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

Gabrielle Fack Julien Grenet Yinghua He

October 2018

(Forthcoming: *The American Economic Review*)

**Abstract**

We propose novel approaches to estimating student preferences with data from matching mechanisms, especially the Gale-Shapley Deferred Acceptance. Even if the mechanism is strategy-proof, assuming that students truthfully rank schools in applications may be restrictive. We show that when students are ranked strictly by some ex-ante known priority index (e.g., test scores), stability is a plausible and weaker assumption, implying that every student is matched with her favorite school/college among those she qualifies for ex post. The methods are illustrated in simulations and applied to school choice in Paris. We discuss when each approach is more appropriate in real-life settings.

JEL Codes: C78, D47, D50, D61, I21

Keywords: *Gale-Shapley Deferred Acceptance Mechanism, School Choice, College Admissions, Stable Matching, Student Preferences*

*Fack: Université Paris-Dauphine, PSL Research University, LEDa and Paris School of Economics, Paris, France (e-mail: gabrielle.fack@dauphine.fr); Grenet: CNRS and Paris School of Economics, 48 boulevard Jourdan, 75014 Paris, France (e-mail: julien.grenet@ens.fr); He: Rice University and Toulouse School of Economics, Baker Hall 246, Department of Economcis MS-22, Houston, TX 77251 (e-mail: yinghua.he@gmail.com). For constructive comments, we thank Nikhil Agarwal, Peter Arcidiacono, Eduardo Azevedo, Estelle Cantillon, Yeon-Koo Che, Jeremy Fox, Guillaume Haeringer, Yu-Wei Hsieh, Adam Kapor, Fuhito Kojima, Jacob Leshno, Thierry Magnac, Ariel Pakes, Mar Reguant, Al Roth, Bernard Salanié, Orie Shelef, Xiaoxia Shi, Matt Shum, Xun Tang, Olivier Tercieux, and seminar/conference participants at ASSA 2016 in San Francisco, Caltech, CES-ENS Cachan/RITM-UPSud, Columbia, Cowles Summer Conferences at Yale, EEA 2015 in Mannheim, ESWC 2015 in Montréal, IAAE 2014 in London, IDEP, IFAU, IFN, LEMNA, Matching in Practice 2014 in Berlin, Monash, Namur, Polytechnique, Rice, Stanford, THEMA, Tilburg, Université Paris-Dauphine, and VATT. The authors are grateful to the staff at the French Ministry of Education (Ministère de l'Éducation Nationale, Direction de l'Évaluation, de la Prospective et de la Performance) and the Paris Education Authority (Rectorat de l'Académie de Paris) for their invaluable assistance in collecting the data. Financial support from the French National Research Agency (l'Agence Nationale de la Recherche) through Projects DesignEdu (ANR-14-CE30-0004) and FDA (ANR-14-FRAL-0005), from the European Union's Seventh Framework Programme (FP7/2007–2013) under the Grant agreement no. 295298 (DYSMOIA), from the Labex OSE, and from the National Science Foundation (award # 1730636) is gratefully acknowledged.

The past decade has seen the Gale-Shapley Deferred Acceptance (DA) becoming the leading centralized mechanism for the placement of students to public schools at every education level, and it is now used by many education systems around the world, including Amsterdam, Boston, Hungary, New York, Paris, and Taiwan.

One of the reasons for the growing popularity of DA is its strategy-proofness (Abdulkadiroğlu and Sönmez, 2003). When applying for admission, students are asked to submit rank-order lists (ROLs) of schools, and it is in their best interest to rank schools truthfully. Students and their parents are thus released from strategic considerations. Consequently, DA also provides policymakers *"with more credible data about school choices, or parent 'demand' for particular schools,"* as argued by Thomas Payzant (former Boston Public Schools superintendent). Indeed, such rank-ordered data contain rich information on student preferences over schools, and are increasingly used in the empirical literature.

Due to the strategy-proofness of DA, one may be tempted to assume that the submitted ROLs reveal students' true preferences over schools. However, this truth-telling assumption can be restrictive in settings where students face only limited uncertainty about their admission outcomes. One such environment is the "strict-priority" setting in which schools rank students by some priority index, e.g., a test score, which is known to students when submitting their ROL. Consider a student who likes a highly selective school but has a low test score. The student may "skip the impossible" and choose not to apply to this school, as she rationally expects a zero admission probability based on available information such as past admission outcomes. This implies that not all students have strong incentives to rank all schools truthfully in their ROLs.[1]

Based on theoretical investigations of student incentive and behavior, we aim to provide empirical approaches to estimating student preferences in the strict-priority setting, which remains largely unexplored in the empirical literature on school choice and college admissions. Our proposed approaches can potentially be applied in many real-life systems, such as those in Table 1, including school choice in Finland, Paris, and Turkey (Panel A) as well as college admissions in Chile, Norway, and Taiwan (Panel B).

The paper's first contribution is to clarify the implications of the truth-telling assumption, which hypothesizes that students always report true preferences. Given the

[1]In contrast, students can be more uncertain about their admission outcomes if (i) schools use lotteries to break ties ex post, or (ii) schools rank students by test scores that are ex ante unknown. In these cases, the aforementioned student may choose to apply to the highly selective school, since uncertainty in priority indices implies that admission probabilities are rarely zero ex ante.

**Table 1:** Centralized School Choice and College Admissions based on the Deferred Acceptance Mechanism with Strict Priority Indices: Examples

| **Country/city** | **Assignment mechanism** | **Choice restrictions** | **Sources** |
|---|---|---|---|
| | | *Panel A. Secondary Education* | |
| Boston (selective schools)[a] | Student-proposing DA | Unrestricted | Abdulkadiroğlu et al. (2014) |
| Chicago (selective schools)[a] | DA (Serial dictatorship)[c] | Up to 6 choices | Pathak and Sönmez (2013) |
| NYC (selective schools)[a] | DA (Serial dictatorship)[c] | Unrestricted | Abdulkadiroğlu et al. (2014) |
| Finland | School-proposing DA | Up to 5 choices | Salonen (2014) |
| Ghana | DA (Serial dictatorship)[c] | Up to 6 choices | Ajayi (2017) |
| Paris | School-proposing DA | Up to 8 choices | Hiller and Tercieux (2014) |
| Romania | DA (Serial dictatorship)[c] | Unrestricted | Pop-Eleches and Urquiola (2013) |
| Singapore | DA (Serial dictatorship)[c] | Up to 6 choices | Teo et al. (2001) |
| Turkey | DA (Serial dictatorship)[c] | Up to 12 choices | Akyol and Krishna (2017) |
| | | *Panel B. Higher Education* | |
| Australia (Victoria) | College-proposing DA | Up to 12 choices | Artemov et al. (2017) |
| Chile | Student-proposing DA | Up to 10 choices | Hastings et al. (2013) |
| Hungary | Student-proposing DA | Unrestricted[b] | Biró (2011) |
| Ireland | College-proposing DA | Up to 10 choices | Chen (2012) |
| Norway | College-proposing DA | Up to 15 choices | Kirkebøen et al. (2016) |
| Spain | Student-proposing DA | Region-specific | Mora and Romero-Medina (2001) |
| Taiwan | College-proposing DA | Up to 100 choices | UAC (2014) |
| Tunisia | College-proposing DA | Up to 10 choices | Luflade (2018) |
| Turkey | College-proposing DA | Up to 24 choices | Saygin (2013) |

*Notes:* [a] For exam schools in Boston, selective enrollment high schools in Chicago, and specialized high schools in NYC, strict priority indices are used in the admission. In contrast, admissions to other schools often do not use strict priority indices. [b] In Hungary, students may apply for any number of programs but they are charged a fee (of approximately 10 euros) for every program after the third application. [c] In all of the countries/cities listed in this table, students' priorities are based on various combinations of grades, entrance/exit exams, and other criteria (aptitude tests, interviews, etc.). When priority indices are not school-specific, i.e., schools/universities rank students in the same way, DA, whether student-proposing or school/college-proposing, is equivalent to serial dictatorship, under which students, in the order of their priority indices, are allowed to choose among the remaining schools or universities.

flourishing empirical literature on the setting in which schools rank students with post-application lotteries (Pathak and Shi, 2014; Abdulkadiroğlu et al., 2017), it is natural to extend those truth-telling-based approaches to the strict-priority setting. Unfortunately, strategy-proofness implies that truth-telling is a *weakly* dominant strategy, leaving open the issue of multiple equilibria because a student may obtain the same outcome by opting for non-truth-telling strategies—as shown in the "skipping the impossible" example above. Making truth-telling even less likely, many applications of DA restrict the length of submittable ROLs, which destroys strategy-proofness (Haeringer and Klijn, 2009).

These arguments are formalized in a theoretical model. Deviating from the literature, we introduce an application cost that students incur when submitting ROLs, and the model therefore has the common real-life applications of DA as special cases. Conditional on both preferences and priorities being private information, we show that for truth-telling to be the unique equilibrium, two conditions are needed: no application cost and large uncertainty in admission outcomes. Neither is easily satisfied in the strict-priority setting. Even without limits on the length of submittable ROLs, students may find it costly to rank a long list of schools. As students know their own priority indices, uncertainty about admission outcomes can also be limited.

Going beyond truth-telling, the paper's second contribution is to propose a set of novel empirical approaches that are theoretically founded. We consider a weaker assumption implied by truth-telling: stability, or justified-envy-freeness, of the matching (Abdulkadiroğlu and Sönmez, 2003), meaning that every student is matched with her favorite feasible school. A school is feasible for a student if its ex post cutoff is below the student's priority index. These cutoffs are well-defined and often observable to the researcher: given the admission outcome, each school's cutoff is the lowest priority index of the students accepted there. Conditional on the cutoffs, stability therefore defines a discrete choice model with personalized choice sets, which is straightforward to analyze empirically.

We show that stability is a plausible assumption, as there exists an equilibrium outcome that is asymptotically stable under certain conditions. When school capacities and the number of students increase proportionally while the number of schools is fixed, the fraction of students not matched with their favorite feasible school tends to zero. Although stability, as an ex post optimality condition, is not guaranteed when students' information is incomplete, we provide numerical evidence suggesting that typical real-life markets are sufficiently large for stability to be almost exactly satisfied.

Based on the theoretical results, we propose a menu of approaches for preference estimation. We start by formalizing the truth-telling assumption under which one can apply rank-ordered models on submitted ROLs. In practice, students rarely rank all available schools, and, therefore, the truth-telling assumption often imposes the exogeneity of the length of a submitted ROL.[2]

Stability, but not asymptotic stability, leads to a discrete choice model with personalized choice sets, so the nonparametric identification in the discrete-choice literature can be applied (e.g., Matzkin, 1993), under the assumption that priority indices and unobserved preference heterogeneity are independent conditional on observables. An advantage of this approach is that it enables estimation with data on admission outcomes only, although ignoring the information in ROLs entails some efficiency loss in the estimation.

We also provide a solution if neither truth-telling nor stability is satisfied: as long

[2]Hence, we distinguish *strict* from *weak* truth-telling. The former assumes that every student ranks all schools truthfully, while the latter requires students to rank their most-preferred schools truthfully and allows them to omit the least-preferred schools.

as students do not play dominated strategies, the submitted ROLs reveal true partial preference orders of schools (Haeringer and Klijn, 2009),[3] allowing us to derive probability bounds for one school being preferred to another. The corresponding moment inequalities can be used for inference (for a survey, see Tamer, 2010). When stability is satisfied and identifies student preferences, these inequalities provide over-identifying information that can improve estimation efficiency (Moon and Schorfheide, 2009).

To guide the choice between these identifying assumptions, we consider several statistical tests, provided that the model is correctly specified and identified. Truth-telling, leading to more restrictions than stability, can be tested against stability using a Hausman-type test (Hausman, 1978) or a test of over-identifying restrictions (Hansen, 1982). Similarly, stability can be tested against undominated strategies: if the outcome is unstable, the stability restrictions are incompatible with the moment inequalities implied by undominated strategies, allowing us to use tests such as Bugni et al. (2015).

Our third contribution is to evaluate the performance of each approach based on simulated and real-life data. Having illustrated the main theoretical results with Monte Carlo simulations, we apply the empirical approaches to school choice data from Paris. There are 1,590 middle school students applying for admissions to 11 academic-track high schools in Paris' Southern District through a version of DA. Schools rank applicants by their academic grades but give priority to low-income students. The findings are more consistent with stability than truth-telling. Our proposed statistical tests reject truth-telling in favor of stability but fail to reject stability against undominated strategies. The tests, however, do not provide definitive proof against truth-telling, since they are conditional on the model's parametric assumptions. Additionally, we provide reduced-form evidence on students' ranking behavior suggesting that some students may have omitted the most selective schools from their ROLs because of low admission probabilities. Moreover, the truth-telling-based estimator is outperformed by the stability-based one when it comes to predicting admission outcomes and student preferences.

To highlight the differences between the proposed approaches and their underlying behavioral assumptions, we summarize the theoretical results and describe the nesting structure of the assumptions in Section 5. We also emphasize the key features of school choice and college admissions in practice that can help researchers to choose the most

[3] An ROL is a true partial preference order if the listed schools are ranked according to true preferences.

appropriate empirical approach to preference estimation.

**Other Related Literature.** Since the seminal work of Abdulkadiroğlu and Sönmez (2003), the theoretical study of student behavior and matching properties under DA has been extensive, and large-market asymptotics are a common analytical tool (see the survey by Kojima, 2015). Closely related to our study is Azevedo and Leshno (2016), who show the asymptotics of stable matchings. Our paper extends theirs to outcomes of Bayesian Nash equilibrium, whereas they assume that students are always truth-telling.

There is a burgeoning literature on preference estimation using centralized school choice data. One strand of this literature uses data from settings in which researchers argue that truth-telling behavior by students is plausible. For example, Hastings et al. (2008) use data from Charlotte-Mecklenburg public school district, and Abdulkadiroğlu et al. (2017) study school choice data from New York City, which is a "lottery" setting.[4] Both papers estimate student preferences under the assumption that students truthfully report their preferences. In the same spirit, assuming students report their true preferences in surveys, Budish and Cantillon (2012) and De Haan et al. (2015) use reported student ordinal preferences to conduct analysis without estimating preferences.

The second strand of the empirical literature explicitly considers students' strategic behavior when estimating student preferences, especially if the mechanism is not strategy-proof, e.g., the (Boston) immediate-acceptance mechanism (Calsamiglia et al., 2014; He, 2015; Hwang, 2014; Kapor et al., 2016; Agarwal and Somaini, 2018). In those settings, observed ROLs are sometimes considered as solutions to the maximization of students' expected utility. Avoiding some difficulties of this strategy-based approach, we instead propose methods that rely on equilibrium outcome of the school choice game.[5]

As to the strict-priority setting, there are only a handful of empirical studies (Burgess et al., 2014; Ajayi, 2017; Akyol and Krishna, 2017). Most of them use ad-hoc solutions to the potential problem of students' non-truth-telling behavior.[6] Akyol and Krishna

[4]The authors perform robustness checks, e.g., only considering students' top three submitted choices.

[5]With assumptions on students' beliefs, the strategy-based approach formulates a discrete choice problem defined on the set of possible ROLs. It faces some challenges: (i) Degenerate admission probabilities can occur, leading to multiple equilibria (He, 2015). (ii) Application costs, especially those related to cognitive load, are often unobservable, necessitating additional assumptions in the maximization of expected utility. (iii) A given ROL is evaluated against a large number of alternative ROLs, sometimes creating computational burden (e.g., there exist $S!/(S-K)!$ lists ranking $1 \leqslant K \leqslant S$ schools).

[6]Analyzing school choice in the U.K., where proximity to schools breaks ties in determining admission to oversubscribed primary schools, Burgess et al. (2014) restrict each student's set of schools to those in close proximity to the student's residence. In the context of admissions to secondary schools in Ghana, where exam scores determine priority, Ajayi (2017) considers a subset of schools with similar selectivity.

(2017) is an exception. Observing the outcome and the cutoffs of high school admissions in Turkey, the authors estimate preferences based on the assumption that every student is assigned to her favorite feasible school, which amounts to assuming stability of the matching. We formalize and clarify this stability assumption, along with other extensions. Although stability is a rather common identifying assumption in the two-sided matching literature (see the surveys by Fox, 2009; Chiappori and Salanié, 2016),[7] it is new in empirical studies of school choice and college admissions.

Lastly, estimation of student preferences with college admissions data is under-explored, often due to the decentralized nature of the admission process. Among centralized admissions, however, there are many applications of the DA mechanism (see Table 1).[8] The specifics of the mechanism have led to numerous studies on the causal effects of education (e.g., Hastings et al., 2015; Kirkebøen et al., 2016), but few on preference estimation. One exception is Kirkebøen (2012) who uses the truth-telling assumption while excluding from a student's choice set every college program at which the student does not meet the formal requirements or is below its previous-year cutoff. Another is Bucarey (2018) who applies our stability-based estimator to evaluate the crowding-out effects of free college tuition for low-income students in Chile.

**Organization of the Paper.** Section 1 presents the model and the theoretical foundation. Section 2 formalizes the empirical approaches and tests, which are illustrated in Monte Carlo simulations in Section 3. School choice in Paris and the empirical results are shown in Section 4. Section 5 discusses practical considerations for applying the approaches to data and outlines some extensions. We conclude in Section 6.

# 1 The Model

To study student behavior, we extend the model in Azevedo and Leshno (2016). An economy, as a school choice/college admissions problem, consists of a finite set of schools/colleges, $\mathcal{S} \equiv \{1, \ldots, S\}$, and a set of students. Student $i$ has a type $\theta_i = (u_i, e_i) \in \Theta \equiv [0,1]^S \times [0,1]^S$, where $u_i = (u_{i,1}, \cdots, u_{i,S}) \in [0,1]^S$ is a vector of von Neumann-Morgenstern (vNM) utilities of being assigned to schools, and $e_i = (e_{i,1}, \cdots, e_{i,S}) \in [0,1]^S$ is a vector of priority indices at schools, a student with a higher index having a higher priority at

[7] This literature usually considers decentralized matching markets; Agarwal (2015) is an exception.
[8] Some centralized college admissions do not use DA, e.g., Brazil (Carvalho et al., 2014).

a school. To simplify notation, we assume that all schools and students are acceptable.[9] Students are matched with schools through a centralized mechanism.

The continuum economy with a unit mass of students is denoted by $E = \{G, q, C\}$, where $G$ is an atomless probability measure over $\Theta$ representing the distribution of student types; $q = (q_1, \cdots, q_S)$ are masses of seats available at each school, where $q_s \in (0, 1)$ for all $s$; lastly, $C$ is an application cost, to be specified shortly. $G$ being atomless implies a measure-zero set of students with indifference in either utilities or priority indices.

