Leveraging Uncertainties to Infer Preferences: Robust Analysis of School Choice

Yeon-Koo Che* Dong Woo Hahm† YingHua He‡

August 1, 2023

Abstract

We study the method for inferring applicants' preferences from school-choice data robustly to payoff-insignificant mistakes that have been documented. The uncertainties applicants face due to the use of lotteries or other sources make certain mistakes more costly than others, thus making some preferences more inferable than others. We propose a novel method that exploits the structure of the uncertainties present in a deferred-acceptance matching environment to reveal applicants' preferences robustly as well as informationally efficiently. We apply our method to estimate the preferences of high-school applicants in Staten Island using New York City's high-school assignment data. The counterfactual analysis of proposed desegregation reforms suggests that their effects can be underestimated if one fails to account for the applicants' inattention/mistakes in preference estimation.

JEL Classification Numbers: C70, D47, D61, D63.
Keywords: School choice, Strategic mistakes, Demand estimation

1 Introduction

The crucial first step in the empirical study of school choice is to infer student preferences from data. The inferences form a basis for demand estimation and enable one to perform a counterfactual analysis of possible policy interventions using the estimates. In principle, the use of strategyproof mechanisms such as the deferred acceptance algorithm, now increasingly common for school choice assignments, should make the inference straightforward, as it gives the dominant-strategy incentive for applicants to reveal their preferences truthfully (Dubins and Freedman, 1981; Roth, 1982; Abdulkadiroglu and Sonmez, 2003). However, recent literature found evidence that a significant fraction of participants makes strategic "mistakes" by not reporting their preferences truthfully even

*Department of Economics, Columbia University, USA. Email: yeonkooche@gmail.com.
†Department of Economics, University of Southern California, USA. Email: dongwooh@usc.edu.
‡Department of Economics, Rice University, USA. Email: yinghua.he@rice.edu.

in strategyproof environments.[1] The vast majority of these mistakes are payoff-irrelevant: they typically involve applicants neglecting to rank, or ranking incorrectly, popular schools that are so out-of-reach for them that even if they had correctly ranked them they would not have obtained them.[2] Yet, even such payoff-insignificant mistakes could bias preference estimates if applicants' rank-order lists (ROLs) are interpreted at face value. For instance, Weak Truth-Telling (WTT), a popular hypothesis that interprets applicants' ROLs as truthful representations of their most preferred choices,[3] would incorrectly infer elite schools as unpopular when many students find them out-of-reach and neglect to rank them on their ROLs.

Fack, Grenet, and He (2019) and Artemov, Che, and He (2023) develop theories for accommodating such mistakes and suggest a method for inferring preferences that are robust to payoff-irrelevant mistakes. They argue that applicants make payoff-irrelevant mistakes in the presence of either application costs or inattention to (small) payoff losses, but nonetheless, virtually all of them are assigned stably, namely, the schools they prefer most among those that are feasible to them given their priority standings. This stability hypothesis proves more robust to mistakes than the truth-telling hypothesis according to simulations, and the estimates based on the former fit admission outcomes and student preferences better than those based on the latter. By now, the method has been adopted by many papers (e.g., Otero, Barahona, and Dobbin, 2021; Combe, Tercieux, and Terrier, 2022; Combe, Dur, Tercieux, Terrier, and Ünver, 2022; Hahm and Park, 2022).

However, the method and its theoretical justification provided by the preceding papers are limited to the environment in which the applicants face no uncertainty about their priorities. In practice, students face priority uncertainties from a number of sources. The biggest source is the lottery used commonly in school assignments. Lotteries are used in the vast majority of the US public school

[1]These mistakes were documented in lab settings (Chen and Sönmez, 2006; Li, 2017; Chen and Kesten, 2019) and high-stake real-world contexts such as school applications (Larroucau and Rios, 2020; Artemov, Che, and He, 2021; Hassidim, Romm, and Shorrer, 2021; Arteaga, Kapor, Neilson, and Zimmerman, 2022; Shorrer and Sóvágó, 2023), and medical resident matching (Rees-Jones, 2017; Rees-Jones and Skowronek, 2018). For instance, Hassidim, Romm, and Shorrer (2021) find in their analysis of Israeli postgraduate psychology program admissions that approximately 19% of students either did not list a scholarship position for a program or ranked a non-scholarship position higher than the corresponding scholarship position against their interests. Artemov, Che, and He (2021) and Shorrer and Sóvágó (2023) find similar findings in college admissions in Australia and Hungary, respectively. Rees-Jones (2017) reports that 17% of the 579 surveyed US medical seniors indicate misrepresenting their preferences in the National Resident Matching Program.

[2]For instance, Artemov, Che, and He (2021) find that even though 35% of the applicants mistakenly forgo lower-tuition options in their application for colleges in Victoria, Australia, only 1-20% of such mistakes had real consequences since such options were "out of reach" for them anyway. Mistakes reported in Shorrer and Sóvágó (2023) and Hassidim, Romm, and Shorrer (2021) are of similar nature. See Table 1 of Artemov, Che, and He (2023).

[3]WTT assumes (a) the observed number of choices on a ROL is exogenous to the student's preference and (b) each student ranks her top preferred schools in the true preference order and those are preferred to all unranked schools.

assignments, sometimes as the only tie-breaker or as a supplement for coarse priorities.[4] Since the applicants in these districts do not know their lottery realizations at the time of application, lotteries introduce uncertainties in the set of schools that will be feasible for them. Second, even when the priorities are strict, they may not be known in advance. A case in point is the New York City (NYC) Specialized High School assignments. The admission is based on the scores the applicants obtain from Specialized High School Admissions Test with little room for ties, but students must submit their ROLs before learning their scores.[5] Finally, uncertainty arises from not knowing other students' priorities or their preferences. While the assignments from past years can provide some guides, this sort of uncertainty is endemic to all systems whether the priorities are strict or not.

These uncertainties raise questions for both the method prescribed by the existing literature and its theoretical foundations. In fact, it is not even clear that the types of payoff-irrelevant mistakes discussed earlier could arise when the applicants face uncertainties. Without knowing which schools are feasible, mistakes become costly, so applicants would become more careful not to make them. However, this logic does not mean that uncertainties make *all* types of mistakes costly. For instance, some schools may still remain out of reach for a student because she may not live in the right geographic zone or she may not have strong academic qualifications, in which case how she ranks these schools would not matter. Second, even if all schools may be within reach, the type of lotteries used in practice, such as Single-Tie-Breaking (STB), leaves room for payoff-irrelevant mistakes.[6] Hence, payoff-irrelevant mistakes do exist even in the presence of uncertainty and preference must be inferred robustly with respect to these mistakes.

At the same time, uncertainties make *some* types of mistakes more costly, and this means that applicants will reveal *some* preferences more reliably in order to avoid such mistakes.[7] The question

[4]In the US, the use of strict priorities in primary and secondary schools is limited to the so-called *screening programs*. The importance of lotteries in student placement has been growing since screening based on academic performances and the residence location is increasingly viewed as causing segregation of student bodies.

[5]Similar uncertainty exists in Chinese college assignments. Some provinces have students apply before taking the exam or after taking the exam but before learning about their scores (Chen and Kesten, 2017).

[6]To illustrate, suppose there are three schools a, b, and c, assigned based (only) on STB lotteries which assign each student based on a single lottery number that applies to all schools: namely, there are cutoffs (in lottery numbers) such that each student is assigned her highest-ranked school among those whose cutoffs are below her lottery number. Suppose that the cutoffs are say ordered always such that $P_a > P_b > P_c$. If a student happens to have preferences b-a-c, ranked in descending order, then no matter her lottery draw, school a is irrelevant for her, since whenever a is feasible to her, b is also feasible to her, and since she prefers b over a. In this case, either dropping a from her ROL or ranking a behind c would entail no payoff loss compared with reporting truthfully as long as she ranks b.

[7]Again, consider the applicant in the previous example with preference b-a-c. Suppose there is no uncertainty and the applicant "knows" that only c is available. Then, no preference is revealed robustly since *any* ROL that lists c will produce the same outcome, i.e., assignment at c. By contrast, given the STB-generated uncertainties, we can learn that $b \succ a, c$.

is *which* preferences can be revealed under *what* uncertainty structure. This is what the current paper studies. Specifically, we develop a novel method called *Transitive Extension of Preferences from Stability*, or TEPS, that exploits the structure of uncertainty facing applicants to infer their preferences robustly to payoff-insignificant mistakes. We then apply TEPS to school choice data from Staten Island, NYC, and conduct counterfactual analyses of the desegregation policies being proposed. We show a robust inference of student preferences is crucial for predicting policy effects.

The theoretical foundation for our method rests on the concept of *robust equilibrium* developed by Artemov, Che, and He (2023), which allows applicants to make payoff-insignificant mistakes. Specifically, we consider applicants in a DA matching market and require applicants to employ strategies that arbitrarily closely attain their best response payoffs as the economy grows large. All applicants playing their dominant strategies of truth-telling is clearly a robust equilibrium; but truth-telling need not be the only robust equilibrium behavior, as applicants may make payoff insignificant mistakes—strategies that entail a payoff loss negligible in a large market. Hence, robust equilibrium is a weakening of the dominant strategies equilibrium. The main implication of the solution concept is Theorem 1, which establishes *asymptotic stability*: as the economy grows large, with high probability, almost all students are assigned stably—namely, they are assigned their favorite schools among those that are feasible given the realized uncertainty.

The asymptotic stability result justifies our TEPS method described as follows. The first step of the procedure is to recover the underlying structure of uncertainty by simulating it.[8] For instance, in the typical DA-STB environment, this means that we conduct the STB-lottery and run DA algorithm in the same manner as the official assignment algorithm. We then invoke asymptotic stability to infer that the assigned school is preferred to all other schools that are feasible given the lottery realization. We repeat this procedure for each lottery draw many times, and obtain preference relations for each draw. The final step is then to link together the preference relations obtained from the many realizations of uncertainty by invoking the transitivity axiom. That is, we take a transitive closure of all preference relations obtained from the pairs of assigned and feasible schools.

The prediction that stable assignments occur for *all* realizations is valid only for a large limit economy; Theorem 1 is consistent with violations of stability in *finite* economies.[9] Our general version of TEPS accommodates such possibilities. Specifically, we define a family of TEPS$^\tau$,

[8]This depends on the exact context of the institutional setting. If the uncertainty arises from other sources, such as test scores that are yet to be realized at the time of application, then one may build an empirical model to simulate the distribution of test scores say based on the observables (which may include various performance measures from preceding schools say).

[9]Payoff relevant mistakes are observed for a significant minority of applicants in Artemov, Che, and He (2021).

indexed by the attention parameter $\tau \in (0, 100]$, such that only those uncertainties (or feasible sets) that are cumulatively most likely up to $\tau\%$ are considered for preference inference, meaning the uncertainties that belong to the bottom $100 - \tau\%$ in likelihood are discarded for being unreliable.

In principle, the level of inattention and tolerance for payoff-relevant mistakes is an empirical matter that depends on the characteristics of the participants. Accordingly, we develop a data-driven testing procedure to select among alternative attention levels, ranging from WTT at one extreme to TEPS$^\tau$ with varying attention levels τ. A Monte Carlo analysis based on simulated data reveals that biased estimation arises when one ignores mistakes made by market participants, and demonstrates that our selection procedure performs well.

Last, we apply our methods to public high school choice data in Staten Island, NYC. We examine how alternative hypotheses used for preference inference perform in describing the real data. In particular, we conduct counterfactual analyses to compare the predicted effects of several desegregation policies based on alternative hypotheses. The goal of this analysis is not just academic. There is a growing concern about the segregation of student bodies in large urban districts such as NYC, and the removal of priorities based on academic qualifications and/or residence locations has been proposed as a method for improving desegregation. How effective such policies will be in attaining the stated goal is a question of significant public interest, for which our methodological advance may take a surprising relevance. The counterfactual analysis of such policy reforms depends on the estimated preferences of target students for the schools that are currently inaccessible to them due to priority restrictions. If any inattention these students exhibited toward these schools is not accounted for, the effectiveness of these reforms could well be understated.

We do find that the effects of the desegregation policies are underestimated if one uses WTT, which neglects students' mistakes/inattention, as the hypothesis for preference inference. Intuitively, WTT assumes that students do not prefer unranked out-of-reach schools, so they would not apply to them even when these schools become within reach under new policies. Second, the removal of geographic and academic screening priorities has a rather modest effect on desegregation even under our selected (TEPS) estimates. This suggests that forces other than school priorities, such as residential segregation, may play an important role in causing school segregation and raises a question on the effectiveness of priorities-removal reforms as a tool for alleviating school segregation.

Related Literature. We contribute to the literature on identifying and estimating student preferences in school choice. We focus on DA and allow students to deviate from truth-telling. Several papers have suggested alternative approaches. Most related to our proposed method, stability has

been considered as an alternative to WTT by Fack, Grenet, and He (2019) and Artemov, Che, and He (2023) (henceforth FGH and ACH, respectively). In particular, ACH uses robust equilibria to capture applicant inattention and justify asymptotic stability, just like the current paper. The crucial difference is that these papers assume that applicants do not face any uncertainties over the feasible set of schools when they apply. As already argued, the ability to allow for uncertainties is important given the ubiquity of lotteries and the endemic nature of finite sample uncertainty. Both the methodology we propose and our theoretical justification are nontrivial.[10] The TEPS procedure of extracting and combining preferences using transitivity and the data-driven method for selecting the right version of TEPS are novel and theoretically well-justified, and they are computationally fast enough to be readily deployed in empirical research. Likewise, our theoretical justification for TEPS, namely Theorem 1, despite the resemblance to the analogous result in ACH, is not a straightforward extension of the latter; as we remark in Section 2, the effect of deviating one's strategy is now stochastic, and "controlling" the stochastic payoff effect presents a new challenge and requires an additional assumption on the structure of the underlying uncertainty.

The current paper is also related to a strand of literature that views each applicant's submitted ROL as an optimal portfolio choice (see, as early examples, He, 2017; Agarwal and Somaini, 2018; Calsamiglia, Fu, and Güell, 2020). In particular, the approach in Agarwal and Somaini (2018), when applied to DA, would be robust to payoff-irrelevant mistakes. However, it would not allow for mistakes with small but positive payoff consequences. Further, when the approach is taken to data, the curse of dimensionality becomes severe as the number of available schools grows.[11] By contrast, our proposed empirical method is computationally easy and robust to the strategical mistakes that are commonly observed in practice.

Some subsequent papers have attempted to deal with the computational issues associated with the portfolio choice approach. Larroucau and Rios (2020) show that it is sufficient to consider a subset of ROLs ("one-shot swaps") to avoid the curse of dimensionality, but this result applies only in settings where the cutoffs, and hence the admission probabilities, are independent across schools (e.g., Multiple Tie-Breaking, or MTB). Idoux (2022) modifies Agarwal and Somaini (2018) by assuming that applicants face application costs, are of limited rationality, and use a heuristic

[10] While the TEPS procedure can be seen as a generalization of the stability hypothesis, it is both novel and theoretically well-founded in a way a naive application of the stability hypothesis is not. For instance, one may directly "implement" the stability hypothesis in the lottery environment by applying it using the lottery that has been used in the actual assignment (if is available in the data) or simply by conducting a lottery once. As we argue in Section 3, such a method is not well justified by our theory and is problematic in other respects.

[11] To verify the optimality of a submitted ROL, one should compare it against all possible ROLs. However, for example, the number of possible ROLs with 10 available schools is nearly 9.9 million.

method to choose their ROLs. She thus avoids the curse of dimensionality and allows schools with small admission probabilities to be unranked, which makes her method robust to payoff-insignifcant mistakes to some degree. However, her method does not allow other common types of mistakes. For example, the ranked schools must be in the true preference order in her model, ruling out the possibility of a student arbitrarily ranking schools with small admission probabilities. Moreover, the method neglects that some ranked schools can be completely irrelevant under STB, including those with high admission probabilities (recall our STB example in Footnote 6.) In contrast, TEPS is robust to all such mistakes as long as the payoff consequence is not too large.

Stability has been a key identifying condition in the two-sided matching literature (see Chiappori and Salanié, 2016, for a survey). There, agents on both sides have preferences that are unknown to the researcher. This setting includes decentralized school choice, such as that in Chile (He, Sinha, and Sun, 2021), in which private schools can rank students freely. On the contrary, in our framework of centralized school choice, school preferences/priorities are known to some extent because schools rank students according to certain pre-specified rules.

2 Theoretical Analysis

We consider a model of a large matching market operated by DA. After introducing the primitives and the robust equilibrium solution concept, we will establish our main theorem.

2.1 Model Primitives

Continuum Economy. We begin with a *continuum economy*, which will serve as a benchmark for finite economies—the object of our central interest. The continuum economy E consists of a finite set of *schools* $C = \{c_1, ..., c_C\}$ and a unit mass of students with types $\widetilde{\theta} \in \widetilde{\Theta}$. Note that, with slight abuse, we use C to denote both the set and cardinality of the schools. Let $\widetilde{C} := C \cup \{\emptyset\}$, where $\emptyset$ corresponds to the outside option (which may involve attending a private school or an outside-district school). The schools have masses $S = (S_1, ..., S_C)$ of seats, where $S_j \in (0, 1]$, and $S_\emptyset = \infty$.

Each student has an **ex ante type** $\widetilde{\theta} = (u, t)$, where $u^{\widetilde{\theta}} = (u_1^{\widetilde{\theta}}, ..., u_C^{\widetilde{\theta}}) \in [\underline{u}, \overline{u}]^C$ are von-Neumann Morgenstern utilities of attending schools C, for some $\underline{u} < 0 < \overline{u}$. By normalization, $u_\emptyset^{\widetilde{\theta}} = 0$ for all $\widetilde{\theta}$. A student has ex-ante priority type t, interpreted as her "intrinsic" priority or belief on her merit. Let T be the set of all ex-ante priority types. Students observe their ex-ante type (u, t) prior to application. Let $\widetilde{\eta}$ denote the probability measure of $\widetilde{\theta} \in \widetilde{\Theta} := [\underline{u}, \overline{u}]^C \times T$.

Given $t \in T$, a student's **scores** $s \in [0, 1]^C$ are drawn according a distribution function $\Phi^t \in \Delta([0, 1]^C)$ that may depend on t. The scores determine a student's ex-post priorities used by

schools for her assignment. Students do not observe these scores when they submit applications.[12] Randomness of scores arises from the use of lotteries as tie-breakers or uncertainty about test scores. Let $S^t := \prod_c [\underline{s}_c^t, \overline{s}_c^t]$ be the support of the scores for type t, where $\underline{s}_c^t$ and $\overline{s}_c^t$ are respectively the infimum and supremum scores for type-t student for school c. We call the pair $\theta = (u, s)$ for a student her **ex-post type**. The measure $\widetilde{\eta}$ of ex ante types, together with $(\Phi^t)_t$, induces a probability measure η on the ex post type $(u, s) \in \Theta := [\underline{u}, \overline{u}]^C \times [0, 1]^C$. We assume that η is atomless. (By contrast, $\tilde{\eta}$ can, and typically will, have atoms.) For our purpose, the atomlessness assumption ensures that indifferences either in student preferences or in schools' scores arise only for a measure zero set of students. In summary, the continuum economy is summarized by $E = [\widetilde{\eta}, S, (\Phi^t)_t]$.

While we establish our main result under a general priority structure in Appendix A, for ease of exposition, here we focus on the following priority structure. The schools C are partitioned into three types/subsets of schools, C_1, C_2, and C_3, so that schools of each subset use distinct priority rules. Specifically, we assume that $T = T_1 \times T_2 \times T_3$ so that schools in C_i apply priorities in $T_i, i = 1, 2, 3$.

(a) *Merit-based assignment with known scores*: Each school in C_1 prioritizes students based on only priorities in $T_1 = [0, 1]^{|C_1|}$, where for each $t \in T_1$, ϕ^t is degenerate with $\underline{s}_c^t = \overline{s}_c^t = t_c$ for all c. This corresponds to the case in which priorities are given by some merit scores which are known to students when they apply. Examples include Australian college admissions and Paris high school assignments. This is the case studied by Artemov, Che, and He (2023).

(b) *Coarse priorities*: We assume that T_2 is finite and ties are broken by lotteries. More precisely, each school $c \in C_2$ has a finite number n_c of intrinsic priorities $\mathcal{T}_c = \{0, \frac{1}{n_c}, ..., \frac{n_c - 1}{n_c}\}$. A student's ex-ante type consists of a profile $t = (t_c)_{c \in C_2} \in T_2 = \prod_{c \in C_2} \mathcal{T}_c$. Then, for each school, $c \in C_2$, the score for a student with type t is then a composite, $s_c = t_c + \lambda_c \frac{1}{n_c}$, where λ_c is the lottery score used for breaking a tie at school c. The lottery scores, $(\lambda_c)_{c \in C_2}$, are either uniform on $[0, 1]^{|C_2|}$ in the case of the multiple tie-breaking (MTB) rule or $\lambda_c = \lambda$, for all $c \in C_2$, with λ drawn uniform randomly from $[0, 1]$, in the case of the single tie-breaking (STB) rule. This feature corresponds to the common structure of priorities used in US public schools.

(c) *Merit-based assignment with unknown scores*: We assume that $T_3 = [0, 1]$, and, for each $t \in T_3$, Φ^t is absolutely continuous with a strictly positive density over scores $s \in [0, 1]$. This corresponds to the case of NYC's Specialized High School assignment, which follows á la serial

[12] In NYC school district, lottery draws have been recently disclosed to the families who request them prior to their application (See here). This is an exception than the norm. Regardless, this revelation of scores does not invalidate our assumption since those scores revealed can be treated as ex-ante priority type t.

dictatorship based on the scores obtained from the SHSAT (Specialized High School Admissions Test). Students submit their ROLs at the time they take the SHSAT. The ex-ante type $t \in [0,1]$ corresponds to the student's belief about the score.

If $C = C_1$, no students face any uncertainties, so the model specializes to that of ACH.

Finite Economies. We are interested in a sequence of finite random economies converging to the continuum economy E in the limit. Specifically, let $F^k = [\widetilde{\eta}^k, S^k, (\Phi^{k,t})_t]$ be a *k-economy* that consists of k students each with type $\widetilde{\theta}$ drawn independently according to $\widetilde{\eta}$, and the vector $S^k = [k \cdot S]/k$ of capacity per student, where $[x]$ is the vector of integers nearest to x (with a rounding down in case of a tie). The distribution $\widetilde{\eta}^k$ is simply an empirical measure of ex-ante types $\widetilde{\theta}$; likewise, we let η^k denote an empirical measure of ex-post types (u, s).

