Online Appendix to

# Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions

Gabrielle Fack Julien Grenet Yinghua He

*The American Economic Review*

October 2018

## List of Appendices

# Appendix A Proofs and Additional Results

Section A.1 collects the proofs and additional results for a finite economy, while those related to asymptotics and the continuum economy are presented in Section A.2.

## A.1 Finite Economy: Proofs from Sections I.B and I.C

**Proof of Proposition 1.**

(i) **Sufficiency**. Without application cost, STT is a dominant strategy (Dubins and Freedman, 1981; Roth, 1982), so we only need to prove it is the unique equilibrium.

Suppose that a non-STT strategy, $\overline{\sigma}$, is another equilibrium. Without loss of generality, let us assume $\overline{\sigma}$ is in pure strategy.

Since STT is a weakly dominant strategy, it implies that, for any $i$ and any $\theta_{-i} \in \Theta^{I-1}$,

$$\sum_{s=1}^{S} u_{i,s} a_s \left(r(u_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right) \geqslant \sum_{s=1}^{S} u_{i,s} a_s \left(\overline{\sigma}(\theta_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right),$$

in which both terms are non-negative given the assumptions on $G$. Moreover, $\overline{\sigma}$ being an equilibrium means that, for any $i$:

$$\sum_{s=1}^{S} u_{i,s} \int a_s \left(r(u_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i}) \leqslant \sum_{s=1}^{S} u_{i,s} \int a_s \left(\overline{\sigma}(\theta_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i}).$$

It therefore must be that, for any $i$ and any $\theta_{-i} \in \Theta^{I-1}$ except a measure-zero set of $\theta_{-i}$,

$$\sum_{s=1}^{S} u_{i,s} a_s \left(r(u_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right) = \sum_{s=1}^{S} u_{i,s} a_s \left(\overline{\sigma}(\theta_i), e_i; \overline{\sigma}(\theta_{-i}), e_{-i}\right). \tag{A.1}$$

Through the following claims, we then show that $\overline{\sigma}$ must be STT, i.e., $\overline{\sigma}(\theta_i) = r(u_i)$.

*Claim 1*: $\overline{\sigma}(\theta_i)$ and $r(u_i)$ have the same top choice.

*Proof of Claim 1*: Given the full support of $G$, there is a positive probability that $i$'s priority indices at all schools are the highest among all students. In this event, $i$ is accepted by $r_i^1$ (her most preferred school) when submitting $r(u_i)$ and accepted by the top choice in $\overline{\sigma}(\theta_i)$ when submitting $\overline{\sigma}(\theta_i)$. As preferences are strict, $\overline{\sigma}(\theta_i)$ must have $r_i^1$ as the top choice to have Equation (A.1) satisfied.

*Claim 2*: $\overline{\sigma}(\theta_i)$ and $r(u_i)$ have the same top two choices.

*Proof of Claim 2*: From Claim 1, we know that $\overline{\sigma}(\theta_i)$ and $r(u_i)$ agree on their top choices. Given $G$'s full support, there is a positive probability that $i$'s type and others' types are such that: (a) $i$'s priority index is the lowest among all students at school $r_i^1$; (b) $i$'s priority index is the highest among all students at all other schools; and (c) all other students have $r_i^1$ as their most preferred school. In this event, by Claim 1, all students rank $r_i^1$ as top choice. Therefore, $i$ is rejected by $r_i^1$, but she is definitely accepted by her second choice. Because STT means she is accepted by $r_i^2$, Equation (A.1) implies that $\overline{\sigma}(\theta_i)$ must also rank $r_i^2$ as the second choice. This proves the claim.

We can continue proving a series of similar claims that $\overline{\sigma}(\theta_i)$ and $r(u_i)$ must agree on top $S$ choices. In other words, $\overline{\sigma}(\theta_i) = r(u_i)$. This proves that there is no non-STT equilibrium, and, therefore, STT is the unique Bayesian Nash equilibrium.

(ii) **Necessity**. The following shows that the zero-application-cost condition is necessary for STT to be an equilibrium strategy for every student type.

Without loss of generality, suppose that $C(S) - C(S-1) > 0$ , which implies that applying to the $S$th choice is costly. Let $\sigma^{STT}$ be the STT strategy. Let us consider students whose $S$th choice in terms of true preferences, $r_i^S$, has a low cardinal value. More specifically, $u_{i,r_i^S} < C(S) - C(S-1)$. For such students, $\sigma^{STT}$ is a dominated strategy, dominated by dropping $r_i^S$ and submitting $(r_i^1, \ldots, r_i^S)$. In other words, $\sigma^{STT}$ is not an equilibrium strategy for these students.

In fact, $\sigma^{STT}$ is not an equilibrium strategy for more student types, given others playing $\sigma^{STT}$. If a student drops an arbitrary school $s$ and submits a partial true preference order $L_i$ of length $(S-1)$, the saved cost of is $C(S) - C(S-1)$, while the associated foregone benefit is at most $u_{i,s} \int a_s \left(L_i, e_i; \sigma_{-i}^{STT}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i})$. The saved cost can exceed the forgone benefit because the latter can be close to zero when $s$ tends to have cutoff much higher than $e_{i,s}$ or when $i$ can be almost certainly accepted by more desirable schools, given that everyone else plays STT. When it is the case, $i$ deviates from STT.

The above arguments can be extended to any non-zero application cost. ■

**Proof of Lemma 1.**

The sufficiency of the first statement is implied by the strategy-proofness of DA and by DA producing a stable matching when everyone is STT. That is, STT is a dominant strategy if $C(|L|) = 0$ for all $L$, which always leads to stability.

To prove its necessity, it suffices to show that there is no dominant strategy when $C(|L|) > 0$ for some $L \in \mathcal{L}$.

If $C(|L|) = +\infty$ for some $L$, we are in the case of the constrained/truncated DA, and it is well known that there is no dominant strategy (see, e.g., Haeringer and Klijn, 2009).

Now suppose that $0 < C(|L|) < +\infty$ for some $L \in \mathcal{L}$. If a strategy ranks fewer than $S$ schools with a positive probability, we know that it cannot be a dominant strategy for the same reason as in the contrained/truncated DA. If a strategy does always rank all schools, then it is weakly dominated by STT. We therefore need to show that STT is not a dominant strategy for all student types, for which we can construct an example where it is profitable for a student to drop some schools from her ROL to save application costs for some profiles of ROLs submitted by other students.

Therefore, there is no dominant strategy when $C(|L|) > 0$ for some $L \in \mathcal{L}$, and hence stability cannot be an equilibrium outcome in dominant strategy.

The second statement is implied by Proposition 1 and that DA produces a stable matching when everyone is STT. ■

**Proof of Proposition 3.**

(i) Suppose that given a realized matching $\hat{\mu}$, there is a student-school pair $(i, s)$ such that $\hat{\mu}(\theta_i) \neq \varnothing$, $u_{i,s} > u_{i,\hat{\mu}(\theta_i)}$, and $e_{i,s} \geqslant P_s(\hat{\mu})$. That is, $i$ is not matched with her favorite feasible school.

Since $i$ is weakly truth-telling, she must have ranked all schools that are more preferred to $\hat{\mu}(\theta_i)$, including $s$. The DA algorithm implies that $i$ must have been rejected by $s$ at some round given that she is accepted by a lower-ranked school $\hat{\mu}(\theta_i)$. As $i$ is rejected by $s$ in some round, the cutoff of $s$ must be higher than $e_{i,s}$. This contradiction rules out the existence of such matchings.

(ii) Given the result in part (i), when every student who has at least one feasible school is matched, everyone must be assigned to her favorite feasible school. Moreover, unmatched students have no feasible school. Therefore, the matching is stable. ■

## A.2 Asymptotics: Proofs and Additional Results

We now present the proofs of results in the main text as well as some additional results on the asymptotics and the continuum economy.

### A.2.1 Matching and the DA Mechanism in the Continuum Economy

We follow Abdulkadiroğlu et al. (2015) and Azevedo and Leshno (2016) to extend the definitions of matching and DA to the continuum economy, $E$.

Similar to that in finite economies, a matching in $E$ is a function $\mu : \Theta \to \mathcal{S} \cup \{\varnothing\}$, such that (i) $\mu(\theta_i) = s$ if student $i$ is matched with $s$; (ii) $\mu(\theta_i) = \varnothing$ if student $i$ is unmatched; (iii) $\mu^{-1}(s)$ is measurable and is the set of students matched with $s$, while $G(\mu^{-1}(s)) \leqslant q_s$; and (iv) for any $s \in \mathcal{S}$, the set $\{\theta \in \Theta : u_{i,\mu(\theta)} \leqslant u_{i,s}\}$ is open.

The last condition is imposed because in the continuum model it is always possible to add a measure-zero set of students to a school without exceeding its capacity. This would generate multiplicities of stable matchings that differ only in sets of measure zero. Condition (iv) rules out such multiplicities. The intuition is that the condition implies that a stable matching always allows an extra measure zero set of students into a school when this can be done without compromising stability.

The DA algorithm works almost the same as in a finite economy. Abdulkadiroğlu et al. (2015) formally define the algorithm, and prove that it converges. A sketch of the mechanism is as follows. At the first step, each student applies to her most preferred school. Every school tentatively admits up to its capacity from its applicants according to its priority order, and rejects the rest if there are any. In general, each student who was rejected in the previous step applies to her next preferred school. Each school considers the set of students it has tentatively admitted and the new applicants. It tentatively admits up to its capacity from these students in the order of its priority, and rejects the rest. The process converges when the set of students that are rejected has zero measure. Although this process might not complete in finite time, it converges in the limit (Abdulkadiroğlu et al., 2015).

### A.2.2 Proofs of Propositions 4 and 5

We start with some intermediate results. Similar to Azevedo and Leshno (2016), we define the convergence of $\{F^{(I)}\}_{I\in\mathbb{N}}$ to $E$ if $q^{(I)}$ converges to $q$ and if $G^{(I)}$ converges to $G$ in the weak-* topology.[A.1] We similarly define the convergence of $\{F^{(I)},\sigma^{(I)}\}_{I\in\mathbb{N}}$ to $(E,\sigma^{\infty})$, additionally requiring the empirical distributions of ROLs prescribed by $\sigma^{(I)}$ in finite economies to converge to those in $E$ prescribed by $\sigma^{\infty}$.

**Lemma A1.** *For a sequence of random economies and equilibrium strategies $\{F^{(I)},\sigma^{(I)}\}_{I\in\mathbb{N}}$ satisfying Assumption 2, $P^{(I)}$, the random cutoff associated with $(F^{(I)},\sigma^{(I)})$, converges to $P(\mu_{(E,\sigma^{\infty})})$ almost surely.*

**Proof of Lemma A1.**

First, we note that the sequence of random economies $\{F^{(I)}\}_{I\in\mathbb{N}}$ converges to $E$ almost surely. By construction, $q^{(I)}$ converges to $q$. Moreover, by the Glivenko-Cantelli Theorem, the empirical distribution functions $G^{(I)}$ converge to $G$ in the weak-* topology almost surely. Therefore, we have that $\{F^{(I)}\}_{I\in\mathbb{N}}$ converges to $E$ almost surely.

Second, we show that $\{F^{(I)},\sigma^{(I)}\}_{I\in\mathbb{N}}$ converges to $(E,\sigma^{\infty})$ almost surely. As $\sigma^{(I)}$ and $\sigma^{\infty}$ map student types to ROLs, $\{F^{(I)},\sigma^{(I)}\}_{I\in\mathbb{N}}$ is a sequence of random economies that are defined with ROLs. A student's "type" is now characterized by $(L_i,e_i)\in\mathcal{L}\times[0,1]^S$. Let $M^{\infty}$ be the probability measure on the modified student types in $(E,\sigma^{\infty})$. That is, for any $\Lambda\subset\mathcal{L}\times[0,1]^S$, $M^{\infty}(\Lambda)=G(\{\theta_i\in\Theta \mid (\sigma^{\infty}(\theta_i),e_i)\in\Lambda\})$. Similarly, $M^{(I)}$ is the empirical distribution of the modified types in the random economy $\{F^{(I)},\sigma^{(I)}\}$. We shall show that $M^{(I)}$ converges to $M^{\infty}$ in the weak-* topology almost surely.

Let $X:\mathcal{L}\times[0,1]^S\to[\underline{x},\overline{x}]\subset\mathbb{R}$ be a bounded continuous function. We also define $M^{(I)}_{\sigma^{\infty}}$ the random probability measure on $\mathcal{L}\times[0,1]^S$ when students play $\sigma^{\infty}$ in random economy $F^{(I)}$. Because the strategy is fixed at $\sigma^{\infty}$ for all $I$, by the same arguments as above (i.e., the convergence of $q^{(I)}$ to $q$ and the Glivenko-Cantelli Theorem), $M^{(I)}_{\sigma^{\infty}}$ converges to $M^{\infty}$ almost surely.

[A.1] The weak-* convergence of measures is defined as $\int Xd\hat{G}^{(I)}\to\int XdG$ for every bounded continuous function $X:[0,1]^{2S}\to\mathbb{R}$, given a sequence of realized empirical distributions $\{\hat{G}^{(I)}\}_{I\in\mathbb{N}}$. This is also known as narrow convergence or weak convergence.

Let $\Theta^{(I)} = \{\theta_i \in \Theta \mid \sigma^{(I)}(\theta_i) \neq \sigma^{\infty}(\theta_i)\}$. We have the following results:

$$\begin{aligned}
&\left|\int X dM^{(I)} - \int X dM^{\infty}\right| \\
\leqslant&\left|\int X dM^{(I)} - \int X dM^{(I)}_{\sigma^{\infty}}\right| + \left|\int X dM^{(I)}_{\sigma^{\infty}} - \int X dM^{\infty}\right| \\
=&\left|\int X(\sigma^{(I)}(\theta_i), e_i) dG^{(I)} - \int X(\sigma^{\infty}(\theta_i), e_i) dG^{(I)}\right| + \left|\int X dM^{(I)}_{\sigma^{\infty}} - \int X dM^{\infty}\right| \\
=&\left|\int_{\theta_i \in \Theta^{(I)}} \left[X(\sigma^{(I)}(\theta_i), e_i) - X(\sigma^{(\infty)}(\theta_i), e_i)\right] dG^{(I)}\right| + \left|\int X dM^{(I)}_{\sigma^{\infty}} - \int X dM^{\infty}\right| \\
\leqslant&(\overline{x} - \underline{x}) G^{(I)}(\Theta^{(I)}) + \left|\int X dM^{(I)}_{\sigma^{\infty}} - \int X dM^{\infty}\right|,
\end{aligned}$$

where the first inequality is due to the triangle inequality; the equalities are because of the definitions of $M^{(I)}$ and $M^{(I)}_{\sigma^{\infty}}$ and because $X(\sigma^{(I)}(\theta_i), e_i) = X(\sigma^{\infty}(\theta_i), e_i)$ whenever $\theta_i \notin \Theta^{(I)}$; the last inequality comes from the boundedness of $X$.

Because $\lim_{I\to\infty} G(\Theta^{(I)}) = 0$ by Assumption 1 and $G^{(I)}$ converges to $G$ almost surely, $\lim_{I\to\infty} G^{(I)}(\Theta^{(I)}) = 0$ almost surely. Moreover, $M^{(I)}_{\sigma^{\infty}}$ converges to $M^{\infty}$ almost surely, and thus the above inequalities implies $\int X dM^{(I)}$ converges to $\int X dM^{\infty}$ almost surely. By the Portmanteau theorem, $M^{(I)}$ converge to $M^{\infty}$ in the weak-* topology almost surely.

This proves $\{F^{(I)}, \sigma^{(I)}\}_{I\in\mathbb{N}}$ converges to $(E, \sigma^{\infty})$ almost surely. By Proposition 3 of Azevedo and Leshno (2016), $P^{(I)}$ converges to $P(\mu_{(E,\sigma^{\infty})})$ almost surely. ■

**Proposition A1.** *Given Assumption 1, in a sequence of random economies and equilibrium strategies $\{F^{(I)}, \sigma^{(I)}\}_{I\in\mathbb{N}}$ satisfying Assumption 2, $\mu(E, \sigma^{\infty}) = \mu^{\infty}$ and thus $\sigma^{\infty}(\theta_i)$ ranks $\mu^{\infty}(\theta_i)$ for all $\theta_i \in \Theta$ except a measure-zero set of student types.*

**Proof of Proposition A1.**

Suppose that the first statement in the proposition is not true, $G(\{\theta_i \in \Theta \mid \mu_{(E,\sigma^{\infty})}(\theta_i) \neq \mu^{\infty}(\theta_i)\}) > 0$ and therefore $P(E, \sigma^{\infty}) \neq P^{\infty}$. Because there is a unique stable matching in $E$, which is the unique equilibrium outcome, by Assumption 1, $\mu_{(E,\sigma^{\infty})}$ is not stable and thus is not an equilibrium outcome.