A random finite economy of size $I$ is denoted by $F^{(I)} = \{G^{(I)}, q^{(I)}, C\}$. $F^{(I)}$ is constructed by independently drawing $I$ students, indexed by $i \in \{1, \cdots, I\}$, from the distribution $G$ and adjusting the numbers of seats to integers. Specifically, $G^{(I)}$ is the random empirical distribution of types for a sample of $I$ students;[10] $q^{(I)} = [q \cdot I]/I$ is the supply of seats per student, where $[x]$ is the vector of integers nearest to $x$ (with a rounding down in case of a tie). We use $\hat{F}^{(I)} = \{\hat{G}^{(I)}, q^{(I)}, C\}$ to denote a realization of $F^{(I)}$.

In the following, we start with $F^{(I)}$ to specify the matching process and to analyze student behavior, because empirical studies deal with finite economies; the extension to the continuum economy $E$ is deferred to Section 1.4.

In a realization of the random economy, $\hat{F}^{(I)}$, schools first announce their capacities, and every student then submits a rank-order list (ROL) of $1 \leqslant K_i \leqslant S$ schools, denoted by $L_i = \left(l_i^1, \ldots, l_i^k, \ldots, l_i^{K_i}\right)$, where $l_i^k \in \mathcal{S}$ is $i$'s $k$th choice. $L_i$ also represents the set of schools being ranked in $L_i$. We define $\succ_{L_i}$ such that $s \succ_{L_i} s'$ if and only if school $s$ is ranked above school $s'$ in $L_i$. The set of all possible ROLs is $\mathcal{L}$, which includes all ROLs ranking at least one school. Student $i$'s true ordinal preference induced by her vNM utilities $u_i$ is denoted by $r(u_i) = \left(r_i^1, \ldots, r_i^S\right) \in \mathcal{L}$.

When submitting an ROL, a student incurs a cost $C\left(|L|\right)$, which depends on the number of schools being ranked in $L$, $|L|$. Furthermore, $C\left(|L|\right) \in [0, +\infty]$ for all $L$ and is weakly increasing in $|L|$. To simplify students' participation decisions, we set $C\left(1\right) = 0$.

Such a cost function flexibly captures many common applications of school choice mechanisms. If $C\left(|L|\right) = 0$ for all $L$, we are in the traditional setting without costs (e.g., Abdulkadiroğlu and Sönmez, 2003); if $C\left(|L|\right) = \infty$ for $|L|$ greater than a constant $\overline{K}$, it

[9]Assuming acceptability of all schools justifies the normalization of $u \in [0, 1]^S$. Although we could extend the preference domain to allow for negative values, this would create the possibility that students avoid being assigned to schools with negative vNM utilities when maximizing expected utility.

[10]For a realized economy with realized student types $(\theta_1, \cdots, \theta_I)$, the realized empirical distribution $\hat{G}^{(I)}$ is defined as $\hat{G}^{(I)}(\theta) = \frac{1}{I}\sum_{i=1}^{I} \mathbb{1}(\theta_i \leqslant \theta)$, $\forall \theta \in \mathbb{R}^{2S}$, where $\mathbb{1}(\cdot)$ is an indicator function.

corresponds to the constrained school choice where students cannot rank more than $\overline{K}$ schools (e.g., Haeringer and Klijn, 2009); when $C(|L|) = \max\{0, (|L| - \underline{K})c\}$, students pay a constant marginal cost $c$ for each choice beyond the first $\underline{K}$ choices, as in Hungarian college admissions (Biró, 2011); lastly, the monotonic cost function may simply reflect that it is cognitively burdensome to rank too many schools.

The student-school match is then solved by a mechanism that takes into account students' ROLs and schools' rankings over students. Our main analysis focuses on the student-proposing Gale-Shapley Deferred Acceptance (DA), leaving the discussion of other variants to Section 5.2. As a computerized algorithm, DA works as follows:

*Round 1.* Every student applies to her first choice. Each school rejects the lowest-ranked students in excess of its capacity and temporarily holds the other students.

Generally, in:

*Round k.* Every student who is rejected in Round $(k-1)$ applies to the next choice on her list. Each school, pooling together new applicants and those it holds from Round $(k-1)$, rejects the lowest-ranked students in excess of its capacity. Those who are not rejected are temporarily held by the schools.

The process terminates after any Round $k$ when no rejections are issued. Each school is then matched with the students it is currently holding.

## 1.1 Information Structure and Decision-Making

In a realization of the finite economy, $\hat{F}^{(I)}$, given its construction, every student's preferences and priority indices are private information, and are i.i.d. draws from $G$, which is common knowledge (but $\hat{G}^{(I)}$, the realization of $G^{(I)}$, remains unknown).

We start by taking student $i$'s point of view. Conditional on others' submitted ROLs and priority indices $(L_{-i},\ e_{-i})$, as well as student $i$'s $(L_i,\ e_i)$, her admission outcome is deterministic, given the algorithm. Specifically, $i$'s admission outcome at school $s$ is:

$$a_s(L_i, e_i; L_{-i}, e_{-i}) \equiv \begin{cases} \mathbb{1}\left(i \text{ is rejected by } l_i^1, \ldots, l_i^k \text{ and accepted by } l_i^{k+1} = s \mid L_i, e_i; L_{-i}, e_{-i}\right) & \text{if } s \in L_i \\ 0 & \text{if } s \notin L_i \end{cases}$$

where $\mathbb{1}(\cdot \mid L_i, e_i; L_{-i}, e_{-i})$ is an indicator function. Moreover, due to the centralized mechanism, a student can receive at most one offer, so $\sum_{s=1}^{S} a_s(L_i, e_i; L_{-i}, e_{-i}) = 0$ or 1.

Of course, $L_{-i}$ and $e_{-i}$ are unknown to $i$ at the time of submitting her ROL, so $i$ takes into account the uncertainty when choosing an action. A pure strategy is $\sigma : \Theta \to \mathcal{L}$. Given $\sigma$, the (ex ante) admission probabilities are $\int a_s\left(\sigma(\theta_i), e_i; \sigma_{-i}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i})$ for all $i$ and $s$, where $\sigma_{-i}(\theta_{-i}) \equiv \{\sigma(\theta_j)\}_{j \neq i}$. We consider a (type-)symmetric equilibrium $\sigma^*$ in pure strategies such that $\sigma^*$ solves the following maximization problem for every $\theta_i$:[11]

$$\sigma^*(\theta_i) \in \underset{\sigma(\theta_i) \in \mathcal{L}}{\arg\max} \left\{ \sum_{s \in \mathcal{S}} u_{i,s} \int a_s\left(\sigma(\theta_i), e_i; \sigma^*_{-i}(\theta_{-i}), e_{-i}\right) dG(\theta_{-i}) - C\left(|\sigma(\theta_i)|\right) \right\}. \tag{1}$$

The existence of pure-strategy Bayesian Nash equilibrium can be established by applying Theorem 4 (Purification Theorem) in Milgrom and Weber (1985), although there can be multiple equilibria. For ease of exposition, the following analysis focuses on pure-strategy equilibrium. We note that while economy $F^{(I)}$ is random, a strategy $\sigma$ is "deterministic" in the sense that it only depends on $(G, I, C)$ but not on the realization of $F^{(I)}$.

We define a realized matching $\hat{\mu}$ as a mapping from $\Theta$ to $\mathcal{S} \cup \{\varnothing\}$ such that: (i) $\hat{\mu}(\theta_i) = s$ if student $i$ is matched with $s$; (ii) $\hat{\mu}(\theta_i) = \varnothing$ if student $i$ is unmatched; and (iii) $\hat{\mu}^{-1}(s)$ is the set of students matched with $s$, while $|\hat{\mu}^{-1}(s)|$ is the number of students matched with $s$ and does not exceed $s$'s capacity.

$\hat{F}^{(I)}$ and $\sigma$ together lead to an ROL profile as inputs into the DA mechanism and result in a matching, $\mu_{\left(\hat{F}^{(I)}, \sigma\right)}$, which is uniquely determined by the mechanism. Note that $\mu_{\left(F^{(I)}, \sigma\right)}$ is a random matching because $F^{(I)}$ is a random economy.

Moreover, the (random) cutoff of school $s$ in random matching $\mu_{\left(F^{(I)}, \sigma\right)}$ is defined as:

$$P_s\left(\mu_{\left(F^{(I)}, \sigma\right)}\right) = \begin{cases} \min\left\{e_{i,s} \mid \mu_{\left(F^{(I)}, \sigma\right)}(\theta_i) = s\right\} & \text{if } |\mu^{-1}_{\left(F^{(I)}, \sigma\right)}(s)| = q_s^{(I)} \\ 0 & \text{if } |\mu^{-1}_{\left(F^{(I)}, \sigma\right)}(s)| < q_s^{(I)} \end{cases}$$

That is, $P_s\left(\mu_{\left(F^{(I)}, \sigma\right)}\right)$ is zero if $s$ does not meet its capacity; otherwise, it is the lowest priority index among all accepted students. The vector of cutoffs is denoted by $P\left(\mu_{\left(F^{(I)}, \sigma\right)}\right)$, and its realization in $\hat{F}^{(I)}$ is $P\left(\mu_{\left(\hat{F}^{(I)}, \sigma\right)}\right)$.

[11] It is innocuous to focus on symmetric equilibrium, because it does not restrict the strategy of any student given that they all have different priority indices (almost surely).

## 1.2 Truth-Telling Behavior in Equilibrium

To assess the plausibility of the truth-telling assumption in empirical studies, we begin by investigating students' truth-telling behavior in equilibrium. A clarification of the concepts is in order. Student $i$ is **weakly truth-telling** (WTT, hereafter) if $\sigma(\theta_i) = (r_i^1, r_i^2, \ldots, r_i^{K_i})$ for $K_i \leqslant S$. That is, $i$ ranks her $K_i$ most-preferred schools by her true preference order but may not rank all schools. If a WTT strategy always truthfully ranks all $S$ schools and thus $\sigma(\theta_i) = r(u_i)$, $i$ is **strictly truth-telling** (STT, hereafter).[12]

We emphasize the difference between WTT and STT because strategy-proofness concerns the latter. However, WTT is often considered in empirical studies because in practice, students rarely rank all available schools, as we shall revisit in Section 2.2.

It is known that DA is strategy-proof when there is no application cost (Dubins and Freedman, 1981; Roth, 1982). That is, when $C(|L|) = 0$ for all $L \in \mathcal{L}$, STT is a weakly-dominant strategy for all students. However, strategy-proofness, or weak dominance of STT, leaves open the possibility of multiple equilibria. Even when all others play STT, there may exist multiple best responses for a given student.[13] It is therefore useful to clarify the conditions under which STT is the unique equilibrium. The following example highlights two sources of equilibrium multiplicity in a complete-information environment.

**Example 1 (Multiple Equilibria under DA without Application Cost).** *Consider a finite economy that has two students* $(i_1, i_2)$*, three one-seat schools* $(s_1, s_2, s_3)$*, but no application cost. As common knowledge, all schools rank* $i_1$ *above* $i_2$*; student* $i_1$*'s preference order is* $(s_1, s_2, s_3)$*, but* $i_2$*'s is* $(s_2, s_1, s_3)$*. There are many equilibria in addition to STT, stemming from two sources: "**irrelevance at the bottom**" and "**skipping the impossible**." Both arise when some admission probabilities are zero.*

*For* $i_1$*, the bottom part of her submitted ROL is irrelevant as long as* $s_1$ *is top-ranked. In fact, any ROL* $(s_1, s', s'')$*, for* $s', s'' \in \{s_2, s_3\} \cup \{\varnothing\}$*, is weakly dominant for* $i_1$*, as she is always accepted by* $s_1$*. For student* $i_2$*, "skipping the impossible" comes into play. She can omit* $s_1$ *from her submitted ROL without affecting her outcome, because* $s_1$ *is always taken by* $i_1$ *in any equilibrium. Making things worse, how she ranks* $s_1$ *is payoff-irrelevant.*

[12]Related to the distinction between STT and WTT, the literature on lab experiments on school choice sometimes also defines truth-telling as being different from STT. For example, Chen and Sönmez (2006) call a student truth-telling under the DA mechanism if she ranks her most-preferred schools up to her district school, at which she has guaranteed admission.

[13]Unfortunately, it is impossible to make STT a strictly-dominant strategy, because it would require STT to be strictly better than all other strategies against all possible action profiles of other students.

One may conjecture that STT might survive as the unique equilibrium when information is incomplete. Indeed, specifying the incompleteness of information, the following proposition provides sufficient conditions and a necessary condition.

**Proposition 1.** *(i)* ***Sufficiency****: STT is the unique Bayesian Nash equilibrium under DA if (i) there is no application cost: $C(|L|) = 0$, $\forall L \in \mathcal{L}$; and (ii) the joint distribution of preferences and priorities $G$ has full support.*
*(ii)* ***Necessity****: For any non-zero application cost, there always exist student types for whom STT is not an equilibrium strategy.*

All proofs can be found in Appendix A. The no-cost condition is violated if students cannot rank as many schools as they wish, or if they suffer a cognitive burden when ranking too many schools. It should also be emphasized that the cost need not be large to make students deviate from STT, because the marginal benefit of ranking an additional school can be close to zero. When a student considers her admission probability at her $k$th choice, she may face a close-to-one probability of being accepted by at least one of her earlier choices. This is in the same spirit as the "irrelevance at the bottom" in Example 1. When the marginal application cost exceeds marginal benefits, STT is no longer a best response, which implies the necessity of the zero-cost condition.

The full-support condition, also considered in Chen and Pereyra (2017), makes all admission probabilities non-zero by introducing uncertainties, and therefore any deviation from STT is costly. This is more plausible when the priority index is determined by an ex post lottery and when the information on others' preferences over schools is less precise.

**Remark 1.** *Proposition 1 specifies when students have incentives to rank* all *schools truthfully, but this result does not extend to WTT. Although it is sometimes used for identification and estimation, the WTT assumption is not supported as an equilibrium.*[14]

We may take one step back and focus on whether students have incentives to order the *ranked* schools truthfully. We call $L_i$, $|L_i| \leqslant S$, a **partial preference order** of schools if $L_i$ respects the true preference order among those ranked in $L_i$. That is, $L_i$ ranks $s$ above $s'$, only if $u_{i,s} > u_{i,s'}$; when $s$ is not ranked in $L_i$, there is no information on how $s$ is ranked relative to any other school according to $i$'s true preferences.

[14]The equilibrium condition, Equation (1), implies that a student may "skip the impossible" by omitting her most-preferred school if the admission probability is close to zero, thus violating WTT.

**Proposition 2.** *Under DA with application cost, if students do not play weakly dominated strategies, a student's submitted ROL is a partial order of her true preferences.*

Proposition 2 can be considered as a corollary of Proposition 4.2 in Haeringer and Klijn (2009), and thus we omit its proof. The key is that a non-partial-preference-order ROL is weakly dominated by the ROL that ranks the same schools according to their true preference order. This result is useful for empirical analysis, as it specifies students' revealed preferences. Section 2.5 formulates how to use this information in estimation.

## 1.3 Matching Outcome: Stability

The above results speak to the plausibility of the truth-telling assumptions, WTT and STT, in empirical studies. In particular, WTT is not theoretically supported as a weakly-dominant strategy even in DA with no application cost; whenever there is any form of application cost, STT is no longer a dominant strategy.

Taking a different perspective, we note that all equilibria lead to the same matching in Example 1. This motivates us to investigate the properties of equilibrium outcomes of DA. Intuitively, the degree of multiplicity in equilibrium outcomes must be smaller than that in equilibrium strategies. In the two-sided matching literature, stability is the leading concept for equilibrium outcome and the main identifying assumption (Chiappori and Salanié, 2016). We investigate whether stability can also be satisfied in all equilibrium outcomes of school choice and college admissions.

Unfortunately, we shall demonstrate that having stability satisfied in all equilibrium outcomes requires similar conditions to those for STT being the unique equilibrium. In fact, whenever there is application cost, stability is not guaranteed in equilibrium either. This is because Bayesian Nash equilibrium implies ex ante optimality of student strategy, while stability requires ex post optimality.

As we study a matching's ex post properties, let us consider $\hat{\mu}$, a realization of the random matching. $(i, s)$ form a **blocking pair** if (i) $i$ prefers $s$ over her matched school $\hat{\mu}(\theta_i)$ while $s$ has an empty seat ($|\hat{\mu}^{-1}(s)| < I \times q_s^{(I)}$), or if (ii) $i$ prefers $s$ over $\hat{\mu}(\theta_i)$ while $s$ has no empty seats ($|\hat{\mu}^{-1}(s)| = I \times q_s^{(I)}$) but $i$'s priority index is higher than its cutoff, $e_{i,s} > \min_{\{j:\hat{\mu}(\theta_j)=s\}}\{e_{j,s}\}$. $\hat{\mu}$ is **stable** if there is no blocking pair. Stability is also known as **elimination of justified envy** in school choice (Abdulkadiroğlu and Sönmez, 2003).

Given a realized matching $\hat{\mu}$, school $s$ is ex post **feasible** for $i$ if $i$'s priority index at $s$ is above $s$'s cutoff, $e_{i,s} \geqslant P_s(\hat{\mu})$. Let $\mathcal{S}(e_i, P(\hat{\mu}))$ be the set of feasible schools for $i$.

With these definitions, combining Lemmata 1 and 2 in Balinski and Sönmez (1999), we reach an important result: *a realized matching $\hat{\mu}$ is stable if and only if every student is matched with her favorite feasible school* (i.e., $\hat{\mu}(\theta_i) = \arg\max_{s \in \mathcal{S}(e_i, P(\hat{\mu}))} u_{i,s}, \forall i$). As the cutoffs of a matching are observed ex post by the researcher, we can define every student's set of feasible schools; stability therefore implies a *discrete choice model with observable, personalized choice sets.* We further formalize this in Section 2.3.

We are interested in stability being satisfied in an outcome of dominant-strategy equilibrium, which would free us from specifying the information structure and from imposing additional equilibrium conditions. The following lemma provides necessary and sufficient conditions, which are similar to those for STT to be the unique equilibrium.

**Lemma 1.** *Under DA, a Bayesian Nash equilibrium in dominant strategy always leads to a stable matching if and only if $C(|L|) = 0$ for all $L$. It is the unique equilibrium outcome if additionally $G$ has full support.*

The "if and only if" statement of the lemma is implied by strategy-proofness of DA without application cost, while the uniqueness statement is a result of Proposition 1.

DA is known to produce a stable matching when students are STT (Gale and Shapley, 1962), but not when students are only WTT. The following results, clarifying the relationship between WTT and stability, have implications for our empirical approaches.