Matching and Mechanism. A *matching* is a mapping $\mu : C \cup \Theta \to 2^{\Theta} \cup (C \cup \Theta)$ that describes how students are assigned to alternative schools based on their ex-post types. It satisfies the usual two-sidedness and consistency requirements as well as "open on the right" as defined in Azevedo and Leshno (2016) (henceforth AL, see p. 1241). The key concept of our inquiry deals with the property known as stability. To describe it, fix an economy, which can be either the continuum economy E or a *realization* of a finite k-economy. Fix any profile $p \in [0,1]^C$, interpreted as **cutoffs** of schools. For each type $\theta = (u, s)$, which consists of a realized utility vector u as well as a realized score vector s, we say a school c is **ex-post feasible** if $s_c > p_c$. Now let $D_c(p)$ denote the measure of students for whom c is the best school among those that are feasible according to their true preferences. A stable matching is often defined in terms of two properties: *individual rationality* and *no blocking*,[13] but here it is more convenient to adopt the following characterization by AL as our definition. A **stable matching** is a matching in which all students are assigned the best feasible schools at cutoffs p which "clear the market" in the sense that $D_c(p) \leq (=) S_c$ for all c (if $p_c > 0$). Importantly, in the k-economy both the cutoffs and the resulting assignment are random, so the stability requires the stated condition to hold with probability one.

Student-proposing deferred acceptance algorithm, or DA in short, takes students' reported ROLs and their scores as input and outputs a stable matching that is Pareto-best for all students given the

[13] Individual rationality requires that no participant (a student or a school) is assigned a partner that is not acceptable. No blocking means that no student-school pair exists such that the student prefers the school over her assignment and the school has either a positive measure of vacant positions or admits a positive measure of students whom the school ranks below that student. In example (b), the scores may be a result of tie-breaking, so no blocking, and therefore stability, should be defined based on ex ante type $\widetilde{\theta}$. Nevertheless, we focus on ex post stability since it, and not the stability, provides a foundation for our empirical method for inferring applicants' preferences. Further, (asymptotic) ex-post stability implies (asymptotic) ex-ante stability, so establishing the former will deliver the latter.

reported ROLs, or the matching that corresponds to the "lowest" market-clearing cutoffs.[14] We call them **DA cutoffs**. As is well-known, DA makes it a weakly dominant strategy for each student to report a truthful ROL. In light of our main motivation, however, we shall allow dominated strategies to be used by students.

2.2 Robust Equilibria

For each k-economy, F^k, we let $\{1, ..., k\}$ index the students present in that economy. We assume that each student observes her own ex-ante type $\widetilde{\theta}$ but not the types of other students. A (mixed) strategy by each student $i \in \{1, ..., k\}$ is denoted by a measurable function $\sigma_i : \widetilde{\Theta} \to \Delta(\mathcal{R})$, where $\mathcal{R}$ is the set of all possible ROLs of $\widetilde{C}$. We use ρ to denote a **truthful reporting strategy (TRS)**: namely, $\rho(\widetilde{\theta})$ ranks school $a \in \widetilde{C}$ ahead of $b \in \widetilde{C}$ if and only if $u_a^{\widetilde{\theta}} > u_b^{\widetilde{\theta}}$.[15]

In order to allow for the type of mistakes observed in recent evidence, we focus on **robust equilibria** developed by ACH. This concept allows applicants to use dominated strategies but requires that the payoff loss resulting from such strategies be arbitrarily small in a large economy. In particular, this solution concept accommodates the commonly documented mistakes of not ranking, or ranking incorrectly, out-of-reach schools.

While we are agnostic about the sources of such mistakes, the current concept captures the decision-maker inattention in a way similar to rational inattention (Matějka and McKay, 2015) and quantile response equilibria (Goeree, Holt, and Palfrey (2002)). Individuals facing decision costs would presumably rationally allocate more attention to choices that have high payoff consequences and less attention to those that have low payoff consequences, so they end up making more mistakes on the latter choices. Unlike these theories, however, such inattention becomes arbitrarily small as the economy grows large.[16]

To operationalize the concept, we need to consider a strategy profile along a sequence of finite economies. Formally, we consider an infinite profile $\sigma := (\sigma_i)_{i \in \mathbb{N}}$ of strategies specified for each player $i \in \mathbb{N}$, with the interpretation that in each k-economy F^k its k-truncation $\sigma^k := (\sigma_1, ..., \sigma_k)$ is employed by its participants as their strategies. Note that we allow students to adopt asymmetric strategies, submitting different ROLs even when they have the same ex ante type. The solution

[14] By the well-known lattice property, the lowest market clearing cutoffs are well defined, given the responsive preferences on both sides.

[15] How ρ breaks a tie is immaterial since η is assumed to be atomless, so a tie will occur with zero probability. Hence, we leave this unspecified.

[16] The idea of payoff-sensitive mistakes is also reminiscent of Myerson (1978)'s proper equilibria. Note, however, the motivations are completely different: the goal here is to allow for such mistakes, whereas proper equilibria (and trembling-hand perfection) refine Nash equilibria by requiring robustness to mistakes.

concept follows:

DEFINITION 1. *An infinite profile σ is a **robust equilibrium** if, for any $\epsilon > 0$, there exists $K \in \mathbb{N}$ such that for $k > K$, its k-truncation σ^k is an interim ϵ-Bayes Nash equilibrium—namely, for i, σ_i gives student i of each type θ a payoff within ϵ of the highest possible (i.e., supremum) payoff she can get by using any strategy when all the others employ σ^k_{-i}.*

The solution concept does not require exact optimality for the strategies rather their near optimality in a sufficiently large economy. Although adding considerable demand on the proof, allowing for asymmetry makes the concept broadly applicable. For our main result, we focus on a slight strengthening of this concept:

DEFINITION 2. *For any $\gamma \in (0, 1)$, a strategy is γ-**regular** if it coincides with TRS with probability at least of γ. A profile σ is a* **regular robust equilibrium** *if it is robust and there exists a $\gamma \in (0, 1)$ such that σ_i is γ-regular for all $i \in \mathbb{N}$.*

2.3 Analysis of Robust Equilibria

As illustrated in the introduction, applicants may not use TRS in a regular robust equilibrium.[17] Hence, one can't use truth-telling as a behavioral restriction. Instead, we establish that the outcome of regular robust equilibrium is asymptotically stable, which will give rise to a method for inferring preferences under the robust equilibrium.

We begin with a few concepts. For any potential cutoffs $p \in [0, 1]^C$, we say a strategy σ_i by student i is **stable-response strategy (SRS) against** p if the student receives the best feasible school relative to the cutoffs with probability one; σ_i is **non-SRS against** p if it is not an SRS against p. In the context of DA, we say a strategy is **stable** if it is stable against the prevailing DA cutoffs, given the strategies. Certainly, TRS is SRS but an SRS need not be TRS, as illustrated by an example in Footnote 6. Next, we define **asymptotic stability** as:

DEFINITION 3. *A profile σ is **asymptotically stable** if the fraction of students who employ SRS in each σ^k converges to 1 in probability as $k \to \infty$.*[18]

[17] While the example in Footnote 6 suggests only one preference type who may deviate from TRS, it is not difficult to imagine that there is a significant fraction of applicants with the same preference type.

[18] More formally, we require that for any $\epsilon > 0$ there exists $K \in \mathbb{N}$ such that in any k-random economy with $k > K$, with the probability of at least $1 - \epsilon$, at least a fraction $1 - \epsilon$ of all students are assigned their best feasible schools given the DA cutoffs p^k.

In words, asymptotic stability means that all except for a probabilistically vanishing fraction of students receive the stable assignment for almost all realizations of uncertainties they face. This will provide a key method for inferring preferences from the choice data generated under the DA algorithm. ACH establishes this result when students face no uncertainty in s (i.e., school types C_1) and under the assumption that η has full support. The full support assumption guarantees the uniqueness of stable matching, which plays a crucial part in the proof. However, it is important to relax full support for our η, as simple STB with no intrinsic priorities and the more realistic scenario of coarse priorities with STB fail full support. This motivates us to weaken the condition:

ASSUMPTION 1 (MARGINAL FULL SUPPORT). *For each $r \in \mathcal{R}$ and for each $c \in C$, $\inf_{s_c \in [0,1]} \bar{\eta}(s_c) > 0$, where*

$$\bar{\eta}(s_c) := \int_{\rho(u)=r, s_{-c}} \eta(u, s_c, s_{-c}) du ds_{-c}$$

is the marginal density for score s_c.

This condition requires that any given ordinal preference type has full support for its *marginal* distribution of score for each school. If the scores have positive *joint* density on the compact support $[0,1]^C$ as is assumed in AL, the condition will hold. Importantly, however, the condition allows for the support of scores to be low dimensional. In particular, Marginal Full Support holds in the case of STB without any intrinsic priorities; in that case, the support of the uncertainty/lottery is diagonal and one-dimensional, yet the marginal of its lottery has full support. With nontrivial intrinsic priorities, the condition states that each ordinal preference type has the full support of priorities. We view the condition to be very mild. We are able to generalize the uniqueness result by Azevedo and Leshno (2016) and Azevedo (2014) under the weaker condition:

LEMMA 1. *If η satisfies Marginal Full Support, then $E = [\eta, S]$ admits a unique stable matching.*

We are now in a position to state the main result.

THEOREM 1. *Suppose η satisfies Marginal Full Support. Then, any regular robust equilibrium is asymptotically stable.*

This theorem states that in a large enough economy, for almost every realization of uncertainty, virtually all students are assigned their most preferred schools among those that are feasible to them given the realized cutoffs, even though the feasible schools are unknown to them when they submit ROLs. This theorem implies that we can invoke stability for each realization of uncertainty; this is precisely what we will do in our TEPS procedure.

In light of the multiplicity in student behavior in robust equilibrium, another question is how to conduct counterfactual analysis, an exercise that is common in market design research. The next corollary of THEOREM 1 justifies that, as far as predicting an assignment outcome is concerned, we can simply assume truth-telling among all students in a counterfactual regime.

COROLLARY 1. *Suppose η satisfies Marginal Full Support. Then, in any regular robust equilibrium, the fraction of students whose assigned school is the one that would arise if everyone employs TRS converges to one in probability as $k \to \infty$.*

REMARK 1 (Relationship with ACH's asymptotic stability result). *Both ACH and the current paper build on the uniform convergence of cutoffs. If the priorities are strict and known in advance as in ACH, the cutoff convergence makes it clear that for most of the applicant types, playing a non-stable response entails a discrete loss, so this leads to no such strategies being used in that setting. But, when priorities may be unknown and random, then the discreteness of the payoff loss from playing a non-stable response strategy is not easy to establish. In fact, the nature of payoff loss depends crucially on the distribution of (unknown) priorities for the applicant. Controlling this distribution requires the properties of the uncertainties embodied in the priorities assumed in (a), (b), and (c). In fact, Appendix A defines a general condition underlying the priority structure that guarantees the discreteness of the payoff loss. (Hence, the theorem is more general than the version stated in the text.) The proof is highly nontrivial: see Lemma 4 and its subsequent use in the proof of Theorem 1′.*

3 Inferring Student Preferences from ROLs: Transitive Extension of Preferences from Stability

Building upon our theoretical results (Theorem 1), we develop a procedure for inferring students' (ordinal) preferences from school choice data, called Transitive Extension of Preferences from Stability (TEPS). We show that it extracts the maximal information about participants' preferences based on our theory.

Suppose we have a typical dataset generated by a centralized DA assignment. Assume that the analyst observes the ROLs submitted by students/applicants and their (intrinsic) priorities as well as the capacities of schools.

3.1 Preference Relations Inferred by Stability and Transitivity

Since we infer the preference of each applicant separately, it is without loss to focus on a single applicant. Assume the applicant has submitted a ROL R and has priorities t. For notational

simplicity, we shall drop the dependence on R and t whenever there is no ambiguity. We assume there is no outside option although outside options can be easily accommodated.[19] Let Ω be the set of all assignment-relevant states, or more precisely, all possible profiles of priorities that our applicant, as well as all other applicants, may have at all schools. In practice, each state $\omega \in \Omega$ corresponds to the lottery draws for all students, the profile of their scores from the entrance exam, and other uncertainties each student faces on the other applicants' preferences, all of which affect the cutoffs of schools.

Given the data, the goal is to identify the set of preference relations $\mathcal{P} := \{(x, y) \in C^2 : x \in C \text{ is inferred more preferred to } y \in C\}$ inferred by that data according to Theorem 1. The precise interpretation of Theorem 1 may depend on the market size and the payoff relevance of mistakes. For expositional ease, we begin with a rather strong interpretation of Theorem 1 that *the matching outcome must be stable for every realization ω of the uncertainty*; namely, no student makes any payoff-relevant mistakes. While this implication is justified only in the limit continuum economy, it serves as a good starting point for describing our method. We will relax this assumption later in Section 3.2.

The TEPS procedure takes the following three steps to produce preference relations $\mathcal{P}$. To illustrate the procedure, consider the example.

EXAMPLE 1. *Suppose that there are six schools, $\{c_0, c_1, c_2, c_3, c_4, c_5\}$ and a student with intrinsic priorities t and a submitted ROL $R = c_4$-c_3-c_2-c_1. Depending on possible realizations of lottery draws, there are four possible events: 1. the set $\{c_3, c_4\}$ is feasible to the student with probability 0.4; 2. the set $\{c_0, c_1\}$ is feasible to the student with probability 0.3; 3. the $\{c_0, c_1, c_2\}$ is feasible to the student with probability 0.25; 4. the set $\{c_1, c_4\}$ is feasible to the student with probability 0.05.*

Step 1. Simulating uncertainties and compiling choice data:

The first step simulates or "traces out" all possible profiles of priorities by simulating uncertainties. For instance, lottery-driven uncertainties can be simulated easily by conducting the lotteries and running DA identically as the underlying process generating the data.[20] Each simulated state $\omega \in \Omega$ leads the realized priorities of students, and given their ROLs, we can obtain the schools'

[19] Our procedure below can be extended straightforwardly with the set of schools redefined to $\widetilde{C} = C \cup \{\emptyset\}$, where $\emptyset$ is an outside option representing an always-feasible alternative, such as non-assignment.

[20] To simulate the student composition uncertainty, we may bootstrap from the observed ROL to create multiple economies/applicant compositions. If the uncertainty arises from other sources, such as test scores that are yet to be realized at the time of application, then one may build an empirical model to simulate the distribution of test scores say based on the observables (which may include various academic performance measures).

cutoffs. From this, we can identify for each student the set $B(\omega)$ of her *feasible schools* in state ω, and her *assigned school*, $\alpha(\omega)$—namely, the highest-ranked choice within $B(\omega)$ according to the student's ROL. Naturally, there can be multiple ω's that lead to the same feasible set and therefore the same assignment. Hence, we consider a partition P_Ω of Ω into equivalence classes, W, such that $B(\omega) = B(\omega')$ if and only if $\omega, \omega' \in W$, for each $W \in P_\Omega$. In the example, we have four equivalence classes of the uncertainties, W_1, W_2, W_3, W_4 in terms of distinct feasible sets, $B_{W_1} = \{c_4, c_3\}$, $B_{W_2} = \{c_1, c_0\}$, $B_{W_3} = \{c_2, c_1, c_0\}$, and $B_{W_4} = \{c_4, c_1\}$. Given her ROL, she would be assigned $\alpha_{W_1} = c_4$, $\alpha_{W_2} = c_1$, $\alpha_{W_3} = c_2$, and $\alpha_{W_4} = c_4$ in the corresponding classes.

We then compile a *choice data* for the student—namely, a pair (B_W, α_W) of a feasible set of schools and an assigned school—for each equivalence class $W \in P_\Omega$.[21] The choice data for the example is depicted in Figure 1.

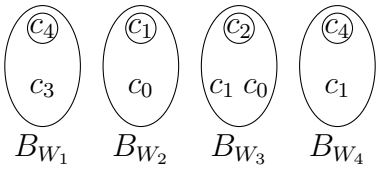


Figure 1: TEPS Step 1: Simulating Uncertainties and Compiling Choice Data

Step 2. Inferring preference relations in each realized uncertainty:

Theorem 1 states that in the limit as the economy grows large, all but the vanishing fraction of students must be stably assigned for every state $\omega \in \Omega$. This implies that, for any given student, her assigned school α_W is her most preferred choice among the set of feasible schools B_W for each $W \in P_\Omega$. This yields preference relations $\widetilde{\mathcal{P}}_W \subset C^2$ for each $W \in P_\Omega$.[22] Let $\widetilde{\mathcal{P}} := \cup_{W \in P_\Omega} \widetilde{\mathcal{P}}_W$ denote the set of all preference relations obtained in this way. We next partition $\widetilde{\mathcal{P}}$ into ordered sets of preference relations $\{\widetilde{\mathcal{P}}_1, \widetilde{\mathcal{P}}_2, ..., \widetilde{\mathcal{P}}_m\}$ such that if $(c, d) \in \widetilde{\mathcal{P}}_i$ and $(c', d') \in \widetilde{\mathcal{P}}_j$, then c is ranked ahead of c' if $i < j$. In our running example, this step produces ordered sets of preference relations, $\widetilde{\mathcal{P}}_1 = \{(c_4, c_3), (c_4, c_1)\}$, $\widetilde{\mathcal{P}}_2 = \{(c_2, c_1), (c_2, c_0)\}$, and $\widetilde{\mathcal{P}}_3 = \{(c_1, c_0)\}$, as represented by three ordered trees with one length, where the roots correspond to the schools the student may be assigned in some realizations of the uncertainties.

Step 3. Extending preference relations by transitivity axiom:

[21] Since we do not consider outside options, we omit (α_W, B_W) if $\alpha_W = \emptyset$, a non-assignment. In case we include an outside option, we should also include (α_W, B_W), where $\alpha_W = \emptyset$.

[22] Recall that each element in $\widetilde{\mathcal{P}}$ is an ordered pair of schools such that c_1 is inferred preferred to c_2 for each $(c_1, c_2) \in \widetilde{\mathcal{P}}$.

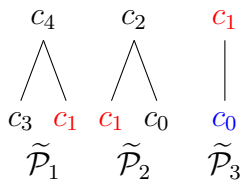


Figure 2: TEPS Step 2: Inferring Preference Relations in Each Realized Uncertainty

The last step links together the preference relations $\widetilde{\mathcal{P}}$ obtained from Stage 2 via the axiom of transitivity. Specifically, we use the following algorithm.[23] Consider first $\widetilde{\mathcal{P}}_m$ and $\widetilde{\mathcal{P}}_{m-1}$, where m is the highest index. The algorithm constructs a new set $\mathcal{P}_{m-1}$, by including $\widetilde{\mathcal{P}}_m \cup \widetilde{\mathcal{P}}_{m-1}$ along with preferences relations obtained by the transitive extension operation: If $(y, z) \in \widetilde{\mathcal{P}}_m$ and $(x, y) \in \widetilde{\mathcal{P}}_{m-1}$—i.e., if the assigned school under $\widetilde{\mathcal{P}}_m$ is revealed inferior to the school that is assigned under $\widetilde{\mathcal{P}}_{m-1}$—then, we declare x to be revealed preferred to z and add (x, z) to $\mathcal{P}_{m-1}$. We then iterate the same procedure on $\mathcal{P}_{m-1}$ and $\widetilde{\mathcal{P}}_{m-2}$ to obtain $\mathcal{P}_{m-2}$, and so on. The repeated iteration returns the final output $\mathcal{P} := \mathcal{P}_1$.

In the example, $\mathcal{P} = \{(c_1, c_0), (c_2, c_1), (c_2, c_0), (c_4, c_1), (c_4, c_3), (c_4, c_0)\}$, depicted in Figure 3 as a a collection of trees.

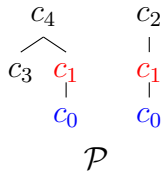


Figure 3: TEPS Step 3: Extending Preference Relations by Transitivity Axiom

Note that TEPS procedure does not infer all preference ranks observed from student's ROL. In the example, TEPS procedure does not infer any relations between c_4 and c_2 despite the ROL listing c_4-c_3-c_2-c_1. The reason is that these two schools are never feasible simultaneously,[24] so their relative rankings are not directly revealed, nor are they inferred indirectly by transitivity. Hence, TEPS does not view this part of the ROL as payoff-relevant or reliable. Also, note that one of the

[23] There are other computationally fast alogorithms for finding a transitive closure of binary relations, based on shortest path search, breadth- or depth-first search, that has polynomial time in the number of choices. See https://en.wikipedia.org/wiki/Transitive_closure.

[24] Such cases occur under MTB which uses distinct lotteries to break ties at different schools. Even with STB, such cases may occur as long as the cutoffs are not degenerate (i.e., in a finite economy.) This is because the relative ranking of schools' cutoffs may change under different realizations of the STB lotteries.

unranked schools, c_5, is never feasible (i.e., out of reach) under any realization of the uncertainties, and as a result, ranking c_5 arbitrarily or even omitting it is payoff-irrelvant for the student. Hence, TEPS does not infer any preference relation regarding c_5. In short, TEPS extracts only a subset of preference relations of what ROL (or WTT, as will be seen later) infers.

Nevertheless, TEPS procedure infers as many preference relations as possible consistent with Theorem 1. To state it formally, fix a student's ROL R and her observable priorities t. For any $\Omega' \subset \Omega$, let $\mathcal{P}^{\text{ST}}(\Omega')$ denote the preference relations that can be inferred by stability on all possible uncertainties in Ω' and the transitivity axiom. Then, we have the following result:

PROPOSITION 1. *Fix a student's ROL R and her observable priorities t. If Step 1 of TEPS simulates the uncertainties $\Omega' \subset \Omega$, then $\mathcal{P} = \mathcal{P}^{ST}(\Omega')$. That is, the TEPS procedure infers the preference relations if and only if they can be inferred by the stability conditions on possible uncertainties $\Omega' \subset \Omega$ and transitivity.*

3.2 Allowing Violations of Stability in TEPS

In principle, Step 1 of the procedure may exhaustively simulate all uncertainties, i.e., by choosing $\Omega' = \Omega$. Imposing stability for all uncertainties is tantamount to assuming that the applicant makes no payoff-relevant mistakes. By THEOREM 1, this is an implication of a robust equilibrium only for the infinite economy. For a finite economy, a robust equilibrium is consistent with violations of stability, or payoff-relevant mistakes, as long as the payoff consequence of the mistakes is sufficiently small. To accommodate these more realistic scenarios, and thus strengthen the robustness of preference inferences, a natural approach is to identify the set of states, or uncertainties $\Omega' \subset \Omega$, that are sufficiently likely and apply TEPS only on these Ω'. Specifically, we proceed as follows.