Recall that $P(E, \sigma^{\infty})$, $P^{\infty}$, $\mu_{(E,\sigma^{\infty})}$, and $\mu^{\infty}$ are constants, although their counterparts in finite economies are random variables. Moreover, $\sigma^{(I)}$ and $\sigma^{\infty}$ are not random either.

For some $\eta, \xi > 0$, we define:

$$\Theta_{(\eta,\xi)} = \left\{ \theta_i \in \Theta \;\middle|\; \begin{array}{c} e_{i,\mu^\infty(\theta_i)} - P^\infty_{\mu^\infty(\theta_i)} > \eta, \\ e_{i,\mu_{(E,\sigma^\infty)}(\theta_i)} - P_{\mu_{(E,\sigma^\infty)}(\theta_i)}(\mu_{(E,\sigma^\infty)}) > \eta, \\ e_{i,s} - P_s(\mu_{(E,\sigma^\infty)}) < -\eta, \text{ for all } s \text{ ranked above } \mu_{(E,\sigma^\infty)}(\theta_i) \text{ by } \sigma^\infty(\theta_i); \\ u_{i,\mu^\infty(\theta_i)} - u_{i,\mu_{(E,\sigma^\infty)}(\theta_i)} > \xi. \end{array} \right\},$$

$\Theta_{(\eta,\xi)}$ must have a positive measure for some $\eta, \xi > 0$ and is a subset of students who can form a blocking pair in $\mu_{(E,\sigma^\infty)}$. Clearly, $\sigma^\infty(\theta_i)$ ranks $\mu_{(E,\sigma^\infty)}(\theta_i)$ but not $\mu^\infty(\theta_i)$ for all $\theta_i \in \Theta_{(\eta,\xi)}$. We further define:

$$\Theta^{(I)}_{(\eta,\xi)} = \Theta_{(\eta,\xi)} \cap \{\theta_i \in \Theta \mid \sigma^{(I)}(\theta_i) \text{ ranks } \mu_{(E,\sigma^\infty)}(\theta_i) \text{ but not } \mu^\infty(\theta_i)\}.$$

By Assumption 2, $\sigma^{(I)}$ converges to $\sigma^\infty$, and thus $\Theta^{(I)}_{(\eta,\xi)}$ converges to $\Theta_{(\eta,\xi)}$ and has a positive measure when $I$ is sufficiently large.

We show below that $\{\sigma^{(I)}\}_{I\in\mathbb{N}}$ is not a sequence of equilibrium strategies. Consider a unilateral deviation for $\theta_i \in \Theta^{(I)}_{(\eta,\xi)}$ from $\sigma^{(I)}(\theta_i)$ to $L_i$ such that the only difference between the two actions is that $\mu_{(E,\sigma^\infty)}(\theta_i)$, ranked in $\sigma^{(I)}(\theta_i)$, is replaced by $\mu^\infty(\theta_i)$ in $L_i$ while $L_i$ is kept as a partial order of $i$'s true preferences.

By Lemma A1, for $0 < \phi < \xi/(1+\xi)$ there exists $n \in \mathbb{N}$ such that, in all $F^{(I)}$ with $I > n$, $i$ is matched with $\mu_{(E,\sigma^\infty)}(\theta_i)$ with probability at least $(1-\phi)$ if submitting $\sigma^{(I)}(\theta_i)$ but would have been matched with $\mu^\infty(\theta_i)$ if instead $L_i$ had been submitted.

Let $\mathrm{EU}(\sigma^{(I)}(\theta_i))$ be the expected utility when submitting $\sigma^{(I)}(\theta_i)$. Then $\mathrm{EU}(\sigma^{(I)}(\theta_i)) \leqslant (1-\phi)u_{i,\mu_{(E,\sigma^\infty)}(\theta_i)} + \phi$ because $\max_s\{u_{i,s}\} \leqslant 1$ by assumption, and $\mathrm{EU}(L_i) \geqslant (1-\phi)u_{i,\mu^\infty(\theta_i)}$. The difference between the two actions is:

$$\begin{aligned} \mathrm{EU}(L_i) - \mathrm{EU}(\sigma^{(I)}(\theta_i)) \geqslant & (1-\phi)u_{i,\mu^\infty(\theta_i)} - (1-\phi)u_{i,\mu_{(E,\sigma^\infty)}(\theta_i)} - \phi \\ \geqslant & (1-\phi)\xi - \phi > 0, \end{aligned}$$

which proves that $\{\sigma^{(I)}\}_{I\in\mathbb{N}}$ is not a sequence of equilibrium strategies. This contradiction further shows that $G(\{\theta_i \in \Theta \mid \mu_{(E,\sigma^\infty)}(\theta_i) \neq \mu^\infty(\theta_i)\}) = 0$ and that $\sigma^\infty(\theta_i)$ ranks $\mu^\infty(\theta_i)$ for all $\theta_i \in \Theta$ except a measure-zero set of student types. ■

We are now ready to prove Proposition 4.

**Proof of Proposition 4.**

Part (i) is implied by Lemma A1 and Proposition A1. Because $F^{(I)}$ converges to $E$ almost surely and $\sigma^{(I)}$ converges to $\sigma^{\infty}$, $P^{(I)}$ converges to $P(\mu_{(E,\sigma^{\infty})})$ almost surely. Moreover, $\mu_{(E,\sigma^{\infty})} = \mu^{\infty}$ except a measure-zero set of students implies that $P(\mu_{(E,\sigma^{\infty})}) = P^{\infty}$. Therefore, $\lim_{I\to\infty} P^{(I)} = P^{\infty}$ almost surely.

To show part (ii), we first define $\Theta^{(I)} = \{\theta_i \in \Theta \mid \sigma^{(I)}(\theta_i) \neq \sigma^{\infty}(\theta_i)\}$. By Assumption 1, $G^{(I)}(\Theta^{(I)})$ converges to zero almost surely. We have the following inequalities:

$$
\begin{aligned}
& G^{(I)}\left(\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \underset{s\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s}\right\}\right) \\
\leqslant & \left| G^{(I)}\left(\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \underset{s\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s}\right\}\right) - G^{(I)}\left(\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}\right\}\right)\right| \\
& + G^{(I)}\left(\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}\right\}\right) \\
\leqslant & G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mathcal{S}(e_i,P^{\infty}) \neq \mathcal{S}(e_i,P^{(I)})\right\}\right) + G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\right\}\right),
\end{aligned}
$$

where the first inequality is due to the triangle inequality; the second inequality is because $\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \arg\max_{s\in\mathcal{S}(e_i,P^{(I)})} u_{i,s}\right\}$ and $\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \arg\max_{s\in\mathcal{S}(e_i,P^{\infty})} u_{i,s}\right\}$ can possibly differ only when $\mathcal{S}(e_i,P^{\infty}) \neq \mathcal{S}(e_i,P^{(I)})$ and because:

$$
\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}\right\} = \left\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\right\}.
$$

Furthermore,

$$
\begin{aligned}
& G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mathcal{S}(e_i,P^{\infty}) \neq \mathcal{S}(e_i,P^{(I)})\right\}\right) \\
= & G^{(I)}\left(\left\{\theta_i \in \Theta \mid \min(P^{\infty}_s, P^{(I)}_s) \leqslant e_{i,s} < \max(P^{\infty}_s, P^{(I)}_s), \exists s \in \mathcal{S}\right\}\right).
\end{aligned}
$$

The right hand side converges to zero almost surely, because almost surely $G^{(I)}$ converges to $G$, which is atomless, and $\lim_{n\to\infty} P^{(I)} = P^{\infty}$ almost surely. Therefore,

$$
\lim_{I\to\infty} G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mathcal{S}(e_i,P^{\infty}) \neq \mathcal{S}(e_i,P^{(I)})\right\}\right) = 0 \text{ almost surely.} \tag{A.2}
$$

Moreover,

$$
\begin{aligned}
&G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\right\}\right) \\
\leqslant &G^{(I)}(\Theta^{(I)}) + G^{(I)}\left(\left\{\theta_i \in \Theta \backslash \Theta^{(I)} \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\right\}\right) \\
\leqslant &G^{(I)}(\Theta^{(I)}) + G^{(I)}\left(\left\{\theta_i \in \Theta \backslash \Theta^{(I)} \mid u_{i,\mu^{(I)}(\theta_i)} > u_{i,\mu^{\infty}(\theta_i)} \ \& \ e_{i,\mu^{(I)}(\theta_i)} \geqslant P^{(I)}_{\mu^{(I)}(\theta_i)}\right\}\right) \\
&+ G^{(I)}\left(\left\{\theta_i \in \Theta \backslash \Theta^{(I)} \mid u_{i,\mu^{(I)}(\theta_i)} < u_{i,\mu^{\infty}(\theta_i)},\ e_{i,\mu^{\infty}(\theta_i)} < P^{(I)}_{\mu^{\infty}(\theta_i)}, \ \& \ e_{i,\mu^{(I)}(\theta_i)} \geqslant P^{(I)}_{\mu^{(I)}(\theta_i)}\right\}\right)
\end{aligned}
$$

In the first inequality, we decompose the student type space into two, $\Theta^{(I)}$ and $\Theta \backslash \Theta^{(I)}$. In the former, students do not adopt $\sigma^{\infty}$, while those in the latter set do and thus rank the school prescribed by $\mu^{\infty}$. The second inequality consider the events when $\mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)$ can possibly happen.

The almost-sure convergence of $G^{(I)}$ to $G$ and that of $P^{(I)}$ to $P^{\infty}$ implies that:

$$
\lim_{I \to \infty} G^{(I)}\left(\left\{\theta_i \in \Theta \backslash \Theta^{(I)} \mid u_{i,\mu^{(I)}(\theta_i)} > u_{i,\mu^{\infty}(\theta_i)} \ \& \ e_{i,\mu^{(I)}(\theta_i)} \geqslant P^{(I)}_{\mu^{(I)}(\theta_i)}\right\}\right) = 0 \text{ almost surely,}
$$

because for any $s$ such that $u_{i,s} > u_{i,\mu^{\infty}(\theta_i)}$, we must have $e_{i,s} < P^{\infty}_s$.

Similarly, almost surely,

$$
\lim_{I \to \infty} G^{(I)}\left(\left\{\theta_i \in \Theta \backslash \Theta^{(I)} \mid u_{i,\mu^{(I)}(\theta_i)} < u_{i,\mu^{\infty}(\theta_i)},\ e_{i,\mu^{\infty}(\theta_i)} < P^{(I)}_{\mu^{\infty}(\theta_i)}, \ \& \ e_{i,\mu^{(I)}(\theta_i)} \geqslant P^{(I)}_{\mu^{(I)}(\theta_i)}\right\}\right) = 0.
$$

Therefore, $G^{(I)}\left(\left\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\right\}\right) = 0$ almost surely. Together with (A.2), it implies that $G^{(I)}\left(\left\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \arg\max_{s \in \mathcal{S}(e_i, P^{(I)})} u_{i,s}\right\}\right)$ converges to zero almost surely. In other words, $\{\mu^{(I)}\}_{I \in \mathbb{N}}$ is asymptotically stable. ■

**Proof of Proposition 5.**

The first statement in part (i) is implied by Proposition 2. Suppose that $i$ is in a blocking pair with some school $s$. It means that the ex post cutoff of $s$ is lower than $i$'s priority index at $s$. Therefore, if $s \in L^{(I)}_i$, the stability of DA (with respect to ROLs) implies that $i$ must be accepted by $s$ or by schools ranked above and thus preferred to $s$. Therefore, $i$ and $s$ cannot form a blocking pair if $s \in L^{(I)}_i$, which proves the second statement in part (i).

Part (iii) is implied by Proposition 4 (part (i)).

To show part (ii), we let $\mathcal{S}_0^{(I)} = \mathcal{S} \backslash L_i^{(I)}$ and therefore:

$$\begin{aligned} B_i^{(I)} &= \Pr(\exists s \in \mathcal{S}_0, u_{i,s} > u_{i,\mu^{(I)}(\theta_i)} \text{ and } e_{i,s} \geqslant P_s^{(I)}) \\ &\leqslant \sum_{s \in \mathcal{S}_0^{(I)}} \Pr\left(e_{i,t} < P_t^{(I)}, \forall t \in L_i^{(I)}, s.t., u_{i,t} > u_{i,s}; e_{i,s} \geqslant P_s^{(I)}\right). \end{aligned}$$

Let $B_{i,s}^{(I)}$ be $\Pr\left(e_{i,t} < P_t^{(I)}, \forall t \in L_i^{(I)}, s.t., u_{i,t} > u_{i,s}; e_{i,s} \geqslant P_s^{(I)}\right)$ for $s \in \mathcal{S}_0^{(I)}$. Since $s \in \mathcal{S}_0$ and $L_i^{(I)}$ is ex ante optimal for $i$ in $F^{(I)}, \sigma^{(I)}$, it implies:

$$\begin{aligned} &\sum_{s \in \mathcal{S}} u_{i,s} \int a_s\left(L_i^{(I)}, e_i; \sigma^{(I)}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i}) - C\left(\left|L_i^{(I)}\right|\right) \\ \geqslant &\sum_{s \in \mathcal{S}} u_{i,s} \int a_s\left(L, e_i; \sigma^{(I)}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i}) - C\left(\left|L_i^{(I)}\right| + 1\right) \end{aligned}$$

where $L$ ranks all schools in $L_i^{(I)}$ and $s$ while respecting their true preference rankings, i.e., adding $s$ to the true partial preference order $L_i^{(I)}$ while keeping the new list a true partial preference order.

For notational convenience, we relabel the schools such that school $k$ is the $k$th choice in $L$ and that $s$ is $k^*$th school in $L$. Those not ranked in $L$ are labeled as $|L_i^{(I)}| + 2, \cdots, S$. It then follows that:

$$\begin{aligned} &C\left(\left|L_i^{(I)}\right| + 1\right) - C\left(\left|L_i^{(I)}\right|\right) \\ \geqslant &\sum_{t=1}^{k^*-1} 0 + B_{i,s}^{(I)} u_{i,s} \\ &+ \sum_{t=k^*+1}^{|L_i^{(I)}|+1} u_{i,t} \cdot \left( \begin{array}{c} \Pr\left(e_{i,t} \geqslant P_t^{(I)}; e_{i,\tau} < P_\tau^{(I)}, \tau = 1, \ldots, t-1\right) - \\ \Pr\left(e_{i,t} \geqslant P_t^{(I)}; e_{i,\tau} < P_\tau^{(I)}, \tau = 1, \ldots, k^*-1, k^*+1, \cdots, t-1\right) \end{array} \right), \\ = &\sum_{t=1}^{k^*-1} 0 + B_{i,s}^{(I)} u_{i,s} \\ &- \sum_{t=k^*+1}^{|L_i^{(I)}|+1} u_{i,t} \cdot \Pr\left(e_{i,t} \geqslant P_t^{(I)}; e_{i,s} \geqslant P_s^{(I)}; e_{i,\tau} < P_\tau^{(I)}, \tau = 1, \ldots, k^*-1, k^*+1, \cdots, t-1\right), \end{aligned}$$

where the zeros in the first term on the right come from the upper invariance of DA. That is, the admission probability at any school ranked above $s$ is the same when $i$ submits either $L_i^{(I)}$ or $L$.