**Proposition 3.** *Suppose that every student is WTT under DA, which may not be an equilibrium. Given a realized matching:*

*(i) whenever a student is assigned, she is matched with her favorite feasible school;*

*(ii) if everyone who has at least one feasible school is assigned, the matching is stable.*

The above results describe the nesting structure of the two assumptions, WTT and stability, although they do not speak to the plausibility of either of them being an equilibrium strategy/outcome. Specifically, WTT is more restrictive, as it implies the no-blocking property among assigned students. We use these results to formulate statistical tests for the choice between WTT and stability in Section 2.4.

## 1.4 Asymptotic Stability in Bayesian Nash Equilibrium

So far, we have shown that neither truth-telling (STT and WTT) nor stability can emerge in equilibrium without some potentially restrictive assumptions. Following the literature on large markets, we study whether stability can be asymptotically satisfied.

We now revisit the continuum economy, $E$, and additionally introduce a sequence of random finite economies $\{F^{(I)}\}_{I \in \mathbb{N}}$ that are constructed from $E$ as before.

The definitions of matching, DA, and stability can be naturally extended to continuum economies as in Abdulkadiroğlu et al. (2015) and Azevedo and Leshno (2016), which are discussed in Appendix A.2.1. These definitions are similar to their counterparts in finite economies. For example, a matching in $E$ when everyone adopts $\sigma$ is $\mu_{(E,\sigma)} : \Theta \to \mathcal{S} \cup \{\varnothing\}$, which satisfies (i) $\mu_{(E,\sigma)}(\theta_i) = s$ when type $\theta_i$ is matched with $s$ and (ii) $G(\mu^{-1}_{(E,\sigma)}(s)) \leqslant q_s$.

It is known that, generically, there exists a unique stable matching in the continuum economy (Azevedo and Leshno, 2016);[15] we impose the conditions for the uniqueness and denote this stable matching in $E$ as $\mu^\infty$ and the corresponding cutoffs as $P^\infty$. To continue our exploration, we make the following assumption:

**Assumption 1.** *Every Bayesian Nash equilibrium of the continuum economy $E$ results in the unique stable matching, $\mu^\infty$.*

A sufficient condition for Assumption 1 is $C(2) > 0$ (i.e., it is costly to apply to more than one school), and when $C(2) = 0$, a sufficient and necessary condition is Ergin acyclicity (Proposition A3 in Appendix A.2.5). An economy is acyclical if no student can block a potential settlement between any other two students without affecting her own match (Ergin, 2002). Appendix A.2.5 gives its formal definition in continuum economies. This condition is satisfied when all schools rank every student by a single priority index.

Because we are interested in equilibrium outcomes, we augment the sequence of economies with equilibrium strategies, $\{F^{(I)}, \sigma^{(I)}\}_{I \in \mathbb{N}}$, where $\sigma^{(I)}$ is a pure-strategy Bayesian Nash equilibrium in $F^{(I)}$ and satisfies the following assumption.

---

[15] A sufficient condition for the uniqueness of stable outcome in $E$ is that $G$ has full support. Even when $G$ does not have full support, the uniqueness can be achieved when $\sum_{s=1}^{S} q_s < 1$. Let $\sigma^{STT}$ be the STT strategy. We define the demand for each school in $(E, \sigma^{STT})$ as a function of cutoffs, $D_s(P \mid E, \sigma^{STT}) = \int \mathbb{1}(u_{i,s} = \max_{s' \in \mathcal{S}(e_i, P)} u_{i,s'}) dG(\theta_i)$. Let $D(P \mid E, \sigma^{STT}) = [D_s(P \mid E, \sigma)]_{s \in \mathcal{S}}$. $E$ admits a unique stable matching if the image under $D(P \mid E, \sigma^{STT})$ of the closure of the set

$$\left\{P \in (0,1)^S : D(P \mid E, \sigma^{STT}) \text{ is not continuously differentiable at } P\right\}$$

has Lebesgue measure zero.

**Assumption 2.** *There exists* $\sigma^\infty$ *such that* $\lim_{I\to\infty} G\left(\left\{\theta_i \in \Theta \mid \sigma^{(I)}(\theta_i) = \sigma^\infty(\theta_i)\right\}\right) = 1$.

A sufficient condition for Assumption 2 is $C(2) > 0$ (Lemma A5 in Appendix A.2.4).[16] Although $F^{(I)}$ is a random economy, $\sigma^{(I)}$ is fixed given the size of the economy. In other words, $\sigma^{(I)}$ remains as an equilibrium strategy in any realization of $F^{(I)}$. Assumption 2 regulates how the equilibria evolve with economy size, which is necessary as there are multiple equilibria. By this assumption, in the sequence $\{\sigma^{(I)}\}_{I\in\mathbb{N}}$, fewer and fewer student types need to adjust their optimal actions when the economy enlarges. Moreover, given Assumption 1, the limit strategy $\sigma^\infty$ leads to $\mu^\infty$ in $E$ (Proposition A1).

### 1.4.1 Asymptotic Stability: Definition and Results

Let the random matching $\mu_{\left(F^{(I)},\sigma^{(I)}\right)}$ be $\mu^{(I)}$, and the associated random cutoffs $P\left(\mu^{(I)}\right)$ be $P^{(I)}$. The following definition formalizes the concept of asymptotic stability.[17]

**Definition 1.** *A sequence of random matchings,* $\{\mu^{(I)}\}_{I\in\mathbb{N}}$*, associated with the sequence of random economies and equilibrium strategies,* $\{F^{(I)}, \sigma^{(I)}\}_{I\in\mathbb{N}}$*, is asymptotically stable if the fraction of students who are matched with their favorite feasible school in a random finite economy (*$F^{(I)}$*) converges to one, almost surely, or, equivalently,*

$$\lim_{I\to\infty} G^{(I)}\left(\{\theta_i \in \Theta \mid \mu^{(I)}(\theta_i) = \underset{s\in\mathcal{S}(e_i,P^{(I)})}{\arg\max}\ u_{i,s}\}\right) = 1, \textit{ almost surely.}$$

We are now ready to introduce our main result.

**Proposition 4.** *In the sequence of random economies and equilibrium strategies,* $\{F^{(I)}, \sigma^{(I)}\}_{I\in\mathbb{N}}$*, if Assumptions 1 and 2 are satisfied, then*

*(i) the random cutoffs converge to those of the stable matching in the continuum economy:* $\lim_{I\to\infty} P^{(I)} = P^\infty$*, almost surely;*

*(ii) the sequence of random matchings,* $\{\mu^{(I)}\}_{I\in\mathbb{N}}$*, is asymptotically stable.*

[16] Allowing $C(2) = 0$, Appendix A.2.4 investigates the properties of equilibrium strategies. The results, Lemmata A2–A4, imply strong restrictions on the sequence of Bayesian Nash equilibria in the direction of satisfying Assumption 2. Specifically, it is shown that a strategy that does not lead to $\mu^\infty$ in the continuum economy cannot survive as an equilibrium when $I \to \infty$. This immediately implies that in sufficiently large economies, every student includes in her ROL the school prescribed by $\mu^\infty$. Moreover, students do not pay a cost to rank more schools in large economies.

[17] We define the probability space, $(\Omega, \mathcal{F}, \mathbb{P})$. Specifically, $\Omega = \prod_{I\in\mathbb{N}} \Theta^I$, and an element in $\Omega$ is denoted by $\omega = (\omega_1, \omega_2, \cdots)$, where $\omega_I$ is a possible realization of student types in the random economy $F^{(I)}$. $\mathcal{F}$ is a Borel $\sigma$-algebra of $\Omega$, and $\mathbb{P}$ is a probability measure from $\mathcal{F}$ to $[0, 1]$.

Part (ii) implies that the fraction of students who are matched with their favorite feasible school, or not in any blocking pair, converges to one almost surely, as the economy grows. This provides justifications for the stability assumption in large markets.[18]

### 1.4.2 Probability of Being in a Blocking Pair for a Given Student

To assess if a matching is likely to be stable, we investigate how the probability that a given student is in a blocking pair changes with economy attributes. The following proposition shows how economy size, application cost, and other factors play a role.

**Proposition 5.** *Suppose student $i$ exists in all economies in the sequence $\{F^{(I)}\}_{I \in \mathbb{N}}$ which is associated with a sequence of Bayesian Nash equilibria in pure strategies $\{\sigma^{(I)}\}_{I \in \mathbb{N}}$.*

*(i) Let $\sigma^{(I)}(\theta_i) = L^{(I)}$; then $L^{(I)}$ is a partial order of $i$'s ordinal preferences. If ex post $i$ forms a blocking pair with $s$, $s$ must not be included in $L^{(I)}$, $s \notin L^{(I)}$.*

*The probability that $i$ is in a blocking pair with any school in the random matching $\mu^{(I)}$, denoted by $B_i^{(I)} = \Pr(\exists s \in \mathcal{S}, u_{i,s} > u_{i,\mu^{(I)}(\theta_i)},$ and $e_{i,s} \geqslant P_s^{(I)})$, satisfies:*

*(ii) $B_i^{(I)}$ is bounded above: $B_i^{(I)} \leqslant |\mathcal{S} \backslash L^{(I)}| \frac{C(|L^{(I)}|+1) - C(|L^{(I)}|)}{\max_{s \in \mathcal{S} \backslash L^{(I)}} u_{i,s}}$;*

*(iii) if $\{\sigma^{(I)}\}_{I \in \mathbb{N}}$ satisfies Assumptions 1 and 2, $B_i^{(I)}$ converges to zero almost surely.*

Because in equilibrium student $i$ reports a partial order of her true preferences, she can only form a blocking pair with a school that she did not rank (part i). Therefore, the probability that $i$ is in a blocking pair decreases whenever it is less costly to rank more schools (part ii). Together, Proposition 5 shows that stability is more plausible when the cost of ranking more schools is lower and/or the economy is large. Moreover, in the case of constrained/truncated DA where there is a limit on the length of ROLs, the higher the number of schools that can be ranked, the more likely stability is to be satisfied.

# 2 Empirical Approaches

Building on the theoretical results, we formalize the estimation of student preferences under different sets of assumptions and propose a series of tests to guide the selection of the appropriate approach. To be more concrete, we consider a logit-type random utility model, although our approaches can be extended to other specifications.

[18] This result, however, does not mean that the probability of a matching being stable converges to one as the market grows. As long as there is at least one blocking pair, a matching is not stable.

This section focuses on a random finite economy $F^{(I)}$ in which $I$ students compete for admissions to $S$ schools. Each school has a positive capacity, and students are assigned through a version of the student-proposing DA. Besides submitted ROLs and admission outcomes, the researcher observes priority indices, student characteristics, and school attributes. Given these observables, we discuss the probability of a student submitting a given ROL or being matched with a given school from the researcher's perspective.

## 2.1 Model Setting and Revealed Preferences

As is traditional and more convenient in empirical analysis, we let the student utility functions take any value on the real line.[19] With some abuse, we still use the same notation for utility functions. To facilitate the analysis, student $i$'s utility from attending schools $s$ is parameterized:

$$u_{i,s} = V_{i,s} + \epsilon_{i,s} = V(Z_{i,s}, \beta) + \epsilon_{i,s}, \tag{2}$$

where $V(\cdot, \cdot)$ is a known function, taking as arguments $Z_{i,s}$, a vector of observable student-school characteristics, and $\beta$, a vector of unknown parameters to be estimated; $\epsilon_{i,s}$ is the unobservable student heterogeneity.

We further define $Z_i \equiv \{Z_{i,s}\}_{s=1}^{S}$, and $\epsilon_i \equiv \{\epsilon_{i,s}\}_{s=1}^{S}$. It is assumed that $\epsilon_i \perp Z_i$ and that $\epsilon_{i,s}$ is i.i.d. over $i$ and $s$ with the type-I extreme value (Gumbel) distribution. Such a formulation rules out outside options, although this assumption can be relaxed.

We also assume that a student's preferences are not affected by other students' school assignments (no peer effects) and that statistics associated with the realized matching, such as cutoffs, do not enter the utility function. This is consistent with the theoretical model in Section 1 and implies that $Z_i$ does not include variables that depend on the ex-post observed matching.

The estimation relies on revealed student preferences in the data, and what information is revealed crucially depends on the imposed assumption—WTT, stability, or undominated stategies. Figure 1 shows an example. WTT takes the submitted ranking as truthful and assumes unranked schools being the least preferable. Stability dictates that a student is matched with her favorable feasible school. Lastly, a submitted ROL

[19] In the theoretical discussion, the utility functions are restricted to be in $[0, 1]$. One can use the inverse of standard normal distribution, $\Phi^{-1}$, to transform them to be on the real line. Note that the expected utility theory cannot be applied to the transformed utility functions; indeed, we do not.

| *Data* | |
|---|---|
| Available schools: | $s_1, \ldots, s_4$ |
| Student $i$'s priority indices: | $e_{i,s} = 0.5$, $\forall s$ |
| Submitted ROL of $i$: | $(s_1, s_3)$ |
| Admission outcome of $i$: | $s_3$ |
| Cutoffs: | $P = (0.8, 0.9, 0.4, 0.2)$ |

$\Longrightarrow$

| *Revealed preferences* | |
|---|---|
| WTT: | $u_{i,s_1} > u_{i,s_3} > u_{i,s}$, for $s = s_2, s_4$ |
| Stability: | $u_{i,s_3} > u_{i,s_4}$ |
| Undominated strategies: | $u_{i,s_1} > u_{i,s_3}$ |

**Figure 1:** Revealed Preferences under Different Assumptions—An Example

reveals the student's partial preference order if no one plays dominated strategies. We now detail how to use this information in the estimation.

## 2.2 Truth-Telling

In the literature on school choice with lotteries, some empirical approaches are based on the truth-telling assumption (Hastings et al., 2008; Abdulkadiroğlu et al., 2017). As similar mechanisms are commonly used in our strict-priority setting, we extend these approaches to our setting and clarify the assumptions embedded within.

We start with WTT instead of STT because in practice students in school choice and college admissions rarely rank the same number of choices (Abdulkadiroğlu et al., 2017; He, 2015; Artemov et al., 2017). Under the assumption of truth-telling without outside option, this can only be consistent with WTT but not STT, because STT requires everyone to rank all schools. We discuss STT with outside options in Appendix A.4.

For notational convenience, we make it explicit that student $i$'s type $\theta_i$ is described by $(u_i, e_i)$. Let $\sigma^W : \mathbb{R}^S \times [0,1]^S \to \mathcal{L}$ be a WTT pure strategy.[20] More precisely, the WTT assumption amounts to the following:

**Assumption (Characterization of Weak Truth-Telling).**
*WTT1. Suppose* $\sigma^W(u_i, e_i) = L = (l^1, \ldots, l^{K_i})$. *$L$ ranks $i$'s top $K_i$ preferred schools according to her true preferences:* $u_{i,l^1} > \cdots > u_{i,l^{K_i}} > u_{i,s'}$ *for all $s'$ not ranked in $L$;*
*WTT2. The number of schools ranked by a student is exogenous:* $u_i \perp |\sigma^W(u_i, e_i)|$, $\forall i$.

We are interested in the choice probability of $L$ conditional on observables, where the uncertainty from the researcher's perspective is due to the utility shocks $(\epsilon_i)$. Note that:

$$\begin{aligned}
&\Pr\left(\sigma^W(u_i, e_i) = L \;\middle|\; Z_i; \beta\right) \\
&= \Pr\left(\sigma^W(u_i, e_i) = L) \;\middle|\; Z_i; \beta; |\sigma^W(u_i, e_i)| = K\right) \times \Pr\left(|\sigma^W(u_i, e_i)| = K \;\middle|\; Z_i; \beta\right),
\end{aligned}$$

[20] Because the preference space is transformed from $[0,1]^S$ to $\mathbb{R}^S$, a strategy is now defined on the transformed type space. Moreover, it will be clear that $\sigma^W$ does not depend on priority indices, $e_i$.

which are calculated by integrating out the unobservables ($\epsilon_i$) in $u_i$. Assumption WTT2 implies that $\Pr\left(|\sigma^W(u_i, e_i)| = K \mid Z_i; \beta\right)$ does not depend on preferences. Thus, in the estimation, it suffices to focus on the following conditional probability:

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

where $s' \nsucc_L s$ indicates that $s'$ is not ranked before $s$ in $L$, including $s$ itself and those excluded from $L$. This rank-ordered (or "exploded") logit model can be seen as a series of conditional logits: one for the top-ranked school ($l^1$) being the most preferred, another for the second-ranked school ($l^2$) being preferred to all schools except $l^1$, and so on.

Let $|\sigma^W|$ be the vector of lengths of all submitted ROLs. The model can be estimated by maximum likelihood estimation (MLE) with the log-likelihood function:

$$\ln L_{TT}\left(\beta \mid \mathbf{Z}, |\sigma^W|\right) = \sum_{i=1}^{I} \sum_{s \in \sigma^W(u_i, e_i)} V_{i,s} - \sum_{i=1}^{I} \sum_{s \in \sigma^W(u_i, e_i)} \ln \Bigg( \sum_{s' \nsucc_{\sigma^W(u_i, e_i)} s} \exp(V_{i,s'}) \Bigg).$$

The estimator is denoted by $\hat{\beta}_{TT}$. Alternatively, the generalized method of moment (GMM) can be employed, for which the moment conditions are derived as in Section 2.5. WTT implies additional restrictions beyond standard discrete choice models (see details in Section 2.4). Thus, the discrete-choice literature (e.g., Matzkin, 1993) implies that student preferences are nonparametrically identified (also see Agarwal and Somaini, 2018).

## 2.3 Stability

We now assume that the matching is stable, which is different from, but in large samples justified by, asymptotic stability. The following analysis abstracts away from the matching mechanism and ignores how stability is obtained. We formulate a stable matching as the outcome of a discrete choice model and clarify the conditions that are needed for identification and estimation.

Consider the matching $\mu$ and the associated cutoffs $P(\mu)$, which are random variables determined by the unobserved utility shocks ($\epsilon$). $\mu$ is the outcome of a discrete choice model with personalized choice set, $\mathcal{S}(e_i, P(\mu))$ (i.e., the set of $i$'s feasible schools). The probability that $i$ is matched with $s$, or chooses $s$ in $\mathcal{S}(e_i, P(\mu))$, is:

$$\Pr\left(s = \mu(u_i, e_i) = \underset{s \in \mathcal{S}(e_i, P(\mu))}{\arg\max}\ u_{i,s} \;\middle|\; Z_i, e_i, \mathcal{S}(e_i, P(\mu)); \beta\right).$$

To proceed, we impose the following assumptions:

**Assumption (Exogeneity of Priority Index and Feasible Set).**

*EXO1. For all $i$, $e_i \perp \epsilon_i | Z_i$: Conditional on observables $Z_i$, student preferences and priority indices are independent.*

*EXO2. For all $i$ and $s$, $\mathbb{1}(e_{i,s} < P_s(\mu)) \perp \epsilon_i | Z_i$, or $\mathcal{S}(e_i, P(\mu)) \perp \epsilon_i | Z_i$: Conditional on observables $Z_i$, a student's preferences and her set of feasible schools are independent.*

Assumption EXO1 implies that, when priority indices ($e_i$) are determined by test scores, no student intentionally under-performs or over-performs in exams.