Recall in Step 1, we partition Ω into equivalence classes in terms of the feasible sets B_W the applicant faces. We now order the feasible sets in descending order of their likelihoods. In the example, the four feasible sets are indexed according to this order, with B_{W_1} being most likely and B_{W_4} being least likely. We then consider the cumulative likelihood of the feasible set starting from the most likely feasible set.

We then apply an "attention parameter" $\tau \in [0, 100]$ in terms of the cumulative likelihood (in percentage) for feasible sets to be included in our TEPS procedure. In our example, applying the threshold of $\tau = 95$ means that we stop at B_{W_3}, as the cumulative likelihood of B_{W_1}, B_{W_2}, and B_{W_3} reach 95%. We then focus our TEPS inference only for $\Omega' = \{B_{W_1}, B_{W_2}, B_{W_3}\}$.[25]

[25]Formally, Step 1 of TEPS with threshold τ always includes the most likely feasible set B_{W_1}, and additionally includes up to $B_{W_{\bar{T}}}$ such that $\sum_{t=1}^{\bar{T}} Prob(W_t) \leq \tau\%$ and $\sum_{t=1}^{\bar{T}+1} Prob(W_t) > \tau\%$.

Let's call the resulting procedure TEPS$^{\tau}$ which produces preference relations $\mathcal{P}^{\tau}$. We label TEPS^{100} and $\mathcal{P}^{100}$ as TEPS^{all} and $\mathcal{P}^{all}$ since they use *all* realizations of the feasible sets. Using this notation, the TEPS procedure and the resulting preference relations in Section 3.1 are TEPS^{all} and $\mathcal{P}^{all}$, respectively. Note that as we decrease the attention parameter τ, TEPS$^{\tau}$ tolerates more violations of stability. When $\tau = 0$, TEPS^{0} only infers a student's preferences based on her most likely feasible set, amounting to her considering only the most likely feasible set. Hence, TEPS^{0} is also called TEPS^{top}.

The logic behind the use of the attention parameter is intuitive and clear. When a student submits her ROL, she may not pay sufficient attention to a low-probability feasible set such as B_{W_4}, which makes the preference "inferred" from such a feasible set to be unreliable. Based on this, our TEPS$^{\tau}$ simply focuses on the most likely feasible sets. In our example, TEPS^{95} produces preference relations $\mathcal{P}^{95}$, depicted in Figure 4, which is a subset of $\mathcal{P}^{100}$ depicted in Figure 3.

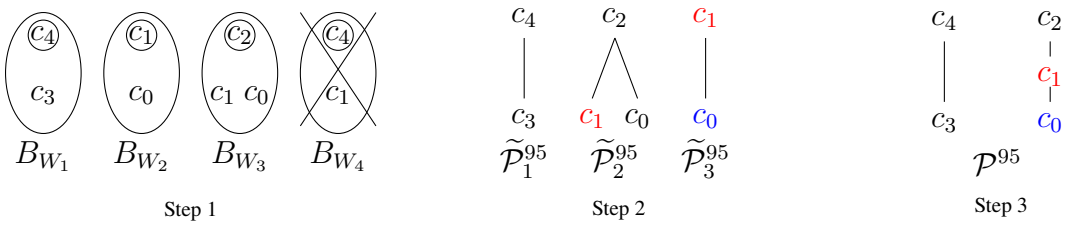


Figure 4: An Example of Allowing Violations of Stability in TEPS: TEPS^{95}

3.3 Comparison of TEPS with Weak Truth-Telling

The literature has traditionally used Weak Truth-Telling (WTT) to infer student preferences (e.g., Abdulkadiroglu, Agarwal, and Pathak, 2017; Laverde, 2022). WTT involves two assumptions: (a) the observed number of choices ranked in any ROL is exogenous to student preferences and (b) every student ranks her top preferred schools according to her preferences and may drop some least preferred schools. Let $\mathcal{P}^{\text{WTT}}$ be the preferences of a student with inferred by WTT. In Example 1, for the student with $R = c_4$-c_3-c_2-c_1, WTT infers $u_{c_4} > u_{c_3} > u_{c_2} > u_{c_1} > u_{c_0}, u_{c_5}$, or $\mathcal{P}^{\text{WTT}} = \{(c_4, c_3), (c_4, c_2), (c_4, c_1), (c_4, c_0), (c_4, c_5), (c_3, c_2), (c_3, c_1), (c_3, c_0), (c_3, c_5), (c_2, c_1), (c_2, c_0), (c_2, c_5), (c_1, c_0), (c_1, c_5)\}$. Crucially, unranked schools, c_0 and c_5 in the example, are inferred as less preferable than all ranked schools in R.

WTT would be justified by the strategyproofness if applicants make no mistakes. In the presence of mistakes, however, WTT may lead to biased estimates of preferences, as have been argued in Fack, Grenet, and He (2019) and Artemov, Che, and He (2023). TEPS$^{\tau}$, for $\tau \in [0, 100]$, address the

problem by offering a flexible menu of methods to identify preferences based on ROLs data in the face of uncertainties and (possible) mistakes. It is important to note that TEPS does not necessarily contradict WTT; rather, it only focuses on the more robustly-inferred information from ROLs than WTT. As one can see, WTT and TEPS$^\tau$ with differing values of τ's form nesting relationships.

PROPOSITION 2. *Fix any student with an ROL R and intrinsic priority t. We have $\mathcal{P}^{top} \subseteq \mathcal{P}^{\tau} \subseteq \mathcal{P}^{\tau'} \subseteq \mathcal{P}^{all} = \mathcal{P}^{ST}(\Omega) \subseteq \mathcal{P}^{WTT}$ for any $0 < \tau < \tau' < 100$.*

The existing literature, motivated by mistakes mentioned earlier, has proposed stability as an alternative to WTT for robustly inferring preferences from ROL data (Fack, Grenet, and He (2019), Artemov, Che, and He (2023)) when there is no uncertainty in priorities. One can view the current method as a natural adaptation of the stability-based method to the environment involving uncertainties. Evidently, the adaptation is nontrivial both in terms of theoretical justification and empirical operationalization.

To illustrate, suppose one adapts the stability restriction rather mechanically by applying it against single ex-post cutoffs obtained from an arbitrary lottery draw. At first glance, this method seems analogous to TEPS^{top}. However, unlike TEPS^{top}, this method is problematic at least for two reasons. First, this ad-hoc adaptation of stability is not well justified by THEOREM 1. The ex-post cutoffs obtained from a single lottery draw could very well be an outlier that is a very low-probability event from the perspective of the applicant. Hence, there is no guarantee that the preference "inferred" from this method reflects the true preference of the applicant. Second, in cases using a tie-breaking lottery, the realized lottery may not be observed, and the researcher may need to draw a lottery and simulate an observed outcome. Since it is impossible to verify how many lottery draws the analysts have made, the method is not immune from an analyst cherry-picking a lottery realization to suit his/her need. By contrast, TEPS^{top} is both well-founded from theoretical justification (THEOREM 1) and is ex-post verifiable.

3.4 Identification and Estimation

The assumption of stability and transitivity, which provide the basis of TEPS, does not uniquely pin down a student's best response or the ROL she submits. In fact, there are typically multiple best responses for a student given others' strategies. One may thus worry that this multiplicity may render our model incomplete in the sense of Tamer (2003). Incompleteness in our context would mean that the mapping from a student's type, (u, t), to TEPS inferred preferences, is a correspondence, which may cause the point identification of the distribution of u to fail.

These concerns do not apply here. We provide some intuition about the completeness, which is formally established in Online Appendix A. Consider a single student of type (u, t). TEPS assumes that she always receives her stable assignment in every realized uncertainty. Integrating over all uncertainties, we obtain a distribution of her possible assignment outcomes that is unique in a large enough economy. Because there are typically multiple ROLs that lead to the same outcome distribution, the mapping from (u, t) to best responses is indeed a correspondence. Yet, TEPS will infer the same set of preferences no matter which best-responding ROL is observed, because the TEPS procedure only depends on the outcome distribution of the student. Hence, the mapping from (u, t) to TEPS inferred preferences is a function, implying that the TEPS hypothesis is complete.

Once students' preferences are inferred, one may employ a parametric function of utilities to fit them with parameter estimation. As is popular in the literature, one may adopt a random utility model based on Extreme Value Type-1 (EVT1) or normal errors and estimate the parameters using maximum likelihood (MLE), Markov Chain Monte Carlo (MCMC) methods such as Gibbs sampler, or Expectation-Maximization (EM) algorithm. We recommend adopting a multinomial probit model (either with or without random coefficients) and estimating it via Gibbs sampler due to its flexibility in drawing (random) cardinal utilities that are consistent with the inferred preferences. For example, when the student's inferred preference is $c_1 \succ c_2$ and $c_1 \succ c_3 \succ c_4$, it is not straightforward to write down the exact likelihood function even with an EVT1 error. However, the inferred preferences provide restrictions on each school's utility, which can be easily incorporated into Gibbs sampler.[26]

3.5 Selecting from among Alternative Preference Hypotheses

Proposition 2 suggests a nested family of alternative preference inference hypotheses. Each hypothesis in the spectrum offers a different way of trading off informational efficiency and robustness of inferences. Which method is best depends on the extent of mistakes being made and their payoff significance. Since the latter issue is empirical in nature, it would be ideal for the selection to be driven by the data.

Specifically, we develop a sequential Wald test method, justified by the nesting structure in Proposition 2. To this end, we assume that TEPS^{top}, the most robust inference, is consistent. Using this as the alternative hypothesis, we test sequentially the consistency of more informationally demanding models, starting with WTT, and then TEPS$^{\tau}$ in the descending order of τ's.

For concreteness, suppose we wish to estimate a parameter vector β in a parametric model, say.

[26] In the example, assuming no outside options, the bounds are given by $u_1 \in (\max\{c_2, c_3\}, \infty), u_2 \in (-\infty, u_1), u_3 \in (u_4, u_1), u_4 \in (-\infty, u_3)$. For each student, one iteration of Gibbs sampler draws the utility of each school sequentially with truncation imposed by the bounds. The procedure is given in the Online Appendix C.

Each method leads to an estimator for β. We let them be the WTT-based estimator ($\widehat{\beta}^{\text{WTT}}$) and the TEPS$^{\tau}$-based estimator ($\widehat{\beta}^{\tau}$). As an example, let us consider TEPS^{top} and 10 other possible values for τ, $\tau_1 = 10, \tau_2 = 20, \ldots, \tau_9 = 90, \tau_{10} = 100$. Note that TEPS$^{\tau_{10}}$=TEPS^{all}. The following procedure selects the empirical method which is consistent and efficient among all consistent estimators that we consider due to the nesting structure in Proposition 2.[27]

Step 1: Test TEPS^{top} vs. WTT We formulate two hypotheses. H_0^{WTT}: $\widehat{\beta}^{\text{WTT}}$ and $\widehat{\beta}^{top}$ are both consistent, while $\widehat{\beta}^{\text{WTT}}$ is efficient; H_1^{WTT}: only $\widehat{\beta}^{top}$ is consistent. Since WTT leads to an efficient estimator under the null, we conduct a Wald test based on the statistic,

$$\left(\widehat{\beta}^{top} - \widehat{\beta}^{\text{WTT}}\right)' \left(V(\widehat{\beta}^{top}) - V(\widehat{\beta}^{\text{WTT}})\right)^{-1} \left(\widehat{\beta}^{top} - \widehat{\beta}^{\text{WTT}}\right)$$

where $V(\cdot)$ denotes the covariance matrix of its argument. Under the null, the test statistic follows a $\chi^2_{|\beta|}$ distribution. Justified by the nesting structure (Proposition 2) and the maintained assumption of TEPS^{top}, if H_0^{WTT} is not rejected, we select $\widehat{\beta}^{\text{WTT}}$ and exit the procedure; otherwise, we perform the next test.

Step 2: Test TEPS^{top} vs. TEPS$^{\tau}$ We start with the largest $\tau = \tau_{10} = 100$. H_0^{τ}: $\widehat{\beta}^{\text{top}}$ and $\widehat{\beta}^{\tau}$ are both consistent, while $\widehat{\beta}^{\tau}$ is efficient; H_1^{τ}: only $\widehat{\beta}^{\text{top}}$ is consistent. We conduct a Wald test based on the statistic,

$$\left(\widehat{\beta}^{\text{top}} - \widehat{\beta}^{\tau}\right)' \left(V(\widehat{\beta}^{top}) - V(\widehat{\beta}^{\tau})\right)^{-1} \left(\widehat{\beta}^{\text{top}} - \widehat{\beta}^{\tau}\right).$$

Under the null, the test statistic follows a $\chi^2_{|\beta|}$ distribution. Justified by the nesting structure (Proposition 2) and the maintained assumption of TEPS^{top}, if H_0^{τ} is not rejected, we select $\widehat{\beta}^{\tau}$; otherwise, we proceed to test the next largest $\tau = \tau_9$. We continue until we reach a value of τ such that H_0^{τ} is not rejected, in which case we select $\widehat{\beta}^{\tau}$, or we reject $H_0^{\tau_1}$, in which case we select $\widehat{\beta}^{top}$.

The order of these steps guarantees the desired size of the tests and makes it appropriate to use the Hausman-type test statistics.

Besides formal statistical tests, some descriptive analyses can be informative as well. For example, we may compare if there are any systematic differences among different groups of schools, e.g., schools ranked in an ROL, ever-feasible unranked schools, and never-feasible unranked schools. WTT implies that there should not be any difference between ever-feasible unranked schools and never-feasible unranked schools, while both groups are worse than all ranked schools. In contrast,

[27] One drawback of the testing procedure is that it relies on the parametric model and may suffer from model misspecification. In that case, if one is willing to consider nonparametric models, the literature on nonparametric tests of nesting models may provide more general tests.

TEPS distinguishes unranked schools by their feasibility status.

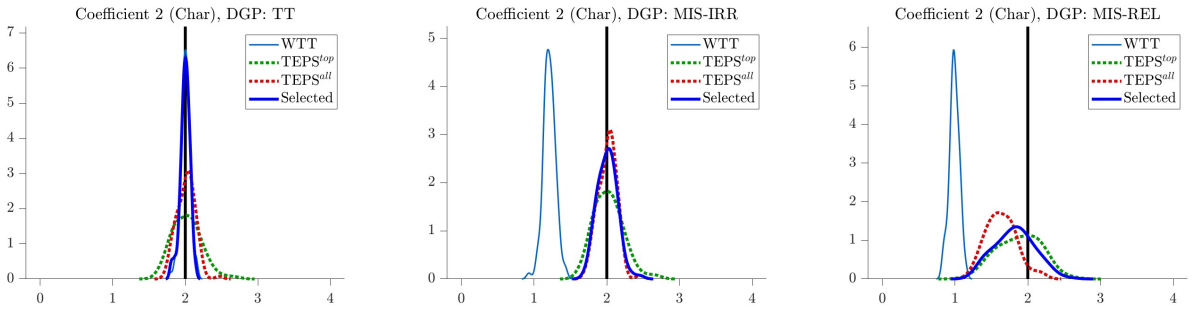
3.6 Performance of TEPS: Monte Carlo Simulations

We analyze the performance of TEPS via Monte Carlo simulations; see Online Appendix B for details. Specifically, we consider 100 copies of finite economies in which 1,000 students apply and are assigned 12 schools through a student-proposing DA algorithm with coarse priorities and a single tie-breaking rule, similar to the NYC high school choice. Each copy contains independent-randomly generated priorities and preferences that follow a parametric random utility model. We consider three data generating processes (DGP); (i) Truth-telling (TT) in which each student submits a truthful ROL; (ii) Payoff Irrelevant Mistakes (MIS-IRR) in which students may skip or flip (i.e., arbitrarily rank) out-of-reach schools, and (iii) Payoff Relevant Mistakes (MIS-REL) in which students may additionally skip schools with small but positive admission chances. For each of 300 scenarios (3 DGPs $\times$ 100 copies), we estimate the preference parameters using Gibbs sampler, based on WTT and TEPS with attention parameters $\tau = 0, 20, 40, 60, 80, 100$. After the estimation, we perform the testing procedure in Section 3.5 to choose the optimal estimate and call it the "selected" estimate.

The results are summarized in Figure 5. First, when the DGP is TT, all estimates are consistent, while the WTT-based estimator has the smallest variance since it uses the maximal, albeit possibly unreliable, information that one can infer from observed ROLs. Second, when some students make (almost) payoff-irrelevant mistakes, the WTT-based estimator is susceptible to strategic mistakes with significant bias. On the other hand, the estimators based on TEPS are robust to payoff-irrelevant mistakes. Third, when some students make payoff-relevant mistakes, TEPS^{all} is no longer consistent because TEPS^{all} does not take students' payoff-relevant mistakes into consideration. On the other hand, TEPS^{top} remains consistent, so do TEPS$^{\tau}$ with $\tau < 100$ (see Online Appendix B.)

Due to these bias-variance tradeoffs, our testing procedure selects WTT-based estimates 94% out of 100 samples when the DGP is TT, but never selects WTT-based estimates when the DGP is MIS-IRR or MIS-REL. Next, when students make no payoff-relevant mistakes (MIS-IRR), the procedure selects TEPS^{all} 87% out of 100 samples since it uses the maximal information among all TEPS-based methods which are robust to payoff-relevant mistakes (Proposition 2). Finally, when students make payoff-relevant mistakes (MIS-REL), the procedure selects TEPS^{80} 68% out of 100 samples since TEPS^{all} may no longer be robust to payoff-relevant mistakes.

(a) Distribution of Estimates



(b) Test Results: Fraction of Each Method Being Selected

Data Generating Process:		TT	MIS-IRR	MIS-REL
Estimation method	WTT	**0.94**	0	0
	TEPS^{top}	0.01	0	0
	TEPS^{20}	0	0	0
	TEPS^{40}	0	0	0.01
	TEPS^{60}	0	0.01	0.14
	TEPS^{80}	0.03	0.12	**0.68**
	TEPS^{all}	0.01	**0.87**	0.17

Figure 5: Monte Carlo Simulations: Performance of TEPS and WTT

Note: See Online Appendix B for the exact description of our Monte Carlo simulation. In panel (a), we plot the kernel density plots of the estimates of β_2 from 100 Monte Carlo samples. The left subfigure corresponds to TT DGP, the center subfigure corresponds to MIS-IRR DGP, and the right subfigure corresponds to MIS-REL DGP. The black vertical line at 2 denotes the true value of the parameter. We depict only WTT, TEPS^{top}, TEPS^{all}, and the Selected estimates for conciseness. In panel (b), we report the fraction of each estimate being chosen as the 'Selected' estimate among 100 Monte Carlo samples, according to the testing procedure in Section 3.5.

4 High School Choice in Staten Island, New York City

We now turn to the study of high school choice in NYC. There are two purposes for this study. First, we deploy and operationalize our framework to actual data, and use it as proof of concept. Specifically, we will put alternative assumptions of applicant inattention/mistakes to estimate student preferences and let the aforementioned testing procedure select the appropriate assumption. Second, and no less importantly, we highlight the importance of choosing the right assumption to infer student preferences by comparing the predicted effects of several proposed desegregation policies based on alternative assumptions on students' mistakes.[28] We show that alternative assumptions lead to different predictions, and ignoring students' mistakes in estimation may likely lead to an

[28]There is a substantive interest among the public and policymakers in diversifying and desegregating student bodies of NYC schools. Recently, the NYC DOE proposed to remove screening based on test scores for admissions and remove geographic preferences based on students' residence (Shapiro, 2021). Some versions of these proposals were implemented and subsequently reversed (Closson, 2022).

underestimation of the predicted effects.

NYC public high school admissions use the student-proposing DA (Abdulkadiroglu, Pathak, and Roth, 2005). Students apply to up to 12 school programs, and each program in a school has its own capacity and independently conducts its admissions.

We focus on students and programs in Staten Island (SI, hereafter) that participated in Round 1, the main round, of NYC public high school choice in the 2016–17 academic year. SI is one of the five boroughs of NYC and can be considered as an independent matching market.[29]

Column (1) of Table 1 presents the summary statistics of the students, and Table 2 describes the programs and schools. Most of the students participate in Round 1 and enroll in their assigned school from that round. In total, we have 3,731 students and 50 programs at 9 schools. The length limitation of the ROL is not binding in SI. The average length was 4.05 and only about 3% of the students in our sample exhausted the list. Online Appendix D gives more details on the data sources and sample restrictions. Almost all middle school students who reside in SI are enrolled in middle schools in SI (about 95%) and apply to high school programs only in SI (see panel D of Table 1). SI is on average richer and has more white students relative to NYC, with a median household income $74,580 vs. $55,191 and the proportion of white students 53% vs. 15%.

4.1 Preference Estimation

To represent student preferences over programs, we adopt a random utility model.[30] We divide the sample into cells and allow the parameters to vary freely across cells à la Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020). Specifically, we construct eight (mutually exclusive and exhaustive) cells based on three binary student characteristics—gender, free or reduced-price lunch (FRPL) status, and Black/Hispanic. Columns (2)–(9) of Table 1 present the summary statistics of the students in each cell.

Let u_{icps} denote the utility of student i in cell c when she is matched with program p at school s. We parameterize the utility as follows:

$$u_{icps} = \alpha_{cs} + \beta_{1,c} D_{is} + \beta_{2,c} \textit{Nearest}_{is} + \beta_{3,c} z_p + \sum_l \gamma_c^l x_i^l Z_s^l + \sum_k \delta_c^k x_i^k z_p^k + \sigma_{\tau(p),c} \epsilon_{icps}, \qquad (1)$$

where α_{cs}'s are cell-specific school fixed effects; D_{is} is the distance from student i's residence to

[29] SI is connected with the rest of the city only by the Staten Island Ferry or the Verrazzano-Narrow Bridge. There are pros and cons of analyzing SI instead of the entire NYC. While the direct implications for policies in NYC may be limited, SI's racial profile and size are closer to the medium size cities in the rest of the United States and hence the implications are more relevant for other US cities.