Also note that $u_{i,s} > u_{i,t}$ for all $t \geqslant k^* + 1$ and that:

$$\sum_{t=k^*+1}^{|L_i^{(I)}|+1} \Pr\left(e_{i,t} \geqslant P_t^{(I)}; e_{i,s} \geqslant P_s^{(I)}; e_{i,\tau} < P_\tau^{(I)}, \tau = 1, \ldots, k^* - 1, k^* + 1, \cdots, t - 1\right) \leqslant B_{i,s}^{(I)}.$$

Besides, $u_{i,k^*+1} \geqslant u_{i,t}$ for all $t = k^* + 2, \cdots, |L_i| + 1$. Therefore, for all $s \in \mathcal{S}_0^{(I)}$,

$$C\left(\left|L_i^{(I)}\right| + 1\right) - C\left(\left|L_i^{(I)}\right|\right) \geqslant B_{i,s}^{(I)} u_{i,s} - B_{i,s}^{(I)} u_{i,k^*+1}$$

This leads to:

$$B_{i,s}^{(I)} \leqslant \frac{C\left(\left|L_i^{(I)}\right| + 1\right) - C\left(\left|L_i^{(I)}\right|\right)}{u_{i,s} - u_{i,k^*+1}} \leqslant \frac{C\left(\left|L_i^{(I)}\right| + 1\right) - C\left(\left|L_i^{(I)}\right|\right)}{u_{i,s}}.$$

Finally, $B_i^{(I)} \leqslant \sum_{s \in \mathcal{S}_0^{(I)}} B_{i,s}^{(I)} \leqslant |\mathcal{S} \backslash L_i| \frac{C(|L_i|+1) - C(|L_i|)}{\max_{s \in \mathcal{S} \backslash L_i^{(I)}} u_{i,s}}$. ■

### A.2.3 Asymptotic Distribution of Cutoffs and Convergence Rates

For the next result, we define the demand for each school in $(E, \sigma)$ as a function of the cutoffs:

$$D_s(P \mid E, \sigma) = \int \mathbb{1}(u_{i,s} = \max_{s' \in \mathcal{S}(e_i, P) \bigcap \sigma(\theta_i)} u_{i,s'}) dG(\theta_i),$$

where $\sigma(\theta_i)$ also denotes the set of schools ranked by $i$; $\mathbb{1}()$ is an indicator function. Let $D(P \mid E, \sigma) = [D_s(P \mid E, \sigma)]_{s \in \mathcal{S}}$.

**Assumption A1.**

*(i) There exists $n \in \mathbb{N}$ such that $\sigma^{(I)} = \sigma^\infty$ for all $I > n$;*

*(ii) $D(\cdot \mid E, \sigma^\infty)$ is $C^1$ and $\partial D(P^\infty \mid E, \sigma^\infty)$ is invertible;*

*(iii) $\sum_{s=1}^S q_s < 1$.*

Part (i) says that $\sigma^\infty$ maintains as an equilibrium strategy in any economy of a size that is above a threshold. This is supported partially by the discussion in Section A.2.4. In particular, when $C(2) > 0$, part (i) is satisfied. $D(\cdot \mid E, \sigma^\infty)$ being $C^1$ (in part ii) holds when $G$ admits a continuous density. In this case, the fraction of students whose demand is affected by changes in $P$ is continuous. Part (iii) guarantees that every school has a positive cutoff in the stable matching of $E$.

Because of Assumption A1, our setting with cardinal preferences can be transformed into one defined by students' ROLs that is identical to that in Azevedo and Leshno (2016). Therefore, some of their results are also satisfied in our setting.

**Proposition A2.** *In a sequence of matchings, $\{\mu^{(I)}\}_{I\in\mathbb{N}}$, of the sequence of random economies and equilibrium strategies, $\{F^{(I)},\sigma^{(I)}\}_{I\in\mathbb{N}}$, satisfying Assumption A1, we have the following results:*

*(i) The distribution of cutoffs in $(F^{(I)},\sigma^{(I)})$ satisfies:*

$$\sqrt{I}\left(P^{(I)}-P^{\infty}\right)\xrightarrow{d}N(0,V(\sigma^{\infty}))$$

*where $V(\sigma^{\infty})=\partial D(P^{\infty}\mid E,\sigma^{\infty})^{-1}\Sigma\left(\partial D(P^{\infty}\mid E,\sigma^{\infty})^{-1}\right)'$, and*

$$\Sigma=\begin{pmatrix} q_1(1-q_1) & -q_1q_2 & \cdots & -q_1q_S \\ -q_2q_1 & q_2(1-q_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & -q_{S-1}q_S \\ -q_Sq_1 & \cdots & -q_Sq_{S-1} & q_S(1-q_S) \end{pmatrix}.$$

*(ii) For any $\eta>0$ and $I>n$, there exist constants $\gamma_1$ and $\gamma_2$ such that the probability that the matching $\mu^{(I)}$ has cutoffs $||P^{(I)}-P^{\infty}||>\eta$ is bounded by $\gamma_1e^{-\gamma_2I}$:*

$$\Pr\left(||P^{(I)}-P^{\infty}||>\eta\right)<\gamma_1e^{-\gamma_2I}.$$

*(iii) Moreover, suppose that $G$ admits a continuous density. For any $\eta>0$ and $I>n$, there exist constants $\gamma_1'$ and $\gamma_2'$ such that, in matching $\mu^{(I)}$, the probability of the fraction of students who can form a blocking pair being greater than $\eta$ is bounded by $\gamma_1'e^{-\gamma_2'I}$:*

$$\Pr\left(G^{(I)}(\{\theta_i\in\Theta\mid\mu^{(I)}(\theta_i)\notin\underset{s\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s}\})>\eta\right)<\gamma_1'e^{-\gamma_2'I}.$$

Parts (i) and (ii) are from Azevedo and Leshno (2016) (Proposition G1 and part 2 of Proposition 3), although our part (iii) is new and extends their part 3 of Proposition 3. Proposition A2 describes convergence rates and thus has implications for empirical approaches based on stability (see Section II.C).

**Proof of Proposition A2 (part iii).**

To show part (iii), we use similar techniques as in the proof for Proposition 3 (part 3) in Azevedo and Leshno (2016). We first derive the following results.

$$
\begin{aligned}
& G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \notin \underset{s \in \mathcal{S}(e_i, P^{(I)})}{\arg\max}\ u_{i,s}\}) \\
=& G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \notin [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \forall s \in \mathcal{S}; \mu^{(I)}(\theta_i) \notin \underset{s \in \mathcal{S}(e_i, P^{\infty})}{\arg\max}\ u_{i,s}\} \\
& + G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}; \mu^{(I)}(\theta_i) \notin \underset{s \in \mathcal{S}(e_i, P^{(I)})}{\arg\max}\ u_{i,s}\}) \\
\leqslant & G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\} \\
& + G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}\})
\end{aligned}
$$

In the first equality, whenever $e_{i,s} \notin [\min(P_s^{\infty}, P_s^{(I_m)}), \max(P_s^{\infty}, P_s^{(I_m)})], \forall s \in \mathcal{S}$, $i$ faces the same set of feasible schools given either $P^{(I)}$ or $P^{\infty}$, $\mathcal{S}(e_i, P^{(I)}) = \mathcal{S}(e_i, P^{\infty})$. Because $\mu^{\infty}$ is stable, $\mu^{\infty}(\theta_i)$ is $i$'s favorite school in $\mathcal{S}(e_i, P^{\infty})$; together with the relaxation of the conditions in the second term, it leads to the inequality.

By Azevedo and Leshno (2016) Proposition 3 (part 3), we can find $\gamma_1'$ and $\gamma_2'$ such that:

$$
\Pr\left(G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\}) > \eta/2\right) < \gamma_1' e^{-\gamma_2' I}/2. \tag{A.3}
$$

Let $\overline{g}$ be the supremum of the marginal probability density of $e_{i,s}$ across all $s$. Denote the set of student types with priority indices which have at least one coordinate close to $P^{\infty}$ by distance $\eta_1/(2S\overline{g})$ (where $\eta_1 = \eta/4$):

$$
\Theta_{\eta_1} = \{\theta_i \in \Theta \mid \exists s \in \mathcal{S}, |e_{i,s} - P_s^{\infty}| \leqslant \eta_1/(2S\overline{g})\}.
$$

Then $G(\Theta_{\eta_1}) \leqslant 2S\overline{g} \cdot \eta_1/(2S\overline{g}) = \eta_1$. The fraction of students in $F^{(I)}$ that have types in $\Theta_{\eta_1}$ is then $G^{(I)}(\Theta_{\eta_1})$. Note that $G^{(I)}(\Theta_{\eta_1})$ is a random variable with mean $G(\Theta_{\eta_1})$. By the Vapnik-Chervonenkis Theorem,[A.2]

$$
\Pr(G^{(I)}(\Theta_{\eta_1}) > 2\eta_1) < \Pr(|G^{(I)}(\Theta_{\eta_1}) - G(\Theta_{\eta_1})| > \eta_1) < \gamma_1' e^{-\gamma_2' I}/4. \tag{A.4}
$$

[A.2] See Azevedo and Leshno (2016) and the references therein for more details on the theorem for its application in our context.

Moreover, by part (ii), we know that:

$$\Pr\left(||P^{(I)} - P^{\infty}|| > \eta_1/(2S\overline{g})\right) < \gamma_1' e^{-\gamma_2' I}/4. \tag{A.5}$$

We can choose $\gamma_1'$ and $\gamma_2'$ appropriately, so that (A.3), (A.4), and (A.5) are all satisfied.

When the event, $||P^{(I)} - P^{\infty}|| > \eta_1/(2S\overline{g})$, does not happen,

$$\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}\} \subset \Theta_{\eta_1}.$$

When neither $||P^{(I)} - P^{\infty}|| > \eta_1/(2S\overline{g})$ nor $G^{(I)}(\Theta_{\eta_1}) > 2\eta_1$ happens,

$$G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}\}) \leqslant 2\eta_1 = \eta/2;$$

the probability that neither of these two events happens is at least $1 - \gamma_1' e^{-\gamma_2' I}/4 - \gamma_1' e^{-\gamma_2' I}/4 = 1 - \gamma_1' e^{-\gamma_2' I}/2$. This implies,

$$\Pr(G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}\}) > \eta/2) < \gamma_1' e^{-\gamma_2' I}/2. \tag{A.6}$$

The events in (A.3) and (A.6) do not happen with probability at least $1 - \gamma_1' e^{-\gamma_2' I}/2 - \gamma_1' e^{-\gamma_2' I}/2 = 1 - \gamma_1' e^{-\gamma_2' I}$; and when they do not happen,

$$\begin{aligned}
&G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \notin \underset{s \in \mathcal{S}(e_i, P^{(I)})}{\arg\max}\ u_{i,s}\}) \\
\leqslant &G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \neq \mu^{\infty}(\theta_i)\} \\
&+ G^{(I)}(\{\theta_i \in \Theta \mid e_{i,s} \in [\min(P_s^{\infty}, P_s^{(I)}), \max(P_s^{\infty}, P_s^{(I)})), \exists s \in \mathcal{S}\}) \\
\leqslant &\eta.
\end{aligned}$$

Therefore,

$$\Pr\left(G^{(I)}(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) \notin \underset{s \in \mathcal{S}(e_i, P^{(I)})}{\arg\max}\ u_{i,s}\}) > \eta\right) < \gamma_1' e^{-\gamma_2' I}.$$

■

### A.2.4 Properties of Equilibrium Strategies in Large Economies

We now discuss the properties of Bayesian Nash equilibria in a sequence of random economies and thus provide some justifications to Assumptions 2 and A1.

We start with Lemma A2 showing that a strategy, which does not result in the stable matching in the continuum economy when being adopted by students in the continuum economy, cannot survive as an equilibrium strategy in sufficiently large economies. This immediately implies that, in finite large economies, every student always includes in her ROL the matched school in the continuum-economy stable matching (Lemma A3). Moreover, students do not pay a cost to rank more schools in large economies (Lemma A4). Lastly, when it is costly to rank more than one school ($C(2) > 0$), in sufficiently large economies, it is an equilibrium strategy for every students to only rank the matched school prescribed by the continuum-economy stable matching.

**Lemma A2.** *If a strategy $\sigma$ results in a matching $\mu_{(E,\sigma)}$ in the continuum economy such that $G(\{\theta_i \in \Theta \mid \mu_{(E,\sigma)}(\theta_i) \neq \mu^{\infty}(\theta_i)\}) > 0$, then there must exist $n \in \mathbb{N}$ such that $\sigma$ is not an equilibrium in $F^{(I)}$ for all $I > n$.*

**Proof of Lemma A2.**

Suppose instead that there is a subsequence of finite random economies $\{F^{(I_n)}\}_{n\in\mathbb{N}}$ such that $\sigma$ is always an equilibrium. Note that we still have $F^{(I_n)} \rightarrow E$ almost surely, and therefore $\{F^{(I_n)}, \sigma\}$ converges to $\{E, \sigma\}$ almost surely.

Given the student-proposing DA, we focus on the student-optimal stable matching (SOSM) in $(F^{(I)}, \sigma)$. By Proposition 3 of Azevedo and Leshno (2016), it must be that $P^{(I_n)} \rightarrow P^{\sigma}$ almost surely, where $P^{(I_n)} = P(\mu_{(F^{(I_n)},\sigma)})$ and $P^{\sigma} = P(\mu_{(E,\sigma)})$.

Because there is a unique equilibrium outcome in $E$, which is also the unique stable matching in $E$ by assumption, $G(\{\theta_i \in \Theta_i \mid \mu_{(E,\sigma)}(\theta_i) \neq \mu^{\infty}(\theta_i)\}) > 0$ in the continuum economy implies that $P^{\sigma}$ is not the cutoffs of $\mu^{\infty}$ ($E$'s stable matching), $P^{\sigma} \neq P^{\infty}$.

Because there is a unique stable matching in $E$ by assumption, $\mu_{(E,\sigma^{\infty})}$ is not stable and thus is not an equilibrium outcome in $E$. There exist some $\eta, \xi > 0$, such that:

$$\Theta_{(\eta,\xi)} = \left\{ \theta_i \in \Theta \;\middle|\; \begin{array}{c} e_{i,\mu^{\infty}(\theta_i)} - P^{\infty}_{\mu^{\infty}(\theta_i)} > \eta, \\ e_{i,\mu_{(E,\sigma)}(\theta_i)} - P^{\sigma}_{\mu_{(E,\sigma)}(\theta_i)} > \eta, \\ e_{i,s} - P^{\sigma}_{s} < -\eta, \text{ for all } s \text{ ranked above } \mu_{(E,\sigma)}(\theta_i) \text{ by } \sigma(\theta_i); \\ u_{i,\mu^{\infty}(\theta_i)} - u_{i,\mu_{(E,\sigma)}(\theta_i)} > \xi. \end{array} \right\}$$

$\Theta_{(\eta,\xi)}$ must have a positive measure for some $\eta, \xi > 0$ and is a subset of students who can form a blocking pair in $\mu_{(E,\sigma)}$. Clearly, $\sigma(\theta_i)$ ranks $\mu_{(E,\sigma)}(\theta_i)$ but not $\mu^\infty(\theta_i)$ for all $\theta_i \in \Theta_{(\eta,\xi)}$.

We show below that $\sigma$ is not an equilibrium strategy in sufficiently large economies. Consider a unilateral deviation for $\theta_i \in \Theta_{(\eta,\xi)}$ from $\sigma(\theta_i)$ to $L_i$ such that the only difference between the two actions is that $\mu_{(E,\sigma)}(\theta_i)$, ranked in $\sigma(\theta_i)$, is replaced by $\mu^\infty(\theta_i)$ in $L_i$ while $L_i$ is kept as a partial order of $i$'s true preferences.

Because $P^{(I_n)} \rightarrow P^\sigma$ almost surely, for $0 < \phi < \xi/(1+\xi)$ there exists $n_1 \in \mathbb{N}$ such that, in all $F^{(I_n)}$ with $I_n > n_1$, $i$ is matched with $\mu_{(E,\sigma)}(\theta_i)$ with probability at least $(1-\phi)$ if submitting $\sigma(\theta_i)$, but would have been matched with $\mu^\infty(\theta_i)$ if instead $L_i$ had been submitted.