Assumption EXO2 deserves some discussion. Most importantly, it *does not require* that cutoffs $P(\mu)$ are conditionally independent of preferences shocks $\epsilon_i$. Instead, it only assumes that the personalized choice set, $\mathcal{S}(e_i, P(\mu))$, is exogenously given, which is necessary for identification in a discrete choice model with personalized choice sets. For instance, if instead $\mathcal{S}(e_i, P(\mu))$ is endogenous and only includes school $s$ when $s$ is $i$'s most preferred school, we lose the identification of $i$'s preferences, because there is no variation in $i$'s choice whenever $s$ is in $\mathcal{S}(e_i, P(\mu))$. Appendix A.5 details such an example, along with a discussion and an example in which the assumption is satisfied.

One may argue that, in a finite market, a student can affect some cutoffs by applying to a school or not, and thus can change the feasibility of some schools. Another concern is that given student preferences, there can be multiple stable matchings. If a single student can select among the stable matchings, Assumption EXO2 is also violated.

These concerns diminish as the economy grows large, because the potential influence on cutoffs by any student decreases and there tends to be a unique stable matching. For instance, Part (i) of Proposition 4 implies that a single student's impact on cutoffs diminishes to zero, almost surely. Moreover, even in small markets, Assumption EXO2 can be satisfied, because the assumption does not require $P(\mu) \perp \epsilon_i | Z_i$. An example is when every school ranks students in the same way, or $e_{i,s} = \overline{e}_i$ for all $s$ and $i$.[21]

[21] In this case, DA is equivalent to serial dictatorship in which students choose the remaining schools one by one in the order determined by their priority indices. There is a unique stable matching for each realization of student types. Moreover, the set of feasible schools for student $i$ is determined

Given the parametric assumptions on utility functions, the corresponding (conditional) log-likelihood function is:

$$\ln L_{ST}\left(\beta \mid \mathbf{Z}, \mathbf{e}, \mathcal{S}(e_i, P(\mu))\right) = \sum_{i=1}^{I}\sum_{s=1}^{S} V_{i,s} \times \mathbb{1}(\mu(u_i, e_i) = s) - \sum_{i=1}^{I} \ln\Big(\sum_{s' \in \mathcal{S}(e_i, P(\mu))} \exp(V_{i,s'})\Big). \quad (3)$$

This estimator is denoted by $\hat{\beta}_{ST}$. Similarly, GMM can be applied, as in Section 2.5.

**Identification.** The above discussion transforms the matching game into a discrete choice model.[22] Therefore, the nonparametric identification arguments for discrete choice models still apply (Matzkin, 1993). An important feature in the stability-based estimation is that students face *personalized* choice sets. As long as the choice sets are determined exogenously (Assumption EXO2), the identification goes through.

Another concern is that a student's priority index may enter her utility functions directly, when, for example, priority indices are determined by test score or student ability. In this case, the stability assumption does not reveal information about low-scoring students' preferences over popular schools, because such schools are often infeasible to them. This may lead to a failure of identifying how test scores determine student preferences.

This problem is mitigated if we have another measure of student ability, as in our empirical exercise. We assume that conditional on student ability, priority indices do not determine preferences and only affect school feasibility. If, additionally, priority indices have full support (i.e., can take any possible value) at each given level of student ability, we can observe some low-ability students having all schools feasible. This restores nonparametric identification in discrete choice models as in Matzkin (1993).

Relative to WTT, the stability assumption uses unambiguously less information from the data (see Figure 1 for an example). WTT utilizes all information implied by the submitted ROLs, while stability only imposes restrictions on admission outcome. One

by the students with higher priority indices. Because preferences are independent across students by assumption, we have $\mathcal{S}(e_i, P(\mu)) \perp \epsilon_i | Z_i$, or $\mathbb{1}(e_{i,s} < P_s(\mu)) \perp \epsilon_i | Z_i$ for all $s$.

It should be noted that $P(\mu) \not\perp \epsilon_i | Z_i$ even in this case. For example, when $i$ chooses $s$ among the feasible schools, the cutoff of $s$ will possibly increase; similarly, $i$ may decrease $s$'s cutoff by choosing a different school. However, we always have $\mathbb{1}(e_{i,s} < P_s(\mu)) \perp \epsilon_i | Z_i$, because $s$ will remain feasible to $i$ either way.

[22] A simplification is that we ignore the restrictions implied by the cutoffs $P(\mu)$, which may lead to efficiency loss in estimation. That is, even when the sets of feasible schools are exogenous to every single student's preferences, $P(\mu)$ is endogenously determined by the model's parameters. However, the additional information in these restrictions may be negligible, since we use the information on the whole matching already. An earlier version of the paper relaxes this assumption and uses the restrictions implied by the cutoffs. Our estimation results from simulated data and school choice data from Paris show that using the cutoff restrictions makes a negligible difference in the estimation results.

may expect that the stability-based approach leads to a loss of information; in particular, we may lose some precision in estimating the substitution patterns when we allow for more flexible random utility models (Berry et al., 2004; Abdulkadiroğlu et al., 2017). Indeed, as we shall see in our Monte Carlo simulations and the analysis of the school choice data from Paris, there is a clear bias-variance tradeoff: stability tolerates non-truth-telling behavior at the cost of yielding less precise estimates.

**Estimation with Asymptotic Stability.** When taking the above results to real-life data, one may be concerned that the matching may not be exactly stable. Indeed, our theoretical results only prove *asymptotic* stability. This raises the question of whether the estimator is still consistent. In Appendix A.3, we show that the MLE with asymptotic stability is consistent (Proposition A4). In a finite economy, the stability-based estimation is incorrectly specified, because some students may not be assigned to their favorite feasible school and their reveal preferences are mis-classified when stability is imposed. However, the fraction of students who are not assigned to their favorite feasible school converges to zero at an exponential rate (part iii of Proposition A2), implying that the mis-classification in revealed preferences vanishes with economy size. By verifying the conditions in Theorem 2.1 of Newey and McFadden (1994), we show that the stability-based estimator is consistent even when the matching is only asymptotically stable.

## 2.4 Testing Truth-Telling against Stability

Having two distinct estimators, $\hat{\beta}_{TT}$ and $\hat{\beta}_{ST}$, makes it possible to test the truth-telling assumption against stability. Maintaining the assumption of identification given stability, we shall see shortly that WTT provides over-identifying restrictions.

Before we present the tests, a few caveats are in order. First, one should check that the conditions for identification (for example, those in Matzkin, 1993) are satisfied before conducting the tests. Second, because the tests are essentially about joint restrictions on the parametric assumptions and the behavioral assumptions, one should be aware of the consequence of model misspecification. Rejecting truth-telling in favor of stability may not provide definitive proof against truth-telling, since the proposed tests are conditional on the model's parametric assumptions. In light of these limitations, it is often useful to provide additional empirical results, such as reduced-form results on student behavior (see, e.g., Section 4.2) and goodness of fit of the estimates (see, e.g., Section 4.4).

**Over-identifying Restrictions.** As summarized in Proposition 3, if every student is WTT and is assigned to a school, the matching is stable. Stability, however, does not imply that students are WTT and is therefore a less restrictive assumption.

To see the additional restrictions from WTT, let us consider student $i$ who submits a $K$-choice list $L$ and is matched with school $s$. Therefore, $s$ must be ranked in $L$. WTT implies the following conditions on the choice probability:

$$\begin{aligned}
&\Pr\left(\sigma^W(u_i, e_i) = L \;\middle|\; Z_i; \beta; |\sigma^W(u_i, e_i)| = K\right) \\
=&\Pr\left(u_{i,l^1} > \cdots > u_{i,l^K} > u_{i,s'}, \forall s' \in \mathcal{S}\backslash L \;\middle|\; Z_i; \beta; |\sigma^W(u_i, e_i)| = K; s = \operatorname*{arg\,max}_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s}\right) \\
&\times \Pr\left(s = \mu(u_i, e_i) = \operatorname*{arg\,max}_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s} \;\middle|\; Z_i; \beta; \mathcal{S}(e_i, P(\mu))\right) \qquad (4)
\end{aligned}$$

This equality uses the fact that the event, $\left(u_{i,l^1} > \cdots > u_{i,l^{K_i}} > u_{i,s'}, \forall s' \in \mathcal{S}\backslash L\right)$, implies $\left(s = \arg\max_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s}\right)$ but not the reverse.[23] This is because $i$'s feasible schools are either ranked below $s$ in $L$ or are omitted from $L$; in either case, WTT requires that $s$ is preferred to any other feasible school. Therefore, the first conditional probability on the right-hand side of the equality cannot always be one. As the restrictions implied by stability are just $\Pr\left(s = \mu(u_i, e_i) = \arg\max_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s} \mid Z_i; \beta; \mathcal{S}(e_i, P(\mu))\right)$, the additional restrictions from WTT are summarized in the first term. When the model is identified under stability, Equation (4) summarizes the over-identifying restrictions.

**Hausman Test.** Our estimator $\hat{\beta}_{TT}$ uses all the restrictions implied by WTT. Therefore, under the null hypothesis that students are WTT, both estimators $\hat{\beta}_{TT}$ and $\hat{\beta}_{ST}$ are consistent but only $\hat{\beta}_{TT}$ is asymptotically efficient. Under the alternative that the matching is stable but not all students are WTT, only $\hat{\beta}_{ST}$ is consistent.

In this setting, the general specification test developed by Hausman (1978) can be

[23] We also make use of the exogeneity of the set of feasible schools (Assumption EXO2) and the exogeneity of the length of submitted ROL (Assumption WTT2). Therefore,

$$\Pr\left(s = \mu(u_i, e_i) = \operatorname*{arg\,max}_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s} \;\middle|\; Z_i; \beta; \mathcal{S}(e_i, P(\mu))\right) = \Pr\left(s = \operatorname*{arg\,max}_{s \in \mathcal{S}(e_i, P(\mu))} u_{i,s} \;\middle|\; Z_i; \beta; |\sigma^W(u_i, e_i)| = K\right).$$

applied by computing the following test statistic:

$$T_H = (\hat{\beta}_{ST} - \hat{\beta}_{TT})'(\hat{\mathbf{V}}_{ST} - \hat{\mathbf{V}}_{TT})^{-1}(\hat{\beta}_{ST} - \hat{\beta}_{TT}),$$

where $(\hat{\mathbf{V}}_{ST} - \hat{\mathbf{V}}_{TT})^{-1}$ is the inverse of the difference between the asymptotic covariance matrices of $\hat{\beta}_{ST}$ and $\hat{\beta}_{TT}$.[24] Under the null hypothesis, $T_H \sim \chi^2(d_\beta)$, where $d_\beta$ is the dimension of $\beta$. If the model is correctly specified and the matching is stable, the rejection of the null hypothesis implies that WTT is violated in the data.

**Testing Over-identifying Restrictions.** The above Hausman test requires that we have a consistent and efficient estimator, $\hat{\beta}_{TT}$. When relying on MLE or GMM, this calls for strong parametric assumptions. An alternative is to construct a test for over-identifying restrictions (Hansen, 1982), which is made feasible because of the nesting structure of WTT and stability due to Proposition 3. Instead of requiring $\hat{\beta}_{TT}$ to be asymptotically efficient, the test for over-identifying restrictions only requires that $\hat{\beta}_{TT}$ utilizes more restrictions than $\hat{\beta}_{ST}$. With Equation (4), we can separate out the additional restrictions and test whether they are satisfied based on Hansen (1982).

**No-blocking among Assigned Students.** The above estimation and tests can be applied even if stability is violated. Part (i) of Proposition 3 states that whenever a WTT student is assigned, she is matched with her favorite feasible school and thus is not in any blocking pair. However, this no-blocking condition can be violated among unassigned students, implying the violation of stability. We can thus re-formulate the above tests as WTT against "no-blocking among assigned students." The estimation based on "no-blocking among assigned students" will exclude unassigned students; it does not create selection bias under the null hypothesis, because the length of every submitted ROL, which determines the probability of being unassigned, is exogenous under WTT.

## 2.5 Undominated Strategies and Stability

The stability-based approach described above is only valid when the matching is stable. However, as we have shown theoretically, stability can fail. Without stability, one may

[24]Since exact stability is assumed, the calculation of $\hat{\mathbf{V}}_{ST}$ does not take into account the sampling variance of cutoffs in a finite economy.

consider the undominated-strategies assumption, under which observed ROLs are students' true partial preference orders. That is, a submitted ROL, $L_i$, respects $i$'s true preference order among the schools ranked in $L_i$ (see, for an example, Figure 1).

These partial orders provide information that can be used to identify student preferences, but only partially, because the econometric structure is now incomplete (Tamer, 2003). In other words, for a student with type $(u_i, e_i)$, the assumption of undominated strategies does not predict a unique ROL for the student. As we shall see, undominated strategies lead to a set of inequality restrictions that can be satisfied by a set of $\beta$'s, instead of a unique vector of $\beta$. Therefore, we lose point identification.

**Moment Inequalities.** Students' submitted ROLs can be used to form *conditional moment inequalities*. Without loss of generality, consider two schools $s_1$ and $s_2$. Since not everyone ranks both schools, the probability of $i$, who adopts the strategy $\sigma(u_i, e_i)$, ranking $s_1$ before $s_2$, i.e., $s_1 \succ_{\sigma(u_i,e_i)} s_2$, is:

$$\begin{aligned}\Pr\left(s_1 \succ_{\sigma(u_i,e_i)} s_2 \mid Z_i; \beta\right) &= \Pr\left(u_{i,s_1} > u_{i,s_2} \text{ and } s_1, s_2 \in \sigma(u_i, e_i) \mid Z_i; \beta\right) \\ &\leqslant \Pr\left(u_{i,s_1} > u_{i,s_2} \mid Z_i; \beta\right) \end{aligned} \tag{5}$$

The first equality is due to undominated strategies, and the inequality defines a lower bound for the conditional probability of $u_{i,s_1} > u_{i,s_2}$. Similarly, an upper bound is:

$$\Pr\left(u_{i,s_1} > u_{i,s_2} \mid Z_i; \beta\right) \leqslant 1 - \Pr\left(s_2 \succ_{\sigma(u_i,e_i)} s_1 \mid Z_i; \beta\right). \tag{6}$$

Inequalities (5) and (6) yield the following conditional moment inequalities:

$$\begin{aligned}\Pr(u_{i,s_1} > u_{i,s_2} \mid Z_i; \beta) - E\left[\mathbb{1}(s_1 \succ_{\sigma(u_i,e_i)} s_2) \mid Z_i; \beta\right] &\geqslant 0; \\ 1 - E\left[\mathbb{1}(s_2 \succ_{\sigma(u_i,e_i)} s_1) \mid Z_i; \beta\right] - \Pr\left(u_{i,s_1} > u_{i,s_2} \mid Z_i; \beta\right) &\geqslant 0.\end{aligned}$$

Similar inequalities can be computed for any school pair and can be generalized to any $n$ schools in $\mathcal{S}$, for $2 \leqslant n \leqslant S$. In the simulations and empirical analysis, we focus on inequalities for pairs. The bounds become uninformative if $n \geqslant 3$, because not many schools are simultaneously ranked by the majority of students. We interact $Z_i$ with the conditional inequalities and obtain $M_1$ unconditional moment inequalities, $(m_1, \ldots, m_{M_1})$.

**Estimation with Moment Inequalities.** For estimation with moment inequalities, one can follow the approach of Andrews and Shi (2013), which is valid for both point and partial identifications. The objective function is a test statistic, $T_{MI}(\beta)$, of the Cramer-von Mises type with the modified method of moments (or sum function). With the unconditional moment inequalities, it is constructed as follows:

$$T_{MI}(\beta) = \sum_{j=1}^{M_1} \left[ \frac{\bar{m}_j(\beta)}{\hat{\sigma}_j(\beta)} \right]_-^2 \tag{7}$$

where $\bar{m}_j(\beta)$ and $\hat{\sigma}_j(\beta)$ are the sample mean and standard deviation of the $j$th moment, $m_j(\beta)$, respectively; and $[\ ]_-$ is such that $[a]_- = \min\{0, a\}$. One can then follow Bugni et al. (2017) to construct *marginal confidence intervals.* For a given coordinate $\beta_k$ of $\beta$, the authors test the hypothesis $H_0 : \beta_k = \beta_0$, for a given $\beta_0 \in \mathbb{R}$. The confidence interval for $\beta_k$'s true value is the convex hull of all $\beta_0$'s at which $H_0$ is not rejected.

While the assumption of undominated strategies seems plausible, it should be noted that the above approach often leads to uninformative confidence intervals of parameters of interest, constrained by the available econometric techniques. However, one can integrate the inequalities with the restrictions implied by stability, when stability is also plausible.

### 2.5.1 Integrating Stability with Undominated Strategies

An important advantage of the stability-based approach is that it only requires data on the admission outcomes. However, submitted ROLs are often observed and can be used to improve estimation efficiency. Under the assumption that stability provides point identification of student preferences, these ROLs provide over-identifying information that can be used together with stability in estimation.

The potential benefits can be illustrated in a simple example. Consider a constrained/truncated DA where students are only allowed to rank up to three schools out of four. With personalized sets of feasible schools under the stability assumption, the preferences over two schools, say $s_1$ and $s_2$, are estimated mainly from the sub-sample of students who are assigned to either of these schools while having priority indices above the cutoffs of both. Yet it is possible that all students include $s_1$ and $s_2$ in their ROLs, even if these schools are not ex post feasible for some students. In such a situation, all students could be used to estimate the preference order of $s_1$ and $s_2$, rather than just a sub-sample. As shown below, this argument can be extended to the case where two or more schools are observed being ranked by a subset of students.

**Moment Equalities.** To integrate the above over-identifying information in ROLs with that from stability, we reformulate the likelihood function described in Equation (3) into *moment equalities*. The choice probability of the matched school can be rewritten as a moment condition by equating theoretical and empirical probabilities:

$$\sum_{i=1}^{I} \Pr\Big(s = \arg\max_{s' \in \mathcal{S}(e_i,P)} (u_{i,s'}) \mid Z_i, P(\mu); \beta\Big) - E\Big(\sum_{i=1}^{I} \mathbb{1}(\mu(u_i, e_i) = s)\Big) = 0, \forall s \in \mathcal{S},$$

where $\mathbb{1}(\mu(u_i, e_i) = s)$ is an indicator function taking the value of one if and only if $\mu(u_i, e_i) = s$. We again interact the variables in $Z$ with the above conditions, leading to $M_2$ moment equalities, $(m_{M_1+1}, \ldots, m_{M_1+M_2})$.