[30] We do not model outside options. Almost all students in our sample (3,724 out of 3,731) were assigned in the data (see Table 1), and less than 2% of them chose not to enroll in their assigned school.

Table 1: Summary Statistics: Sample Means across Students

	(1) Total	(2) Cell 1	(3) Cell 2	(4) Cell 3	(5) Cell 4	(6) Cell 5	(7) Cell 6	(8) Cell 7	(9) Cell 8
Panel A. Cell characteristics									
Female	0.49	no	yes	no	yes	no	yes	no	yes
Free or reduced-price lunch (FRPL)	0.53	no	no	yes	yes	no	no	yes	yes
Black/Hispanic	0.36	no	no	no	no	yes	yes	yes	yes
Panel B. Student characteristics									
Asian	0.10	0.10	0.09	0.25	0.22	0.00	0.00	0.00	0.00
White	0.53	0.90	0.90	0.73	0.75	0.00	0.00	0.00	0.00
Black	0.12	0.00	0.00	0.00	0.00	0.30	0.27	0.33	0.34
Hispanic	0.25	0.00	0.00	0.00	0.00	0.70	0.74	0.67	0.66
Median Income ($1,000)	74.58	85.21	86.09	77.72	78.42	65.97	66.56	59.55	57.32
Grade 7 ELA	310.87	317.38	326.83	306.23	316.91	299.04	319.93	289.67	301.48
Grade 7 Math	310.21	326.72	326.11	313.36	315.34	296.88	308.43	285.54	287.27
Panel C. Admission outcomes									
Matched with top choice	0.57	0.61	0.60	0.56	0.55	0.55	0.54	0.54	0.56
Unassigned	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00
Panel D. Submitted ROLs									
# of SI Choices	4.05	4.03	3.97	4.01	4.06	4.14	3.69	4.11	4.20
# of Choices (all)	4.06	4.05	3.98	4.02	4.07	4.16	3.70	4.12	4.21
Observations	3731	710	743	471	455	178	136	545	493

Table 2: Summary Statistics: Sample Means across Programs/Schools

	Program		School	
	Mean	Std. Dev.	Mean	Std. Dev.
Capacity	134.06	(215.47)	744.78	(551.49)
9th Grade Size	71.30	(108.79)	396.78	(293.48)
% High Performer (ELA)	31.08	(26.82)	34.12	(15.01)
% High Performer (Math)	26.21	(25.46)	25.83	(12.19)
% White	36.32	(22.14)	43.93	(21.80)
% Asian	7.94	(7.39)	7.61	(3.40)
% Black	19.93	(13.82)	17.51	(12.79)
% Hispanic	34.45	(16.76)	29.57	(12.72)
% Free or Reduced-Price Lunch (FRPL)	63.40	(18.08)	60.50	(13.65)
$\mathbb{1}$(STEM)	0.20	(0.40)		
Observations	50		9	

Note: Programs that are 'For Continuing 8th Graders' and zoned do not have capacity restrictions. We count the total number of students who are eligible at those programs and report that as their capacity.

school s;[31] $Nearest_{is}$ is a binary variable which equals 1 if D_{is} is the smallest for i among all schools and 0 otherwise; x_i is a vector of student characteristics to be interacted with program characteristics z_p and school characteristics Z_s. Specifically, x_i includes 7th grade standardized ELA and Math

[31] We use the exact address for schools. For students, we use the centroid of the census tract that a student resides in as her residence. We calculate the line distance between the two using the Haversine formula.

scores, their average, and median neighborhood income. z_p and Z_s include the proportions of high performers in 7th grade standardized ELA and Math (a high performer is defined as being above the 75th percentile), the size of the 9th grade, the proportion of each race/ethnicity group, the proportion of students on FRPL (calculated at program and school levels, respectively), and an indicator as to whether program p's focus is in STEM fields. ϵ_{icps} is i.i.d. standard normal conditional on the above observables. $\sigma_{\tau(p),c}$ allows for heterogeneous variances of the unobserved idiosyncratic preference shocks based on program type $\tau(p)$ (STEM, or others), and we normalize $\sigma^2_{others,c} = 1, \forall c$.

We estimate equation (1) separately for the eight cells. In terms of observables, Table 1 shows that there is substantial heterogeneity across the cells. For example, the median neighborhood income of cell 1, which consists of male, non-FRPL, non-Black/Hispanic students, is $85,207, much higher than $57,321 for cell 8, consisting of female, FRPL, Black/Hispanic students. For each cell, we employ a Gibbs sampling procedure for the Hierarchical Bayesian estimation, which is described in detail in Online Appendix C.

4.1.1 Comparing Assumptions for Inferring Student Preferences

The full preference estimates are reported in Online Appendix F.1. For TEPS-based estimators, we consider TEPS^{top}, TEPS^{10}, TEPS^{20}, $\cdots$, TEPS^{90}, and TEPS^{all}. Below, we present the test results (Section 3.5) and descriptive statistics that are consistent with the test results. We then evaluate how well each set of estimates fit the data.

Test Results: Selected Estimates. The estimates *selected* by our testing procedure in Section 3.5 are reported in Table 3. It shows that WTT is never selected, implying that students in our data across all the cells tend to make some, possibly payoff-insignificant, mistakes. Furthermore, the selected version ranges from TEPS^{20} to TEPS^{all} across the cells, indicating that there is some heterogeneity in the degree to which such mistakes are made in terms of students' observable characteristics.

Table 3: Test Results: Selected Estimates

Cells	1	2	3	4	5	6	7	8
Female	no	yes	no	yes	no	yes	no	yes
Free or reduced-price lunch (FRPL)	no	no	yes	yes	no	no	yes	yes
Black/Hispanic	no	no	no	no	yes	yes	yes	yes
Selected Estimates	TEPS^{70}	TEPS^{all}	TEPS^{40}	TEPS^{20}	TEPS^{70}	TEPS^{80}	TEPS^{50}	TEPS^{20}

Descriptive Analysis. Recall TEPS distinguishes unranked programs (in a student's ROL) depending on whether they are feasible or infeasible, whereas WTT makes no such distinction; i.e., it treats all unranked schools as if they are less preferred than ranked schools. To see which

treatment makes more sense, it is instructive to inspect whether the characteristics of programs vary systematically based on their feasibility status. Figure 6 reports the program's percentage of high-performing students in the average score[32] by their feasibility status.

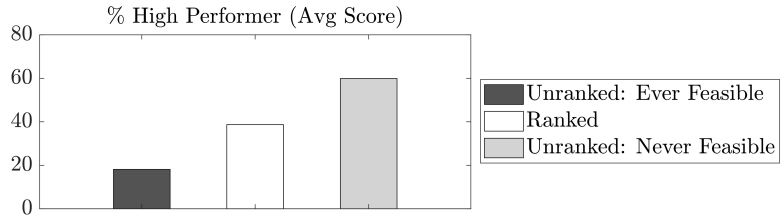


Figure 6: Characteristics of Ranked and Unranked Programs by Feasibility Status

Notes: For each student, we classify the programs into three types—ever-feasible-unranked, never-feasible-unranked, and ranked. Ranked programs are those included in the student's ROL, and Stages 1 and 2 of the TEPS^{all} procedure determine the feasibility of each unranked program. We focus on the fraction of high-performing students (measured by the average of ELA and math scores) in a program. In Online Appendix F.2, we report results on additional program characteristics. The figure reports the average across all students for each type of program.

There is a clear order among the three types of programs: ever-feasible unranked programs have the smallest percentages of high performers, never-feasible unranked programs have the largest percentages, and ranked programs lie in the middle. The fact the ever-feasible unranked programs has a smaller percentage than the never-feasible unranked programs is likely driven mechanically by the definition of feasibility, but the fact that the former has a higher percentage than ranked programs is noteworthy and is consistent with the selection of TEPS, as opposed to WTT, in all cells.[33]

Model Fit. To see how alternative estimates based on the different assumptions fit the data, we predict the matchings and calculate the average characteristics of the assigned programs. Specifically, we use each set of estimates to simulate students' ordinal preferences, let them report truthfully to DA, and obtain a matching for each lottery draw. Corollary 1 validates the assumption that students are truth-telling in the analysis of matching *outcomes*. For each program characteristic, we calculate the average over 40,000 simulations for each set of estimates. Table 4 reports how close the average is to the data. Note that we use the estimates from all eight cells together for these exercises.

We report the model fit based on WTT, TEPS^{all}, TEPS^{top}, and the selected estimates for conciseness. Two patterns are present. First, the WTT-based estimates do not fit the data well, while TEPS-based and the selected estimates closely replicate the data. This reflects that WTT is vulnerable to students' payoff-insignificant mistakes and infers potentially unreliable information

[32] A student is said to be a high performer if her average of ELA and math scores is above the 75th percentile.

[33] We obtain the same conclusion when we look at other program characteristics such as the percentages of high performers in ELA and math separately and the proportion of non-FRPL students (see Online Appendix F.2.)

Table 4: Model Fit: Average Characteristics of Assigned Programs

	Black and Hispanic Students					White, Asian, and Other Students				
	Data	Difference b/t model prediction & data				Data	Difference b/t model prediction & data			
		WTT	TEPS^{all}	TEPS^{top}	Selected		WTT	TEPS^{all}	TEPS^{top}	Selected
% High Performer (ELA)	27.66	-1.12	0.24	0.01	0.11	46.00	-0.36	-0.25	-0.39	-0.33
		(0.46)	(0.49)	(0.51)	(0.52)		(0.26)	(0.24)	(0.27)	(0.28)
% High Performer (Math)	20.66	-0.76	0.30	0.09	0.21	39.68	-0.56	-0.49	-0.44	-0.51
		(0.43)	(0.43)	(0.47)	(0.49)		(0.27)	(0.23)	(0.25)	(0.27)
% FRPL (Program Level)	67.06	1.33	0.55	0.75	0.86	49.86	1.02	0.63	0.68	0.62
		(0.31)	(0.36)	(0.43)	(0.40)		(0.23)	(0.17)	(0.21)	(0.19)
% White	34.19	-2.58	-0.88	-1.27	-1.30	58.41	-1.06	-0.58	-0.89	-0.65
		(0.45)	(0.48)	(0.57)	(0.55)		(0.33)	(0.26)	(0.29)	(0.26)
% Asian	7.29	0.16	0.47	0.48	0.52	9.27	-0.09	-0.05	-0.13	-0.12
		(0.15)	(0.17)	(0.17)	(0.16)		(0.08)	(0.09)	(0.09)	(0.09)
% Black	21.40	0.82	-0.04	0.13	0.23	9.75	0.23	0.20	0.46	0.30
		(0.29)	(0.32)	(0.36)	(0.34)		(0.16)	(0.17)	(0.18)	(0.16)
% Hispanic	35.91	1.59	0.42	0.63	0.53	21.61	0.83	0.36	0.49	0.41
		(0.33)	(0.34)	(0.39)	(0.38)		(0.22)	(0.17)	(0.19)	(0.18)
Size of 9th Grade	135.54	-15.89	-3.08	2.69	1.38	287.77	-13.32	-3.18	-7.57	-5.98
		(4.50)	(4.75)	(5.16)	(5.02)		(3.96)	(2.70)	(2.93)	(2.79)
1(STEM)	15.67	-0.89	-1.89	-2.81	-1.73	14.56	0.77	0.84	1.13	0.84
		(0.98)	(1.76)	(3.28)	(1.98)		(0.56)	(0.79)	(1.39)	(0.96)

Notes: We sample 200 draws from the posterior distribution of each parameter and, for each draw, draw 200 sets of lotteries and run DA $200 \times 200 = 40,000$ times. The mean and standard deviations across the preference estimates draws are reported.

about their preferences. Second, the predictions from TEPS^{top} tend to be less precise (i.e., have higher standard deviations) than those based on TEPS^{all} as well as the selected estimates (recall that TEPS$^{\tau}$ with large enough attention parameter τ were selected in Table 3). This is because TEPS^{top} uses less information about student preferences than other TEPS-based estimates.

4.2 Counterfactual Analysis

We now proceed to evaluate the effects of desegregation policies in NYC based on alternative estimates. Motivated by the desegregation policies that were recently debated and partially implemented by the NYC DOE, we consider the following three counterfactual scenarios:[34]

(i) No screening: no programs screen students based on their academic performances, which means that a lottery is the only tie-breaker to distinguish among applicants who belong to the same priority group.

(ii) No zoning: no programs prioritize students based on their residence. This amounts to removing zone priorities at zoned programs in SI.

(iii) No priorities: remove all priority rules so that admissions are based only on a single tie-breaking

[34]Similar policies are studied in Idoux (2022) for middle schools and Hahm and Park (2022) in a dynamic setting.

lottery.

We evaluate the desegregation effect of each scenario and compare it to the observed matching in the data given the current regime. To do so, for each set of estimates (WTT; TEPS$^{\tau}$ with $\tau = top, 10, \cdots, 90, all$; or selected), we randomly select 200 draws of u_{icps} for each (i, c, p, s) from its respective posterior distribution and let each student submit an ROL representing her true ordinal preferences per COROLLARY 1. We then draw 200 sets of random lotteries, construct students' priority scores under each scenario, and run DA to obtain counterfactual outcomes. We report the average of each measure we consider across 200×200=40,000 simulation results for each scenario and each set of estimates. For conciseness, we only report the results from using the WTT estimates, TEPS^{top} estimates, and the selected estimates.

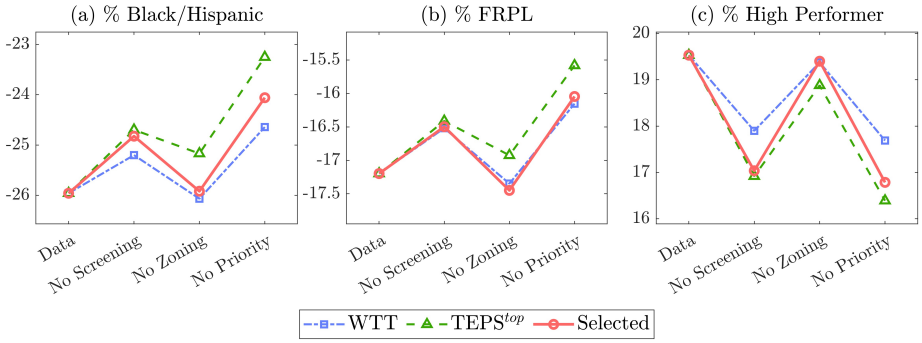


Figure 7: Racial Gap in Characteristics of Assigned Programs

We evaluate the effects of the three policies on the racial gap between Black or Hispanic students and others. Figure 7 focuses on the average characteristics of the assigned program.[35] The current regime produces a sizable racial gap. On average, a Black or Hispanic student is assigned to a program with 57% Black or Hispanic students, while a student of other racial and ethical groups is assigned to a program with only 31% of Black or Hispanic, leading to a gap of -26 percentage points (panel a). Similarly significant gaps are present in terms of the proportion of FRPL students (panel b) and high performers (panel c).

We find two prominent patterns in the predicted policy effects: 1) WTT-based estimates which neglect students' mistakes/inattention may underestimate the policy effects, and 2) the overall predicted effects are modest which suggests that forces other than priority rules may be more

[35]For each student in each racial group, we calculate three observable characteristics of the assigned program—proportion of Black and Hispanic students, proportion of FRPL students, and average 7th-grade standardized test score. We then calculate the mean across all students in each racial group and report it in Figure 7.

important in desegregating schools.

First, notably, the effects of the policies are underestimated if one uses the WTT-based estimates for the counterfactual analysis. For example, WTT would predict that the decrease in the racial gap in the percentage of Black or Hispanic in the assigned program is only half (about 1 percentage point) compared to what the selected estimates predict. This underestimation is expected. WTT neglects students' mistakes/inattention and assumes that students do not prefer unranked out-of-reach schools, so they would not apply to them even when these schools become within reach under a counterfactual policy.

Second, given the drastic nature of the contemplated policies, one may view the predicted effects, even under our selected estimates, to be rather small. While our selected estimates show that the no-priority policy reduces the racial gaps most significantly, the magnitude is modest.[36] This may be because there are other forces, for example, residential segregation, that may play an important role in school segregation (Laverde, 2022; Park and Hahm, 2022). In Online Appendix F.2.2, we present descriptive evidence that is supportive of such a hypothesis. The current finding raises an important question on the effectiveness of removing priorities as a tool for achieving desegregation.

5 Conclusion

This paper explores what we can learn about students' preferences when they make payoff-insignificant mistakes in a school choice setting under DA. We study a general DA matching market in which students may face uncertainties about their outcomes, and consider a robust equilibrium—a weakening of the dominant strategies equilibrium—in which no applicant makes payoff-significant mistakes. We show that a robust equilibrium is asymptotically stable—the proportion of students assigned their favorite feasible schools converges to one in probability. Based on the theoretical finding, we develop a novel, three-step procedure called the *Transitive Extension of Preferences from Stability* that exploits the structure of the uncertainties present in a matching environment to robustly infer students' preferences. We show that it infers the maximal information on students' preferences justified by stability and transitivity of preferences. We then apply our method to estimate students' preferences from NYC's high school assignment data and evaluate the effects of several desegregation policies. The empirical results illustrate that the truth-telling assumption can lead to biased estimates and policy implications when students make payoff-insignificant mistakes,

[36]For example, the racial gap in the percentage of Black or Hispanic in the assigned program is reduced by 2 percentage points. Dropping screening alone narrows the gap by only 1 percentage point, and dropping zoning alone reduces the gap only by a negligible amount.

while our method is robust to such mistakes.

References

Abdulkadiroglu, A., N. Agarwal, and P. A. Pathak (2017): "The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match," *American Economic Review*, 107(12), 3635–89.

Abdulkadiroglu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 364–367.

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020): "Do parents value school effectiveness?," *American Economic Review*, 110(5), 1502–39.

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

Agarwal, N., and P. Somaini (2018): "Demand analysis using strategic reports: An application to a school choice mechanism," *Econometrica*, 86(2), 391–444.

Arteaga, F., A. J. Kapor, C. A. Neilson, and S. D. Zimmerman (2022): "Smart matching platforms and heterogeneous beliefs in centralized school choice," *The Quarterly Journal of Economics*, 137(3), 1791–1848.

Artemov, G., Y.-K. Che, and Y. He (2021): "Strategic 'Mistakes': Implications for Market Design Research," *Working Paper*.

Artemov, G., Y.-K. Che, and Y. He (2023): "Stable Matching with Mistaken Agents," *Journal of Political Economy Microeconomics*.

Azevedo, E. M. (2014): "Imperfect competition in two-sided matching markets," *Games and Economic Behavior*, 83, 207–223.

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

Calsamiglia, C., C. Fu, and M. Güell (2020): "Structural estimation of a model of school choices: The boston mechanism versus its alternatives," *Journal of Political Economy*, 128(2), 642–680.

Chen, Y., and O. Kesten (2017): "Chinese college admissions and school choice reforms: A theoretical analysis," *Journal of Political Economy*, 125(1), 99–139.

——— (2019): "Chinese college admissions and school choice reforms: An experimental study," *Games and Economic Behavior*, 115, 83–100.

Chen, Y., and T. Sönmez (2006): "School choice: an experimental study," *Journal of Economic Theory*, 127(1), 202–231.

Chiappori, P.-A., and B. Salanié (2016): "The Econometrics of Matching Models," *Journal of Economic Literature*, 54(3), 832–861.

Closson, T. (2022): "In a Reversal, New York City Tightens Admissions to Some Top Schools," *New York Times*.

Combe, J., U. M. Dur, O. Tercieux, C. Terrier, and M. U. Ünver (2022): "Market Design for Distributional Objectives in (Re) assignment: An Application to Improve the Distribution of Teachers in Schools," .

Combe, J., O. Tercieux, and C. Terrier (2022): "The Design of Teacher Assignment: Theory and Evidence," *The Review of Economic Studies*, 89(6), 3154–3222.

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

Dubins, L. E., and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *The American Mathematical Monthly*, 88(7), 485–494.

Fack, G., J. Grenet, and Y. He (2019): "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economics Review*, 109(4), 1486–1529.

Gelman, A., and D. B. Rubin (1992): "Inference from iterative simulation using multiple sequences," *Statistical science*, 7(4), 457–472.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2002): "Quantal Response Equilibrium and Overbidding in Private-Value Auctions," *Journal of Economic Theory*, 104(1), 247–272.

Hahm, D. W., and M. Park (2022): "A Dynamic Framework of School Choice: Effects of Middle Schools on High School Choice," *Working Paper*.

Hassidim, A., A. Romm, and R. I. Shorrer (2021): "The limits of incentives in economic matching procedures," *Management Science*, 67(2), 951–963.

He, Y. (2017): "Gaming the Boston School Choice Mechanism in Beijing," Toulouse School of Economics and Rice University, Unpublished mimeo.

He, Y., S. Sinha, and X. Sun (2021): "Identification and Estimation in Many-to-one Two-sided Matching without Transfers," *arXiv preprint arXiv:2104.02009*.

Idoux, C. (2022): "Integrating New York City Schools: The Role of Admission Criteria and Family Preferences," *Working Paper*.

Larroucau, T., and I. Rios (2020): "Do "short-list" students report truthfully? Strategic behavior in the Chilean college admissions problem," *Working Paper*.

Laverde, M. (2022): "Distance to Schools and Equal Access in School Choice Systems," Discussion paper.

Li, S. (2017): "Obviously strategy-proof mechanisms," *American Economic Review*, 107(11), 3257–87.

Matějka, F., and A. McKay (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 105(1), 272–298.

Myerson, R. B. (1978): "Refinements of the Nash equilibrium concept," *International journal of game theory*, 7, 73–80.

Otero, S., N. Barahona, and C. Dobbin (2021): "Affirmative action in centralized college admission systems: Evidence from Brazil," Discussion paper, Working paper.

Park, M., and D. W. Hahm (2022): "Location Choice, Commuting, and School Choice," *Working Paper*.

Rees-Jones, A. (2017): "Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match," *Games and Economic Behavior*.

Rees-Jones, A., and S. Skowronek (2018): "An experimental investigation of preference misrepresentation in the residency match," *Proceedings of the National Academy of Sciences*, 115(45), 11471–11476.

Roth, A. E. (1982): "The economics of matching: Stability and incentives," *Mathematics of operations research*, 7(4), 617–628.