Let $\mathrm{EU}(\sigma(\theta_i))$ be the expected utility when submitting $\sigma(\theta_i)$. Then $\mathrm{EU}(\sigma(\theta_i)) \leqslant (1 - \phi)u_{i,\mu_{(E,\sigma)}(\theta_i)} + \phi$, because $\max_s\{u_{i,s}\} \leqslant 1$ by assumption, and $\mathrm{EU}(L_i) \geqslant (1-\phi)u_{i,\mu^\infty(\theta_i)}$. The difference between the two actions is:

$$\begin{aligned}\mathrm{EU}(L_i) - \mathrm{EU}(\sigma(\theta_i)) \geqslant &(1-\phi)u_{i,\mu^\infty(\theta_i)} - (1-\phi)u_{i,\mu_{(E,\sigma)}(\theta_i)} - \phi \\ \geqslant &(1-\phi)\xi - \phi > 0,\end{aligned}$$

implying that $\sigma$ is not an equilibrium strategy in $F^{(I_n)}$ for $I_n > n_1$. This contradiction further implies that there exist $n \in \mathbb{N}$ such that $\sigma$ is not an equilibrium strategy in all $F^{(I)}$ with $I > n$. ■

**Lemma A3.** *If a strategy $\sigma$ is such that $G(\{\theta_i \in \Theta \mid \sigma(\theta_i)$ does not rank $\mu^\infty(\theta_i)\}) > 0$, then there must exist $n \in \mathbb{N}$ such that $\sigma$ is not an equilibrium in $F^{(I)}$ for all $I > n$.*

**Proof of Lemma A3.**

Note that $G(\{\theta_i \in \Theta \mid \sigma(\theta_i)$ does not rank $\mu^\infty(\theta_i)\}) > 0$ implies $G(\{\theta_i \in \Theta \mid \mu_{(E,\sigma)}(\theta_i) \neq \mu^\infty(\theta_i)\}) > 0$, because $i$ cannot be matched with $\mu^\infty(\theta_i)$ if $\sigma(\theta_i)$ does not rank $\mu^\infty(\theta_i)$. Lemma A2 therefore implies the statement in this lemma. ■

Lemmata A2 and A3 imply that, in large enough economies, there exist equilibrium strategies with which every student ranks her matched school prescribed by $\mu^\infty$. The following lemma further bounds the number of choices that a student ranks.

**Lemma A4.** *Suppose $C(K) = 0$ and $C(K+1) > 0$ for $1 \leqslant K \leqslant (S-1)$. Consider a strategy $\sigma$ such that $\sigma(\theta_i)$ ranks at least $K+1$ schools for all $\theta_i \in \Theta' \subset \Theta$ and $G(\Theta') > 0$.*

*In the sequence of random economies, $\{F^{(I)}\}_{I \in \mathbb{N}}$, there exists $n \in \mathbb{N}$ such that $\sigma$ is not an equilibrium strategy in any economy $F^{(I)}$ for $I > n$.*

**Proof of Lemma A4.**

By Lemma A3, we only need to consider all $\sigma$ that rank $\mu^\infty(\theta_i)$ for $\theta_i$. Otherwise, the statement is satisfied already.

Let $C(K+1) = \xi$. By Proposition 3 of Azevedo and Leshno (2016), it must be that $P^{(I)} \to P^\sigma$ almost surely in the sequence $\{F^{(I)}, \sigma\}_{I \in \mathbb{N}}$, where $P^{(I)} = P(\mu_{(F^{(I)},\sigma)})$ and $P^\sigma = P(\mu_{(E,\sigma)})$. For $0 < \phi < 2\xi$, there must exit $n \in \mathbb{N}$ such that $i$ is matched with $\mu^\infty(\theta_i)$ with probability at least $1 - \phi$ in $F^{(I)}$ for all $I > n$.

Let $\mathrm{EU}(\sigma(\theta_i))$ be the expected utility when submitting $\sigma(\theta_i)$. We compare this strategy with any unilateral deviation $L_i$ that keeps ranking $\mu^\infty(\theta_i)$ but drops one of the other ranked schools in $\sigma(\theta_i)$.

Then $\mathrm{EU}(\sigma(\theta_i)) \leqslant (1-\phi)u_{i,\mu^\infty(\theta_i)} + \phi - \xi$, where the right side assumes that $i$ obtains the highest possible utility (equal to one) whenever not being matched with $\mu_\infty(\theta_i)$. Moreover, $\mathrm{EU}(L_i) \geqslant (1-\phi)u_{i,\mu^\infty(\theta_i)} + \xi$. The difference between the two actions is:

$$\mathrm{EU}(L_i) - \mathrm{EU}(\sigma(\theta_i)) \geqslant 2\xi - \phi > 0,$$

which proves that $\sigma$ is an equilibrium strategy in $F^{(I)}$ for $I > n$. ■

Moreover, when $C(2) > 0$, we can obtain even sharper results:

**Lemma A5.** *Suppose $C(2) > 0$ (i.e., it is costly to rank more than one school), and $\sigma(\theta_i) = (\mu^\infty(\theta_i))$ (i.e., only ranking the school prescribed by $\mu^\infty$) for all student types. In a sequence of random economies $\{F^{(I)}\}_{I \in \mathbb{N}}$, there exists $n \in \mathbb{N}$ such that $\sigma$ is a Bayesian Nash equilibrium in $F^{(I)}$ for all $I > n$.*

**Proof of Lemma A5.** This is implied by Lemmata A3 and A4 (when $K = 1$). ■

### A.2.5 Equilibrium and Stable Matching

In a finite economy with complete information, it is known that a matching in equilibrium can be unstable (Haeringer and Klijn, 2009). They further show that, in finite economies, DA with constraints implements stable matchings in Nash equilibria if and only if the student priority indices at all schools satisfy the so-called Ergin acyclicity condition (Ergin, 2002). We extend this result to the continuum economy and to a more general class of DA mechanisms where application costs, $C(|L|)$, are flexible.

**Definition A1.** *In a continuum economy, we fix a vector of capacities, $\{q_s\}_{s=1}^{S}$, and a distribution of priority indices, $H$. An Ergin cycle is constituted of distinct schools $(s_1, s_2)$ and subsets of students $\{\Theta_1, \Theta_2, \Theta_3\}$ (of equal measure $q_0 > 0$ ), whose elements are denoted by $\theta_1$, $\theta_2$, and $\theta_3$, respectively, and whose "identities" are $i_1$, $i_2$, and $i_3$, such that the following conditions are satisfied:*

*(i) Cycle condition: $e_{i_1,s_1} > e_{i_2,s_1} > e_{i_3,s_1}$, and $e_{i_3,s_2} > e_{i_1,s_2}$, for all $i_1$, $i_2$, and $i_3$.*

*(ii) Scarcity condition: there exist (possibly empty) disjoint sets of agents $\Theta_{s_1}$, $\Theta_{s_2} \in \Theta\backslash\{\Theta_1, \Theta_2, \Theta_3\}$ such that $e_{i,s_1} > e_{i_2,s_1}$ for all $\theta_i \in \Theta_{s_1}$, $|\Theta_{s_1}| = q_{s_1} - q_0$; $e_{i,s_2} > e_{i_1,s_2}$ for all $i \in \Theta_{s_1}$, and $|\Theta_{s_2}| = q_{s_2} - q_0$.*

*A priority index distribution $H$ is Ergin-acyclic if it allows no Ergin cycles.*

This acyclicity condition is satisfied if all schools rank students in the same way. With this, we extend Theorem 6.3 in Haeringer and Klijn (2009) to the continuum economy.

**Proposition A3.** *In the continuum economy $E$:*

*(i) If $C(2) = 0$, every (pure-strategy) Bayesian Nash equilibrium results in a stable matching if and only if the economy satisfies Ergin-acyclicity (Haeringer and Klijn, 2009).*

*(ii) If $C(2) > 0$, all (pure-strategy) Bayesian Nash equilibrium outcomes are stable.*

*Proof.* To prove parts (i) and (ii), we use the proof of Theorem 6.3 in Haeringer and Klijn (2009) and, therefore, that of Theorem 1 in Ergin (2002). They can be directly extended to the continuum economy under more general DA mechanisms. We notice the following:

(a) The continuum economy can be "discretized" such that each subset of students can be treated as a single student. When doing so, we do not impose restrictions on the sizes of the subsets, as long as they have a positive measure. This allows us to use the derivations in the aforementioned proofs.

(b) The flexibility in the cost function of ranking more schools does not impose additional restrictions. As we focus on equilibrium, for any strategy with more than one school listed, we can find a one-school list that has the same or higher payoff. Indeed, many steps in the aforementioned proofs involve such a trick.

■

## A.3 Consistency of the Preference Estimator under the Assumption of Asymptotic Stability

We provide a proof of the consistency of MLE under the assumption of asymptotic stability. The same proof can be extended to the corresponding GMM estimator.

Let us consider a sequence of random economies and strategies that satisfies Assumptions 1 and 2. The associated matchings and cutoffs are $\{\mu^{(I)}, P^{(I)}\}_{I\in\mathbb{N}}$. We further assume that $\lim_{I\to\infty} P^{(I)} = P^{\infty}$, almost surely, and that $\{\mu^{(I)}\}_{I\in\mathbb{N}}$ is asymptotically stable.[A.3]

In this section, we follow the notation in Newey and McFadden (1994) and define:

$$Q_0(\beta|P^{\infty}) = E_{\{Z,e\}}\left(\ln\left[\Pr\left(\mu^{\infty}(u_i, e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\, u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i, P^{\infty}); \beta\right)\right]\right),$$

where the expectation is taken over $Z_i, e_i$. Recall that both $\mu^{\infty}$ and $P^{\infty}$ are deterministic.

We also define the following regularity conditions.

**Assumption A2.** *Suppose that the data are i.i.d., that $\beta_0$ is the true parameter value, and that the sequence $\{\mu^{(I)}, P^{(I)}\}_{I\in\mathbb{N}}$ has $\lim_{I\to\infty} P^{(I)} = P^{\infty}$, almost surely, and $\{\mu^{(I)}\}_{I\in\mathbb{N}}$ being asymptotically stable. We impose the following regularity conditions:*

*(i) $Q_0(\beta|P^{\infty})$ is continuous in $\beta$ and uniquely maximized at $\beta_0$.*

*(ii) $\beta \in \mathcal{B}$, which is compact.*

*(iii) At any $\beta \in \mathcal{B}$ for almost all $Z_i$ and $e_i$,* $\Pr\left(\mu^{\infty}(u_i, e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\, u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i, P^{\infty}); \beta\right)$ *is bounded away from zero and continuous.*

[A.3] Recall that under the stability assumption, only students who have at least two feasible schools contribute to the estimation. If one has zero or one feasible school, her match (or her choice) does not reveal any information about her preferences. To simplify the notations below, we implicitly assume that a student's probability of being matched with the school prescribed by the match in a finite economy or the continuum economy is one, whenever she has no feasible school.

*(iv)* $E_{\{Z,e\}}\left(\sup_{\beta\in\mathcal{B}}\left|\ln\left[\Pr\left(\mu^{\infty}(u_i,e_i)=\underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}\mid Z_i,e_i,\mathcal{S}(e_i,P^{\infty});\beta\right)\right]\right|\right)<\infty.$

*(v)* $G(\Theta_\delta)E_{(Z,e)}\left(\sup_{\beta\in\mathcal{B},s\in\mathcal{S},P\in\mathcal{P}(e_i,s)}\left|\ln\frac{\Pr\left(s=\underset{s\in\mathcal{S}(e_i,P)}{\arg\max}\ u_{i,s}|Z_i,e_i,\mathcal{S}(e_i,P);\beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i)=\underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}|Z_i,e_i,\mathcal{S}(e_i,P^{\infty});\beta\right)}\right|\mid(u_i,e_i)\in\Theta_\delta\right)$ *converges to zero as* $\delta\to 0$*, where* $\Theta_\delta\equiv\{(u_i,e_i)\in\Theta: e_{i,s}\in(P_s^{\infty}-\delta,P_s^{\infty}+\delta),\exists s\in\mathcal{S}\}$ *for* $\delta>0$ *and* $\mathcal{P}(e_i,s)\equiv\{P\in[0,1]^S: s\in\mathcal{S}(e_i,P)\}$ *is the set of all possible cutoffs making* $s$ *feasible to* $i$*.*

Conditions (i) and (ii) are standard for identification of the model; conditions (iii) and (iv) are satisfied in common applications of discrete choice models, including logit and probit models with or without random coefficients.

Condition (v) extends condition (iv). Without loss in our setting, we assume $G$ admits a marginal density of $e_i$, and thus $G(\Theta_\delta)\to 0$ as $\delta\to 0$. Condition (v) is then satisfied if the conditional expectation in (v) is either bounded or grows to infinity at a slower rate than $1/G(\Theta_\delta)$ when $\delta\to 0$. This is satisfied in the aforementioned discrete choice models that have full-support utility shocks, as choice probabilities are bounded away from zero almost surely given that $\beta\in\mathcal{B}$.[A.4]

To proceed, we define $\widehat{Q}_I(\beta|P^{(I)})$ as the average of log-likelihood based on stability when the economy is of size $I$. That is,

$$\widehat{Q}_I(\beta|P^{(I)})=\frac{1}{I}\sum_{i=1}^{I}\sum_{s=1}^{S}\mathbb{1}(\mu^{(I)}(u_i,e_i)=s)\ln\left[\Pr\left(s=\underset{s'\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s'}\mid Z_i,e_i,\mathcal{S}(e_i,P^{(I)});\beta\right)\right].$$

$\widehat{Q}_I(\beta|P^{(I)})$ is possibly incorrectly specified because $\mu^{(I)}$ may not be exactly stable. That is, some students may not be matched with their favorite feasible school in $\mu^{(I)}$.

Correspondingly, we also define:

$$\widehat{Q}_I(\beta|P^{\infty})=\frac{1}{I}\sum_{i=1}^{I}\ln\left[\Pr\left(\mu^{\infty}(u_i,e_i)=\underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}\mid Z_i,e_i,\mathcal{S}(e_i,P^{\infty});\beta\right)\right].$$

In this definition, with the same economy as used in $\widehat{Q}_I(\beta|P^{(I)})$, we construct the hypothetical matching $\mu^{\infty}$ and cutoff $P^{\infty}$. Recall that both $\mu^{\infty}$ and $P^{\infty}$ are deterministic. $\widehat{Q}_I(\beta|P^{\infty})$ is then

[A.4]Equivalently, this requires that a choice probability is strictly positive for almost all $Z$. This is also true in the usual models with random coefficients; random coefficients often have full support on the real line and therefore lead to strictly positive choice probabilities.

the average of the log-likelihood function of this hypothetical dataset. $\widehat{Q}_I(\beta|P^\infty)$ is correctly specified because in matching $\mu^\infty$, every student is matched with her favorite feasible school given $P^\infty$.

The following lemma shows that the MLE estimator would be consistent if we could have access to the hypothetical dataset and use $\widehat{Q}_I(\beta|P^\infty)$.

**Lemma A6.** *When conditions (i)-(iv) in Assumption A2 are satisfied,*
*(i)* $\sup_{\beta\in\mathcal{B}} |\widehat{Q}_I(\beta|P^\infty) - Q_0(\beta|P^\infty)| \xrightarrow{p} 0$, *and*
*(ii)* $\tilde{\beta}_I$ *is consistent (i.e.,* $\tilde{\beta}_I \xrightarrow{p} \beta_0$*), where* $\tilde{\beta}_I = \arg\max_{\beta\in\mathcal{B}} \widehat{Q}_I(\beta|P^\infty)$.

*Proof.* Note that

$$\begin{aligned}&\left|\ln\left[\Pr\left(\mu^\infty(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^\infty)}{\arg\max}\, u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^\infty);\beta\right)\right]\right| \\ \leqslant &\sup_{\beta\in\mathcal{B}}\left|\ln\left[\Pr\left(\mu^\infty(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^\infty)}{\arg\max}\, u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^\infty);\beta\right)\right]\right|.\end{aligned}$$

Implied by condition (iv) of Assumption A2, the right-hand-side of the inequality has a finite first moment. Together with conditions (ii) and (iii), this implies that the conditions in Lemma 2.4 of Newey and McFadden (1994) are satisfied, leading to part (i) of the above lemma. By Theorem 2.1 of Newey and McFadden (1994), part (ii) is also satisfied. ∎

**Lemma A7.** *Given Assumption A2,* $\sup_{\beta\in\mathcal{B}} |\widehat{Q}_I(\beta|P^{(I)}) - \widehat{Q}_I(\beta|P^\infty)| \xrightarrow{p} 0$.

*Proof.* Lemma A3 shows that in sufficiently large economies, every student except a measure-zero set includes in her ROL the school prescribed by $\mu^\infty$. For each student, whenever $\mathcal{S}(e_i, P^{(I)}) = \mathcal{S}(e_i, P^\infty)$, $\mu^{(I)}(u_i,e_i) = \mu^\infty(u_i,e_i)$.