**Estimation with Moment (In)equalities.** To obtain consistent point estimates with both equality and inequality moments (henceforth, moment (in)equalities), we augment the test statistic in Equation (7) to incorporate the $M_2$ unconditional moment equalities:

$$T_{MEI}(\beta) = \sum_{j=1}^{M_1} \left[\frac{\bar{m}_j(\beta)}{\hat{\sigma}_j(\beta)}\right]_-^2 + \sum_{j=M_1+1}^{M_1+M_2} \left[\frac{\bar{m}_j(\beta)}{\hat{\sigma}_j(\beta)}\right]^2. \tag{8}$$

We denote the point estimate $\hat{\beta}_{MEI}$, which minimizes $T_{MEI}(\beta)$, and we can take the same approach as in Bugni et al. (2017) to construct marginal confidence intervals for $\beta$.

## 2.6 Testing Stability against Undominated Strategies

Given the identification of student preferences under stability, the moment inequalities add over-identifying information. This constitutes a test of stability, provided that students do not play dominated strategies. More precisely, if both assumptions are satisfied, the moment (in)equalities in Section 2.5.1 should yield a point estimate that fits the data relatively well; otherwise, there should not exist a point $\beta$ that satisfies all moment (in)equalities. Formally, we follow the specification test in Bugni et al. (2015).

It should be noted that, for the above test, we maintain the undominated-strategies assumption, which may raise concerns, because students could make mistakes as documented in several real-life contexts; moreover, untrue partial preference ordering is not dominated under the school-proposing DA. We revisit these issues in Section 5.2.

The discussion in Section 2.5 provides another test of the undominated-strategies assumption, which also relies on the non-emptyness of the identified set under the null hypothesis (Bugni et al., 2015). That is, if there is no value of $\beta$ satisfying the moment inequalities, the undominated-strategies assumption is not satisfied. Unfortunately, the available methods of moment (in)equalities tend to result in conservative confidence sets of parameters, which implies that this test may lack power.

# 3 Results from Monte Carlo Simulations

To illustrate the proposed estimation approaches and tests, we carry out Monte Carlo simulations, the details of which are consigned to Appendix C.

Bayesian Nash equilibrium of the school choice problem is simulated in two settings where $I$ students compete for admission to 6 schools with per capita capacities $\{q_s\}_{s=1}^{6} = \{0.1, 0.1, 0.05, 0.1, 0.3, 0.3\}$.[25] The first is the **constrained/truncated DA** where students are allowed to rank up to $K$ schools ($K < 6$). The second setting, labelled as **DA with cost**, allows students to rank as many schools as they wish but imposes a constant marginal cost $c$ per additional school in the list after the first choice.

Student preferences over schools follow a random utility model:

$$u_{i,s} = \alpha_s - d_{i,s} + 3(a_i \cdot \bar{a}_s) + \epsilon_{i,s}, \tag{9}$$

where $\alpha_s$ is school $s$'s fixed effect; $d_{i,s}$ is the distance from student $i$'s residence to $s$; $a_i$ is $i$'s ability; $\bar{a}_s$ is school $s$'s quality; and $\epsilon_{i,s}$ is randomly drawn from the type-I extreme value distribution. Student priority indices are constructed such that (a) $i$'s priority index at each school is correlated with her ability $a_i$ (correlation coefficient 0.7) and (b) $i$'s priority indices at any two schools $s$ and $s'$ are also correlated (correlation coefficient 0.7).

Several lessons can be drawn from these simulations. The first is that in both settings,

[25] Appendix C.2 describes the details on solving equilibrium. In general, there are multiple equilibria. We focus on the one that is found by an algorithm iterating over the following steps: (i) for each candidate ROL (a true partial preference order of the schools), every student calculates the admission probability at each school by comparing her priority indices to the cutoff distribution; (ii) each student selects the ROL that maximizes her expected utility; (iii) the matchings across $M$ simulation samples jointly lead to an updated cutoff distribution; (iv) students update the admission probabilities based on the updated distribution. The initial cutoff distribution is the empirical cutoff distribution with strictly truth-telling students, and steps (i)–(iv) are repeated until a fixed point in the cutoff distribution is found.

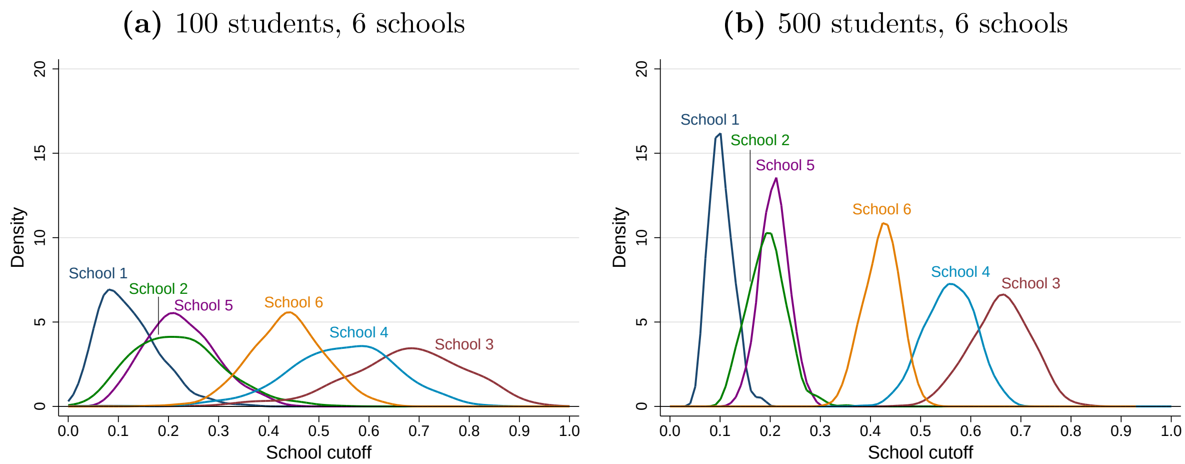


**Figure 2:** Monte Carlo Simulations: Impact of Economy Size on the Equilibrium Distribution of Cutoffs (Constrained/Truncated DA)

*Notes:* This figure shows the marginal distribution of school cutoffs in equilibrium under the constrained/truncated DA (ranking 4 out of 6 schools) when varying the number of students, $I$, who compete for admission to 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. See Appendix C for details on the Monte Carlo simulations.

the distribution of school cutoffs is close to jointly normal and degenerates as school capacities and the number of students increase proportionally while holding constant the number of schools (Figure 2); the matching is almost stable (i.e., almost every student is assigned to her favorite feasible school) even in moderately-sized economies. By contrast, WTT is often violated among the majority of the students, even when they can rank 4 out of 6 schools (constrained DA) or when the cost of including an extra school is negligibly small (DA with cost).[26] When the application cost increases, equilibrium strategies may prescribe that many students rank fewer than 6 schools even though they are allowed to rank all of them. Based on these results, observing that only a few students make full use of their ranking opportunities may not be viewed as a compelling argument in favor of truth-telling when the application cost is a legitimate concern.

The second insight is that stability leads to estimates much closer to the true values than WTT. Table 2 reports the results from estimation under each of the following assumptions: (i) weak truth-telling (columns 2–4); (ii) stability (columns 5–7); and (iii) stability and undominated strategies (columns 8–10). Panel A is for the constrained/truncated DA where students are allowed to rank up to 4 schools; Panel B is for the DA with a marginal application cost equal to $10^{-6}$.

[26] Consistent with Proposition 5, our simulations show that the fraction of students who are matched with their favorite feasible school decreases with the application cost. However, students with justified envy are rare unless students face very large application costs (see Figure C4 in the Appendix C.3).

**Table 2:** Monte Carlo Results (500 Students, 6 Schools, 500 Samples)

| | | Identifying assumptions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Weak Truth-telling | | | Stability of the matching | | | Stability and undominated strategies | | |
| | True value (1) | Mean (2) | SD (3) | CP (4) | Mean (5) | SD (6) | CP (7) | Mean (8) | SD (9) | CP (10) |
| *Panel A. Constrained/Truncated DA (ranking up to 4 out of 6 schools)* | | | | | | | | | | |
| Parameters | | | | | | | | | | |
| School 2 | 0.50 | −0.13 | 0.06 | 0.00 | 0.51 | 0.29 | 0.94 | 0.50 | 0.28 | 1.00 |
| School 3 | 1.00 | −2.08 | 0.14 | 0.00 | 1.05 | 0.58 | 0.96 | 1.02 | 0.57 | 1.00 |
| School 4 | 1.50 | −1.29 | 0.12 | 0.00 | 1.54 | 0.52 | 0.96 | 1.52 | 0.50 | 1.00 |
| School 5 | 2.00 | 0.56 | 0.07 | 0.00 | 2.02 | 0.31 | 0.96 | 2.01 | 0.29 | 1.00 |
| School 6 | 2.50 | 0.23 | 0.12 | 0.00 | 2.53 | 0.45 | 0.96 | 2.51 | 0.43 | 1.00 |
| Own ability × school quality | 3.00 | 9.40 | 0.64 | 0.00 | 2.97 | 2.29 | 0.96 | 3.05 | 2.26 | 1.00 |
| Distance | −1.00 | −0.71 | 0.08 | 0.08 | −1.01 | 0.20 | 0.95 | −1.01 | 0.20 | 1.00 |
| Summary statistics (averaged across Monte Carlo samples) | | | | | | | | | | |
| Average length of submitted ROLs | | | 4.00 | | | | | | | |
| Fraction of weakly truth-telling students | | | 0.39 | | | | | | | |
| 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 100% 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). | | | | | | | |
| *Panel B. DA with application cost (constant marginal cost $c = 10^{-6}$)* | | | | | | | | | | |
| Parameters | | | | | | | | | | |
| School 2 | 0.50 | 0.41 | 0.09 | 0.88 | 0.51 | 0.29 | 0.94 | 0.49 | 0.28 | 1.00 |
| School 3 | 1.00 | 0.57 | 0.16 | 0.23 | 1.05 | 0.58 | 0.96 | 1.00 | 0.53 | 1.00 |
| School 4 | 1.50 | 1.17 | 0.15 | 0.37 | 1.54 | 0.52 | 0.96 | 1.49 | 0.48 | 1.00 |
| School 5 | 2.00 | 1.74 | 0.11 | 0.32 | 2.02 | 0.30 | 0.96 | 1.99 | 0.29 | 1.00 |
| School 6 | 2.50 | 2.24 | 0.14 | 0.50 | 2.54 | 0.45 | 0.96 | 2.48 | 0.41 | 1.00 |
| Own ability × school quality | 3.00 | 2.19 | 0.72 | 0.77 | 2.96 | 2.29 | 0.96 | 3.16 | 2.29 | 1.00 |
| Distance | −1.00 | −0.93 | 0.09 | 0.88 | −1.01 | 0.20 | 0.95 | −1.00 | 0.20 | 1.00 |
| Summary statistics (averaged across Monte Carlo samples) | | | | | | | | | | |
| Average length of submitted ROLs | | | 4.60 | | | | | | | |
| Fraction of weakly truth-telling students | | | 0.79 | | | | | | | |
| 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 37% 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 set 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 two data generating processes for an economy in which 500 students compete for admission to 6 schools: a constrained/truncated DA where students are allowed to rank up to 4 schools out of 6 (Panel A); an unconstrained DA where students can rank as many schools as they wish, but incur a constant marginal cost $c = 10^{-6}$ for including an extra school in their ROL beyond the first choice (Panel B). Under assumptions (i) and (ii), student preferences are estimated using maximum likelihood estimation. Under assumption (iii), they are 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). See Appendix C for details on the Monte Carlo simulations.

The WTT-based estimator ($\hat{\beta}_{TT}$) is severely biased (column 2). Particularly in Panel A, we note that low-ability students' valuation of the most popular schools (e.g., School 6) tends to be underestimated, because such schools are more likely to be omitted from these students' ROLs due to their low admission probabilities. This bias is also present among small schools (e.g., Schools 3 and 4), which are often left out of ROLs because their cutoffs tend to be higher than those of equally desirable but larger schools.

By contrast, the average of the stability-based estimates ($\hat{\beta}_{ST}$) is reasonably close to the true values. Its standard deviations, however, are larger than those obtained under WTT. This efficiency loss is a direct consequence of ignoring the information content of ROLs.[27] The Hausman test strongly rejects WTT in favor of stability.

The estimator based on moment (in)equalities ($\hat{\beta}_{MEI}$), which integrates stability with information in ROLs, is also consistent (column 8). Moreover, the test based on moment (in)equalities never rejects the null hypothesis that stability is consistent with undominated strategies. A limitation of this approach, however, is that the currently available methods for conducting inference based on moment (in)equalities are typically conservative. As a result, the 95 percent marginal confidence intervals based on moment (in)equalities cover the true values too often (coverage probability, or CP, is close to one).

# 4 School Choice in Paris

Since 2008, the Paris Education Authority (*Rectorat de Paris*) assigns students to public high schools based on a version of the school-proposing DA called AFFELNET (Hiller and Tercieux, 2014). At the district level, student priority indices are not school-specific (as detailed below) and the mechanism is equivalent to a serial dictatorship.

Towards the end of the Spring term, final-year middle school students who are admitted to the upper secondary academic track (*Seconde Générale et Technologique*)[28] are requested to submit an ROL of up to 8 public high schools to the Paris Education Authority. Students' priority indices are determined as follows:

(i) Students' academic performance during the last year of middle school is graded on a scale of 400 to 600 points.

[27] Appendix C.4 further quantifies the efficiency loss in simulations with strictly truth-telling students.

[28] In the French educational system, students are tracked at the end of the final year of *collège* (equivalent to middle school), at the age of 15, into vocational or academic upper secondary education.

(ii) Paris is divided into four districts. Students receive a "district" bonus of 600 points at each school located in their home district. Thus, students applying to a within-district school have full priority over out-of-district applicants to the same school.

(iii) Low-income students are awarded a bonus of 300 points.[29] As a result, these students are given full priority over all other students from the same district.

The DA algorithm is run at the end of the academic year to determine school assignment for the following academic year. Unassigned students can participate in a supplementary round of admissions by submitting a new ROL of schools among those with remaining seats, the assignment mechanism being the same as for the main round.

Note that the mechanism would be strategy-proof if there were no constraints on the length of ROLs, because it is equivalent to serial dictatorship. Nonetheless, under the current mechanism, it is still a dominated strategy to submit an ROL that is not a partial order of true preferences (Proposition 2).

## 4.1 Data

For our empirical analysis, we use data from Paris' Southern District (*Sud*) and study the behaviors of 1,590 within-district middle school students who applied for admission to the district's 11 public high schools for the academic year 2013-14. Owing to the 600-point "district" bonus, this district is essentially an independent market.[30]

Along with socio-demographic characteristics and home addresses, our data contain all the relevant variables to replicate the matching algorithm, including the school capacities, the submitted ROLs, and the priority indices (converted into percentiles between 0 and 1). Individual examination results for the *Diplôme national du brevet* (DNB)—a national exam that all students take at the end of middle school—are used to construct different measures of academic ability (French, math, and composite score), which are normalized as percentiles between 0 and 1. Note that the DNB exam scores are not used in the computation of the student priority index, which is based on the grades obtained throughout the final year of middle school. The DNB scores therefore provide additional

[29] The low-income status is conditional on a student applying for and being granted the means-tested low-income financial aid in the last year of middle school. A family with two children would be eligible in 2013 if its taxable income was below 17,155 euros. The aid ranges from 135 to 665 euros per year.

[30] Out-of-district applicants could affect the availability of school seats in the supplementary round, but this is of little concern since, in the district, only 22 students were unassigned at the end of the main round (for the comparison between assigned and unassigned students, see Appendix Table E1).

measures of student ability.[31] Table 3 reports students' characteristics, choices, and admission outcomes. Almost half of the students are of high socioeconomic status (SES), while 15 percent receive the low-income bonus. 99 percent are assigned to a within-district school in the main admission round, but only half obtain their first choice.

Table 4 presents summary statistics for the 11 high schools. Columns 1–4 show a high degree of stratification among the schools, both in terms of the average ability of students enrolled in 2012 and of their social background (measured by the fraction of high SES students). Columns 5–8 describe school choice in 2013. The district's total capacity (1,692 seats) is unevenly distributed across schools: the smallest school has 62 seats while the largest has 251. School cutoffs in 2013 are strongly correlated with school quality. The last column shows the fraction of submitted ROLs in which each school is ranked. The least popular three schools are each ranked by less than 24 percent of students, and two of them remain under-subscribed (Schools 1 and 3) and thus have a zero cutoff. Consistent with our Monte Carlo results, smaller schools are omitted by more students, even if they are of high quality. Likewise, a sizeable fraction of students (20 percent) do not rank the best-performing school (School 11) in their ROLs.

Enrollment data further reveals a high level of compliance with the assignment outcome. Among the assigned students, 96 percent attend the school they were matched with (Appendix Table E1), about 1 percent attend a public high school different from their assignment school, and less than 3 percent opt out to enroll in a private school.

## 4.2 Evaluating the Assumptions: Reduced-Form Evidence

To evaluate the WTT and stability assumptions, we investigate if students are less likely to rank schools at which they expect low admission probabilities. Similar to "skipping the impossible" as in Example 1, this behavior would be inconsistent with WTT.

Figure 3 focuses on the district's four most selective schools (as measured by their cutoffs). For each school, we separately plot the fraction of students who rank it in their ROL as a function of their distance to the school cutoff, measured by the difference (using the original scale in points) between the student's priority index and the cutoff.[32]

[31] See Appendix B for a description of the data sources and Appendix Figure E1 for a map.

[32] We restrict the sample for a school to students whose score is no more than 50 points away from its cutoff. Due to the low-income bonus of 300 points, low-income students' priority indices are always well above the cutoffs. They are therefore not considered in the analysis.