Shapiro, E. (2021): "New York City Will Change Many Selective Schools to Address Segregation," *New York Times*.

Shorrer, R. I., and S. Sóvágó (2023): "Dominated Choices in a Strategically Simple College Admissions Environment," *Journal of Political Economy Microeconomics*.

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

A Proofs from Sections 2 and 3

We first introduce a general priority structure that nests the priority structure assumed in the main text as a special case. We then establish asymptotic stability using that general priority structure in Theorem 1′ below, which will in turn imply Theorem 1.

A.1 A General Priority Structure

Here, we consider a more general priority structure than those consisting of (a), (b), and (c), assumed in the text. Recall the ex-ante priority types T. Let $T_d := \{t \in T : \underline{s}_c^t = \overline{s}_c^t, \forall c\}$ be the types whose scores are all degenerate, and let $T_n := T \setminus T_d$ denote types whose scores for some schools are non-degenerate. We assume that T is compact, that T_n is closed, and that there exists some $\kappa > 0$ such that $\overline{s}_c^t - \underline{s}_c^t > \kappa$ if $\overline{s}_c^t - \underline{s}_c^t > 0$. Further, for each $t \in T_n$, Φ^t is absolutely continuous and admits density ϕ^t on S^t. As in the text, let $\widetilde{\eta}$ denote the probability measure of $\widetilde{\theta} \in \widetilde{\Theta} := [\underline{u}, \overline{u}]^C \times T$. In summary, the continuum economy is summarized by $E = [\widetilde{\eta}, S, (\Phi^t)_t]$.

In addition to Marginal Full Support, we introduce two additional conditions, which will be later shown to hold under the priority structure satisfying (b) and (c).

Assumption 2 (Finite Atoms [at Extremal ex-post Scores]). *The distribution of $(\underline{s}_c^t, \overline{s}_c^t)_t$, viewed as a random function of t, has at most a finite number of atoms.*

This assumption can be seen as weakening the assumption of AL model. First, for model (a), with no uncertainty on ex-post scores, we have $T = T_d$, with $\underline{s}_c^t = \overline{s}_c^t$ for all $t \in T$. In this case, the atomlessness of ex-post measure η (which is also assumed in AL) implies that the infimum and supremum scores have no point mass, so Finite Atoms will be trivially satisfied. However,

this condition does allow point mass for infimum and supremum scores but only finitely many of them; this is the case of coarse priorities in (b). A mass of students belong to each of finitely many priorities, so there will be a finite number of atoms in the infima and suprema of their ex-post scores.

For the last condition, for any $\delta \in (0,1)$, we say a school c is δ-**feasible** for type t given p if $\overline{s}_c^t - \delta > p_c$ and δ-**infeasible** for type t given p if $\underline{s}_c^t + \delta < p_c$. Plainly, δ-feasibility and δ-infeasibility mean feasibility and infeasibility, respectively, with probabilities that are bounded away from zero.

Assumption 3 (Rich Uncertainty). *Fix any $p \in [0,1]^C$ and any $\delta < \bar{\delta}$, for some $\bar{\delta} > 0$. There exists $\beta(\delta) > 0$ such that, for any $t \in T$, for any δ-feasible schools $a, b \in C$, and for any set $C^t \subset C \setminus \{a,b\}$ of δ-infeasible schools,*

$$\Pr\{s_c^t < p_c, \forall c \in C^t, \text{ and } s_a^t > p_a, s_b^t > p_b\} > \beta(\delta), \tag{A.1}$$

whenever $\Pr\{s_c^t < p_c, \forall c \in C^t, s_a^t > p_a\} > 0$ *and* $\Pr\{s_c^t < p_c, \forall c \in C^t, s_b^t > p_b\} > 0$.[A-1]

The last condition, which we argue is quite weak, is most technical and thus requires unpacking. In words, the condition states that whenever school a is feasible or b is feasible, while all δ-infeasible schools in C^t are infeasible, each with positive probability, then a and b are *both simultaneously* feasible and C^t are infeasible with probability at least of $\beta(\delta)$, for some $\beta(\delta) > 0$.[A-2] Essentially, this rules out the case in which a student's scores for a and b are extremely negatively correlated. Hence, the condition is quite weak, and as we show below, is satisfied in all realistic environments mentioned above.

Lemma 2. *Rich Uncertainty holds for priority structures, (a), (b), and (c).*

Proof. For case (a), $\underline{s}_c^t = \overline{s}_c^t$ for all $t \in T$, so the probability in the LHS of Equation (A.1) is equal to one. Hence, Rich Uncertainty holds trivially. We thus focus on (b) and (c).

Structure (b): STB version. The cutoff p_c for each school c corresponds to a lottery threshold $\ell_c^t \in [0,1]$, possibly dependent on the type t of the student. To see this, the cutoff can be expressed as: $p_c = \hat{t}_c + \frac{\lambda_c}{n_c}$, where the $\hat{t}_c$ is the intrinsic priority level that one needs for admission at c and λ_c is the lottery cutoff score for c. If $\overline{s}_c^t - \delta > p_c$ for some $\delta > 0$, either the student has an intrinsic priority level of $\hat{t}_c$ or a higher priority level. In the former, $s_c > p_c$ amounts to having a lottery draw above λ_c, so $\ell_c^t = \lambda_c$. In the latter, $s_c > p_c$ regardless of the lottery draw, so $\ell_c^t = 0$. Similarly, if $\underline{s}_c^t + \delta < p_c$ for some $\delta > 0$, either the student has an intrinsic priority level of $\hat{t}_{c,i}$ or a lower priority

[A-1] Rich Uncertainty is vacuously satisfied if schools with the specified restrictions do not exist. Also, if $C^t = \emptyset$, the required condition reduces to $\Pr\{s_a^t > p_a, s_b^t > p_b\} > \beta(\delta)$ whenever $\Pr\{s_a^t > p_a\} > 0$ and $\Pr\{s_b^t > p_b\} > 0$.

[A-2] One non-obvious condition is the existence of a lower bound probability $\beta(\delta)$ that is independent of t. The condition is still reasonable given the compactness of T and T_n.

level. In the former, $s_c < p_c$ amounts to having a lottery draw below λ_c, in which case $\ell_c^t = \lambda_c$. In the latter, $s_c < p_c$ regardless of the lottery draw, so $\ell_c^t = 1$.

Given the STB structure, the LHS of (A.1) then reduces to

$$\Pr\left\{\max\{\ell_a^t, \ell_b^t\} < \lambda^t < \min_{c \in C^t} \ell_c^t\right\}, \tag{A.2}$$

namely, the probability that the student draws an STB lottery number $\lambda^t \in (\max\{\ell_a^t, \ell_b^t\}, \min_{c\in C^t} \ell_c^t)$. Further, if $\Pr\{s_c^t < p_c, \forall c \in C^t, s_a^t > p_a\} > 0$ and $\Pr\{s_c^t < p_c, \forall c \in C^t, s_b^t > p_b\} > 0$, it must be that $\max\{\ell_a^t, \ell_b^t\} < \min_{c\in C^t} \ell_c^t$. Note that $\max\{\ell_a^t, \ell_b^t\} = \lambda_i$ for some $i \in \{a, b, x\}$ and $\min_{c\in C^t} \ell_c^t = \lambda_j$ for some $j \in C^t \cup \{y\}$ such that $\lambda_i < \lambda_j$, where $\lambda_x := 0$ and $\lambda_y := 1$. Hence, (A.2) equals

$$\begin{aligned}\Pr\left\{\max\{\ell_a^t, \ell_b^t\} < \lambda^t < \min_{c \in C^t} \ell_c^t\right\} &= \Pr\{\lambda_i < U < \lambda_j\} \\ &\geq \min\left\{|\lambda_c - \lambda_{c'}| : c, c' \in \hat{C}, \lambda_c \neq \lambda_{c'}\right\} =: \beta(\delta),\end{aligned}$$

where U is uniform random variable on $[0, 1]$ and $\hat{C} := C \cup \{x, y\}$. Importantly, this lower bound $\beta(\delta)$ does not depend on t or the particular pair (a, b) or other $c \in C^t$. It only depends on $(\lambda_c)_c$, which is determined uniquely by p.

Model (b): MTB version. The approach is similar to that of STB. In particular, the first part (mapping cutoffs p to lottery cutoffs $\lambda = (\lambda_c)_c$) is exactly the same. The difference is that the LHS of (A.1) is now equal to

$$\Pr\left\{\lambda_c^t < \ell_c^t, \forall c \in C^t, \ell_i^t < \lambda_i^t, \forall i = a, b\right\}, \tag{A.3}$$

where λ_i^t is t's MTB draw for school i. Clearly, given the hypothesis, this probability must be positive, which implies that $\ell_i^t < 1$ for $i = a, b$ and $\ell_c^t > 0$ for $c \in C^t$. Since each $\ell_j^t = \lambda_j$ for some $j \in C \cup \{x, y\}$, we have

$$\Pr\left\{\lambda_c^t < \ell_c^t, \forall c \in C^t, \ell_i^t < \lambda_i^t, \forall i = a, b\right\} \geq \Pr\left\{U^{(C-2:1)} < \lambda_* \text{ and } U^{(2:2)} > \lambda^*\right\} =: \beta(\delta),$$

where $\lambda_* := \min\{\lambda_i : i \in C \cup \{x, y\}, \lambda_i > 0\}$, $\lambda^* := \max\{\lambda_i : i \in C \cup \{x, y\}, \lambda_i < 1\}$, and $U^{(n:m)}$ is the m-th highest value of n independent draws of $U[0, 1]$. The proof is complete upon noting that the lower bound $\beta(\delta)$ is independent of t and of particular pair (a, b) or other $c \in C^t$.

Model (c): The argument is similar to that of Model (b). We first consider the NYC-exam school model where the ex post score is one-dimensional. The argument is similar to that of Model of (b)-STB. In particular, one can show that the LHS of (A.1) is lower bounded by the probability

that one's score lies within a certain interval:

$$\Pr\left\{\max\{p_a^t, p_b^t\} < s^t < \min_{c\in C^t} p_c^t\right\}. \tag{A.4}$$

Similarly to (b), if $\Pr\{s_c^t < p_c, \forall c \in C^t, s_a^t > p_a\} > 0$ and $\Pr\{s_c^t < p_c, \forall c \in C^t, s_b^t > p_b\} > 0$, then we must have $\max\{p_a^t, p_b^t\} < \min_{c\in C^t} p_c^t$. Hence, (A.4) is lower bounded by

$$\beta(\delta) := \varrho \cdot \min\{\delta \wedge |p_c - p_{c'}| : c, c' \in \hat{C}, p_c \neq p_{c'}\} > 0,$$

where $\hat{C} := C \cup \{x, y\}$ with $p_x := 0, p_y := 1$ and $\varrho := \min_{t\in T} \min_{s\in[\underline{s}^t, \overline{s}^t]} \phi^t(s) > 0$. The positivity of ϱ follows from the full and compact support assumption for each t and the compactness of T. In words, the probability lower bound is given by the shortest length of distinct score cutoffs or by δ, whichever is smaller. The relevance of δ here comes from the fact that, due to the δ-rejectability or δ-acceptability of the schools, the support of t's score in the interval $(\max\{p_a^t, p_b^t\}, \min_{c\in C^t} p_c^t)$ spans at least the length of δ, whenever that interval exceeds δ in length. The proof is complete upon noting that the lower bound $\beta(\delta)$ is independent of t and of particular pair (a, b) or other $c \in C^t$.

We next consider a model where ex post scores are not perfectly correlated. The argument is similar to that of Model of (b)-MTB. Again, one can show that

$$\begin{aligned}
&\Pr\{s_c^t < p_c, \forall c \in C^t, \text{ and } s_a^t > p_a, s_b^t > p_b\} \\
\geq &\Pr\{s_c^t < p_*, \forall c \in C^t, \text{ and } s_a^t > p^*, s_b^t > p^*\} \\
\geq &\hat{\varrho}^C \cdot (\delta \wedge p_*)^{C-2}(\delta \wedge (1-p^*))^2 =: \beta(\delta),
\end{aligned}$$

where $p_* := \min\{p_i : i \in C \cup \{x, y\}, p_i > 0\}$, $p^* := \max\{p_i : i \in C \cup \{x, y\}, p_i < 1\}$, and $\hat{\varrho} := \min_{t\in T} \min_{s\in S^t} \phi^t(s) > 0$. Again, the appearance of δ in the lower bound follows from the fact that the support of t's score for $i = a, b$ in the interval $[0, p_i]$ and its scores for $c \in C^t$ in the interval $[p_c, 1]$ each spans δ in length, whenever each interval exceeds δ in length, again due to the δ-acceptability and δ-rejectability of these schools. The proof is complete upon noting that the lower bound $\beta(\delta)$ is independent of t and of particular pair (a, b) or other $c \in C^t$. □

A.2 Proof of Lemma 1

Proof. We first show that the following condition, called *strict gross substitutes*, guarantees the uniqueness: η is such that for any $p, p' \in [0, 1]^C$ with $p < p'$,

$$(SGS) \qquad \sum_{c\in J} D_c(p') < \sum_{c\in J} D_c(p),$$

where $J := \{c \in C : p_c < p'_c\}$. Suppose to the contrary there are two stable matchings characterized by two cutoffs p and p'. By the lattice property, we can assume without loss that $p' > p$. Since both p and p' clear the markets and since $p'_c > p_c \geq 0$ for each $c \in J$, for each $c \in J$,

$$D_c(p) \leq S_c \text{ and } D_c(p') = S_c.$$

Summing across $c \in J$, we get

$$\sum_{c \in J} D_c(p') \geq \sum_{c \in J} D_c(p),$$

which contradicts (SGS).

To prove the statement, it now suffices to show that η satisfies (SGS). To this end, fix any $p, p' \in [0,1]^C$ with $p < p'$ and the corresponding set $J = \{c \in C : p_c < p'_c\}$. Observe that any ex-post type θ who demands a school in J at cutoffs p' never switches its demand to a school in $C \setminus J$ or to the outside option $\emptyset$, when the cutoffs shift to p. Hence, any such type continues to demand a (possibly different) school in J at cutoffs p. Next, consider the type (u, s) with $u_c > 0 > u_{c'}$, $c \in J$, $\forall c' \neq c$, and $s_c \in (p_c, p'_c)$. By Marginal Full Support, there is a positive mass of such types. These types could not demand any school at p' but now demand c at p. Collecting our observations, we conclude that $D_c(p') < D_c(p)$. Since the same result holds for each $c \in J$, (SGS) holds. □

A.3 Proof of Asymptotic Stability of Robust Equilibria

We now prove our main result under the general priority structure satisfying Marginal Full Support, Finite Atoms, and Rich Uncertainty:

THEOREM 1′. *Suppose η satisfies Marginal Full Support, Finite Atoms, and Rich Uncertainty. Then, any regular robust equilibrium is asymptotically stable.*

Note that THEOREM 1 follows from THEOREM 1′ in light of the fact that the priority structures in the former satisfy Finite Atoms and Rich Uncertainty.

Before proceeding with the proof, we need to perform a few preliminary analyses. Specifically, for each k-economy F^k, we study the strategy profile for that economy, or the k-truncation of σ, denoted by $\sigma^k := (\sigma_1, ..., \sigma_k)$. A later analysis requires us to consider the consequence of an arbitrary student i deviating to truthful reporting ρ. We denote the resulting profile by $\sigma_{(i)} := (\rho, \sigma_{-i})$ which is obtained by replacing the i-th component of σ with ρ. Likewise, $\sigma^k_{(i)}$ denotes the k-truncation of $\sigma_{(i)}$. Note that if $i > k$, then $\sigma^k_{(i)} = \sigma^k$. Let $P^k_{(i)}$ denote the cutoffs that would prevail if the students employ $\sigma^k_{(i)}$. We first establish a desirable limit behavior of $(P^k_{(i)})_{i \in \mathbb{N}_0}$ as $k \to \infty$, where $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

LEMMA 3. *Let σ be any γ-regular strategy profile. There exists a subsequence $\left\{F^{k_\ell}\right\}_\ell$ such that*

$$\sup_{0\le i\le k_\ell}\|P^{k_\ell}_{(i)}-\overline{p}\| \xrightarrow{p} 0 \text{ as } \ell\to\infty.$$

Proof. The proof follows exactly the same argument as ACH, upon noting that the uniqueness of stable matching follows from (SGS) (instead of ACH's full support assumption). □

Next, define a set

$$\tilde{\Theta}^\delta := \left\{(u,t)\in\tilde{\Theta} : \inf_{c,c'\in\tilde{C},c\neq c'}|u_c-u_{c'}|>\delta;\ \forall c,\ \overline{s}^t_c=\underline{s}^t_c \text{ or } \overline{s}^t_c\neq\bar{p}_c\Rightarrow|\overline{s}^t_c-\bar{p}_c|>\delta;\ \underline{s}^t_c\neq\bar{p}_c\Rightarrow|\underline{s}^t_c-\bar{p}_c|>\delta\right\}.$$

The Finite Atoms condition, together with atomlessness of η, implies that $\tilde{\eta}(\tilde{\Theta}^\delta)\to 1$ as $\delta\to 0$.[A-3]

Further, we introduce a few notations. Define $B_\nu(p):=\{p'\in[0,1]^C : ||p'-p||<\nu\}$. Consider a student $\tilde{\theta}$ and fix any arbitrary economy with other students playing some arbitrary reporting strategies. If she adopts any arbitrary reporting strategy $\sigma'_i(\tilde{\theta})$, this induces set of assignment probability $(X'_c)_{c\in\tilde{C}}$, where X'_c is the probability of the student being assigned to a school that she weakly prefers to c. Suppose she switches to TRS which induces assignment probability $(X^*_c)_{c\in\tilde{C}}$, where X^*_c represents the probability of the student being assigned a school that she weakly prefers to c. The strategyproofness of DA implies that TRS yields better assignment than any other strategy in the sense of first-order stochastic dominance:

FACT 1. *For each $c\in\tilde{C}$, $X^*_c-X'_c\ge 0$.*

Naturally, we define the **probability gain** from the switch to be $\max_{c\in\tilde{C}}(X^*_c-X'_c)$. The following preliminary result is useful.

LEMMA 4. *Fix any $\bar{p}\in[0,1]^C$ and any $\delta\in(0,\frac{1}{2}\min\{\kappa,\zeta\})$, where $\zeta:=\inf_{c,c'\in C\cup\{x,y\},\bar{p}_c\neq\bar{p}_{c'}}|\bar{p}_c-\bar{p}_{c'}|$. Then, there exist $\nu(\delta)>0$ and $\alpha(\delta)>0$ such that deviating from any non-SRS against p to TRS yields a probability gain of at least $\alpha(\delta)$, and thus a payoff gain of at least $\delta\alpha(\delta)$, for any student with type $\tilde{\theta}\in\Theta^\delta$, provided that non-SRS and TRS induce cutoffs p' and p'' both in $B_{\nu(\delta)}(p)$.*

Proof. Let $\beta(\delta)$ be the probability lower bound defined in Rich Uncertainty. Set $\nu(\delta)=\min\{\frac{\delta}{4},\frac{\beta(\delta)}{4\upsilon}\}$, where $\upsilon:=\max_{t\in T_n,s\in S^t}\phi^t(s)$. Fix a student with type $\tilde{\theta}=(u,t)\in\tilde{\Theta}^\delta$. Without loss, we index schools $\tilde{C}=\{1,...,C+1\}$ (including the outside option $\emptyset$) so that $u_j>u_{j'}$ if and only if $j<j'$.

[A-3] We note that the limit set $\tilde{\Theta}^-:=\tilde{\Theta}-\cup_{\delta>0}\tilde{\Theta}^\delta$ contains no mass point. This is seen by the fact that the only possibility of point mass in $\tilde{\Theta}^-$ may arise from a point mass occuring at some $\overline{s}^t_c=\underline{s}^t_c=p^t_c$ for some c. (Note that any type t with $\overline{s}^t_c>\underline{s}^t_c$ belongs to $\cup_{\delta>0}\tilde{\Theta}^\delta$, so it does not beling to $\tilde{\Theta}^-$.) But if there were such a point mass, then there must be a positive mass at score $s_c=p^t_c$, which contradicts the atomlessness of η.

Let C_j denote the set of schools that student $\tilde{\theta}$ prefers to j. Let $C^F \subset \tilde{C}$ be a set of *feasible* schools that the student would get with positive probabilities when she plays SRS against $\bar{p}$. It means that for each $j \in C^F$, we have

$$\Pr\{s_c < \bar{p}_c, \forall c \in C_j, s_j > \bar{p}_j\} > 0. \tag{A.5}$$

There could be a feasible, and less preferred, school outside C^F that she is not assigned but would have been with positive probability if she had ranked it sufficiently favorably. Since $\tilde{\theta} \in \tilde{\Theta}^\delta$, all these feasible schools are in fact δ-acceptable.

Now take any non-SRS σ' against p for that student. The strategy must rank some less-preferred, feasible—and thus δ-acceptable—school b ahead of some school $a \in C^F$, where she prefers a to b, and let a be the most preferred school in C^F that suffers from such a ranking-reversal by the non-SRS strategy. Let C_a be the schools that the student prefers to a among $\tilde{C}$. For the reversal to make a difference (which follows from σ' being non-SRS), we must have

$$\Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_b > \bar{p}_b\} > 0. \tag{A.6}$$

Suppose as hypothesized that σ' in the hypothesized economy induces a cutoff $p' \in B_{\nu(\delta)}$. We wish to compute the probability X'_a of the student receiving a or better according to the true preferences. Let $C' \subset C_a$ be the set of schools among C_a that σ' ranks ahead of a. Then,

$$\begin{aligned}
X'_a \leq & \Pr\{\exists c \in C' \text{ s.t. } s_c > p'_c\} + \Pr\{s_c < p'_c, \forall c \in C', s_b < p'_b, s_a > p'_a\} \\
\leq & \Pr\{\exists c \in C_a \text{ s.t. } s_c > p'_c\} + \Pr\{s_c < p'_c, \forall c \in C_a, s_b < p_b, s_a > p'_a\} \\
\leq & \Pr\{\exists c \in C_a \text{ s.t. } s_c > p_c\} + \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_b < \bar{p}_b, s_a > \bar{p}_a\} + \upsilon\|\bar{p}' - \bar{p}\| \\
\leq & \Pr\{\exists c \in C_a \text{ s.t. } s_c > \bar{p}_c\} + \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_b < \bar{p}_b, s_a > \bar{p}_a\} + \frac{\beta(\delta)}{4},
\end{aligned}$$

where the first inequality holds since it is possible that σ' ranks some other less preferred school other than b ahead of a, the second follows from the fact that $C' \subset C_a$, the third follows from the fact that for type $\tilde{\theta} \in \tilde{\Theta}^\delta$, $B_{\delta/4}(\bar{p})$ contains no point mass of scores and that $p', p \in B_{\nu(\delta)}(\bar{p}) \subset B_{\delta/4}(\bar{p})$, and the last inequality follows from the fact that $\nu(\delta) \leq \frac{\beta(\delta)}{4\upsilon}$ and $p' \in B_{\nu(\delta)}(\bar{p})$.