Therefore, for any $\beta \in \mathcal{B}$, in sufficiently large economies,

$$\begin{aligned}&\left|\widehat{Q}_I(\beta|P^{(I)}) - \widehat{Q}_I(\beta|P^\infty)\right| \\ =&\left|\frac{1}{I}\sum_{i:\mathcal{S}(e_i,P^{(I)})\neq\mathcal{S}(e_i,P^\infty)}\sum_{s=1}^{S}\mathbb{1}(\mu^{(I)}(u_i,e_i)=s)\ln\frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\, u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P^{(I)});\beta\right)}{\Pr\left(\mu^\infty(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^\infty)}{\arg\max}\, u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^\infty);\beta\right)}\right|\end{aligned}$$

$$\leqslant \frac{1}{I} \sum_{i:\mathcal{S}(e_i,P^{(I)})\neq\mathcal{S}(e_i,P^{\infty})} \sum_{s=1}^{S} \mathbb{1}(\mu^{(I)}(u_i,e_i)=s) \left| \ln \frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P^{(I)}); \beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^{\infty}); \beta\right)} \right|$$

$$\leqslant \frac{1}{I} \sum_{i:\mathcal{S}(e_i,P^{(I)})\neq\mathcal{S}(e_i,P^{\infty})} \sup_{\beta\in\mathcal{B},s\in\mathcal{S},P\in\mathcal{P}(e_i,s)} \left| \ln \frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P)}{\arg\max}\ u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P); \beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^{\infty}); \beta\right)} \right|$$

$$\leqslant \underbrace{\frac{1}{I} \sum_{(u_i,e_i)\in\Theta_\delta} \sup_{\beta\in\mathcal{B},s\in\mathcal{S},P\in\mathcal{P}(e_i,s)} \left| \ln \frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P)}{\arg\max}\ u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P); \beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^{\infty}); \beta\right)} \right|}_{(*)}$$

$$+ \underbrace{\frac{1}{I} \sum_{\substack{(u_i,e_i)\in\Theta\backslash\Theta_\delta: \\ \mathcal{S}(e_i,P^{(I)})\neq\mathcal{S}(e_i,P^{\infty})}} \sup_{\beta\in\mathcal{B},s\in\mathcal{S},P\in\mathcal{P}(e_i,s)} \left| \ln \frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P)}{\arg\max}\ u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P); \beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^{\infty}); \beta\right)} \right|}_{(**)} \tag{A.7}$$

Since the above results apply to all $\beta \in \mathcal{B}$, inequality (A.7) is also satisfied if we replace the first line with $\sup_{\beta\in\mathcal{B}} \left|\widehat{Q}_I(\beta|P^{(I)}) - \widehat{Q}_I(\beta|P^{\infty})\right|$.

By the law of large numbers, as $I \to \infty$, $(*)$ in inequality (A.7) converges almost surely to $G(\Theta_\delta)E_{(Z,e)}\left(\sup_{\substack{\beta\in\mathcal{B},s\in\mathcal{S}, \\ P\in\mathcal{P}(e_i,s)}} \left|\ln \frac{\Pr\left(s=\underset{s'\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s'}|Z_i,e_i,\mathcal{S}(e_i,P^{(I)});\beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i)=\underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s}|Z_i,e_i,\mathcal{S}(e_i,P^{\infty});\beta\right)}\right| \mid (u_i,e_i)\in\Theta_\delta\right)$, which, as a function of $\delta$, converges to zero if $\delta \to 0$ (condition v of Assumption A2). This implies that, for $\eta_1 > 0$, there exists $\delta_{\eta_1}$ such that for all $\delta < \delta_{\eta_1}$, we have

$$G(\Theta_\delta)E_{(Z,e)}\left(\sup_{\substack{\beta\in\mathcal{B},s\in\mathcal{S}, \\ P\in\mathcal{P}(e_i,s)}} \left|\ln \frac{\Pr\left(s = \underset{s'\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s'} \mid Z_i, e_i, \mathcal{S}(e_i,P^{(I)}); \beta\right)}{\Pr\left(\mu^{\infty}(u_i,e_i) = \underset{s\in\mathcal{S}(e_i,P^{\infty})}{\arg\max}\ u_{i,s} \mid Z_i, e_i, \mathcal{S}(e_i,P^{\infty}); \beta\right)}\right| \mid (u_i,e_i)\in\Theta_\delta\right) < \eta_1/2.$$

By inequality (A.7) and the law of large numbers, we can choose $n_1$ such that for any $I > n_1$, $\sup_{\beta\in\mathcal{B}} \left|\widehat{Q}_I(\beta|P^{(I)}) - \widehat{Q}_I(\beta|P^{\infty})\right| < \eta_1/2 + \eta_1/2 + (**)|_{\delta<\delta_{\eta_1}} = \eta_1 + (**)|_{\delta<\delta_{\eta_1}}$ almost surely, where $(**)|_{\delta<\delta_{\eta_1}}$ indicates the last term in inequality (A.7) evaluated at $\delta < \delta_{\eta_1}$.

Moreover, by assumption, $\lim_{I\to\infty} P^{(I)} = P^{\infty}$, almost surely. Given $\delta < \delta_{\eta_1}$ and for any $\eta_2$, there exists $n_2$ such that for $I > n_2$, $\Pr\left(\left\|P^{(I)} - P^{\infty}\right\| \leqslant \delta\right) > 1 - \eta_2$. When this happens (i.e., $\|P^{(I)} - P^{\infty}\| \leqslant \delta$), ($**$) in inequality (A.7) evaluated at $\delta$ is zero because $\{(u_i, e_i) \in \Theta \backslash \Theta_\delta : \mathcal{S}(e_i, P^{(I)}) \neq \mathcal{S}(e_i, P^{\infty})\}$ is empty.

Therefore, for any $\eta_1$ and $\eta_2$, there exist $n_1$ and $n_2$ such that whenever $I > \max\{n_1, n_2\}$, $\Pr\left(\sup_{\beta \in \mathcal{B}} \left|\widehat{Q}_I(\beta|P^{(I)}) - \widehat{Q}_I(\beta|P^{\infty})\right| > \eta_1\right) < \eta_2$, which proves the lemma. ■

**Proposition A4.** *When Assumption A2 is satisfied, $\widehat{\beta}_I$ is consistent (i.e., $\widehat{\beta}_I \xrightarrow{p} \beta_0$), where $\widehat{\beta}_I = \arg\max_{\beta \in \mathcal{B}} \widehat{Q}_I(\beta|P^{(I)})$.*

*Proof.* Assumption A2, Lemma A6 (part i), and Lemma A7 imply that $\widehat{Q}_I(\beta|P^{(I)})$ satisfies the conditions in Theorem 2.1 of Newey and McFadden (1994). Hence, $\widehat{\beta}_I \xrightarrow{p} \beta_0$. ■

## A.4 Estimation with Strict Truth-Telling and Outside Option

The following discussion supplements Section II.B in which we present how weak truth-telling (WTT) can be applied to data on school choice and college admissions and what assumptions it entails. However, assuming the length of submitted ROL is exogenous (Assumption WTT2) may seem restrictive. An alternative way to relax this assumption is to introduce an outside option and to make some school unacceptable to some students.

Suppose that $i$'s utility for her outside option is denoted by $u_{i,0} = V_{i,0} + \epsilon_{i,0}$, where $\epsilon_{i,0}$ is a type I extreme value. We then augment the type space of each student with the outside option and let $\sigma^S : \mathbb{R}^{(S+1)} \times [0,1]^S \to \mathcal{L}$ be an STT pure strategy defined on the augmented preference space. More precisely, one version of the STT assumption contains the following two assumptions:

**Assumption (Strict Truth-Telling with Outside Option).**
*STT1. $\sigma^S(u_i, u_{i,0}, e_i)$ ranks all $i$'s acceptable schools according to her true preferences.*
*STT2. Students do not rank unacceptable schools: $u_{i,0} > u_{i,s'}$ for all $s'$ not ranked by $\sigma^S(u_i, u_{i,0}, e_i)$.*

Given these two assumptions, similar to the case with WTT, either MLE or GMM can be applied based on the following choice probabilities:[A.5]

$$
\begin{aligned}
&\Pr\left(\sigma^S(u_i, u_{i,0}, e_i) = L \;\middle|\; Z_i; \beta\right) \\
=&\Pr\left(u_{i,l^1} > \cdots > u_{i,l^K} > u_{i,0} > u_{i,s'}\ \forall\, s' \in \mathcal{S}\backslash L \;\middle|\; Z_i; \beta\right) \\
=&\frac{\exp(V_{i,0})}{\exp(V_{i,0}) + \sum_{s'\notin L} \exp(V_{i,s'})} \prod_{s\in L} \left( \frac{\exp(V_{i,s})}{\exp(V_{i,0}) + \sum_{s' \nsucc_L s} \exp(V_{i,s'})} \right).
\end{aligned}
$$

Recall that $s' \nsucc_L s$ indicates that $s'$ is not ranked before $s$ in $L$, which includes $s$ itself and the schools not ranked in $L$.

Assumptions STT1 and STT2 can be justified as an equilibrium outcome when there is no application cost. However, there may be an issue of multiple equilibria created by unacceptable schools. Namely, if a student can always decline to enroll at an unacceptable school, she may not mind including or excluding that school in her ROL and being assigned to it (He, 2015).

## A.5 Assumption EXO2 for the Stability-Based Estimator

The necessity of Assumption EXO2 can be seen in a following modified utility function:

$$\overline{u}_{i,s} \equiv u_{i,s} - \infty \times \mathbb{1}(e_{i,s} < P_s(\mu)) = V(Z_{i,s}, \beta) - \infty \times \mathbb{1}(e_{i,s} < P_s(\mu)) + \epsilon_{i,s},$$

where $-\infty \times \mathbb{1}(\{e_{i,s} < P_s(\mu))$ is zero for feasible schools but equal to $-\infty$ for infeasible ones, thus making them always less desirable. With $\mathbb{1}(e_{i,s} < P_s(\mu))$ being personalized "prices," a stable matching is then equivalent to discrete choice based on the modified utility functions (He et al., 2015, 2018). That is, a realized matching $\hat{\mu}$ is stable if and only if $\hat{\mu}(\theta_i) = \hat{\mu}(u_i, e_i) = \arg\max_{s\in\mathcal{S}} \overline{u}_{i,s}$. Although utility shocks can depend on $Z_i$ for identification in usual discrete choice models (Matzkin, 1993), $\mathbb{1}(e_{i,s} < P_s(\mu))$ is special. Conditional on $\mathbb{1}(e_{i,s} < P_s(\mu))$ and $Z_i$, $i$'s ordinal preferences and thus $i$'s choice may lack variation without Assumption EXO2. This is shown in the following example.

[A.5] For an example imposing the STT assumption, see He and Magnac (2016) in which the authors observe students ranking all available options and have information on the acceptability of each option.

**An example with Assumption EXO2 violated.** Let $I = 3$ and $S = 3$, each with one seat. Students have the same preferences, $(u_{i,1}, u_{i,2}, u_{i,3}) = (0.9, 0.6, 0.3)$ for $i \in \{1, 2, 3\}$; the priority index vectors $(e_{i,1}, e_{i,2}, e_{i,3})$ are $(0.8, 0.5, 0.8)$ for $i = 1$, $(0.5, 0.8, 0.3)$ for $i = 2$, and $(0.3, 0.3, 0.5)$ for $i = 3$.

Suppose students are strictly truth-telling. Therefore, the matching is stable. The cutoffs are $P = (0.8, 0.8, 0.5)$, which leads to $\mathcal{S}(e_1, P) = \{1, 3\}$. However, if $(u_{i,1}, u_{i,2}, u_{i,3}) = (0.6, 0.9, 0.3)$ for $i = 1$, then $P' = (0.5, 0.5, 0.5)$ and $\mathcal{S}(e_1, P') = \{1, 2, 3\}$. Therefore, for $i = 1$, $\mathcal{S}(e_i, P(\mu)) \not\perp \epsilon_i | Z_i$. If the data generating process is as such, conditional on student 1's set of feasible schools, we never observe $u_{1,1} > u_{1,2}$, or school 1 being chosen over school 2 when both are feasible.

**An example with Assumption EXO2 satisfied.** Let us consider the following example with Ergin cyclicity where each school has one seat.

| School priority ranking (high to low) | Student ordinal preferences (more to less preferred) |
|---|---|
| $s_1$: $i_1$, $i_3$, $i_2$ | $i_1$: $s_2$, $s_1$, $s_3$ |
| $s_2$: $i_2$, $i_1$, $i_3$ | $i_2$: $s_1$, $s_2$, $s_3$ |
| $s_3$: $i_2$, $i_1$, $i_3$ | $i_3$: $s_1$, $s_2$, $s_3$ |

When everyone is strictly truth-telling, $i_1$'s set feasible schools $\mathcal{S}(e_{i_1}, P) = \{s_1, s_3\}$. No matter how $i_1$'s ordinal preferences change, $i_1$'s feasible schools do not change, as long as the matching is stable.[A.6] Therefore, given others' preferences, Assumption EXO2 is satisfied for $i_1$.

[A.6] In any stable matching, $i_2$ is assigned to $s_2$; otherwise, either $i_3$ or $i_2$ would have justified envy. Therefore, $s_2$ is not feasible to $i_1$. Both $s_1$ and $s_3$ are feasible to $i_1$ in any stable matching, because $i_1$ has a higher priority at both schools than $i_3$, while $i_2$'s assignment is fixed at $s_2$.

# Appendix B Data

## B.1 Data Sources

For the empirical analysis, we use three administrative data sets on Parisian students, which are linked using an encrypted version of the French national student identifer (*Identifiant National Élève*).

(i) **Application Data**: The first data set was provided to us by the Paris Education Authority (*Rectorat de Paris*) and contains all the information necessary to replicate the assignment of students to public academic-track high schools in the city of Paris for the 2013-2014 academic year. This includes the schools' capacities, the students' ROLs of schools, and their priority indices at every school. Moreover, it contains information on students' socio-demographic characteristics (age, gender, parents' SES, low-income status, etc.), and their home addresses, allowing us to compute distances to each school in the district.

(ii) **Enrollment Data**: The second data set is a comprehensive register of students enrolled in Paris' middle and high schools during the 2012–2013 and 2013–2014 academic years (*Base Elèves Académique*), which is also from the Paris Education Authority. This data set allows to track students' enrollment status in all Parisian public and private middle and high schools.

(iii) **Exam Data**: The third data set contains all Parisian middle school students' individual examination results for a national diploma, the *Diplôme national du brevet* (DNB), which students take at the end of middle school. We obtained this data set from the statistical office of the French Ministry of Education (*Direction de l'Évaluation, de la Prospective et de la Performance du Ministère de l'Éducation Nationale*).

## B.2 Definition of Variables

**Priority Indices.** Students' priority indices at every school are recorded as the sum of three main components: (i) students receive a "district" bonus of 600 points on each of the schools in their list which are located in their home district; (ii) students' academic performance during the last year of middle school is graded on a scale of 400 to 600 points;

(iii) students from low-income families are awarded an additional bonus of 300 points. We convert these priority indices into percentiles between 0 and 1.

**Student Scores.** Based on the DNB exam data set, we compute several measures of student academic performance, which are normalized as percentiles between 0 and 1 among all Parisian students who took the exam in the same year. Both French and math scores are used, and we also construct the students' composite score, which is the average of the French and math scores. Note that students' DNB scores are different from the academic performance measure used to calculate student priority indices as an input into the DA mechanism. Recall that the latter is based on the grades obtained by students throughout their final year of middle school.

**Socio-Economic Status.** Students' socio-economic status is based on their parents' occupation. We use the French Ministry of Education's official classification of occupations to define "high SES": if the occupation of the student's legal guardian (usually one of the parents) belongs to the "very high SES" category (company managers, executives, liberal professions, engineers, academic and art professions), the student is coded as high SES, otherwise she is coded as low SES.[A.7]

## B.3 Construction of the Main Data Set for Analyses

In our empirical analysis, we use data from the Southern District of Paris (*District Sud*). We focus on public middle school students who are allowed to continue their studies in the academic track of upper secondary education and whose official residence is in the Southern District. We exclude those with disabilities, those who are repeating the first year of high school, and those who were admitted to specific selective tracks offered by certain public high schools in Paris (e.g., music majors, bilingual courses, etc.), as these students are given absolute priority in the assignment over other students. This leads to the exclusion of 350 students, or 18 percent of the total, the majority of whom are grade repeaters. Our data thus include 1,590 students from 57 different public middle schools, with 96 percent of students coming from one of the district's 24 middle schools.