**Table 3:** High School Applicants in the Southern District of Paris: Summary Statistics

| Variable | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|
| | | *Panel A. Student characteristics* | | | |
| Age | 15.0 | 0.4 | 13 | 17 | 1,590 |
| Female | 0.51 | 0.50 | 0 | 1 | 1,590 |
| French score | 0.56 | 0.25 | 0.00 | 1.00 | 1,590 |
| Math score | 0.54 | 0.24 | 0.01 | 1.00 | 1,590 |
| Composite score | 0.55 | 0.21 | 0.02 | 0.99 | 1,590 |
| High SES | 0.48 | 0.50 | 0 | 1 | 1,590 |
| With low-income bonus | 0.15 | 0.36 | 0 | 1 | 1,590 |
| | | *Panel B. Choices and outcomes* | | | |
| Number of choices within district | 6.6 | 1.3 | 1 | 8 | 1,590 |
| Assigned to a within-district school | 0.99 | 0.12 | 0 | 1 | 1,590 |
| Assigned to first choice school | 0.56 | 0.50 | 0 | 1 | 1,590 |
| | | *Panel C. Attributes of first-choice school* | | | |
| Distance (km) | 1.52 | 0.93 | 0.01 | 6.94 | 1,590 |
| Mean student French score | 0.62 | 0.11 | 0.32 | 0.75 | 1,590 |
| Mean student math score | 0.61 | 0.13 | 0.27 | 0.78 | 1,590 |
| Mean student composite score | 0.61 | 0.12 | 0.31 | 0.77 | 1,590 |
| Fraction high SES in school | 0.53 | 0.15 | 0.15 | 0.71 | 1,590 |
| | | *Panel D. Attributes of assigned school* | | | |
| Distance (km) | 1.55 | 0.89 | 0.06 | 6.94 | 1,568 |
| Mean student French score | 0.56 | 0.12 | 0.32 | 0.75 | 1,568 |
| Mean student math score | 0.54 | 0.14 | 0.27 | 0.78 | 1,568 |
| Mean student composite score | 0.55 | 0.13 | 0.31 | 0.77 | 1,568 |
| Fraction high SES in school | 0.48 | 0.15 | 0.15 | 0.71 | 1,568 |

*Notes:* This table provides summary statistics on the choices of middle school students from the Southern District of Paris who applied for admission to the district's 11 public high schools for the academic year starting in 2013, based on administrative data from the Paris Education Authority (*Rectorat de Paris*). 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 $[0, 1]$. The composite score is the average of the scores in French and math. The correlation coefficient between French and math scores is 0.50. School attributes, except distance, are measured by the average characteristics of students enrolled in each school in the previous year (2012).

**Table 4:** High Schools in the Southern District of Paris: Summary Statistics

| | School attributes (2012) | | | | Assignment outcomes (2013) | | | |
|---|---|---|---|---|---|---|---|---|
| School | Mean French score (1) | Mean math score (2) | Mean composite score (3) | Fraction high SES students (4) | Capacity (5) | Count (6) | Admission cutoffs (7) | Fraction ROLs ranking it (8) |
| School 1 | 0.32 | 0.31 | 0.31 | 0.15 | 72 | 19 | 0.000 | 0.22 |
| School 2 | 0.36 | 0.27 | 0.32 | 0.17 | 62 | 62 | 0.015 | 0.23 |
| School 3 | 0.37 | 0.34 | 0.35 | 0.16 | 67 | 36 | 0.000 | 0.14 |
| School 4 | 0.44 | 0.35 | 0.39 | 0.46 | 140 | 140 | 0.001 | 0.59 |
| School 5 | 0.47 | 0.44 | 0.46 | 0.47 | 240 | 240 | 0.042 | 0.83 |
| School 6 | 0.47 | 0.46 | 0.46 | 0.32 | 171 | 171 | 0.069 | 0.71 |
| School 7 | 0.58 | 0.54 | 0.56 | 0.56 | 251 | 251 | 0.373 | 0.91 |
| School 8 | 0.58 | 0.66 | 0.62 | 0.30 | 91 | 91 | 0.239 | 0.39 |
| School 9 | 0.65 | 0.62 | 0.63 | 0.66 | 148 | 148 | 0.563 | 0.83 |
| School 10 | 0.68 | 0.66 | 0.67 | 0.49 | 237 | 237 | 0.505 | 0.92 |
| School 11 | 0.75 | 0.78 | 0.77 | 0.71 | 173 | 173 | 0.705 | 0.80 |

*Notes:* This tables provides summary statistics on the attributes of high schools in the Southern District of Paris and on the outcomes of the 2013 assignment round, based on administrative data from the Paris Education Authority (*Rectorat de Paris*). School attributes in 2012 are measured by the average characteristics of the schools' enrolled students in 2012–2013. 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 $[0, 1]$. The composite score is the average of the scores in French and math. The correlation coefficient between school-average math and French scores is 0.97.

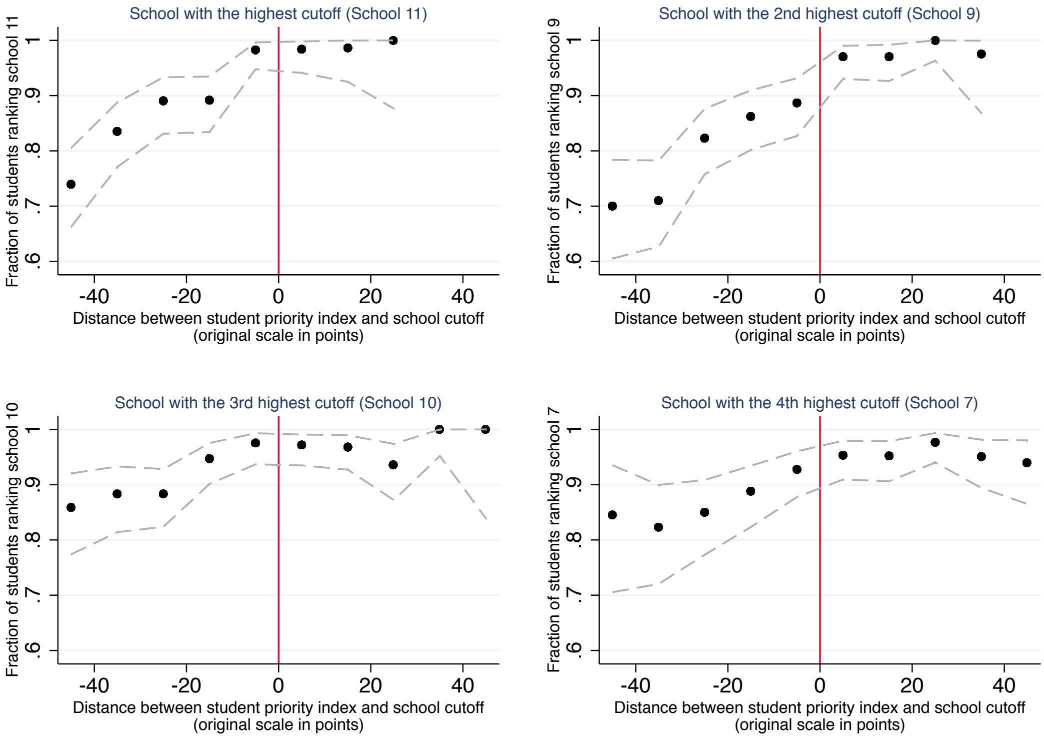


**Figure 3:** Fraction of Students Ranking Each of the Four Most Selective Schools in the Southern District of Paris, by Distance to School Cutoff

*Notes:* The results are calculated with data from the Paris Education Authority on students who applied to the 11 high schools of the Southern District in 2013. The figure shows the ranking behavior of students as a function of the distance (using the original scale in points) between each school's cutoff and students' priority index. For each school, the sample only includes students with a priority index within $-50$ and $+50$ points of the cutoff, and students are grouped into bins of 10-point width. Bins with less than 10 observations are excluded. Each point represents the fraction of students in a given bin who rank the school in their list. The dotted lines show the 95 percent confidence interval. Low-income students are not included because the low-income bonus of 300 points places them well above the cutoffs.

Each plot shows that almost all students with a priority index above a school's cutoff include that school in their ROL, whereas the fraction of students ranking the school decreases rapidly when the priority index falls below the cutoff. Irrespective of strategic considerations, one might expect high priority students to have a stronger preference for the most selective schools—since priorities are positively correlated with academic performance—and hence to rank them more often. However, the kink around the cutoffs is consistent with students omitting the most selective schools from their ROL because of the low admission probabilities. In Appendix D.1, we show that the kink-shaped relationship between student priority index and their ranking behavior is robust to controlling for potential determinants of preferences, including distance to school and the student's DNB exam scores in French and math. Recall that DNB scores are not used to calculate the priority indices. These results cannot be easily reconciled with truth-telling behavior.

The evidence in Figure 3 suggests the potential influence of expected admission probabilities on student ranking behavior. At the time of application, students know their academic grades and low-income status but not their priority ranking nor the ex post cutoffs.[33] They can, however, gather information on past cutoffs to assess admission probabilities. While we do not have direct information on students' beliefs, Figure 4 shows that the current year (2013) cutoffs are similar to those from the previous year (2012).[34] This lends support to the assumption that students have some ability to predict their admission probabilities. Although not a necessary condition for the matching to be stable, this feature makes the stability assumption more likely to be satisfied.

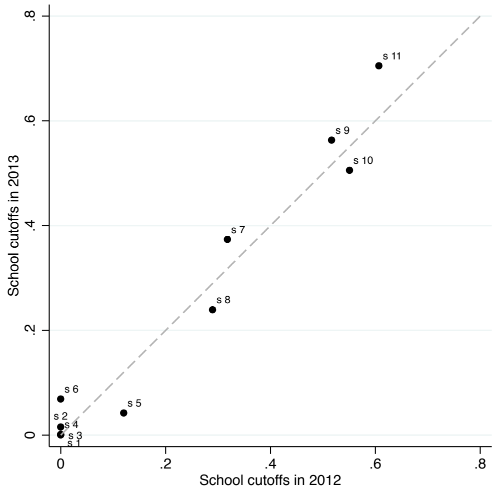


**Figure 4:** School Cutoffs in 2012 and 2013

*Notes:* The results are calculated with administrative data from the Paris Education Authority. Each dot represents a school, with its cutoff in 2013 on the Y-axis and the one in 2012 on the X-axis. The dashed line denotes the 45-degree line.

## 4.3 Estimation and Test Results

We parameterize student $i$'s utility of being matched with school $s$ as follows:

$$u_{i,s} = \alpha_s - d_{i,s} + Z'_{i,s}\gamma + \lambda\epsilon_{i,s}, \; s = 1, \dots, 11; \tag{10}$$

[33]This uncertainty in both priority ranking and cutoffs may explain why some students find it optimal to rank multiple schools, given that the cost of ranking up to 8 choices is arguably negligible.

[34]The comparison could not be performed for earlier years due to the modifications in the computation of the priority index and the small changes in the set of available schools.

where $\alpha_s$ is the school fixed effect, $d_{i,s}$ is the distance to $s$ from $i$'s residence, and $Z_{i,s}$ is a vector of student-school-specific observables. As observed heterogeneity, $Z_{i,s}$ includes two variables that capture potential non-linearities in the disutility of distance and control for potential behavioral biases towards certain schools: "closest school" is a dummy variable equal to one if $s$ is the closest to student $i$ among all 11 schools; "high school co-located with middle school" is another dummy that equals one if high school $s$ and the student's middle school are co-located at the same address.[35] To account for students' heterogeneous valuation of school quality, interactions between student scores and school scores are introduced separately for French and math, as well as an interaction between own SES and the fraction of high SES students in the school. These school attributes are measured among the entering class of 2012, whereas our focus is on students applying for admission in 2013. We normalize the variables in $Z_{i,s}$ so that each school's fixed effect can be interpreted as the mean valuation, relative to School 1, of a non-high-SES student who has median scores in both French and math, whose middle school is not co-located with that high school, and for whom the high school is not the closest to her residence.

The error term $\epsilon_{i,s}$ is assumed to be an i.i.d. type-I extreme value, and the variance of unobserved heterogeneity is $\lambda^2$ multiplied by the variance of $\epsilon_{i,s}$. The effect of distance is normalized to $-1$, and, therefore, the fixed effects and $\gamma$ are all measured in terms of willingness to travel. As a usual position normalization, $\alpha_1 = 0$. We do not consider outside options because of students' almost perfect compliance with the assignment outcome.

Using the same procedures as in the Monte Carlo simulations (described in Appendix C), we obtain the results summarized in Table 5, where each column reports estimates under a given set of identifying assumptions: (i) weak truth-telling (column 1); (ii) stability (column 2); and (iii) stability with undominated strategies (column 3).[36]

The results provide clear evidence that the WTT-based estimates (column 1) are rather different from the others. Specifically, a downward bias is apparent for popular

[35] There are five such high schools in the district.

[36] For the estimates in column 3, we use the method of moment (in)equalities where inequalities are constructed as described in Section 2.5. Determined by $Z_{i,s}$, we interact French score, math score, and distances to Schools 1 and 2 with the conditional moments. Although one could use more variables, e.g., SES status and distance to other schools, they provide little additional variation. In principle, the assumption of undominated strategies alone implies partial identification (Section 2.5). Because stability is not rejected by our test, we do not present results based on this approach (available upon request). We note that the marginal confidence intervals from moment inequalities only turn out to be wide in our empirical setting, and hence are relatively uninformative. The possible reasons are that the empirical bounds for the probability of a preference ordering over a pair of schools are fairly wide, and that the available methods to conduct inference based on moment inequalities are typically conservative.

**Table 5:** Estimation Results under Different Sets of Identifying Assumptions

| | Identifying assumptions | | |
|---|---|---|---|
| | Weak Truth-telling | Stability of the matching | Stability and undominated strategies |
| | (1) | (2) | (3) |
| | *Panel A. School fixed effects* | | |
| School 2 | −0.71 | 1.46 | 1.21 |
| | [−1.17, −0.24] | [0.64, 2.28] | [0.14, 2.29] |
| School 3 | −2.12 | 1.03 | 0.84 |
| | [−2.66, −1.58] | [0.19, 1.86] | [−0.56, 2.01] |
| School 4 | 3.31 | 2.91 | 2.90 |
| | [2.75, 3.86] | [2.07, 3.76] | [2.36, 3.39] |
| School 5 | 5.13 | 4.16 | 4.16 |
| | [4.41, 5.84] | [3.22, 5.10] | [3.71, 4.49] |
| School 6 | 4.87 | 4.24 | 4.30 |
| | [4.21, 5.54] | [3.29, 5.18] | [3.73, 4.82] |
| School 7 | 7.33 | 6.81 | 6.24 |
| | [6.47, 8.18] | [5.65, 7.98] | [5.76, 7.28] |
| School 8 | 1.59 | 4.46 | 4.27 |
| | [1.10, 2.08] | [3.46, 5.47] | [2.98, 5.26] |
| School 9 | 6.84 | 7.77 | 6.57 |
| | [6.07, 7.61] | [6.55, 8.99] | [5.84, 7.26] |
| School 10 | 7.84 | 7.25 | 6.44 |
| | [6.94, 8.75] | [6.01, 8.49] | [5.87, 7.05] |
| School 11 | 5.35 | 7.28 | 5.61 |
| | [4.62, 6.08] | [6.06, 8.51] | [4.98, 7.33] |
| | *Panel B. Covariates* | | |
| Closest school | −0.37 | −0.19 | −0.15 |
| | [−0.63, −0.11] | [−0.47, 0.10] | [−0.75, 0.57] |
| High school co-located | 2.54 | 1.76 | 1.54 |
| with middle school | [2.02, 3.07] | [1.19, 2.32] | [0.17, 3.12] |
| Student French score [×10] | 0.20 | 0.18 | 0.23 |
| × school French score [×10] | [0.16, 0.23] | [0.13, 0.24] | [0.10, 0.35] |
| Student math score [×10] | 0.30 | 0.27 | 0.30 |
| × school math score [×10] | [0.26, 0.34] | [0.21, 0.32] | [0.18, 0.40] |
| High SES | 6.79 | 4.92 | 8.12 |
| × fraction high SES in school | [5.62, 7.97] | [3.31, 6.54] | [4.18, 12.55] |
| Scaling parameter | 3.09 | 1.33 | 1.50 |
| | [2.79, 3.38] | [1.16, 1.50] | [1.20, 1.64] |
| Number of students | 1,590 | 1,568 | 1,590 |

*Notes:* This table reports the estimates of the parameters in Equation (10) for the Southern District of Paris, with the coefficient on distance being normalized to −1. The point estimates in columns 1 and 2 are based on maximum likelihood, whereas those in column 3 are based on moment equalities and inequalities, with 95 percent confidence intervals in brackets. *Model selection tests*: A Hausman test, testing (1) against (2), rejects WTT in favor of stability (p-value < 0.01); a test based on moment equalities and inequalities does not reject the null hypothesis that stability is consistent with undominated strategies at the 5 percent significance level.

schools that are not ranked by many students, such as Schools 8 and 11. School 8, which is omitted by 61 percent of students, is deemed by WTT to be less desirable than all the schools included in the ROL, which leads to a low estimated fixed effect. Similarly, the fixed effect estimate of School 11, one of the most popular schools, varies substantially across the identifying assumptions. The under-estimation is mitigated when the model is estimated under a different assumption (columns 2 and 3). Provided that the model is correctly specified, the Hausman test rejects WTT in favor of stability (p-value $< 0.01$); the test based on moment (in)equalities does not reject the null hypothesis that stability is consistent with undominated strategies at the 5 percent significance level.

The results show that "closest school" has no significant effect, but students significantly prefer co-located schools. Compared with low-score students, those with high French (math) scores have a stronger preference for schools with higher French (math) scores. Moreover, high SES students prefer schools that have admitted a larger fraction of high SES students in the previous year (2012).

Although the WTT-based estimates of the coefficients of covariates (Panel B) are not markedly different from the stability-based estimates, one cannot conclude that the WTT assumption produces reasonable results, as shown by the estimates of fixed effects. To provide a better evaluation, we now compare the estimators by their model fit.

## 4.4 Goodness of Fit

In three dimensions (cutoffs, assignment, and revealed preferences), we compare the observed values to those predicted by the estimates from Table 5. This comparison reveals that the stability-based estimates fit the data well, as opposed to those based on WTT (see Appendix D.2 for computational details).

Specifically, Figure 5 and Appendix Table D2 show that the stability-based estimates (with or without undominated strategies) predict cutoffs close to the observed ones.[37] By contrast, WTT substantially under-predicts the cutoffs of the most popular schools.

Panel A of Table 6 compares each student's predicted assignment to the observed one. The stability-based estimates have 33–38 percent success rates, whereas the WTT-based estimates accurately predict only 22 percent of the assignments. In Panel B, we take as

[37] It should be emphasized that the stability-based estimation does not try to fit cutoffs directly, neither does it restrict a student's preferences over infeasible schools. The difference in predicted cutoffs between stability and WTT is solely due to their differences in predicting preferences.