Suppose next the student switches to TRS and as result faces $p'' \in B_{\nu(\delta)}(p)$ as cutoffs. The probability of getting a or better schools from the switch is given by:

$$\begin{aligned}
X^*_a = & \Pr\{\exists c \in C_a \text{ s.t. } s_c > p''_c\} + \Pr\{s_c < p''_c, \forall c \in C_a, s_a > p''_a\} \\
\geq & \Pr\{\exists c \in C_a \text{ s.t. } s_c > \bar{p}_c\} + \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_a > \bar{p}_a\} - \upsilon\|p'' - \bar{p}\|
\end{aligned}$$

$$\geq \Pr\{\exists c \in C_a \text{ s.t. } s_c > \bar{p}_c\} + \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_a > \bar{p}_a\} - \frac{\beta(\delta)}{4}.$$

Again, the second inequality follows from the fact that for type $\tilde{\theta} \in \tilde{\Theta}^\delta$, $B_{\delta/4}(\bar{p})$ contains no point mass of scores and that $p'', p \in B_{\nu(\delta)}(\bar{p}) \subset B_{\delta/4}(\bar{p})$, and the last inequality follows from the fact that $\nu(\delta) \leq \frac{\beta(\delta)}{4v}$ and $p'' \in B_{\nu(\delta)}(\bar{p})$.

Consequently, the probability gain from switching from σ' to TRS is at least:

$$\begin{aligned}
\max_{c \in \tilde{C}}(X^*_c - X'_c) \geq & X^*_a - X'_a \\
& \geq \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_a > \bar{p}_a\} - \Pr\{s_c < p_c, \forall c \in C_a, s_b < \bar{p}_b, s_a > \bar{p}_a\} - \frac{\beta(\delta)}{2} \\
& = \Pr\{s_c < \bar{p}_c, \forall c \in C_a, s_a > \bar{p}_a, s_b > \bar{p}_b\} - \frac{\beta(\delta)}{2} \\
& \geq \beta(\delta) - \frac{\beta(\delta)}{2} = \frac{\beta(\delta)}{2}.
\end{aligned}$$

The last inequality follows from Rich Uncertainty. To see this, note first that each school in $C_a \setminus C^F$ is infeasible at p to $\tilde{\theta} = (u, t)$; then, since $\tilde{\theta} \in \tilde{\Theta}^\delta$, it is δ-rejectable at p for $\tilde{\theta}$. Consider next any $c \in C^F \cap C_a$. The school is feasible given p to $\tilde{\theta} = (u, t)$, but we must have $\underline{s}^t_c < \bar{p}$; otherwise, the student would not have been assigned a with positive probability under SRS given p. Then, since $\tilde{\theta} \in \tilde{\Theta}^\delta$, $\underline{s}^\delta_c < \bar{p} - \delta$, so it is δ-rejectable. Finally, recall that both a and b are feasible for $\tilde{\theta}$; again since $\tilde{\theta} \in \tilde{\Theta}^\delta$, both a and b are δ-acceptable for $\tilde{\theta}$. Therefore, (A.5) and (A.6) imply the inequality (A.1), leading to that last inequality.

Setting $\alpha(\delta) = \frac{\beta(\delta)}{2}$, we have established that the switch results in the probability gain of at least α. Since $\tilde{\theta} \in \tilde{\Theta}^\delta_p$, the associated payoff gain is

$$\sum_i (X^*_i - X^*_{i-1}) u^{\tilde{\theta}}_i - \left[\sum_i (X'_i - X'_{i-1}) u^{\tilde{\theta}}_i\right] \geq \sum_i (X^*_i - X'_i)(u^{\tilde{\theta}}_i - u^{\tilde{\theta}}_{i+1}) \geq (X^*_a - X'_a)(u^{\tilde{\theta}}_a - u^{\tilde{\theta}}_{a+1}) > \alpha\delta,$$

where $X^*_0 = X'_0 \equiv 0$. We note that this bound does not depend on $\tilde{\theta} \in \tilde{\Theta}^\delta$. □

Proof of Theorem 1. For any sequence $\{F^k\}$ induced by E, fix any arbitrary regular robust equilibrium $\{(\sigma^k_{1 \leq i \leq k})\}_k$. The strategies induce a random ROL, R_i, for each player i. We prove that the fraction of students who play non-SRS in $\left\{\left(\sigma^k_i\right)_{1 \leq i \leq k}\right\}_k$ converges in probability to zero as $k \to \infty$. Suppose not. Then, there exists $\varepsilon > 0$ and a subsequence of finite economies $\left\{F^{k_j}\right\}_j$ such that

$$\Pr\left(\text{The fraction of students playing SRS against } p^{k_j} \text{ is no less than } 1 - \varepsilon\right) < 1 - \varepsilon. \qquad (*)$$

By Lemma 3, there exists a sub-subsequence of economies $\left\{F^{k_{j_\ell}}\right\}_\ell$, such that the associated cutoffs $p^{k_{j_\ell}}$ converge to $\bar{p}$ in probability, where $\bar{p}$ are the deterministic cutoffs stated in Lemma 3.

We choose $\delta > 0$ small enough so that $\eta(\tilde{\Theta}^\delta) > (1-\varepsilon)^{1/3}$ (this can be done since η is absolutely continuous). We then choose $\nu(\delta)$ and $\alpha(\delta)$ according to Lemma 4.

By WLLN, we know that $\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta)$ converges to $\eta(\tilde{\Theta}^\delta)$ in probability, and therefore there exists L_1 such that for all $\ell > L_1$ we have

$$\Pr\left(\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \geq (1-\varepsilon)^{1/2}.$$

For each economy $F^{k_{j_\ell}}$, consider the event

$$A^{k_{j_\ell}} := \left\{ \sup_{0 \leq i \leq k_{j_\ell}} ||P_{(i)}^{k_{j_\ell}} - \bar{p}|| < \nu(\delta) \right\}.$$

Since $P_{(i)}^{k_{j_\ell}} \xrightarrow{p} \bar{p}$ uniformly over $i \in \mathbb{N}_0$, there exists L_2 such that, for all $\ell > L_2$ s.t. for all $\ell > L_2$ we have

$$\Pr\left(A^{k_{j_\ell}}\right) \geq \max\left\{(1-\varepsilon)^{1/6}, 1-(1-\varepsilon)^{1/2}\left[(1-\varepsilon)^{1/3} - (1-\varepsilon)^{1/2}\right]\right\}. \qquad (**)$$

Because $\left\{\left(\sigma_i^k\right)_{1 \leq i \leq k}\right\}_k$ is a robust equilibrium, there exists L_3 s.t. for all $\ell > L_3$ the strategy profile $(\sigma_i^{k_{j_\ell}})_{i=1}^{k_{j_\ell}}$ is a $\delta \cdot \alpha(\delta)\left[(1-\varepsilon)^{1/6} - (1-\varepsilon)^{1/3}\right]$-BNE for economy $F^{k_{j_\ell}}$.

By WLLN, there exists a sufficiently large $\hat{L}$ such that $\hat{L}$ i.i.d. Bernoulli random variables with probability $(1-\varepsilon)^{1/3}$ have a sample mean greater than $(1-\varepsilon)^{1/2}$ with probability no less than $(1-\varepsilon)^{1/3}$. Define next L_4 such that $\ell > L_4$ implies $(1-\varepsilon)^{1/2} k_{j_\ell} > \hat{L}$.

Now we fix an arbitrary $\ell > \max\{L_1, L_2, L_3, L_4\}$ and wish to show that in economy $F^{k_{j_\ell}}$

$$\Pr\left(\text{The fraction of students playing SRS against } p^{k_{j_\ell}} \text{ is no less than } 1-\varepsilon\right) \geq 1-\varepsilon,$$

which would contradict $(*)$ and complete the proof.

We first prove that in economy $F^{k_{j_\ell}}$, a student with $\tilde{\theta} \in \tilde{\Theta}^\delta$ plays SRS against $\bar{p}$ with probability no less than $(1-\varepsilon)^{1/3}$. To see this, suppose to the contrary that there exists some student i and some type $\tilde{\theta} \in \tilde{\Theta}^\delta$ such that

$$\Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays SRS against } \bar{p}\right) < (1-\varepsilon)^{1/3}.$$

Then deviating to truthful reporting will give student i with type $\tilde{\theta} \in \tilde{\Theta}^\delta$ at least a gain of

$$\mathbb{E}\left[i\text{'s gain from deviation} \,\middle|\, \sigma_i^{k_{j_l}} \text{ plays non-SRS against } p^{k_{j_\ell}}\right] \Pr\{\sigma_i^{k_{j_\ell}} \text{ plays non-SRS against } p^{k_{j_\ell}}\}$$

$$
\begin{aligned}
&\geq \mathbb{E}\left[i\text{'s gain from deviation } \middle| \{\sigma_i^{k_{j_l}} \text{ plays non-SRS against } p^{k_{j_\ell}}\} \wedge A^{k_{j_l}} \right] \\
&\quad \cdot \Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays non-SRS against } p^{k_{j_\ell}} \textit{ and} \text{ event } A^{k_{j_\ell}} \right) \\
&\geq \delta\alpha(\delta) \cdot \Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays non-SRS against } p^{k_{j_\ell}} \textit{ and} \text{ event } A^{k_{j_\ell}} \right) \\
&\geq \delta\alpha(\delta) \cdot \Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays non-SRS against } \bar{p} \textit{ and} \text{ event } A^{k_{j_\ell}} \right) \\
&\geq \delta\alpha(\delta) \left[\Pr\left(A^{k_{j_\ell}} \right) - \Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays SRS against } \bar{p} \right) \right] \\
&\geq \delta\alpha(\delta) \left[(1-\varepsilon)^{1/6} - (1-\varepsilon)^{1/3} \right],
\end{aligned}
$$

where the inequalities are explained as follows. The first inequality holds since the gains from the deviation is nonnegative whether the event $A^{k_{j_\ell}}$ occurs or not. The second inequality results from Lemma 4, upon noting that by Lemma 3 the cutoffs under the non-SRS and those that would prevail if i deviates to TRS are within $\nu(\delta)$-distance from $\bar{p}$. The third inequality holds since, given our choice δ, conditional on event $A^{k_{j_\ell}}$, a non-SRS against $\bar{p}$ is a non-SRS against $p^{k_{j_\ell}}$. The above inequalities contradict $\ell > L_3$, which implies that the strategy profile $\left(\sigma_i^{k_{j_\ell}}\right)_{i=1}^{k_{j_\ell}}$ is a $\delta\alpha(\delta)\left[(1-\varepsilon)^{1/6} - (1-\varepsilon)^{1/3}\right]$-BNE for the economy $F^{k_{j_\ell}}$.

Therefore, in economy $F^{k_{j_\ell}}$, for each student $i = 1, \ldots, k_{j_\ell}$ and each $\tilde{\theta} \in \tilde{\Theta}^\delta$, we have

$$
\begin{aligned}
&\Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays SRS against } \bar{p} \middle| \, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2} \right) \\
&= \Pr\left(\sigma_i^{k_{j_\ell}}(\tilde{\theta}) \text{ plays SRS against } \bar{p} \right) \geq (1-\varepsilon)^{1/3}, \qquad (***)
\end{aligned}
$$

where the first equality holds because student i's random report according to her mixed strategy is independent of random draws of the students' type.

Then we have

$$
\begin{aligned}
&\Pr\left(\begin{array}{c} \text{The fraction of students with } \tilde{\theta} \in \tilde{\Theta}^\delta \\ \text{playing SRS against } \bar{p} \text{ is no less than } (1-\varepsilon)^{1/2} \end{array} \middle| \, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2} \right) \\
&\geq \Pr\left(\begin{array}{c} \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \cdot k_{j_\ell} \text{ i.i.d. Bernoulli random variables with} \\ \text{probability } (1-\varepsilon)^{1/3} \text{ have a sample mean no less than } (1-\varepsilon)^{1/2} \end{array} \middle| \, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2} \right) \\
&\geq \Pr\left(\begin{array}{c} \hat{L} \text{ i.i.d. Bernoulli random variables with probability } (1-\varepsilon)^{1/3} \\ \text{have a sample mean no less than } (1-\varepsilon)^{1/2} \end{array} \right) \\
&\geq (1-\varepsilon)^{1/3}, \qquad (A.7)
\end{aligned}
$$

where the first inequality follows from (***) and that σ_i's are independent across students conditioning on the event $\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}$, and the second inequality holds since, for $\ell > L_4$, $\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}$ implies that $\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \cdot k_{j_l} > \hat{L}$.

Comparing the finite economy random cutoff $p^{k_{j_\ell}}$ with the deterministic cutoff $\bar{p}$, we have

$$
\begin{aligned}
&\Pr\left(\begin{array}{c}\text{The fraction of students with } \tilde{\theta} \in \tilde{\Theta}^\delta \\ \text{playing SRS against } p^{k_{j_\ell}} \text{ is no less than } (1-\varepsilon)^{1/2}\end{array}\,\middle|\, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \\
&\geq \Pr\left(\begin{array}{c}\text{The fraction of students with } \tilde{\theta} \in \tilde{\Theta}^\delta \\ \text{playing SRS against } \bar{p} \text{ is no less than } (1-\varepsilon)^{1/2}, \\ \text{and event } A^{k_{j_\ell}}\end{array}\,\middle|\, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \\
&\geq \Pr\left(\begin{array}{c}\text{The fraction of students with } \tilde{\theta} \in \tilde{\Theta}^\delta \\ \text{playing SRS against } \bar{p} \text{ is no less than } (1-\varepsilon)^{1/2}\end{array}\,\middle|\, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \\
&- \Pr\left(A^{k_{j_\ell}} \text{ does not occur}\middle|\, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \\
&\geq (1-\varepsilon)^{1/3} - \frac{1 - \Pr\left(A^{k_{j_\ell}}\right)}{\Pr\left(\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right)} \\
&\geq (1-\varepsilon)^{1/3} - \frac{(1-\varepsilon)^{1/2}\left[(1-\varepsilon)^{1/3} - (1-\varepsilon)^{1/2}\right]}{(1-\varepsilon)^{1/2}} \\
&= (1-\varepsilon)^{1/2}.
\end{aligned}
$$

The first inequality follows since in event $A^{k_{j_\ell}}$, the strategy $\sigma_i(\theta)$ is SRS against $P^{k_{j_\ell}}$ if and only if $\sigma_i(\theta)$ is SRS against $\bar{p}$ for type $\theta \in \tilde{\Theta}^\delta$. The third inequality follows from (A.7). The fourth inequality follows from (**).

The construction of L_1 implies $\Pr\left(\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \geq (1-\varepsilon)^{1/2}$, so finally we have in economy $F^{k_{j_\ell}}$,

$$
\begin{aligned}
&\Pr\left(\text{The fraction of students playing SRS against } p^{k_{j_\ell}} \text{ is no less than } 1-\varepsilon\right) \\
&\geq \Pr\left(\begin{array}{c}\text{At least } (1-\varepsilon)^{1/2} \text{ of students with } \tilde{\theta} \in \tilde{\Theta}^\delta \text{ play SRS against } p^{k_{j_\ell}} \\ \textit{and } \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\end{array}\right) \\
&= \Pr\left(\eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \cdot \Pr\left(\begin{array}{c}\text{At least } (1-\varepsilon)^{1/2} \text{ of students} \\ \text{with } \tilde{\theta} \in \tilde{\Theta}^\delta \text{ play SRS against } p^{k_{j_\ell}}\end{array}\,\middle|\, \eta^{k_{j_\ell}}(\tilde{\Theta}^\delta) \geq (1-\varepsilon)^{1/2}\right) \\
&\geq (1-\varepsilon)^{1/2} \cdot (1-\varepsilon)^{1/2} \\
&= 1-\varepsilon,
\end{aligned}
$$

where the last inequality follows from the above string of inequalities. Therefore, we have obtained a contradiction to $(*)$, and the statement of Theorem 1 follows. □

A.4 Proof of Proposition 1

Proof. First, as TEPS only uses stability and transitivity to infer preferences, it must be that $\mathcal{P} \subseteq \mathcal{P}^{\text{ST}}(\Omega')$. Second, we show that for any $(c, c') \in \mathcal{P}^{\text{ST}}(\Omega')$, $(c, c') \in \mathcal{P}$. As c' is inferred worse than c by stability and transitivity, there must be a sequence of schools, $c^1, \ldots, c^J$ for $1 < J \leq C$ with $c^1 = c$ and $c^J = c'$, such that c^{j-1} and c^j, for $1 < j \leq J$, are both feasible in some realized uncertainty in Ω' when c^{j-1} is the assigned school. Since TEPS uses all uncertainties in Ω' in Stage 1, it must be that $(c_{j-1}, c_j) \in \widetilde{\mathcal{P}}$ for $1 < j \leq J$ in Stage 2. By transitivity at Stage 3 of TEPS, $(c, c') = (c_1, c_J) \in \mathcal{P}$. We thus have $\mathcal{P} = \mathcal{P}^{\text{ST}}(\Omega')$. □

A.5 Proof of Proposition 2

Proof. First, for any $0 < \tau < \tau' < 100$, $\mathcal{P}^{top} \subseteq \mathcal{P}^{\tau} \subseteq \mathcal{P}^{\tau'} \subseteq \mathcal{P}^{all}$ directly follows from Stage 1 of each TEPS$^{\tau}$, and $\mathcal{P}^{all} = \mathcal{P}^{\text{ST}}(\Omega)$ directly follows from Proposition 1 . We only need to prove that $\mathcal{P}^{all} \subseteq \mathcal{P}^{\text{WTT}}$. Consider some $(c, c') \in \mathcal{P}^{all}$. We show that $(c, c') \in \mathcal{P}^{\text{WTT}}$.

First, since TEPS^{all} infers only an assigned school in some realized uncertainty is preferred to other schools, and an unranked school can never be a student's assigned school, c should be a ranked school in R. Next, notice that $\mathcal{P}^{all}$ can never include any (c, c') such that c' is ranked above c in R. To see this, assume to the contrary that $(c, c') \in \mathcal{P}^{all}$ and that c' is ranked above c in R. Then at Stage 1 of TEPS^{all}, there must be a sequence of schools, $c^1, \ldots, c^J$ for $1 < J \leq C$ with $c^1 = c$ and $c^J = c'$, such that c^{j-1} and c^j, for $1 < j \leq J$, are both feasible in some realized uncertainty when c^{j-1} is the assigned school. Then, for each $1 < j \leq J$, it must be that c^{j-1} is ranked above c^j, which implies that $c^1 = c$ is ranked above $c^J = c'$ in R, a contradiction.

Hence, c' is either (i) a school ranked lower than c on R, or (ii) an unranked school. Since WTT infers that c is more preferable than *every* school ranked below c and *every* unranked school, it should be that $(c, c') \in \mathcal{P}^{\text{WTT}}$. Therefore, $\mathcal{P}^{all} \subseteq \mathcal{P}^{\text{WTT}}$. □

Online Appendix to

Leveraging Uncertainties to Infer Preferences: Robust Analysis of School Choice

(Not for Publication)

Yeon-Koo Che Dong Woo Hahm YingHua He

August 2023

A Multiple Equilibria, Completeness, and Coherency

This appendix shows that our TEPS^{all} procedure does not suffer from incompleteness or incoherence in the sense of Tamer (2003). Incompleteness in our context would mean that the mapping from a student's type, (u,t), to TEPS inferred preferences is a correspondence, which may cause the point identification of the distribution of u to fail. Meanwhile, incoherency would imply that the model does not have a well-defined likelihood for TEPS inferred preferences given exogenous variables, implying certain logical inconsistency.

We start with some definitions. Recall that we consider a student with submitted ROL R and intrinsic priorities t. We also make it explicit that the preferences inferred by TEPS^{all} depend on R and t, i.e., $\mathcal{P}^{all}(R,t)$.

We say ROL R' is *consistent* with $\mathcal{P}^{all}(R,t)$ if R' satisfies two conditions: (i) every ever-assigned school $c=\alpha_W$ for some $W\subseteq\Omega$ is included in R', and (ii) for any c' ranked above c in R', c' is *not* inferred worse than than c, i.e., $(c,c')\notin\mathcal{P}^{all}(R,t)$. This implies that we may exclude a never-feasible school from R' or insert it in R' at any position; however, for an ever-feasible school that is included in R', its position in R' must respect the inferred preferences. Let $\mathcal{R}^*(R,t)$ be the set of all ROLs that are consistent with $\mathcal{P}^{all}(R,t)$. It can be verified that $\mathcal{R}^*(R,t)$ includes the student's true preference order given the assumption of stability and transitivity.

Further, let $\mathcal{U}^*(R,t)\subseteq[\underline{u},\overline{u}]^C$ be all the utility types that are consistent with $\mathcal{P}^{all}(R,t)$. That is, if $u\in\mathcal{U}^*(R,t)$, the associated true preference order $\rho(u)$ is in $\mathcal{R}^*(R,t)$.

Proposition A.1. *Assume that in every realized uncertainty, each student cannot change her own set of feasible schools and that the stable matching is unique and achieved. Suppose that a student of type (u,t) submits an ROL R. We have the following results:*

(i) ***Equivalent class of ROLs.*** *When her ROL R is replaced by R', the student receives her stable assignment in every realized uncertainty if and only if $R'\in\mathcal{R}^*(R,t)$.*

(ii) ***Completeness.*** *$\mathcal{P}^{all}(R,t) = \mathcal{P}^{all}(R',t)$, $\forall R' \in \mathcal{R}^*(R,t)$. That is, given stability, TEPS^{all} infers a unique set of preference relations for the student even if she submitted any ROL in $\mathcal{R}^*(R,t)$.*
(iii) ***Coherency.*** *TEPS^{all} infers $\mathcal{P}^{all}(R,t)$ if and only if $u \in \mathcal{U}^*(R,t)$.*

Proof. We prove the three statements one by one.
(i) We first prove sufficiency and then necessity.