[A.7] There are four official categories: low SES, medium SES, high SES, and very high SES.

# Appendix C Monte Carlo Simulations

This appendix provides details on the Monte Carlo simulations that we perform to assess our empirical approaches and model selection tests. Section C.1 specifies the model, Section C.2 describes the data generating processes, Section C.3 reports a number of summary statistics for the simulated data, and Section C.4 discusses the main results.

## C.1 Model Specification

**Economy Size.** We consider an economy where $I = 500$ students compete for admission to $S = 6$ schools. The vector of school capacities is specified as follows:

$$I \cdot \{q_s\}_{s=1}^{6} = \{50, 50, 25, 50, 150, 150\}.$$

Setting the total capacity of schools (475 seats) to be strictly smaller than the number of students (500) simplifies the analysis by ensuring that each school has a strictly positive cutoff in equilibrium.

**Spatial Configuration.** The school district is stylized as a disc of radius 1 (Figure C1). The schools (represented by red circles) are evenly located on a circle of radius 1/2 around the district centroid; the students (represented by blue circles) are uniformly distributed across the district area. The cartesian distance between student $i$ and school $s$ is denoted by $d_{i,s}$.

**Student Preferences.** To represent student preferences over schools, we adopt a parsimonious version of the random utility model described in Section II.A. Student $i$'s utility from attending school $s$ is specified as follows:

$$u_{i,s} = 10 + \alpha_s - d_{i,s} + \gamma(a_i \cdot \bar{a}_s) + \epsilon_{i,s}, \ s = 1, \ldots, 6; \tag{A.8}$$

where $10 + \alpha_s$ is school $s$'s fixed effects; $d_{i,s}$ is the walking distance from student $i$'s residence to school $s$; $a_i$ is student $i$'s ability; $\bar{a}_s$ is school $s$'s quality; and $\epsilon_{i,s}$ is an error term that is drawn from a type-I extreme value distribution. Setting the effect of distance to $-1$ ensures that other coefficients can be interpreted in terms of willingness to travel.

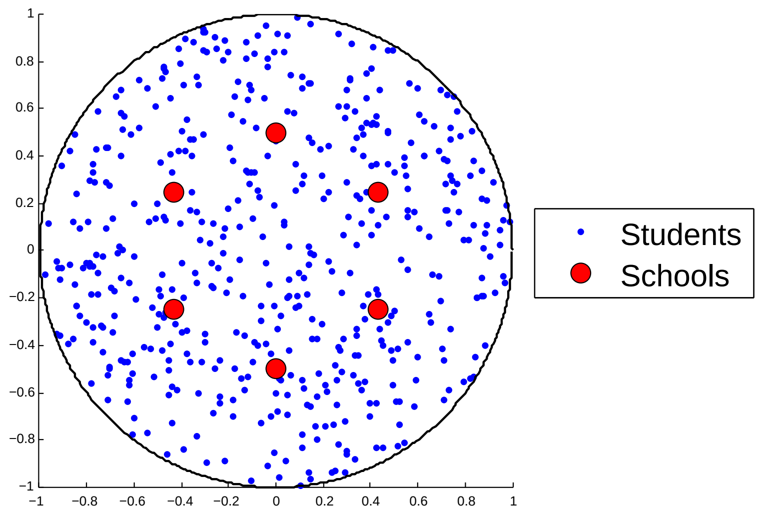


**Figure C1:** Monte Carlo Simulations: Spatial Distribution of Students and Schools

*Notes:* This figure shows the spatial configuration of the school district considered in one of the Monte Carlo samples, for the case with 500 students and 6 schools. The school district is represented as a disc of radius 1. The small blue and large red circles show the location of students and of schools, respectively.

The school fixed effects above the common factor, 10, are specified as follows:

$$\{\alpha_s\}_{s=1}^{6} = \{0, 0.5, 1.0, 1.5, 2.0, 2.5\}$$

Adding the common value of 10 for every school ensures that all schools are acceptable in the simulated samples.

Students' abilities $\{a_i\}_{i=1}^{I}$ are randomly drawn from a uniform distribution on the interval $[0, 1]$. School qualities $\{\bar{a}_s\}_{s=1}^{S}$ are exogenous to students' idiosyncratic preferences $\epsilon_{i,s}$. The procedure followed to ascribe values to the schools' qualities is discussed at the end of this section.

The positive coefficient $\gamma$ on the interaction term $a_i \cdot \bar{a}_s$ reflects the assumption that high-ability students value school quality more than low-ability students. In the simulations, we set $\gamma = 3$.

**Priority Indices.** Students are ranked separately by each school based on a school-specific index $e_{i,s}$. The vector of student priority indices at a given school $s$, $\{e_{i,s}\}_{i=1}^{I}$ is constructed as correlated random draws with marginal uniform distributions on the interval [0,1], such that: (i) student $i$'s index at each school is correlated with her ability $a_i$ with a correlation coefficient of $\rho$; (ii) $i$'s indices at any two schools $s_1$ and $s_2$ are also correlated with correlation coefficient $\rho$. When $\rho$ is set equal to 1, a student has the same priority at all schools. When $\rho$ is set equal to zero, her priority indices at the different schools are uncorrelated. For the simulations presented in this appendix, we choose $\rho = 0.7$. It is assumed that student know their priority indices but not their priority ranking at each school.

**School Quality.** To ensure that school qualities $\{\bar{a}_s\}_{s=1}^{S}$ are exogenous to students' idiosyncratic preferences, while being close to those observed in Bayesian Nash equilibrium of the school choice game, we adopt the following procedure: we consider the unconstrained student-proposing DA where students rank all schools truthfully; students' preferences are constructed using random draws of errors and a common prior about the average quality of each school; students rank schools truthfully and are assigned through the DA mechanism; each school's quality is computed as the average ability of students assigned to that school; a fixed-point vector of school qualities, denoted by $\{\bar{a}_s^*\}_{s=1}^{S}$, is found; the value of each school's quality is set equal to mean value of $\bar{a}_s^*$ across the samples.

The resulting vector of school qualities is:

$$\{\bar{a}_s\}_{s=1}^{6} = \{0.28, 0.39, 0.68, 0.65, 0.47, 0.61\}$$

## C.2 Data Generating Processes

The simulated data are constructed under two distinct data generating processes (DGPs).

**DGP 1: Constrained/Truncated DA.** This DGP considers a situation where the student-proposing DA is used to assign students to schools but where the number of schools that students are allowed to rank, $K$, is strictly smaller than the total number of available schools, $S$. For expositional simplicity, students are assumed to incur no cost when ranking exactly $K$ schools. Hence:

$$C(|L|) = \begin{cases} 0 & \text{if } |L| \leqslant K \\ +\infty & \text{if } |L| > K \end{cases}$$

In the simulations, we set $K = 4$ (students are allowed to rank up to 4 schools out of 6).

**DGP 2: Unconstrained DA with Cost.** This DGP considers the case where students are not formally constrained in the number of schools they can rank but nevertheless incur a constant marginal cost, denoted by $c(>0)$, each time they increase the length of their ROL by one, if this list contains more than one school. Hence:

$$C(|L|) = c \cdot (|L| - 1),$$

where the marginal cost $c$ is strictly positive. In the simulations, we set $c = 10^{-6}$.

For each DGP, we adopt a two-stage procedure to solve for a Bayesian Nash equilibrium of the school choice game.

**Stage 1: Distribution of Cutoffs Under Unconstrained DA.** Students' "initial" beliefs about the distribution of school cutoffs are based on the distribution of cutoffs that arises when students submit unrestricted truthful rankings of schools under the standard DA. Specifically:

(i) For $m = 1, \cdots, M$, we independently generate sample $m$ by drawing students' geographic coordinates, ability $a_i^{(m)}$, school-specific priority indices $e_{i,s}^{(m)}$, and idiosyncratic preferences $\epsilon_{i,s}^{(m)}$ over the $S$ schools for all $I$ students. We then calculate $u_{i,s}^{(m)}$ for all $i = 1, \cdots, I$, $s = 1, \cdots, S$, and $m = 1, \cdots, M$.

(ii) Student $i$ in sample $m$ submits a complete and truthful ranking $r(u_i^{(m)})$ of the schools; i.e., $i$ is strictly truth-telling.

(iii) After collecting $\{r(u_i^{(m)})\}_{i=1}^I$, the DA mechanism assigns students to schools taking into account their priority indices in sample $m$.

(iv) Each matching $\mu^{(m)}$ in sample $m$ determines a vector of school cutoffs $P^{(m)} \equiv \{P_s^{(m)}\}_{s=1}^S$.

(v) The cutoffs $\{P^{(m)}\}_{m=1}^M$ are used to derive the empirical distribution of school cutoffs under the unconstrained DA, which is denoted by $\hat{F}^0(\cdot \mid \{P^{(m)}\}_{m=1}^M)$.

In the simulations, we set $M = 500$.

**Stage 2: Bayesian Nash Equilibrium.** For each DGP, the $M$ Monte Carlo samples generated in Stage 1 are used to solve the Bayesian Nash equilibrium of the school choice game. Specifically:

(i) Each student $i$ in each sample $m$ determines all possible true partial preference orders $\{L_{i,n}^{(m)}\}_{n=1}^{N}$ over the schools, i.e., all potential ROLs of length between 1 and $K$ that respect $i$'s true preference ordering $R_{i,m}$ of schools among those ranked in $L_{i,n}^{(m)}$; for each student, there are $N = \sum_{k=1}^{K} S!/[k!(S-k)!]$ such partial orders. Under the constrained/truncated DA (DGP 1), students consider only true partial preference orderings of length $K$ ($< S$), i.e., 15 candidate ROLs when they rank exactly 4 schools out of 6;[A.8] under the unconstrained DA with cost (DGP 2), students consider all true partial orders of length up to $S$, i.e., 63 candidate ROLs when they can rank up to 6 schools.

(ii) For each candidate ROL $L_{i,n}^{(m)}$, student $i$ estimates the (unconditional) probabilities of being admitted to each school by comparing her indices $e_{i,s}$ to the distribution of cutoffs. Initial beliefs on the cutoff distribution are based on $\hat{F}^0(\cdot \mid \{P^{(m)}\}_{m=1}^{M})$, i.e., the empirical distribution of cutoffs under unconstrained DA with strictly truth-telling students.

(iii) Each student selects the ROL $L_i^{(m)*}$ that maximizes her expected utility, where the utilities of each school are weighted by the probabilities of admission according to her beliefs.

(iv) After collecting $\{L_i^{(m)*}\}_{i=1}^{I}$, the DA mechanism is run in sample $m$.

(v) The matchings across the $M$ samples jointly determine the "posterior" empirical distribution of school cutoffs, $\hat{F}^1(\cdot \mid \cdot)$.

(vi) Students use $\hat{F}^1(\cdot \mid \cdot)$ as their beliefs, and steps (ii) to (v) are repeated until a fixed point is found, which occurs when the posterior distribution of cutoffs ($\hat{F}^t(\cdot \mid \cdot)$) is consistent with students' beliefs $\hat{F}^{t-1}(\cdot \mid \cdot)$. The equilibrium beliefs are denoted by $\hat{F}^*(\cdot \mid \cdot)$.

The simulated school choice data are then constructed based on a new set of $M$ Monte Carlo samples, which are distinct from the samples used to find the equilibrium distribution of cutoffs. In each of these new Monte Carlo samples, submitted ROLs are students' best

[A.8] This is without loss of generality, because in equilibrium the admission probability is non-degenerate and it is, therefore, in students' best interest to rank exactly 4 schools.

response to the equilibrium distribution of cutoffs ($\hat{F}^*(\cdot \mid \cdot)$). The school choice data consist of students' priority indices, their submitted ROLs, the student-school matching, and the realized cutoffs in each sample.

## C.3 Summary Statistics of Simulated Data

We now present some descriptive analysis on the equilibrium cutoff distributions and the 500 Monte Carlo samples of school choice data that are simulated for each DGP.

**Equilibrium Distribution of Cutoffs.** The equilibrium distribution of school cutoffs is displayed in Figure C2 separately for each DGP. In line with the theoretical predictions (Proposition A2), the marginal distribution of cutoffs is approximately normal. Because both DGPs involve the same profiles of preferences and produce almost identical matchings, the empirical distribution of cutoffs under the constrained/truncated DA (left panel) is very similar to that observed under the unconstrained DA with cost (right panel).

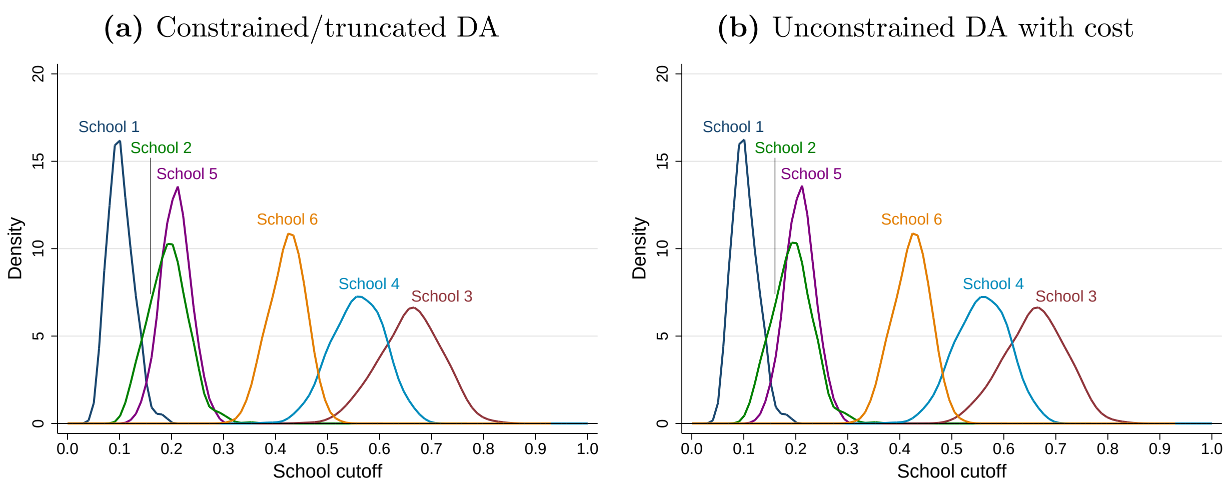


**Figure C2:** Monte Carlo Simulations: Equilibrium Distribution of School Cutoffs (6 schools, 500 students)

*Notes:* This figure shows the equilibrium marginal distribution of school cutoffs under the constrained/truncated DA (left panel) and the DA with cost (right panel) in a setting where 500 students compete for admission to 6 schools. With 500 simulated samples, the line fits are from a Gaussian kernel with optimal bandwidth using MATLAB's ksdensity command.

School cutoffs are not strictly aligned with the school fixed effects, since cutoffs are also influenced by school size. In the simulations, small schools (e.g., Schools 3 and 4) tend to have higher cutoffs than larger schools (e.g., Schools 5 and 6) because, in spite of being less popular, they can be matched only with a small number of students, which pushes their cutoffs upward.[A.9]

Figure C3 reports the marginal distribution of cutoffs in the constrained/truncated DA for various economy sizes. The simulations show that as the number of seats and the number of students increase proportionally while holding the number of schools constant, the distribution of school cutoffs degenerates and becomes closer to a normal distribution.

**Summary Statistics.** Table C1 shows some descriptive statistics of the simulated data from both DGPs. The reported means are averaged over the 500 Monte Carlo samples.

All students under the constrained/truncated DA submit ROLs of the maximum allowed length (4 schools). Under the unconstrained DA with cost, students are allowed to rank as many schools as they wish but, due to the cost of submitting longer lists, they rank 4.6 schools on average.

Under both DGPs, all school seats are filled, and, therefore, 95 percent of students are assigned to a school. Weak truth-telling is violated under the constrained/truncated DA, since less than half of submitted ROLs rank truthfully students' most-preferred schools. Although less widespread, violations of WTT are still observed under the unconstrained DA with cost, since about 20 percent of students do not truthfully rank their most-preferred schools. By contrast, almost every student is matched with her favorite feasible school under both DGPs.