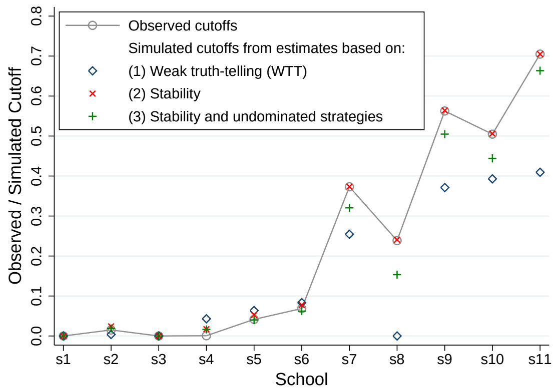


**Figure 5:** Goodness of Fit: Observed vs. Simulated Cutoffs

*Notes:* This figure compares the cutoffs observed for the 11 high schools of Paris' Southern District in 2013 to those simulated with the three sets of estimates in Table 5. The simulated cutoffs are averaged over 300 simulated samples. See Appendix D.2 for details.

given the schools that a student has included in her submitted ROL, and compute the probability of observing this particular preference order among the ranked schools. The observed order of students' top two choices has a mean predicted probability of 60 or 62 percent based on the stability-based estimates, higher than the 55 percent achieved by the WTT-based estimates. We next consider the observed order of a student's full list of choices. Again, the stability-based estimates outperform those based on WTT, with an average predicted probability between 2.2 and 2.5 percent for the former versus 1.2 percent for the latter. The predictive power of the stability-based estimates along the two measures in Panel B is noteworthy because the prediction is partly out of sample.[38]

# 5 Summary and Discussion

As a summary of the results, we clarify when each approach is more appropriate for empirical analysis. We also discuss whether the results can be extended to the school-proposing DA, the case with non-equilibrium behavior, and settings beyond school choice.

## 5.1 Choosing among the Approaches: A Summary

In the preference estimation with real-life data from centralized school choice and college admissions, some practical considerations should be taken into account. Recall that we

[38]In the data, 54 percent of students ranked at least one infeasible school among their top two choices (34 percent ranked one infeasible school, while 20 percent ranked two). The average fraction of infeasible schools among all submitted choices is 30 percent.

**Table 6:** Goodness-of-Fit Measures Based on Different Sets of Identifying Assumptions

| | Estimates from | | |
|---|---|---|---|
| | Weak Truth-telling | Stability of the matching | Stability and undominated strategies |
| | (1) | (2) | (3) |
| *Panel A. Simulated vs. observed assignment (300 simulated samples)* | | | |
| Mean predicted fraction of students assigned to observed assignment | 0.220 | 0.383 | 0.326 |
| | (0.011) | (0.010) | (0.012) |
| *Panel B. Predicted vs. observed partial preference order of given schools* | | | |
| Mean predicted probability that a student prefers the top-ranked school to the 2nd-ranked in her submitted ROL | 0.553 | 0.615 | 0.595 |
| Mean predicted probability that a student's partial preference order among the schools in her ROL coincides with the submitted rank order | 0.012 | 0.025 | 0.022 |

*Notes:* This table reports two sets of goodness-of-fit measures comparing the observed outcomes to those predicted under the different sets of identifying assumptions as in Table 5, for the high school assignment of students in the Southern District of Paris. Panel A compares students' observed assignment with their predicted assignment in 300 simulated samples. In all simulations, we vary only the utility shocks, which are kept common across columns 1–3 (see Appendix D.2 for details). Predicted and observed assignments are compared by computing the average predicted fraction of students who are assigned to their observed assignment school, with standard deviations across the simulation samples reported in parentheses; in other words, this is the average fraction of times each student is assigned to her observed assignment in the 300 simulated samples. Panel B uses two measures to compare students' observed partial preference order of given schools (revealed in their submitted ROL) with the prediction, among students who rank at least two schools: (i) 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) mean predicted probability that a student's partial preference order among the schools in her ROL coincides with the submitted rank order. Due to the logit specification, those probabilities can be calculated without simulation.

focus on the strict-priority setting in which students are ranked based on strict priority indices that are ex ante known privately. Building on the results from our theoretical and empirical analyses, this section emphasizes some of the key market features that deserve careful examination when one decides which approach to use in a given context.

**The Nesting Structure of Identifying Assumptions.** Our results imply that the identifying assumptions follow a nesting structure, as depicted in Figure 6.

Truth-telling is a natural candidate identifying assumption because of DA's strategy-proofness. However, strict truth-telling (i.e., students truthfully rank all schools) is not an equilibrium, if students cannot rank all schools at no cost (Proposition 1). In real-life data, students seldom rank all schools, which calls for a weaker version of the truth-telling assumption. As clarified in the theoretical analysis, weak truth-telling (i.e., students truthfully rank their most preferred schools and omit some least preferred ones) does not follow directly from strategy-proofness, as it requires additional assumptions such as the length of ROLs being independent of preferences.

Stability is an even weaker assumption on student behavior, while still allowing for the

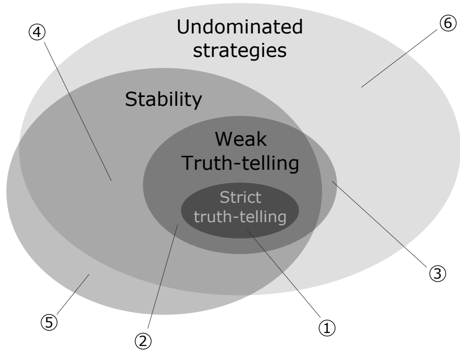


**Figure 6:** Nesting Structure of Identifying Assumptions

*Notes:* This figure shows the nesting structure of the identifying assumptions that can be used to analyze data generated by DA and its variants in the strict-priority setting. The numbered areas correspond to different combinations of identifying assumptions: ① strict truth-telling; ② weak truth-telling and stability; ③ weak truth-telling without stability; ④ stability and undominated strategies; ⑤ stability without undominated strategies; ⑥ undominated strategies without stability.

identification of preferences. It states that every student is assigned to her favorite ex post feasible school, and is always satisfied when students are strictly truth-telling. Although stability is not guaranteed in all Bayesian Nash equilibria, even when students are weakly truth-telling, it is asymptotically satisfied when the economy grows large (Proposition 4).

The third candidate identifying assumption is that students do not play dominated strategies (Proposition 2), so that submitted ROLs reveal students' partial preference orders of schools. Weak truth-telling is a special case of this more general assumption, whereas stability may hold even if students play dominated strategies.

**The Choice of Empirical Approaches.** When choosing among the candidate identifying assumptions, consideration should be given to the features of the problem under study, as well as the available data. For each assumption, Table 7 summarizes the features making it more plausible, the required data, and some discussion about identification and estimation. Truth-telling is more likely to be satisfied when students can rank as many schools as they wish at no cost, and face large uncertainty about each school's exact ranking of students. Conditional on students' submitted ROLs being observed, preferences can be estimated using either MLE or GMM. The choice between weak truth-telling and strict truth-telling depends on whether students rank all schools (Section 2.2) and on the importance of outside options (Appendix A.4).

When students face some cost of ranking more schools (e.g., if the length of submittable ROLs is restricted), stability can be a more plausible assumption than truth-telling.

**Table 7:** Summary of Empirical Approaches

| Identifying assumption | What makes the assumption more plausible? | Required data | Identification and estimation |
|---|---|---|---|
| **Weak Truth-Telling**: Schools in a submitted ROL are ranked in true preference order and omitted ones are less preferred. | (1) No cost of ranking more schools, e.g., no restriction on the length of submittable ROLs and choice set not being too large. (2) At the time of application, each student knows her own priority index but not others', and the distribution of priority indices has a large variance. | Submitted ROLs | Point identification<br><br>Estimation by, e.g., MLE/GMM |
| **Stability of the matching**: Every student is assigned to her favorite feasible school. Priority indices and unobserved preference heterogeneity are conditionally independent. | Stability is satisfied if truth-telling holds and (almost) everyone is assigned. Otherwise, it is more likely to be true when (1) market is large (many students, big schools); (2) students are less constrained when applying to more schools; (3) students face limited uncertainty about how schools rank them at the time of application; (4) students know more about others' preferences; or (5) cutoffs are easy to predict. | Admission outcome, school capacities, priority indices | Point identification<br><br>Estimation by, e.g., MLE/GMM |
| **Undominated strategies**: Submitted ROLs are true partial preference orders. | (1) No "safety school" so that "irrelevance at the bottom" of one's ROL is less likely. (2) No "impossible school" so that students do not rank impossible school arbitrarily. | Submitted ROLs | Partial identification<br><br>Estimation using moment inequalities |
| **Stability and Undominated strategies**: See the conditions laid out separately for stability and undominated strategies. | See the conditions laid out separately for stability and undominated strategies. | Submitted ROLs, admission outcome, school capacities, priority indices | Point identification<br><br>Estimation using moment equalities and moment inequalities |

*Notes:* This table describes the empirical approaches to analyses of data generated by DA and its variants in the strict-priority setting. In addition, there are two tests available: (i) weak truth-telling can be tested against stability ($H_0$: both weak truth-telling and stability are satisfied; $H_1$: only stability is satisfied), e.g., using the Hausman (1978) or Hansen (1982) tests; (ii) stability can be tested against undominated strategies ($H_0$: both stability and undominated strategies are satisfied; $H_1$: only the undominated-strategies assumption is satisfied) using the approach in Bugni et al. (2015).

This assumption is more likely to hold when the market is larger (i.e., many students and many seats per school), when students are less constrained in applying to multiple schools (e.g., longer ROLs), when they are less uncertain about each school's ranking of all students at the time of application, when they know more about others' preferences, or when it is easier for them to predict school cutoffs (Proposition 5). Our Monte Carlo simulations additionally provide numerical evidence suggesting that stability is a plausible assumption even when students face non-negligible application costs (Appendix C.3).

Estimating preferences based on stability uses information on the admission outcome, the school capacities, and the priority indices, but has the advantage of not requiring data on submitted ROLs. However, it is necessary to assume the conditional independence between priority index and unobserved preference heterogeneity. Compared to truth-telling, the main cost of the stability-based approach is its limited power to identify rich substitution patterns, because the information content of ROLs is discarded.

Weak truth-telling does not always imply stability, but it does imply no-blocking among all *assigned* students (Proposition 3). Therefore, weak truth-telling can be tested against stability (or no-blocking among assigned students) using the Hausman (1978) and Hansen (1982) tests. It should be emphasized that these tests do not provide definitive proof against truth-telling unless the model is correctly specified and identified.

If it is believed that neither truth-telling nor stability is likely to be satisfied, preferences can still be partially identified under the assumption that students do not play dominated strategies. This assumption is more plausible when no school is either "safe" or "impossible" for students, making it less likely that students rank some schools in an arbitrary manner. Submitted ROLs can then be used to form conditional moment inequalities that partially identify preferences.

When the conditions for both stability and undominated-strategies assumptions are jointly satisfied, the moment inequalities from the latter assumption provide over-identifying information that can be integrated with the stability assumption to estimate preferences based on all of the available data (ROLs, matching outcome, school capacities, and priority indices). Additionally, the stability assumption can be tested against the undominated-strategies assumption using the specification test in Bugni et al. (2015).

## 5.2 Discussion and Extension

**The School-Proposing DA.** Our main results can be extended to the school-proposing DA, which is also commonly used in practice (see Table 1). Under this mechanism, schools "propose" to students following the order of student priority indices. Proposition 2 no longer holds; that is, students might have incentives not to report a true partial preference order (Haeringer and Klijn, 2009). Nonetheless, the asymptotic stability result (Proposition 4) is still valid, as its proof does not rely on Proposition 2. Indeed, the matching can be stable in equilibrium (Haeringer and Klijn, 2009).

To summarize, if the market under the school-proposing DA has features making the matching stable (see Table 7), we can formulate identification and estimation of student preferences based on stability. However, the truth-telling assumption no longer has theoretical support, as the school-proposing DA is not strategy-proof for students (Roth, 1982). The approach based on undominated strategies does not apply either, since there are no dominated strategies under this mechanism (Haeringer and Klijn, 2009).

**Non-Equilibrium Strategies.** We have thus far assumed that everyone plays an equilibrium strategy with a common prior. More realistically, some students could have different information and make mistakes when strategizing.

Indeed, a growing number of studies find that strategic mistakes are common even in strategy-proof environments. Laboratory experiments show that a significant fraction of subjects do not report their preferences truthfully in strategy-proof mechanisms (Chen and Sönmez, 2006). More relevantly, mistakes occur in real-world contexts, e.g., the admissions to Israeli graduate programs in psychology (Hassidim et al., 2016), the medical resident match in the U.S. (Rees-Jones, 2018), and the Australian university admissions (Artemov et al., 2017). Without estimating preferences, these studies show that a non-negligible fraction of participants make unambiguous mistakes in their ROLs.

However, the vast majority of these mistakes are not payoff relevant. In other words, although some students play dominated strategies, the matching is still close to stable, corresponding to area 5 in Figure 6. Based on these observations, the results in Artemov et al. (2017) imply that, as identifying restrictions, assuming stability can be more robust and more plausible than the assumption of undominated strategies.

**Beyond School Choice and College Admissions.** Although the analysis has focused on school choice and college admissions, our results can apply to certain assignment procedures based on DA. Let us call agents on the two sides "applicants" and "recruiters," respectively. The key requirement is that when applying, applicants have sufficiently precise information on how recruiters rank them and that researchers observe how recruiters exactly rank applicants.[39] Examples include the assignment of teachers to schools in France (Combe et al., 2016) and the Scottish Foundation Allocation Scheme matching medical school graduates with training programs (Irving, 2011). The estimation approaches discussed in Section 2 could be implemented in these settings.

# 6 Conclusion

We present novel approaches to estimating student preferences with school choice or college admissions data generated by the popular Deferred Acceptance mechanism when applicants are ranked strictly by some ex-ante known priority index. We provide theoretical and empirical evidence showing that, in this commonly observed setting, it can be rather restrictive to assume that students truthfully rank schools when applying for

[39] Without information on how either side ranks the other, it becomes the classical two-sided matching, and additional assumptions are needed for identification and estimation (Chiappori and Salanié, 2016).

admission. Instead, stability (or justified-envy-freeness) of the matching provides rich identifying information, while being a weaker assumption on student behavior. Assuming that students do not play dominated strategies, we also discuss methods with moment inequalities, which can be useful with or without stability. A series of tests are proposed to guide the selection of the appropriate approach.

The estimation and testing methods are illustrated with Monte Carlo simulations. When applied to school choice data from Paris, our results are more consistent with stability than with truth-telling. Reduced-form evidence on ranking behavior suggests that some students omit the most selective schools from their list because of low admission probabilities. Provided that the model is correctly specified, our proposed tests reject truth-telling but not stability. Compared with our preferred estimates based on stability (with or without imposing undominated strategies), assuming truth-telling leads to an under-estimation of preferences for popular or small schools. Moreover, the stability-based estimators outperform the truth-telling-based estimator in predicting matching outcomes and student preferences.

Our approaches are applicable to many school choice and college admissions systems around the world, as well as to other matching schemes such as teacher assignment in France and medical matching in Scotland.

# References

**Abdulkadiroğlu, Atila and Tayfun Sönmez**, "School Choice: A Mechanism Design Approach," *American Economic Review*, 2003, *93* (3), 729–747.

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

**Abdulkadiroğlu, Atila, Joshua D. Angrist, and Parag A. Pathak**, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 2014, *82* (1), 137–196.

**_ , Nikhil Agarwal, and Parag A. Pathak**, "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match," *American Economic Review*, 2017, *107* (12), 3635–89.

**Agarwal, Nikhil**, "An Empirical Model of the Medical Match," *American Economic Review*, 2015, *105* (7), 1939–1978.

**_ and Paulo Somaini**, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 2018, *86* (2), 391–444.

**Ajayi, Kehinde**, "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," 2017. IED Discussion Paper 259.

**Akyol, Şaziye Pelin and Kala Krishna**, "Preferences, Selection, and Value Added: A Structural Approach," *European Economic Review*, 2017, *91*, 89–117.

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

**Artemov, Georgy, Yeon-Koo Che, and Yinghua He**, "Strategic 'Mistakes': Implications for Market Design Research," 2017. Manuscript.

**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.

**Balinski, Michel and Tayfun Sönmez**, "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 1999, *84*, 73–94.

**Berry, Steven, James Levinsohn, and Ariel Pakes**, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

**Biró, Péter**, "University Admission Practices—Hungary," 2011. MiP Country Profile 5; matching-in-practice.eu.

**Bucarey, Alonso**, "Who Pays for Free College? Crowding Out on Campus," 2018. Manuscript.

**Budish, Eric and Estelle Cantillon**, "Strategic Behavior in Multi-Unit Assignment Problems: Theory and Evidence from Course Allocation," *American Economic Review*, 2012, *102* (5), 2237–71.

**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.

**Burgess, Simon, Ellen Greaves, Anna Vignoles, and Deborah Wilson**, "What Parents Want: School Preferences and School Choice," *The Economic Journal*, 2014, *125* (587), 1262–1289.

**Calsamiglia, Caterina, Chao Fu, and Maia Güell**, "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives," 2014. Barcelona GSE Working Paper No. 811.

**Carvalho, José-Raimundo, Thierry Magnac, and Qizhou Xiong**, "College Choice Allocation Mechanisms: Structural Estimates and Counterfactuals," 2014. IZA Discussion Paper No. 8550.

**Chen, Li**, "University Admission Practices—Ireland," 2012. MiP Country Profile 8; matching-in-practice.eu.

**_ and Juan Sebastián Pereyra**, "Self-selection in School Choice," 2017. Manuscript.

**Chen, Yan and Tayfun Sönmez**, "School Choice: An Experimental Study," *Journal of Economic theory*, 2006, *127* (1), 202–231.

**Chiappori, Pierre-André and Bernard Salanié**, "The Econometrics of Matching Models," *Journal of Economic Literature*, 2016, *54* (3), 832–861.

**Combe, Julien, Olivier Tercieux, and Camille Terrier**, "The Design of Teacher Assignment: Theory and Evidence," 2016. Manuscript.

**De Haan, Monique, Pieter A. Gautier, Hessel Oosterbeek, and Bas Van der Klaauw**, "The Performance of School Assignment Mechanisms in Practice," 2015. IZA Discussion Paper No. 9118.

**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.

**Fox, Jeremy T.**, "Structural Empirical Work Using Matching Models," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, 2009.

**Gale, David E. and Lloyd S. Shapley**, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69* (1), 9–15.

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

**Hansen, Lars Peter**, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

**Hassidim, Avinash, Assaf Romm, and Ran I. Shorrer**, "'Strategic' Behavior in a Strategy-proof Environment," 2016. Manuscript.