Sufficiency. Suppose $R' \in \mathcal{R}^*(R,t)$. In any given realized uncertainty, by the definition of $\mathcal{R}^*(R,t)$, R' must rank the assigned school above any other simultaneously feasible schools, while the assigned school is included in R'. The student must be assigned the same school regardless of submitting R or R', because she cannot change her set of feasible schools in this realized uncertainty. Hence, she always obtains her stable assignment.

Necessity. Suppose that R and R' give the student the same stable assignment in every realized uncertainty. We argue that R' is in $\mathcal{R}^*(R,t)$. For any c that is the assigned school in a realized uncertainty, c must be ranked in R'. Therefore, to show $R' \in \mathcal{R}^*(R,t)$, we only need to prove that for any c' ranked above c in R', c' is *not* inferred worse than c, or $(c,c') \notin \mathcal{P}^{all}(R,t)$. Suppose on the contrary that c' ranked above c in R' and $(c,c') \in \mathcal{P}^{all}(R,t)$. For TEPS to infer that c' is worse than c given that R is submitted, there must exist $c^1,\dots,c^J$ for $1 < J \leq C$ with $c^1 = c$ and $c^J = c'$ such that, for each $1 < j \leq J$, there exists a realized uncertainty in which c^{j-1} is the assigned school while c^j is feasible. This implies that $c^1,\dots,c^{J-1}$ must be included in R'. Since $c^J = c'$ is ranked above $c^1 = c$ in R', R' must rank c^{j^*} above c^{j^*-1} for some $1 < j^* \leq J$. Because the student cannot influence her own feasible set, there must exist a realized uncertainty in which the student's assignment when submitting R' is c^{j^*}, in contrast to c^{j^*-1} which is the assignment when submitting R. This contradiction lets us conclude that R' must be consistent with $\mathcal{P}^{all}(R,t)$, or equivalently, $R' \in \mathcal{R}^*(R,t)$.

(ii) By part (i), the set of ever-assigned schools is the same under R or R'. Hence, for any c which is never assigned to the student, there does not exist $(c,c'), \forall c' \in C$ such that $(c,c') \in \mathcal{P}^{all}(R,t)$ or $(c,c') \in \mathcal{P}^{all}(R',t)$.

Consider c' such that $(c,c') \in \mathcal{P}^{all}(R,t)$ for some c that is assigned to the student in some realization of uncertainty. When R is submitted, there must exist $c^1,\dots,c^J$ for $1 < J \leq C$ with $c^1 = c$ and $c^J = c'$ such that, for each $1 < j \leq J$, there exists a realized uncertainty in which c^{j-1} is the assigned school while c^j is feasible. Together with the assumption that i cannot change her own feasible set, part (i) implies that, in each of those same realized uncertainties associated with R, c^{j-1}

remains the assigned school while c^j is still feasible for each $1 < j \leq J$ even when R' is submitted. Hence, $(c, c') \in \mathcal{P}^{all}(R', t)$. Similarly, we can show that for any c' such that $(c, c') \in \mathcal{P}^{all}(R', t)$ for some c that is assigned to the student in some realization of uncertainty, $(c, c') \in \mathcal{P}^{all}(R, t)$.

Taken together, $\mathcal{P}^{all}(R, t) = \mathcal{P}^{all}(R', t)$.

(iii) We first prove sufficiency and then necessity.

Sufficiency. For $u \in \mathcal{U}^*(R, t)$, to guarantee the unique stable matching in each realized uncertainty, part (i) implies that the student must submit an ROL in $\mathcal{R}^*(R, t)$. By part (ii), TEPS infers $\mathcal{P}^{all}(R, t)$ for the student.

Necessity. Suppose that TEPS infers $\mathcal{P}^{all}(R, t)$ for the student while $u \notin \mathcal{U}^*(R, t)$. By the definition of $\mathcal{U}^*(R, t)$, the student's true preference order, $\rho(u)$, is not in $\mathcal{R}^*(R, t)$. Part (i) implies that the student's assignment from submitting $\rho(u)$ is not the same as her assignment from submitting R, while R gives her the unique stable assignment. This implies that reporting truthfully leads to an unstable assignment for her, contradicting the property of the DA mechanism.

□

The assumptions in Proposition A.1 are justified by our Theorem 1 in large markets. When the market becomes large, a student's impact on cutoffs and thus her own feasible sets becomes negligible and the matching in any realized uncertainty is virtually stable and unique.

Part (i) implies that, although stability does not predict a unique ROL for every student, it predicts a unique class of outcome-equivalent ROLs, $\mathcal{R}^*(R, t)$. Further, part (ii) shows that TEPS^{all} maps a student's true preferences into a unique set of inferred preferences, regardless which ROL from $\mathcal{R}^*(R, t)$ the student submits. Finally, part (iii) indicates that our model is coherent and that the likelihood of a set of TEPS^{all} inferred preferences can be written as the likelihood of the student's cardinal preferences satisfying a set of conditions.

Figure A.1 illustrates the logic behind the proposition. Following part (i) of Proposition A.1, we divide the action space into equivalent classes of ROLs that lead to the same distribution of stable assignments (middle panel). As part (ii) implies, each equivalence class of ROLs uniquely maps into a set of preference relations inferred by TEPS (right panel). Finally, although the mapping between the utility space and the action space is a correspondence, the mapping between the utility space and the space of possible sets TEPS inferred preferences is a function as implied by part (iii).

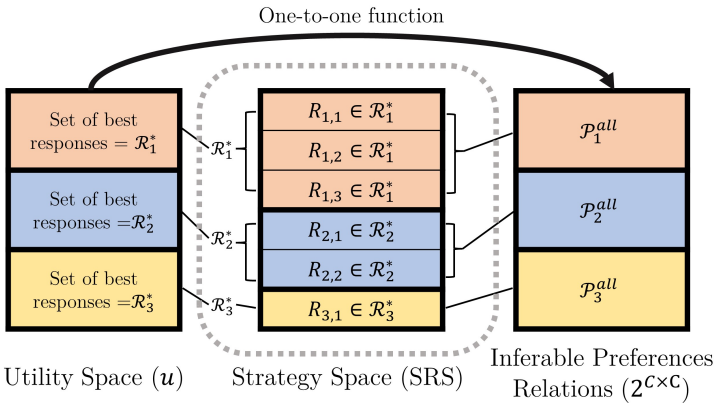


Figure A.1: Completeness and Coherency

B Performance of Transitive Extension of Preferences from Stability: Monte Carlo Simulation

This section describes the Monte Carlo simulations that we perform to analyze the implications of our theoretical results.

B.1 Model Specification

Consider a finite economy in which $k = 1000$ students apply to $C = 12$ schools for admission. The vector of school capacities is specified as follows:

$$\{S_c\}_{c=1}^{12} = \{110, 50, 100, 100, 50, 100, 100, 50, 100, 100, 50, 100\}$$

The total capacity is set to be larger than the total number of students by 10 in order to ensure all students are matched to some school in the simulations.

The economy is located in an area within a circle of radius 1. Students are uniformly distributed in the circle and schools are evenly located on another circle of radius 1/2 around the center. Denote the Euclidean distance between student i and school c as $d_{i,c}$.

Students are matched with schools through a student-proposing DA algorithm with single tie-breaking (DA-STB), similar to the NYC high school choice. Students submit an ROL of schools that can include any number of available schools. We assume all schools are acceptable to all students; hence, student i submits an ROL including all 12 schools if she truthfully reports her preferences.

School priorities over students are coarse. Each school has 4 categories of school-specific

(intrinsic) priority groups 0, 1, 2 and 3, with a larger number indicating a higher priority. Denote the priority group that student i belongs at school c as $t_{i,c} \in \{0, 1, 2, 3\}$. Therefore, school c prioritizes student i over i' if $t_{i,c} > t_{i',c}$. A student's priority at a school is drawn independently and uniformly from the four groups. All students are eligible/acceptable at each school. Each student knows her priority group at each school at the time of submitting ROL.

To break ties in priorities, every student is assigned a random lottery number drawn from *Uniform*$[0, 1]$, l_i for all i. Lottery numbers are not known at the time of submitting ROL. The score of student i at school c is $s_{i,c} = \frac{t_{i,c}+l_i}{4} \in [0, 1]$. School c prioritizes student i over i' if and only if $s_{i,c} > s_{i',c}$.

Student preferences over schools follow a random utility model without an outside option. Student i's utility from being matched with school c is specified as follows:

$$u_{i,c} = \beta_1 \times c + \beta_2(D_i \times A_c) + \beta_3 d_{i,c} + \beta_4 Small_c + \epsilon_{i,c}, \quad \forall i, c \tag{B.1}$$

where $\beta_1 \times c$ is school c's baseline quality; $d_{i,c}$ is the distance between student i's and school c; $D_i = 1$ or 0 is student i's type (e.g., disadvantaged or not); $A_c = 1$ or 0 is school c's type (e.g., known for resources for disadvantaged students); $Small_c = 1$ if $S_c = 50$, 0 otherwise; and $\epsilon_{i,c}$ is distributed as $N(0, \sigma_c^2)$ where $\sigma_c^2 = 1$ for $c = 1, \cdots, 6$ and $\sigma_c^2 = 2$ for $c = 7, \cdots, 12$. $\epsilon_{i,c}$ are independent across all i and c.

The type of school c, A_c, equals 1 if c is an odd number and otherwise 0. The type of student i, D_i, is 1 with probability 2/3 among the lowest priority group of school 1 ($t_{i,1} = 0$); and $D_i = 0$ for all students in highest three priority groups ($t_{i,1} \in \{1, 2, 3\}$).

The coefficients of interest are $(\beta_1, \beta_2, \beta_3, \beta_4)$ which are fixed at $(0.3, 2, -1, 0)$ in the simulations. By this specification, schools with larger indices are of a higher quality, and $Small_c$ does not play a role in student preferences over schools. The purpose of estimation is to recover these coefficients and therefore the distribution of preferences.

B.2 Data Generating Process

Each simulation sample contains an independent preference profile obtained by randomly drawing D_i and $\{t_{i,c}, d_{i,c}, \epsilon_{i,c}\}_c$ for all i from the distributions specified above. In all samples, school capacities and school types (A_c and $Small_c$) are kept constant.

The first set of simulation samples, *cutoff* samples, are used to simulate the joint distribution of the 12 schools' cutoffs (in terms of score, $s_{i,c}$) by letting every student submit an ROL ranking all schools according to her true preferences. To do so, we simulate 100 samples each consisting of 12

schools and 1,000 students. Each sample contains 1,000 sets of independent draws of tie-breaking lotteries l_i. After running the DA algorithm, we calculate the cutoffs in each simulation sample with each draw of lottery. Figure B.2 shows the marginal distribution of each school's cutoff from $100 \times 1000 = 100,000$ simulated realizations. Note that schools with smaller capacities tend to have higher cutoffs. For example, school 11 with 50 seats often has the highest cutoff, although school 12 with 100 seats has the highest baseline quality. Since every student is guaranteed a seat at some school, school 1 which has the lowest baseline quality has a cutoff equal to 0 (not depicted in the graph).

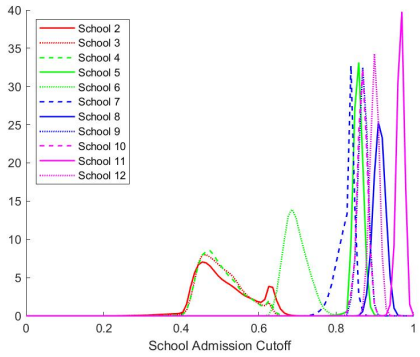


Figure B.2: Simulated Cutoff Distribution

To generate data on student behavior and admission outcomes for preference estimation, we simulate another 100 samples, the *estimation* samples, with new independent draws of T_i and $\{p_{i,c}, d_{i,c}, \epsilon_{i,c}\}_c$. For each of the 100 estimation samples, we calculate the distribution of cutoffs which represents the uncertainties in cutoffs that students face. By drawing a set of tie-breaking lotteries, we simulate the uncertainty due to random tie-breaking.[A-1] The resulting cutoff distributions are used in the Transitive Extension of Preferences from Stability procedure to determine the feasible sets for each realization of the uncertainty. The *estimation* samples are used for the estimation, and in each of them, we consider three types of data generating processes (DGPs) with different student behaviors.

(i) **TT (Truth-Telling)** : Every student submits an ROL ranking all 12 schools according to her true preferences.

(ii) **MIS-IRR ((Almost) Payoff Irrelevant Mistakes)**: A fraction of students skip schools with

[A-1] One may additionally simulate uncertainties in cutoffs due to the finiteness of the economy. To do so, we can randomly generate economies in the spirit of the bootstrap by randomly resampling a set of students, and then use a set of tie-breaking lotteries for each resample.

which they are never matched according to the simulated distribution of cutoffs. For a given student, a skipped school can have a high (expected) cutoff and thus be "out of reach" (i.e., never feasible.) Alternatively, the school may also have a low cutoff, but the student is always accepted by one of her more-preferred schools. To specify the fraction of skippers, we first randomly choose about 25.6% of the students to be never-skippers who always rank all schools truthfully. All other students are potential skippers, and we make all of them skip. Students with $T_i = 1$ are more likely to skip than those with $T_i = 0$, as their scores tend to be lower: 95.6 percent of $T_i = 1$ are potential skippers, compared to 70.1 percent of $T_i = 0$. Finally, we introduce flips by adding back the most preferred school (according to each skipper's true preference) on the end of the ROL if it is never-feasible and thus skipped. It is important to note that some of the mistakes may turn out to be payoff relevant, because a never-matched school is determined by the simulated cutoff distributions (which may not exhaust all uncertainties in cutoffs).

(iii) **MIS-REL (Payoff Relevant Mistakes)**: In addition to MIS-IRR, i.e., given all the potential skippers have skipped the never-matched schools, we now let them make payoff-relevant mistakes. That is, students skip some of the schools with which they have a small chance of being matched according to the simulated distribution of cutoffs. Recall that the joint distribution of cutoffs is only simulated once under the assumption that everyone is truth-telling. We specify a threshold and make the skippers omit the schools at which they have an admission probability lower than the threshold, where the threshold is equal to 10 percent. We allow for flips in the same fashion as in MIS1.

In summary, for each of the 100 estimation samples, we simulate the matching games 3 times: TT, MIS-IRR, and MIS-REL. Table B.1 summarizes the scenarios under each data generating process (DGP) we consider. Each entry reported is a percentage that is averaged over the 100 estimation samples. The fraction of skippers increases from 0% in TT to 74.4% in MIS-IRR and MIS-REL. As a result, weak truth-telling is exactly satisfied only in TT, and the fraction of students reporting preferences consistent with WTT is only 26.6% in MIS-IRR and 28.8% in MIS-REL.

B.3 Estimation and Results

The random utility model described by Equation (B.1) is estimated under two methods, WTT and TEPS using Gibbs Sampler, where the procedure is described in the Appendix C. For TEPS estimators, we use TEPS^{top}, TEPS^{20}, TEPS^{40}, TEPS^{60}, TEPS^{80}, and TEPS^{all}. Table B.2 presents

Table B.1: Skips and Mistakes in Monte Carlo Simulations (%)

	Scenarios: DGP w/ Different Student Strategies		
	Truth-Telling	Payoff Irrelevant Mistakes	Payoff Relevant Mistakes
	TT	MIS-IRR	MIS-REL
WTT: *Weak-Truth-Telling*	100	26.6	28.8
Matched w/ favorite feasible school	100	100	96.2
Skippers	0	74.4	74.4

the mean and standard deviation of the posterior mean of each parameter across the 100 samples.[A-2]

We evaluate the performance of the two sets of estimators along two dimensions. The first is the bias-variance tradeoff, focusing on β_2 which measures how students of type $T_i = 1$ value schools of type $A_c = 1$. The second dimension is the performance of the TEPS estimators relative to the TEPS estimators that use smaller set of inferred preferences, which are both robust to (some) strategic mistakes.

Table B.2: Estimation with Different Identifying Assumptions: Monte Carlo Results

DGPs	Identifying Condition	Quality ($\beta_1 = 0.3$)			Interaction ($\beta_2 = 2$)			Distance ($\beta_3 = -1$)			Small ($\beta_4 = 0$)		
		mean	s.d.	$\sqrt{MSE}$	mean	s.d.	$\sqrt{MSE}$	mean	s.d.	$\sqrt{MSE}$	mean	s.d.	$\sqrt{MSE}$
TT	WTT	0.30	0.00	0.00	2.00	0.06	0.06	-1.00	0.03	0.03	0.00	0.02	0.02
	TEPS^{top}	0.30	0.02	0.02	2.02	0.20	0.20	-1.03	0.11	0.12	0.00	0.07	0.07
	TEPS^{20}	0.30	0.02	0.02	2.02	0.20	0.20	-1.02	0.11	0.12	0.00	0.07	0.07
	TEPS^{40}	0.30	0.02	0.02	2.02	0.20	0.20	-1.03	0.12	0.12	0.00	0.07	0.07
	TEPS^{60}	0.30	0.02	0.02	2.03	0.20	0.20	-1.02	0.11	0.11	-0.01	0.07	0.07
	TEPS^{80}	0.30	0.01	0.01	2.02	0.17	0.17	-1.02	0.10	0.10	0.00	0.06	0.06
	TEPS^{all}	0.30	0.01	0.01	2.01	0.12	0.12	-1.01	0.07	0.07	0.00	0.04	0.04
	Selected	0.30	0.01	0.01	2.00	0.07	0.07	-1.00	0.04	0.04	0.00	0.03	0.03
MIS-IRR	WTT	0.08	0.00	0.22	1.21	0.08	0.79	-0.49	0.04	0.51	-0.19	0.02	0.19
	TEPS^{top}	0.30	0.02	0.02	2.02	0.20	0.20	-1.03	0.11	0.12	0.00	0.07	0.07
	TEPS^{20}	0.30	0.02	0.02	2.02	0.20	0.20	-1.03	0.11	0.12	0.00	0.07	0.07
	TEPS^{40}	0.30	0.02	0.02	2.02	0.20	0.20	-1.03	0.11	0.12	0.00	0.07	0.07
	TEPS^{60}	0.30	0.02	0.02	2.03	0.20	0.20	-1.02	0.11	0.11	-0.01	0.07	0.07
	TEPS^{80}	0.30	0.01	0.01	2.02	0.17	0.17	-1.02	0.09	0.10	0.00	0.06	0.06
	TEPS^{all}	0.30	0.01	0.01	2.01	0.13	0.13	-1.01	0.07	0.07	0.00	0.04	0.04
	Selected	0.30	0.01	0.01	2.01	0.13	0.13	-1.01	0.08	0.08	0.00	0.04	0.04
MIS-REL	WTT	0.13	0.00	0.17	0.99	0.07	1.01	-0.44	0.03	0.56	-0.11	0.02	0.11
	TEPS^{top}	0.29	0.02	0.02	1.88	0.29	0.31	-0.98	0.14	0.14	0.01	0.09	0.09
	TEPS^{20}	0.29	0.02	0.02	1.88	0.29	0.31	-0.97	0.14	0.14	0.01	0.09	0.09
	TEPS^{40}	0.29	0.02	0.02	1.88	0.29	0.31	-0.97	0.14	0.14	0.01	0.09	0.09
	TEPS^{60}	0.29	0.02	0.02	1.87	0.28	0.31	-0.96	0.13	0.14	0.01	0.09	0.09
	TEPS^{80}	0.28	0.02	0.03	1.82	0.25	0.30	-0.92	0.12	0.14	0.00	0.10	0.10
	TEPS^{all}	0.24	0.01	0.06	1.64	0.21	0.42	-0.77	0.08	0.24	-0.07	0.10	0.12
	Selected	0.28	0.03	0.04	1.82	0.27	0.32	-0.91	0.13	0.16	-0.01	0.09	0.09

Note: The results are from the 100 Monte Carlo samples.

[A-2]We iterate through the MCMC 100,000 times and discard the first 75,000 for mixing. We calculated the Potential Scale Reduction Factor (PSRF) (Gelman and Rubin (1992)) to ensure enough convergence of the posterior distributions.

Bias-Variance Tradeoff Figure B.3 plots the distributions of the estimates of β_2 given each DGP; the true value of β_2 is 2. The figures plot WTT, TEPS^{top}, and TEPS^{all}.

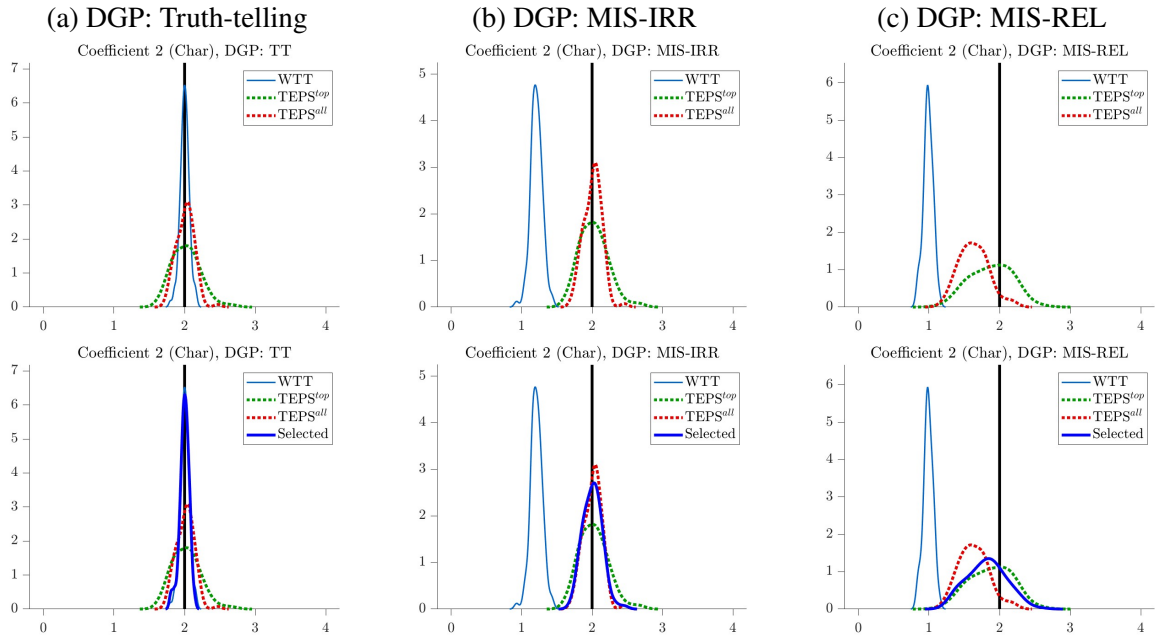


Figure B.3: Distribution of Estimates based on WTT and TEPS ($\beta_2 = 2$)

Note: We plot the kernel density plots of the estimates of β_2 from 100 Monte Carlo samples. The first column corresponds to when the data generating process is TT, the second column corresponds to when the data generating process is MIS-IRR, and the third column corresponds to when the data generating process is MIS-REL. The black vertical line at 2 denotes the true value of the parameter.