**Comparative Statics.** To explore how the cost of ranking more schools affects weak truth-telling and ex post stability in equilibrium, we simulated data for DGP 2 (DA with cost) using different values of the cost parameter $c$, while keeping the other parameters at their baseline values.[A.10]

[A.9] Note that this phenomenon is also observed if one sets $\gamma = 0$, i.e., when students' preferences over schools do not depend on the interaction term $a_i \cdot \bar{a}_s$.

[A.10] We performed a similar exercise for DGP 1 (constrained/truncated DA) by varying the number of schools that students are allowed to rank. The results (available upon request) yield conclusions similar to those based on DGP 2 (DA with cost).

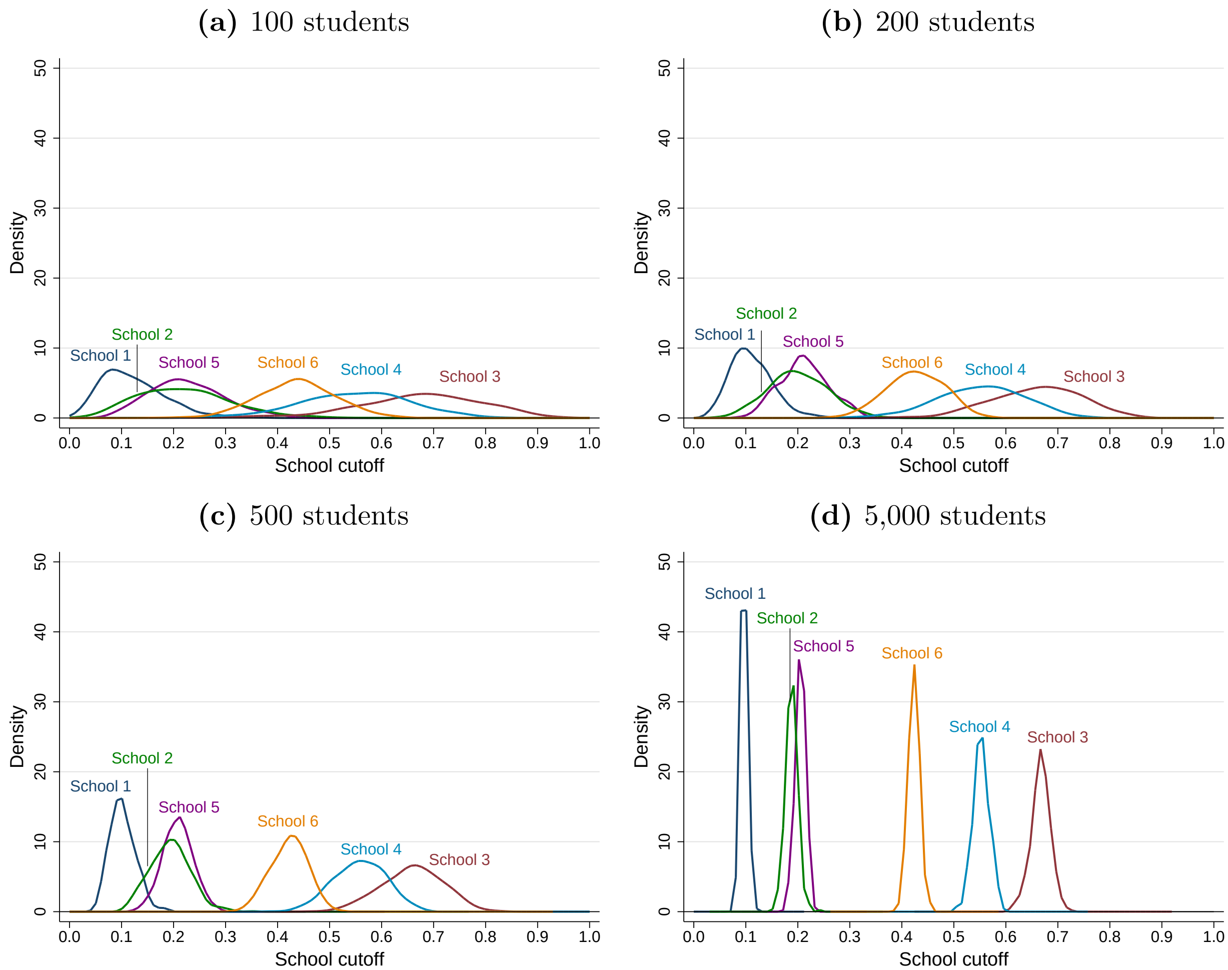


**Figure C3:** Monte Carlo Simulations: Impact of Economy Size on the Equilibrium Distribution of Cutoffs (Constrained/Truncated DA)

*Notes:* This figure shows the equilibrium marginal distribution of school cutoffs under the constrained/truncated DA (ranking 4 out of 6 schools) when varying the number of students, $I$, who compete for admission into 6 schools with a total enrollment capacity of $I \times 0.95$ seats. Using 500 simulated samples, the line fits are from a Gaussian kernel with optimal bandwidth using MATLAB's ksdensity command.

For each value of the cost parameter, we simulated 500 samples of school choice data and computed the following statistics by averaging across samples: (i) average length of submitted ROLs; (ii) average fraction of weakly truth-telling students; and (iii) average fraction of students assigned to their favorite feasible school.

The results of this comparative statics exercise are displayed in Figure C4. They confirm that, in our simulations, stability is a weaker assumption than WTT whenever students face a cost of ranking more schools: the share of students assigned to their favorite feasible school (blue line) is always larger than the share of WTT students (red line). Consistent with the

**Table C1:** Monte Carlo Simulations: Summary Statistics

| | Data generating process | |
|---|---|---|
| | Constrained/truncated DA<br>(1) | Unconstrained DA with cost<br>(2) |
| | *Panel A. Outcomes* | |
| Average length of submitted ROLs | 4.00 | 4.60 |
| | (0.000) | (0.054) |
| Assigned to a school | 0.950 | 0.950 |
| | (0.000) | (0.000) |
| Weakly truth-telling | 0.391 | 0.792 |
| | (0.022) | (0.018) |
| Assigned to favorite feasible school | 1.000 | 1.000 |
| | (0.001) | (0.000) |
| | *Panel B. Parameters* | |
| Number of students | 500 | 500 |
| Number of schools | 6 | 6 |
| Number of simulated samples | 500 | 500 |
| Maximum possible length of ROL | 4 | 6 |
| Marginal application cost ($c$) | 0 | $10^{-6}$ |

*Notes:* This table presents summary statistics of simulated data under two DGPs: (i) constrained/truncated DA (column 1): students are only allowed to rank 4 schools out of 6; and (ii) unconstrained DA with cost (column 2): students can rank as many schools as they like, but incur a constant marginal cost of $c = 10^{-6}$ per extra school included in their ROL beyond the first choice. Standard deviations across the 500 simulation samples are in parentheses.

predictions from Section I.D.2, the fraction of students who are matched with their favorite feasible school decreases with the marginal cost of ranking more schools (parameter $c$). In our simulations, violations of this assumption are very rare, except in the extreme case where students face a large marginal application cost $c$ equal to 1 (in which case students rank only 1.3 school on average).

## C.4 Results

**Estimation and Testing.** With the simulated data at hand, student preferences described by Equation (A.8) are estimated under different sets of identifying assumptions: (i) weak truth-telling; (ii) stability; and (iii) stability and undominated strategies. Estimates under assumption (i) are based on a rank-ordered logit model using maximum likelihood. Estimates under assumption (ii) are obtained from a conditional logit model where each student's choice set is restricted to the ex post feasible schools and where the matched school is the chosen

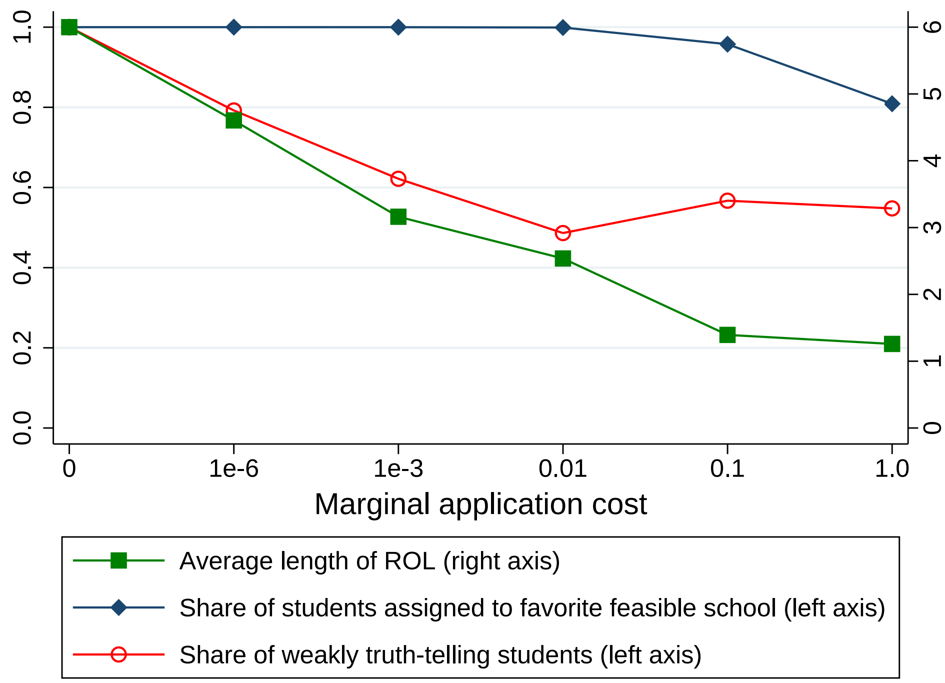


**Figure C4:** Monte Carlo Simulations: Impact of the Marginal Cost of Applying to Schools on Equilibrium Outcomes (500 Students, 6 Schools)

*Notes:* This figure presents summary statistics of simulated data under unconstrained DA with cost (DGP 2), in which students can rank as many schools as they like, but incur a constant marginal cost $c$ per extra school included in their ROL beyond the first. The data are simulated using different values of the marginal cost parameter $c$, while maintaining the other parameters at their baseline values. For each value of the cost parameter, 500 samples of school choice data are simulated. The following statistics are computed by averaging across samples: (i) average length of submitted ROLs; (ii) average fraction of weakly truth-telling students; (iii) average fraction of students matched with favorite feasible school.

alternative.[A.11] Finally, estimates under assumption (iii) are based on Andrews and Shi (2013)'s method of moment (in)equalities, using the approach proposed by Bugni et al. (2017) to construct the marginal confidence intervals for the point estimates.[A.12]

The results from 500 Monte Carlo samples are reported and discussed in the main text (Table 2). They are consistent with the theoretical predictions for both the constrained/truncated DA (Panel A) and the unconstrained DA with cost (Panel B).

[A.11]Our stability-based estimator is obtained using maximum likelihood. It can be equivalently obtained using a GMM estimation with moment equalities defined by the first-order conditions of the log-likelihood function.

[A.12]The conditional moment inequalities are derived from students' observed orderings of all 15 possible pairs of schools (see Section II.E). The variables that are used to interact with these conditional moment inequalities and thus to obtain the unconditional ones are student ability ($a_i$), distance to School 1 ($d_{i,1}$) and distance to School 2 ($d_{i,2}$), which brings the total number of moment inequalities to 120.

**Efficiency Loss from Stability-Based Estimates.** The efficiency loss from estimating the model under stability is further explored by comparing the truth-telling-based and stability-based estimates in a setting where students are strictly truth-telling. To that end, we generate a new set of 500 Monte Carlo samples using the unconstrained DA DGP, after setting the marginal application cost $c$ to zero. In this setting, all students submit truthful ROLs that rank all 6 schools. The estimation results, which are reported in Table C2, show that while both truth-telling-based and stability-based estimates are close to the true parameters values, the latter are much more imprecisely estimated than the former (column 6 vs. column 3): the stability-based estimates have standard deviations 2.5 to 3.8 times larger than the TT-based estimates. Note, however, that the efficiency loss induced by the stability assumption is considerably reduced when combing stability and undominated strategies (column 9 vs. column 3): the standard deviations of estimates based on the moment (in)equality approach are only 1.3 to 1.9 larger than their truth-telling counterparts.

Reassuringly, the Hausman test rejects truth-telling against stability in exactly 5 percent of samples, which is the intended type-I error rate. This test can therefore serve as a useful tool to select the efficient truth-telling-based estimates over the less efficient stability-based estimates when both assumptions are satisfied.

**Table C2:** Monte Carlo Results: Unconstrained DA (500 Students, 6 Schools, 500 Samples)

| | | Identifying assumptions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Weak Truth-telling | | | Stability of the matching | | | Stability and undominated strategies | | |
| | True value | Mean | SD | CP | Mean | SD | CP | Mean | SD | CP |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Parameters | | | | | | | | | | |
| School 2 | 0.50 | 0.50 | 0.10 | 0.94 | 0.51 | 0.29 | 0.94 | 0.50 | 0.12 | 1.00 |
| School 3 | 1.00 | 1.01 | 0.16 | 0.95 | 1.05 | 0.58 | 0.96 | 1.00 | 0.22 | 1.00 |
| School 4 | 1.50 | 1.52 | 0.15 | 0.95 | 1.54 | 0.52 | 0.96 | 1.51 | 0.21 | 1.00 |
| School 5 | 2.00 | 2.02 | 0.11 | 0.95 | 2.02 | 0.30 | 0.96 | 2.02 | 0.14 | 1.00 |
| School 6 | 2.50 | 2.52 | 0.14 | 0.96 | 2.54 | 0.45 | 0.96 | 2.53 | 0.19 | 1.00 |
| Own ability × school quality | 3.00 | 2.98 | 0.66 | 0.95 | 2.96 | 2.29 | 0.96 | 3.08 | 0.99 | 1.00 |
| Distance | −1.00 | −1.00 | 0.08 | 0.96 | −1.01 | 0.20 | 0.95 | −1.02 | 0.16 | 1.00 |

| Summary statistics (averaged across Monte Carlo samples) | |
|---|---|
| Average length of submitted ROLs | 6.00 |
| Fraction of weakly truth-telling students | 1.00 |
| Fraction of students assigned to favorite feasible school | 1.00 |

| Model selection tests | |
|---|---|
| Truth-telling ($H_0$) vs. Stability ($H_1$): | $H_0$ rejected in 5% of samples (at 5% significance level). |
| Stability ($H_0$) vs. Undominated strategies ($H_1$): | $H_0$ rejected in 0% of samples (at 5% significance level). |

*Notes:* This table reports Monte Carlo results from estimating students' preferences under different sets of identifying assumptions: (i) weak truth-telling; (ii) stability; (iii) stability and undominated strategies. 500 Monte Carlo samples of school choice data are simulated under the following data generating process for an economy in which 500 students compete for admission to 6 schools: an unconstrained DA where students can rank as many schools as they wish, with no cost for including an extra school in their ROL. Under assumption (iii), the model is estimated using Andrews and Shi (2013)'s method of moment (in)equalities. Column 1 reports the true values of the parameters. The mean and standard deviation (SD) of point estimates across the Monte Carlo samples are reported in columns 2, 5 and 8, and in columns 3, 6 and 9, respectively. Columns 4, 7 and 10 report the coverage probabilities (CP) for the 95 percent confidence intervals. The confidence intervals in models (i) and (ii) are the Wald-type confidence intervals obtained from the inverse of the Hessian matrix. The marginal confidence intervals in model (iii) are computed using the method proposed by Bugni et al. (2017). Truth-telling is tested against stability by constructing a Hausman-type test statistic from the estimates of both approaches. Stability is tested against undominated strategies by checking if the identified set of the moment(in)equality model is empty, using the test proposed by Bugni et al. (2015).

# Appendix D Additional Results and Goodness of Fit

Section D.1 of this appendix presents additional results on students' ranking behavior (extending Section IV.B in the main text). Section D.2 describes the goodness-of-fit statistics that we use to compare the estimates of student preferences under different sets of identifying assumptions (Section IV.D in the main text).

## D.1 Additional Results on Students' Ranking Behavior

The reduced-form evidence presented in Section IV.B of the main text suggests that students' ranking behavior could be influenced by their expected admission probabilities, as the fraction of students ranking a selective school is shown to be close to one for students with a priority index above the school cutoff, while decreasing rapidly when the priority index falls below the cutoff.