**Hastings, Justine S., Christopher A. Neilson, and Seth D. Zimmerman**, "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile," 2013. NBER Working Paper No. 19241.

**_ , _ , and _** , "The Effects of Earnings Disclosure on College Enrollment Decisions," 2015. NBER Working Paper No. 21300.

**_ , Thomas Kane, and Douglas Staiger**, "Heterogeneous Preferences and the Efficacy of Public School Choice," 2008. Manuscript.

**Hausman, Jerry A.**, "Specification Tests in Econometrics," *Econometrica*, 1978, *46* (6), 1251–1271.

**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.

**Hiller, Victor and Olivier Tercieux**, "Choix d'Écoles en France: une Évaluation de la Procedure Affelnet," *Revue Économique*, 2014, *65*, 619–656.

**Hwang, Sam Il Myoung**, "A Robust Redesign of High School Match," 2014. Manuscript.

**Irving, Rob**, "Matching Practices for Entry-Labor Markets—Scotland," 2011. MiP Country Profile 3; matching-in-practice.eu.

**Kapor, Adam, Christopher A. Neilson, and Seth Zimmerman**, "Heterogeneous Beliefs and School Choice," 2016. Manuscript.

**Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad**, "Field of Study, Earnings, and Self-Selection," *Quarterly Journal of Economics*, 2016, *113* (3), 1057–1111.

**Kirkebøen, Lars**, "Preferences for Lifetime Earnings, Earnings Risk and Monpecuniary Attributes in Choice of Higher Education," 2012. Statistics Norway Discussion Papers No. 725.

**Kojima, Fuhito**, "Recent Developments in Matching Theory and their Practical Applications," 2015. Manuscript.

**Luflade, Magaux**, "The value of Information in Centralized School Choice Systems," 2018. Manuscript.

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

**Milgrom, Paul R. and Robert J. Weber**, "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, 1985, *10* (4), 619–632.

**Moon, Hyungsik R. and Frank Schorfheide**, "Estimation with Overidentifying Inequality Moment Conditions," *Journal of Econometrics*, 2009, *153* (2), 136–154.

**Mora, Ricardo and Antonio Romero-Medina**, "Understanding Preference Formation in a Matching Market," 2001. Universitad Carlos III de Madrid Working Paper No. 01-59 (Economics Series).

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

**Pathak, Parag A. and Peng Shi**, "Demand Modeling, Forecasting, and Counterfactuals, Part I," 2014. NBER Working Paper No. 19859.

_ **and Tayfun Sönmez**, "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation," *American Economic Review*, 2013, *103* (1), 80–106.

**Pop-Eleches, Cristian and Miguel Urquiola**, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2013, *103* (4), 1289–1324.

**Rees-Jones, Alex**, "Suboptimal Behavior in Strategy-Proof Mechanisms: Evidence from the Residency Match," *Games and Economic Behavior*, 2018, *108*, 317–330.

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

**Salonen, Mikko A. A.**, "Matching Practices for Secondary Schools—Finland," 2014. MiP Country Profile 19; matching-in-practice.eu.

**Saygin, Perihan O.**, "Gender Differences in Preferences for Taking Risk in College Applications," *Economics of Education Review*, 2013, *52*, 120–133.

**Tamer, Elie**, "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 2003, *70* (1), 147–165.

_ , "Partial Identification in Econometrics," *Annual Review of Economics*, 2010, *2* (September), 167–195.

**Teo, Chung-Piaw, Jay Sethuraman, and Wee-Peng Tan**, "Gale-Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications," *Management Science*, 2001, *47* (9), 1252–1267.

**UAC (University Admission Committee, Taiwan)**, *University Admission Guide* (in Chinese), 2014. `www.uac.edu.tw/103data/103recruit.pdf` (accessed 09/06/2017).

(For Online Publication)

Appendix to

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

Gabrielle Fack Julien Grenet Yinghua He

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 1.2 and 1.3

**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

[A.1] The weak-* convergence of measures is defined as $\int X d\hat{G}^{(I)} \to \int X dG$ 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.

$(E, \sigma^{\infty})$, 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.

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:

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

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,

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

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 $e_{i,s} < P_s^{\infty}$.

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 $G^{(I)}\left(\{\theta_i \in \Theta : \mu^{(I)}(\theta_i) \notin \arg\max_{s \in \mathcal{S}(e_i, P^{(I)})} u_{i,s}\}\right)$ 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:

$$
\begin{aligned}
&C\left(\left|L_i^{(I)}\right| + 1\right) - C\left(\left|L_i^{(I)}\right|\right) \\
\geqslant &\sum_{t=1}^{k^*-1} 0 + B_{i,s}^{(I)} u_{i,s}
\end{aligned}
$$

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

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 $\gamma_1 e^{-\gamma_2 I}$:*

$$\Pr\left(||P^{(I)} - P^\infty|| > \eta\right) < \gamma_1 e^{-\gamma_2 I}.$$

*(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 2.3).

**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]

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

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

### 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 $F^{(I_n)} \to E$ 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 $P^{(I_n)} \to P^{\sigma}$ 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 $P^{(I_n)} \to P^{\sigma}$ 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.*

*(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)

[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.

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 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* $\tilde{\beta}_I = \arg\max_{\beta \in \mathcal{B}} \widehat{Q}_I(\beta|P^{\infty})$.

[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.

*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,* $\sup_{\beta\in\mathcal{B}}|\hat{Q}_I(\beta|P^{(I)})-\hat{Q}_I(\beta|P^{\infty})|\xrightarrow{p}0$.

*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,

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

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

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 $G(\Theta_\delta)E_{(Z,e)}\left(\sup_{\substack{\beta\in\mathcal{B},s\in\mathcal{S},\\ P\in\mathcal{P}(e_i,s)}}\left|\ln\frac{\Pr\left(s=\arg\max_{s'\in\mathcal{S}(e_i,P^{(I)})} u_{i,s'}|Z_i,e_i,\mathcal{S}(e_i,P^{(I)});\beta\right)}{\Pr\left(\mu^\infty(u_i,e_i)=\arg\max_{s\in\mathcal{S}(e_i,P^\infty)} u_{i,s}|Z_i,e_i,\mathcal{S}(e_i,P^\infty);\beta\right)}\right| \mid (u_i,e_i)\in\Theta_\delta\right)$, 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 $G(\Theta_\delta)E_{(Z,e)}\left(\sup_{\substack{\beta\in\mathcal{B},s\in\mathcal{S},\\ P\in\mathcal{P}(e_i,s)}}\left|\ln\frac{\Pr\left(s=\arg\max_{s'\in\mathcal{S}(e_i,P^{(I)})} u_{i,s'}|Z_i,e_i,\mathcal{S}(e_i,P^{(I)});\beta\right)}{\Pr\left(\mu^\infty(u_i,e_i)=\arg\max_{s\in\mathcal{S}(e_i,P^\infty)} u_{i,s}|Z_i,e_i,\mathcal{S}(e_i,P^\infty);\beta\right)}\right| \mid (u_i,e_i)\in\Theta_\delta\right)<\eta_1/2$. 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., $\left\|P^{(I)}-P^\infty\right\|\leqslant\delta$), $(**)$ 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 2.2 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:

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

[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.

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.

**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 2.1. Student $i$'s utility from attending school $s$ is specified as follows:

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

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.

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

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

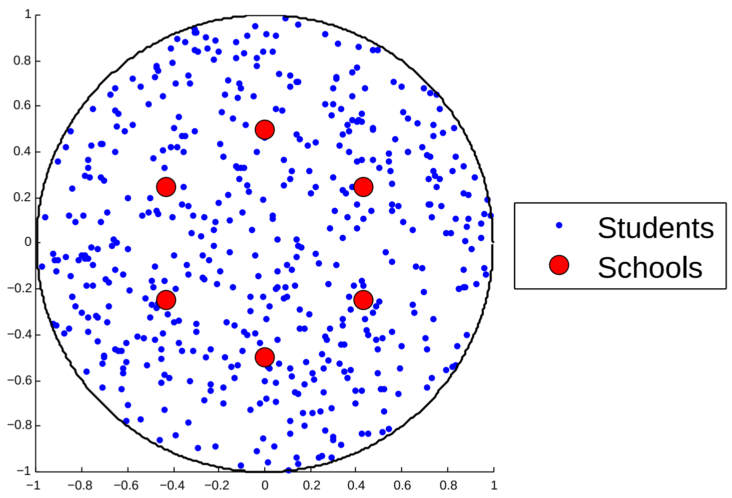


**Figure C1:** Monte Carlo Simulations: Spatial Distribution of Students and Schools

*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.

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 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

[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.

(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).

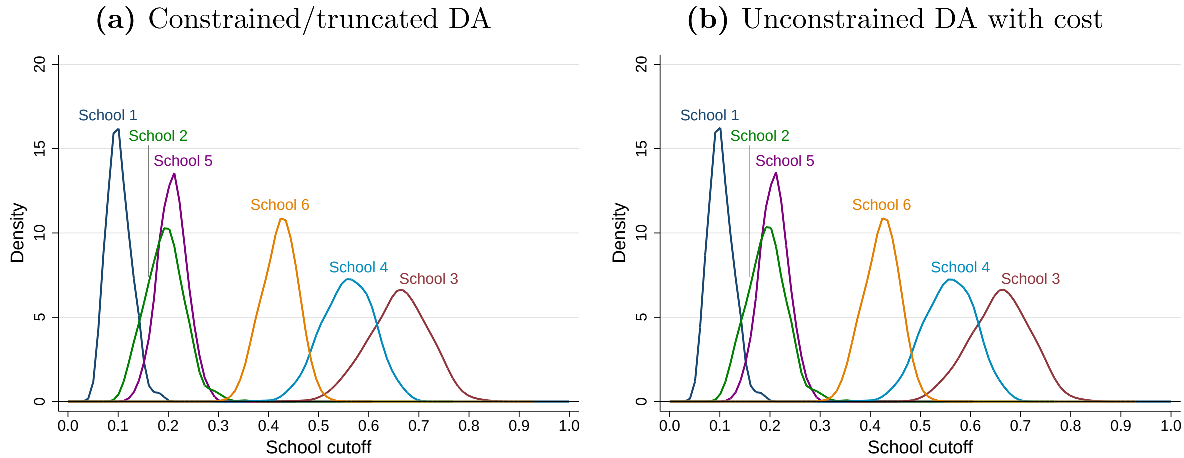


**Figure C2:** Monte Carlo Simulations: Equilibrium Distribution of School Cutoffs (6 schools, 500 students)

*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.

[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 $a_i \cdot \bar{a}_s$.

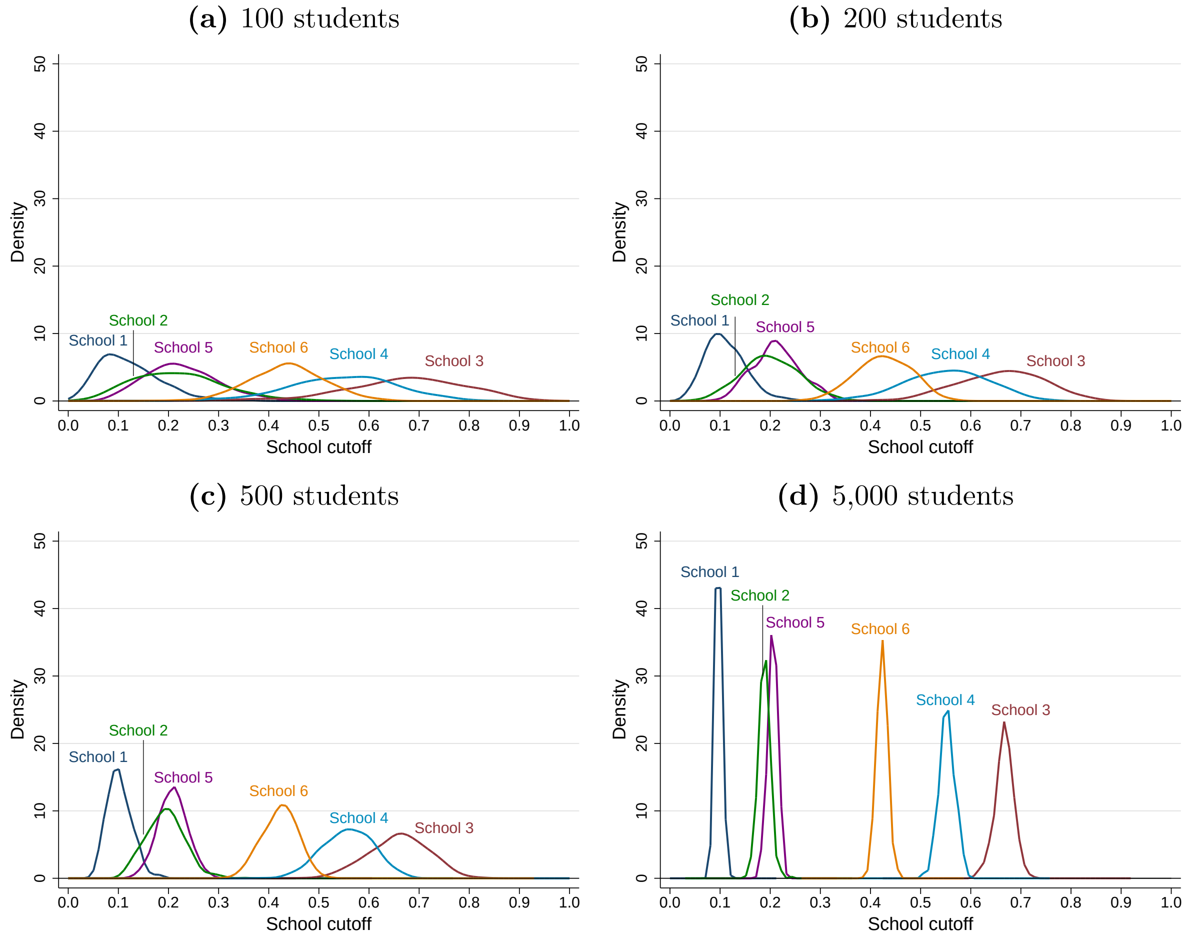


**Figure C3:** Monte Carlo Simulations: Impact of Economy Size on the Equilibrium Distribution of Cutoffs (Constrained/Truncated DA)

*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.

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.

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

| | Data generating process | |
|---|---|---|
| | Constrained/truncated DA (1) | Unconstrained DA with cost (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.

**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]

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

[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).

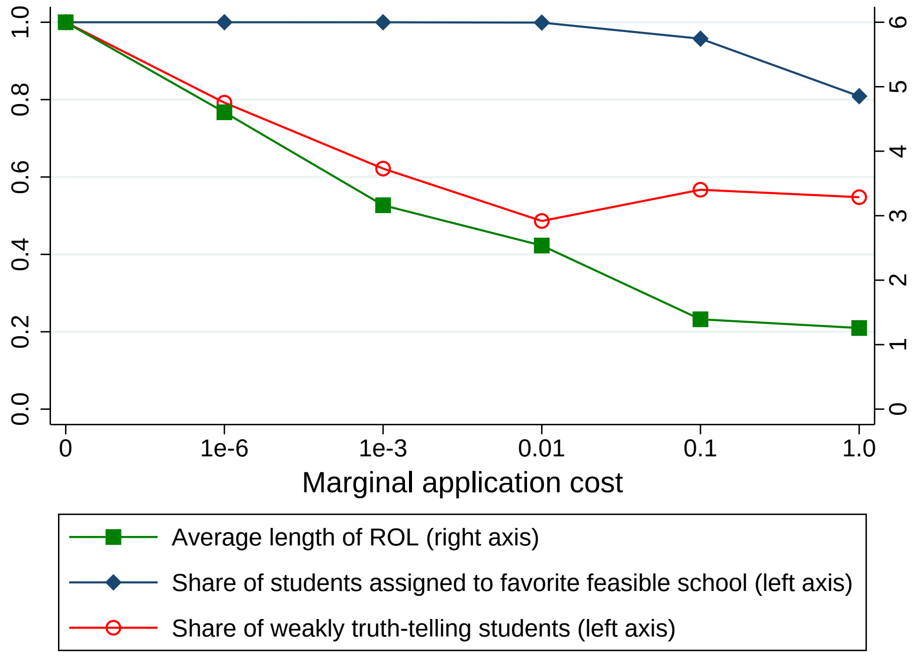


**Figure C4:** Monte Carlo Simulations: Impact of the Marginal Cost of Applying to Schools on Equilibrium Outcomes (500 Students, 6 Schools)

*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.

line). Consistent with the predictions from Section 1.4.2, 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 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).

**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.

[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 2.5). 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.

**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 (1) | Mean (2) | SD (3) | CP (4) | Mean (5) | SD (6) | CP (7) | Mean (8) | SD (9) | CP (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 4.2 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 4.4 in the main text).

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

The reduced-form evidence presented in Section 4.2 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 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.

[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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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

| | | Cutoffs in simulated samples with estimates from | | |
|---|---|---|---|---|
| | Observed cutoffs | 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

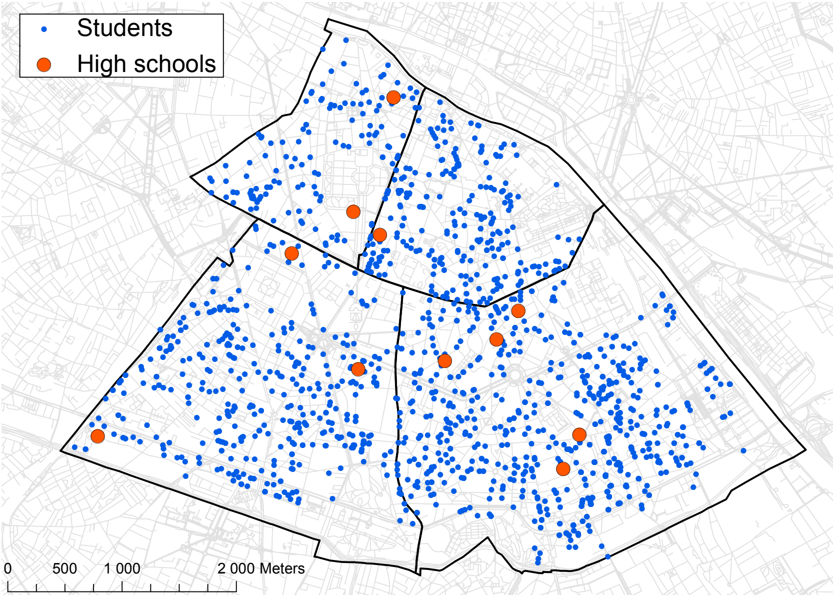


**Figure E1:** The Southern District of Paris for Public High School Admissions

*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.