We extend this analysis by evaluating whether the pattern in Figure 3 is robust to controlling for potential determinants of student preferences. In particular, since the decision to rank a selective school might be influenced by the student's ability, we investigate whether the correlation between the priority index and the probability of ranking a selective school is still present once we control for the student's DNB scores in French and math.[A.13]

Specifically, we estimate the following linear probability model separately for each of the four schools with the highest cutoffs in the Southern District of Paris:

$$y_{i,s} = \delta_0 + \delta_1 \cdot \mathbb{1}\{e_{i,s} < P_s\} \times (e_{i,s} - P_s) + \delta_2 \cdot \mathbb{1}\{e_{i,s} \geqslant P_s\} \times (e_{i,s} - P_s) + \delta_3 \cdot \mathbb{1}\{e_{i,s} \geqslant P_s\} + Z'_{i,s}\pi + \epsilon_{i,s}, \quad \text{(A.9)}$$

where $y_{i,s}$ is an indicator function that takes the value of one if student $i$ included school $s$ in her ROL, and zero otherwise; the coefficients $\delta_1$ and $\delta_2$ allow the linear relationship between student $i$'s priority index at school $s$ ($e_{i,s}$) and the probability of ranking that school to differ on either side of the school cutoff ($P_s$), while the coefficient $\delta_3$ allows for a discontinuous jump in the ranking probability at the cutoff; $Z_{i,s}$ is a vector of student-school-specific characteristics, which includes the student's DNB exam scores in French and math, an indicator for having a high SES background, the distance to school $s$ from $i$'s place of

[A.13] Note that students' DNB scores are different from the academic performance measure that is used to calculate student priority indices as an input into the DA mechanism (see Appendix B).

**Table D1:** Correlation between Student Priority Index and Probability of Ranking the Most Selective Schools in the Southern District of Paris

| | Dependent variable: School $s$ is ranked by student | | | | | |
|---|---|---|---|---|---|---|
| | $s$ = School 11 (school with the highest cutoff) | | | $s$ = School 9 (school with second highest cutoff) | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| PRIORITY INDEX (ORIGINAL SCALE IN POINTS / 100) | | | | | | |
| (Priority index − school cutoff)×$\mathbb{1}${priority index < cutoff} | 0.549*** | 0.428*** | 0.430*** | 0.381*** | 0.213*** | 0.236*** |
| | (0.052) | (0.064) | (0.064) | (0.073) | (0.078) | (0.078) |
| (Priority index − school cutoff)×$\mathbb{1}${priority index ⩾ cutoff} | 0.083 | −0.074 | 0.074 | 0.043 | −0.156* | 0.061 |
| | (0.077) | (0.079) | (0.102) | (0.077) | (0.085) | (0.089) |
| $\mathbb{1}${priority index ⩾ cutoff} | −0.022 | −0.025 | −0.015 | 0.068** | 0.058** | 0.045* |
| | (0.022) | (0.022) | (0.023) | (0.026) | (0.026) | (0.026) |
| STUDENT TEST SCORES | | | | | | |
| French score | | 0.054 | 0.124 | | 0.153*** | 0.524*** |
| | | (0.050) | (0.209) | | (0.049) | (0.202) |
| Math score | | 0.179*** | 0.370* | | 0.196*** | 0.375* |
| | | (0.053) | (0.214) | | (0.051) | (0.211) |
| French score (squared) | | | −0.100 | | | −0.388** |
| | | | (0.167) | | | (0.160) |
| Math score (squared) | | | −0.254 | | | −0.268 |
| | | | (0.173) | | | (0.171) |
| OTHER COVARIATES | | | | | | |
| High SES student | | | 0.094*** | | | 0.092*** |
| | | | (0.020) | | | (0.019) |
| Distance to School (in km) | | | −0.051*** | | | −0.069*** |
| | | | (0.014) | | | (0.010) |
| Closest school | | | 0.010 | | | −0.090*** |
| | | | (0.036) | | | (0.030) |
| School co-located with student's middle school | | | 0.014 | | | - |
| | | | (0.028) | | | - |
| Number of students | 1,344 | 1,344 | 1,344 | 1,344 | 1,344 | 1,344 |
| Adjusted R-squared | 0.123 | 0.132 | 0.164 | 0.086 | 0.108 | 0.164 |
| $F$-TEST: JOINT SIGNIFICANCE OF THE THREE COEFFICIENTS ON PRIORITY INDEX | | | | | | |
| $F$-stat | 79.04 | 18.14 | 21.48 | 51.61 | 10.62 | 13.32 |
| $p$-value | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

*Notes:* Results are calculated with administrative data from the Paris Education Authority (*Rectorat de Paris*) for students from the Southern District who applied for admission to public high schools for the academic year starting in 2013. Columns 1–3 report estimates from of a linear probability model describing the probability that a student ranks the school with the highest cutoff (School 11) as a function of her priority index, her test scores in French and math, and additional student-specific characteristics. Columns 4–6 report estimates for the probability of ranking the school with the second highest cutoff (school 9). The empirical specification allows for the effect of the priority index to vary depending on whether the student is above or below the school's cutoff, and allows for a discontinuous jump in the ranking probability at the cutoff. French and math scores are from the exams of the *Diplôme national du brevet* (DNB) in middle school and are measured in percentiles and normalized to be in $[0, 1]$. Low-income students are not included in the sample due to the low-income bonus of 300 points placing them well above the cutoffs. Heteroskedasticity-robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

residence, an indicator for school $s$ being co-located with the student's middle school, and an indicator for school $s$ being the closest to her residence.

For reasons of space, Table D1 only reports the OLS estimates of model (A.9) for the two schools with the highest cutoffs, i.e., School 11 (columns 1–3) and School 9 (columns 4–6). The results for the third and fourth most selective schools (Schools 10 and 7, respectively) yield similar conclusions and are available upon request.

Columns 1 and 4 report estimates of the model without the covariates, $Z_{i,s}$, and can be viewed as the regression version of the graphs displayed in the upper panel of Figure 3 in the main text. Columns 2 and 5 add controls for the student's exam scores in French and math. Columns 3 and 7 use a more flexible specification that controls for a quadratic function of French and math scores, and includes the full set of covariates. Note that low-income students are not included in the the estimation sample for the same reason as in Figure 3, because the low-income bonus places them well above the cutoff.

Table D1 confirms that the kink-shaped relationship between student priority index and the probability of ranking the district's most selective schools is robust to controlling for students' academic performance and other observable characteristics. Across all specifications, the probability of ranking School 11 or School 9 increases significantly with the student's priority index, up to the point where the school becomes ex post feasible; above the cutoff, student ranking behavior is essentially uncorrelated with the value of the priority index.

Overall, these reduced-form results suggest that students' submitted choices are influenced by their priority index, in ways that seem uncorrelated with their underlying preferences. This type of behavior cannot be easily reconciled with weak truth-telling.

## D.2 Goodness of Fit

The goodness-of-fit statistics reported in Panel A of Table 5 in the main text are based on simulation techniques (Panel A), whereas those reported in Panel B use closed-form expressions for the choice probabilities (due to the logit specification). We use these goodness-of-fit measures to compare the predictive performance of the preference estimates obtained under different sets of identifying assumptions.

### D.2.1 Simulation-Based Goodness-of-Fit Measures

To compare different estimators' ability to predict school cutoffs and students' assignment, we use several simulation-based goodness-of-fit statistics. We keep fixed the estimated coefficients and $Z_{i,s}$, and draw utility shocks as type-I extreme values. This leads to the simulated utilities for every student in 300 simulation samples. When studying the WTT-based estimates, we let students submit their top 8 schools according to their simulated preferences; the matching is obtained by running DA. For the other sets of estimates, because stability is assumed, we focus on the unique stable matching in each sample, which is calculated using students' priority indices and simulated ordinal preferences.

**Predicted Cutoffs.** Observed school cutoffs are compared to those simulated using the different estimates. The results, which are averaged over the 300 simulated samples, are reported in Table D2, with standard deviations across the samples in parentheses (see Figure 5 in the main text for a graphical representation).

**Predicted Assignment.** Students' observed assignment is compared to their simulated assignment by computing the average predicted fraction of students who are assigned to their observed assignment school; in other words, this is the average fraction of times each student is assigned to her observed assignment in the 300 simulated samples, with standard deviations across the simulation samples reported in parentheses. The results are reported in Panel A of Table 6 in the main text.

### D.2.2 Predicted vs. Observed Partial Preference Order

Our second set of goodness of fit measures involves comparing students' observed partial preference order (revealed by their ROL) with the predictions based on different sets of identifying assumptions. We use two distinct measures: (i) the mean predicted probability that a student prefers the top-ranked school to the 2nd-ranked in her submitted ROL, which is averaged across students; and (ii) the mean predicted probability that a student's partial preference order among the schools in her ROL coincides with the submitted rank order. Because of the type-I extreme values, we can exactly calculate these probabilities. The results are reported in Panel B of Table 6 in the main text.

**Table D2:** Goodness of Fit: Observed vs. Simulated Cutoffs

| | Observed cutoffs | Cutoffs in simulated samples with estimates from | | |
|---|---|---|---|---|
| | | Weak Truth-telling | Stability of the matching | Stability and undominated strategies |
| | (1) | (2) | (3) | (4) |
| School 1 | 0.000 | 0.000 | 0.000 | 0.000 |
| | | (0.000) | (0.000) | (0.000) |
| School 2 | 0.015 | 0.004 | 0.024 | 0.019 |
| | | (0.006) | (0.012) | (0.013) |
| School 3 | 0.000 | 0.000 | 0.000 | 0.000 |
| | | (0.000) | (0.000) | (0.000) |
| School 4 | 0.001 | 0.043 | 0.017 | 0.017 |
| | | (0.015) | (0.007) | (0.008) |
| School 5 | 0.042 | 0.064 | 0.053 | 0.040 |
| | | (0.016) | (0.012) | (0.013) |
| School 6 | 0.069 | 0.083 | 0.077 | 0.062 |
| | | (0.025) | (0.022) | (0.024) |
| School 7 | 0.373 | 0.254 | 0.373 | 0.320 |
| | | (0.020) | (0.010) | (0.012) |
| School 8 | 0.239 | 0.000 | 0.241 | 0.153 |
| | | (0.001) | (0.023) | (0.047) |
| School 9 | 0.563 | 0.371 | 0.564 | 0.505 |
| | | (0.033) | (0.017) | (0.023) |
| School 10 | 0.505 | 0.393 | 0.506 | 0.444 |
| | | (0.029) | (0.011) | (0.014) |
| School 11 | 0.705 | 0.409 | 0.705 | 0.663 |
| | | (0.040) | (0.009) | (0.013) |

*Notes:* This table compares the cutoffs, observed for the 11 high schools of the Southern District of Paris in 2013, to the average cutoffs simulated under various identifying assumptions as in Table 5. The reported values for the simulated cutoffs are averaged over 300 simulated samples, and the standard deviations across the samples are reported in parentheses. In all simulations, we vary only the utility shocks, which are kept common across columns 2–4.

# Appendix E Supplementary Figure and Table

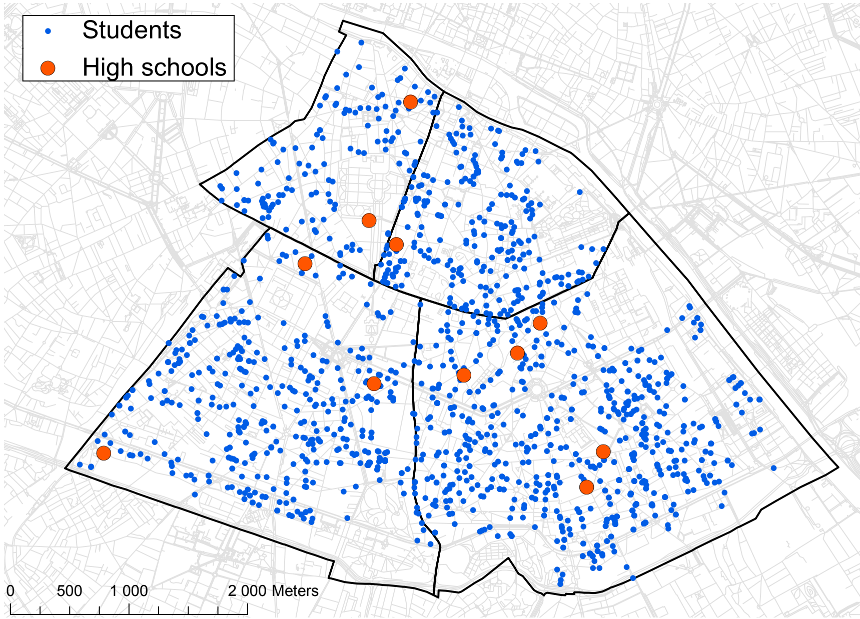


**Figure E1:** The Southern District of Paris for Public High School Admissions

*Notes:* The Southern District of Paris covers four of the city's 20 *arrondissements* (administrative divisions): 5th, 6th, 13th and 14th. The large red circles show the location of the district's 11 public high schools (*lycées*). The small blue circles show the home addresses of the 1,590 students in the data.

**Table E1:** Assigned and Unassigned Students in the Southern District of Paris

| | Assigned students | Unassigned students |
|---|---|---|
| *Panel A. Student characteristics* | | |
| Age | 15.0 | 15.0 |
| Female | 0.51 | 0.45 |
| French score | 0.56 | 0.45 |
| Math score | 0.54 | 0.47 |
| Composite score | 0.55 | 0.46 |
| High SES | 0.48 | 0.73 |
| With low-income bonus | 0.16 | 0.00 |
| *Panel B. Enrolment outcomes* | | |
| Enrolled in assignment school | 0.96 | |
| Enrolled in another public school | 0.01 | 0.65 |
| Enrolled in a private school | 0.03 | 0.35 |
| Number of students | 1,568 | 22 |

*Notes:* The summary statistics are based on administrative data from the Paris Education Authority (*Rectorat de Paris*), for students who applied to the 11 high schools of Paris's Southern District for the academic year starting in 2013. All scores are from the exams of the *Diplôme national du brevet* (DNB) in middle school and are measured in percentiles and normalized to be in $[0, 1]$. Enrollment shares are computed for students who are still enrolled in the Paris school system at the beginning of the 2013-2014 academic year (97 percent of the initial sample). Students unassigned after the main round have the possibility of participating in a supplementary round, but with choices restricted to schools with remaining seats.

# References


**Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda**, "Expanding 'Choice' in School Choice," *American Economic Journal: Microeconomics*, 2015, *7* (1), 1–42.

**Andrews, Donald W. K. and Xiaoxia Shi**, "Inference Based on Conditional Moment Inequalities," *Econometrica*, 2013, *81* (2), 609–666.

**Azevedo, Eduardo M. and Jacob D. Leshno**, "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, 2016, *124* (5), 1235–1268.

**Bugni, Federico A., Ivan A. Canay, and Xiaoxia Shi**, "Specification Tests for Partially Identified Models Defined by Moment Inequalities," *Journal of Econometrics*, 2015, *185* (1), 259–282.

**_ , _ , and _** , "Inference for Subvertors and Other Functions of Partially Identified Parameters in Moment Inequality Models," *Quantitative Economics*, 2017, *8* (1), 1–38.

**Dubins, Lester E. and David A. Freedman**, "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 1981, *88* (7), 485–494.

**Ergin, Haluk I.**, "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 2002, *70* (6), 2489–2497.

**Haeringer, Guillaume and Flip Klijn**, "Constrained School Choice," *Journal of Economic Theory*, 2009, *144* (5), 1921–1947.

**He, Yinghua**, "Gaming the Boston School Choice Mechanism in Beijing," 2015. TSE Working Paper No. 15-551.

**_ and Thierry Magnac**, "A Pigouvian Approach to Congestion in Matching Markets," 2016. Manuscript.

**_ , Antonio Miralles, Marek Pycia, and Jianye Yan**, "A pseudo-market approach to allocation with priorities," *American Economic Journal: Microeconomics*, 2018, *10* (3), 272–314.

**_ , Sanxi Li, and Jianye Yan**, "Evaluating Assignment without Transfers: A Market Perspective," *Economics Letters*, 2015, *133*, 40–44.

**Matzkin, Rosa L.**, "Nonparametric Identification and Estimation of Polychotomous Choice Models," *Journal of Econometrics*, 1993, *58* (1-2), 137–168.

**Newey, Whitney K. and Daniel McFadden**, "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

**Roth, Alvin E.**, "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 1982, *7* (4), 617–628.