There are a few notable patterns. First, when the DGP is TT, all estimates are consistent, while the WTT-based estimator having the smallest variance as shown in Panel (a). This is expected, since no students make strategic mistakes violating the WTT or stability assumptions. Furthermore, the WTT-based estimator having the smallest variance is not only true under DGP TT but also under other DGPs. Intuitively, this is due to the fact that WTT uses the maximal (but possibly unreliable) information that one can infer from observed ROLs.

Second, Panel (b) shows the results from MIS-IRR in which some students make (almost) payoff-irrelevant mistakes. The WTT-based estimator is susceptible to strategic mistakes. For example, the WTT-based estimates have a mean 1.21 (standard deviation 0.08). That is, the WTT-based estimator is no longer consistent and the bias is sizable. On the other hand, the estimators based on TEPS are robust to payoff-irrelevant mistakes.

Finally, Panel (c) show the results from MIS-REL in which some students make payoff-relevant mistakes. Recall that when students make payoff relevant mistakes, first, WTT is not satisfied and

thus WTT-based estimator is inconsistent, and second, stability is also not 100% satisfied and thus TEPS^{all} which does not take care of payoff relevant mistakes is also inconsistent. However, the bias of TEPS^{all} remains a lot smaller than that of the WTT-based estimator, since the violation of WTT (74.4%) is much more severe than the violation of stability (3.8%) (see Table B.1).

The bias-variance tradeoffs are summarized by the square-root of the mean squared errors in Table B.2. In DGP TT, where all estimators are consistent, the WTT-based estimator attains the minimum mean squared error, as it has the minimum variance where biases of all estimators are close to zero. However, in DGP MIS-IRR with payoff-irrelevant mistakes, the mean squared error of the WTT-based estimator is larger than that of any other estimator. This is due to the fact that even though the WTT-based estimator has the smallest variance, it is significantly inconsistent. The mean squared errors of the WTT-based estimator are even larger with DGP MIS-REL with payoff-relevant mistakes.

Performance of the TEPS estimator. We now compare the relative performance among TEPS-based estimators. As discussed earlier, to the extent that stability is satisfied, all TEPS-based estimators are consistent. However, the procedures by which the TEPS estimators are constructed imply that TEPS^{top} uses less (but potentially more reliable) information contained in the observed ROLs compared to other TEPS-based estimators. Therefore, TEPS$^{\tau}$, $\tau = 20, 40, 60, 80, 100$, should have higher precision compared to TEPS^{top}. Similarly, TEPS$^{\tau}$ is expected to have higher precision compared to TEPS$^{\tau'}$ for all $\tau' < \tau$.

Figure B.3 shows the results, in which we only present TEPS^{top} and TEPS^{all} for clear comparison. TEPS^{all} is more precise than TEPS^{top}, having a higher level of concentration around the true value $\beta_2 = 2$. As reported in Table B.2, TEPS$^{\tau}$, $\tau = 20, 40, 60, 80, 100$ always have (weakly) smaller standard deviations compared to TEPS^{top} and TEPS$^{\tau'}$ for all $\tau' > \tau$. For example, in MIS-IRR in which approximately 74.4% students make payoff-irrelevant mistakes, the standard deviation of TEPS^{top} is 0.20, while that of TEPS^{all} is 0.13.

Next, in DGP MIS-REL where students skip some schools with a positive admission probability and hence make payoff-relevant mistakes, the choice of TEPS threshold (i.e., how much payoff irrelevant mistakes to be tolerated) becomes important. The bias of TEPS^{all} is larger than TEPS^{top}, TEPS^{20}, TEPS^{40}, TEPS^{60}, and TEPS^{80}. For example, for $\beta_2 = 2$, the bias of TEPS^{all} estimator is 0.36 where those of TEPS^{top} is 0.12 on average (Table B.2). As expected, the bias decreases as we tolerate less payoff-relevant mistakes (i.e., as we decrease τ in TEPS$^{\tau}$).

Table B.3: Determining the Selected Estimator: Test Results (at the 5% significance level)

Data Generating Process:		TT	MIS-IRR	MIS-REL
Estimation method	WTT	**0.94**	0	0
	TEPS^{top}	0.01	0	0
	TEPS^{20}	0	0	0
	TEPS^{40}	0	0	0.01
	TEPS^{60}	0	0.01	0.14
	TEPS^{80}	0.03	0.12	**0.68**
	TEPS^{all}	0.01	**0.87**	0.17

Note: The results are from the 100 estimation samples.

Choosing among the estimation methods. Recall the main motivation which led us to introduce the TEPS-based estimator is to "correct" the strategic mistakes that students might make in a school choice environment in order to obtain robust estimates. Without the information on how students make mistakes, we follow the procedure described in Section 3.5 to determine the selected estimator that corrects (almost) all the mistakes (thus consistent) and uses the maximum information from the data (thus has the highest precision among all consistent estimators we consider.) Table B.3 reports the test results based on the testing procedure with size equal to 0.05, and the resulting 'chosen' estimates are reported in the second row of Figure B.3 and 'Selected' rows of Table B.2. As expected, WTT is chosen 94% out of 100 estimation samples when the DGP is TT since all estimators are consistent while WTT uses the maximum information (Proposition 2). When the DGP is MIS-IRR, TEPS^{all} is chosen 87%. Note that TEPS^{top}, TEPS^{20}, TEPS^{40} are never chosen since all TEPS-based estimators are consistent, and TEPS$^{\tau}$'s with larger τ use more information relative to that with smaller τ's. Finally, when the DGP is MIS-REL, TEPS^{80} is chosen 68%. Note that the testing procedure selects TEPS$^{\tau}$ with $\tau < 100$ most of the time since students make payoff-relevant mistakes in MIS-REL and TEPS^{all} cannot handle payoff-relevant mistakes and thus is inconsistent.

C Markov Chain Monte Carlo Procedure for Preference Estimation

C.1 Setup

There are k students competing for admissions to C schools/programs. Each school c has a type $\tau(c) \in \{1, \cdots, \bar{T}\}$ where $\bar{T} \leq C$. Denote the number of school with type τ by C_τ. WLOG, let schools be ordered in the increasing order of type. Student i's utility when being admitted to school c is given by,

$$U_{i,c} = X_{i,c}\beta + \epsilon_{i,c}, \tag{C.2}$$

where $\epsilon_{i,c}$ is i.i.d. $N(0, \sigma^2_{\tau(c)})$.[A-3] Then $\Sigma \equiv Var(\epsilon_i)$ is a diagonal matrix with $(\sigma^2_1, \cdots, \sigma^2_1, \cdots, \sigma^2_{\bar{T}}, \cdots, \sigma^2_{\bar{T}})$ on the diagonal. We use Σ wherever possible for notational simplicity.

Let $\mathcal{P}_i^{method}$ be the set of all preference relations inferred by some method (for example, WTT, TEPS^{top}, or TEPS$^{\tau}$ for some $\tau \in (0, 100]$) for student i. The procedure we describe below applies to any method for inferring preference relation from choice data. When there is no ambiguity, we simply use $\mathcal{P}_i$. Denote the schools that are inferred to be less preferred than c as $\mathcal{L}_{i,c}$ and those are inferred to be preferred to c as $\mathcal{M}_{i,c}$:

$$\mathcal{L}_{i,c} = \{c' : (c, c') \in \mathcal{P}_i \text{ i.e., } c' \text{ is revealed to be less preferred than } c\}$$
$$\mathcal{M}_{i,c} = \{c' : (c', c) \in \mathcal{P}_i \text{ i.e., } c' \text{ is revealed preferred to } c\}$$

C.2 A Gibbs sampling procedure

We specify the following diffuse priors:

$$\beta \sim N(0, A^{-1})$$
$$\sigma^2_\tau \sim IW(\nu_\tau, V_{0\tau}), \;\; \tau = 1, 2, \cdots, \bar{T} - 1$$

where

$$A^{-1} = 100 \cdot I_{dim(\beta)},$$
$$\nu_\tau = 3 + C_\tau, \;\; V_{0\tau} = 3 + C_\tau, \;\; \forall \tau$$

We go through the following iterative process, a Gibbs sampler.

Initialization.

1. Draw Σ^0 from its prior distribution.
2. Draw β^0 from its prior distribution.
3. Draw U^0: Following the index of students, $i = 1, \ldots, k$, we draw $\{U^0_{i,c}\}_c$ sequentially as follows: for a given i,

   (a) start with $c = 1$. Draw $U^0_{i,1}$ from $N(X_{i,1}\beta^0, \Sigma^0_{1,1})$;

   (b) for each $c = 2, \ldots, C$, draw $U^0_{i,c}$ from $N(X_{i,c}\beta^0, \Sigma^0_{c,c})$ with truncations imposed by $\mathcal{P}_i$. To be

[A-3] Note that σ^2_τ for some τ has to be normalized to 1 for identification. WLOG, we set $\sigma^2_{\bar{T}} = 1$ for the following.

specific, let

$$\widetilde{\mathcal{L}}_{i,c} = \{c' : c' < c \text{ and } (c, c') \in \mathcal{P}_i \text{ i.e., } c' \text{ is revealed to be less preferred than } c\} \subseteq \mathcal{L}_{i,c}$$
$$\widetilde{\mathcal{M}}_{i,c} = \{c' : c' < c \text{ and } (c', c) \in \mathcal{P}_i \text{ i.e., } c' \text{ is revealed preferred to } c\} \subseteq \mathcal{M}_{i,c}$$

That is, $\widetilde{\mathcal{L}}_{i,c}$ ($\widetilde{\mathcal{M}}_{i,c}$) is the set of schools whose utility is already drawn,[A-4] and at the same time less (more) preferred than school c in $\mathcal{P}_i$ respectively. Then, we draw $U^0_{i,c}$ with truncations from below at $\underline{u}^0_c$, and from above at $\bar{u}^0_c$ where

$$\underline{u}^0_c = \begin{cases} \max\{U_{i,c'} : c' \in \widetilde{\mathcal{L}}_{i,c}\} & \text{if } \widetilde{\mathcal{L}}_{i,c} \neq \emptyset \\ -\infty & \text{if } \widetilde{\mathcal{L}}_{i,c} = \emptyset \end{cases}, \quad \bar{u}^0_c = \begin{cases} \min\{U_{i,c'} : c' \in \widetilde{\mathcal{M}}_{i,c}\} & \text{if } \widetilde{\mathcal{M}}_{i,c} \neq \emptyset \\ \infty & \text{if } \widetilde{\mathcal{M}}_{i,c} = \emptyset \end{cases}$$

Note that if both $\underline{u}^0_c$ and $\bar{u}^0_c$ are finite, we draw from two-sided truncated distribution, and if only either one of them is finite, we draw from one-sided truncated distribution, and if none of them is finite, we draw from the untruncated distribution.

Iteration $r \geq 1$.

1. Following the index of students, $i = 1, \ldots, k$, we draw $\{U^r_{i,c}\}_c$ sequentially as follows.
   For a given i, for $c = 1, \ldots, C$, draw $U^r_{i,c}$ from $N(X_{i,c}\beta^{r-1}, \Sigma^{r-1}_{c,c})$ with truncations imposed by $\mathcal{P}_i$. To be specific, we draw $U^r_{i,c}$ with truncations from below at $\underline{u}^r_c$, and from above at $\bar{u}^r_c$ where

$$\underline{u}^r_c = \begin{cases} \max\left\{\{U^r_{i,c'} : c' \in \widetilde{\mathcal{L}}_{i,c}\} \cup \{U^{r-1}_{i,c'} : c' \in \mathcal{L}_{i,c} \setminus \widetilde{\mathcal{L}}_{i,c}\}\right\} & \text{if } \mathcal{L}_{i,c} \neq \emptyset \\ -\infty & \text{if } \mathcal{L}_{i,c} = \emptyset \end{cases}$$
$$\bar{u}^r_c = \begin{cases} \min\left\{\{U^r_{i,c'} : c' \in \widetilde{\mathcal{M}}_{i,c}\} \cup \{U^{r-1}_{i,c'} : c' \in \mathcal{M}_{i,c} \setminus \widetilde{\mathcal{M}}_{i,c}\}\right\} & \text{if } \mathcal{M}_{i,c} \neq \emptyset \\ \infty & \text{if } \mathcal{M}_{i,c} = \emptyset \end{cases}$$

   where $\mathcal{L}_{i,c}, \widetilde{\mathcal{L}}_{i,c}, \mathcal{M}_{i,c}, \widetilde{\mathcal{M}}_{i,c}$ are defined above. In words, for example, if c is revealed preferred to some school by $\mathcal{P}_i$ (i.e., $\mathcal{L}_{i,c} \neq \emptyset$), the lower bound $\underline{u}^r_c$ is given by the maximum among the r-th draws of utilities of schools that are less preferred than c and are already drawn in the current step (i.e., $\{U^r_{i,c'} : c' \in \widetilde{\mathcal{L}}_{i,c}\}$) and the $(r-1)$-th draws of utilities of schools that are less preferred than c but not drawn in the current step (i.e., $\{U^{r-1}_{i,c'} : c' \in \mathcal{L}_{i,c} \setminus \widetilde{\mathcal{L}}_{i,c}\}$.) $\bar{u}^r_c$ is defined analogously but with schools that are more preferred than c by $\mathcal{P}_i$.

[A-4] Note that we draw $c = 1, \cdots, C$ sequentially so that $U^0_{i,c'}, c' < c$ is already drawn in this step when drawing $U^0_{i,c}$.

2. Draw β^r from the posterior distribution $N(\tilde{\beta}, V)$[A-5] where

$$\begin{aligned} V &= (X^{*\top} X^* + A)^{-1} \\ \tilde{\beta} &= V X^{*\top} U^* \\ (\Sigma^{r-1})^{-1} &= \Lambda^\top \Lambda \\ X_i^* &= \Lambda^\top X_i \\ U_i^* &= \Lambda^\top U_i^r \\ X &= \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}, \quad U^r = \begin{bmatrix} U_1^r \\ \vdots \\ U_k^r \end{bmatrix} \end{aligned}$$

and X_i is C by $dim(\beta)$ matrix and $U_i \in \mathbb{R}^C$.

This is a standard result for Bayesian regression with normal errors.

3. For each $\tau = 1, \cdots, \bar{T} - 1$, draw $(\sigma_\tau^2)^r$ from the posterior distributions $IW(\nu_\tau + kC(\tau), V_{0\tau} + S_\tau)$ where

$$\begin{aligned} S_\tau &= \sum_{c:\tau(c)=\tau} \sum_{i=1}^{k} \varepsilon_{i,c}^r (\varepsilon_{i,c}^r)^\top \\ \varepsilon_{i,c}^r &= U_{i,c}^r - X_{i,c}\beta^r \end{aligned}$$

where $\varepsilon_{i,c}^r$, $U_{i,c}^r$ and $X_{i,c}$ denote the part of ε_i^r, U_i^r and X_i corresponding to school c.

4. Save and pass (β^r, Σ^r, U^r) to the next iteration.

D Data on NYC High School Choice

D.1 Institutional Background

NYC public high school system consists of two sectors: specialized high schools and regular high schools. There are 9 specialized high schools in NYC.[A-6] We do not consider 9 specialized

[A-5] Note that $\tilde{\beta}$ is specific to the r-th iteration where we omit the dependence on r for notational simplicity.

[A-6] They are Stuyvesant High School; Brooklyn Technical High School; Bronx High School of Science; High School of American Studies at Lehman College; The Brooklyn Latin School; High School for Mathematics, Science and Engineering at City College; Queens High School for the Sciences at York College; Staten Island Technical High School; and Fiorello H. LaGuardia High School of Music & Art and Performing Arts. These schools, except for Fiorello H. LaGuardia High School, use Specialized High School Admission Test (SHSAT) as the sole criterion of admission, which is a required exam for students wanting to attend any of the specialized high school. Fiorello H. LaGuardia High

high schools in our analysis, because they use different admission methods from the regular high schools and students can submit a separate ROL of specialized high schools.

Regular high schools are traditional public schools and have six types of programs differentiated in their admission method: *Unscreened*, *Limited unscreened*, *Screened*, *Audition*, *Educational option*, and *Zoned* programs. Multiple programs of different types may be offered by a single school. *Unscreened* programs admit students by a random lottery number attached to each student. *Limited unscreened* programs operate similarly as *Unscreened* programs but give higher priority to students who attended an information session or open houses. *Screened* programs as well as *Audition* programs rank students by individual assortment of criteria. For example, *Screened* programs use several criteria such as final report-card grade, statewide standardized test scores, and attendance and punctuality. *Audition* programs hold school/program specific auditions to admit students. *Educational option* is a mixture of unscreened and screened programs. They have a purpose of serving students at diverse academic performance levels and divide students into high (the highest 16%), middle (the middle 68%) and low (the lowest 16%) levels in terms of English Language Arts (ELA) score. 50% of the seats in each group are filled using school-specific criteria similarly as a *Screened* program and the other 50% are filled randomly similarly as an *Unscreened* program. *Zoned* programs give priority or guarantee admission to students who apply and live in the zoned area of the school.

Each program has its own eligibility and priority group criteria. For example, in academic year 2016–17, Young Women's Leadership School in Astoria opened its seats only to female students, i.e., being a female student was the eligibility criterion. Besides, they gave the highest admission priority to continuing eighth graders, then to students or residents in Queens, and then to other NYC residents.

The number of priority groups given to students are a lot smaller than the number of applicants to each program. Hence students who apply to programs that do not actively rank students —*Unscreened*, *Limited Unscreened*, *Zoned* and the unscreened part of *Educational option*—are often in the same priority group. Hence the ties need to be broken for the SPDA algorithm to run. For this purpose, a random lottery number is drawn and attached to each student which is used to break ties at all programs that require tie-breaking in the same fashion—single tie-breaking (STB) rule. The lottery number is unknown to the student at the moment of application.

The timeline of the admission process is as follows (Corcoran and Levin (2011)). In October

School uses audition as its admission criterion.

and November, students may apply to specialized high schools for which they should take SHSAT or audition at LaGuardia High School. In December, they are required to submit up to 12 ranked non-specialized high school programs regardless of their application status to specialized high schools. In March, the SPDA algorithms for specialized high schools (SHSAT takers) and non-specialized high schools (all students) are separately run, which is called Round 1. The Department of Education sends each student a letter with their offer from regular schools and offer from specialized high schools, if any. If a student receives offers from both regular and specialized high schools, he/she must choose one. All students who accept a Round 1 offer has a finalized admission decision. If a student did not submit an application in Round 1, did not receive an offer in Round 1, or wants to apply to a program with availability, he/she can participant in Round 2. Round 2 takes place in March and operates with students who submitted Round 2 applications and programs that still have seat availability. Round-2 offers automatically replace Round 1 offers, if any. Students who are unassigned in Rounds 1 and 2 or reject the assignment go to the administrative round in which students are administratively assigned a school on a case-by-case basis.

D.2 Data Source

The main data that we use is the administrative data from the New York City Department of Education (NYC DOE) for academic year 2016–17. There are four sets of data used to construct information on the applicants. First, high school application (HSAP) data contains the submitted ROLs and student information such as ELA and math standardized test scores, English-language Learner (ELL) status, and student priorities at programs (including priority rank, priority criteria, and eligibility). Second, June biographic data provides comprehensive student biographic information, including ethnicity, gender, disability status, as well as information on attendance and punctuality. Third, standardized test data contains more detailed information on statewide standardized exams. Fourth, zoned DBN data provides information on the zoned school of each student, census tract and school district of each student's residence. Finally, Middle School Course and Grade data contains all of the courses and information on credits and grades for each student in a given year. There exists a unique scrambled student ID variable that enables merging all NYC DOE dataset while personally identifying a student is impossible. Lastly, we use information on each census tract and zip code in NYC obtained from 2016 American Community Survey 5-Year Estimates from US Census database.

School information is constructed using the NYC High School Directory that is published every year before the application process starts. This includes each program's capacity in the previous year, the number of students who applied in the previous year, eligibility and priority

criteria, accountability data such as progress report, graduation rate and college enrollment rate, and types of language classes provided. Other variables about the current 9th graders, such as ethnicity composition and fraction of high performing students, are constructed using the high school application data from the previous year, academic year 2015–2016.

D.3 Sample Restrictions

We focus on students from Staten Island. There are two different samples—one for tracing out the uncertainties, the other for estimation of student preferences: *Priority Construction* sample and *Estimation* sample. They differ because of missing values in some variables. *Priority Construction* sample is used to reconstruct priority scores of each student at each school/program. *Estimation* sample is a strict subset of *Priority Construction* sample, and used for preference estimation and counterfactual simulations.

First, *Priority Construction* sample consists of students who applied to at least one Staten Island school/program. Among such 4,824 students, 785 did not have information on variables needed to reconstruct priority, and are dropped leaving us 4,039 students. Next, *Estimation* sample consists of students who applied to at least one Staten Island school/program, went to a middle school in Staten Island, and resided in Staten Island at the point of application. Among such 4,480 students, 741 did not have information on variables needed for priority reconstruction or estimation, and are dropped. Finally, 8 students had invalid ROLs such as containing an invalid program code on their ROLs and are dropped, leaving us 3,731 students.

Finally, we adjust the capacities of each school/program whenever we restrict our sample. Specifically, we treat Round 1 assignment in the data (or Round 2 assignment if applicable) of each dropped student as fixed whenever we resimulate DA using the restricted sample. This is in order to ensure that we do not overestimate the probabilities of school/programs being feasible to each student in our TEPS procedure.

E TEPS for NYC High School Choice

In this section, we describe the TEPS (Transitive Extension of Preferences from Stability) procedure for NYC high school choice.

E.1 Constructing Priority Scores

Before describing how to simulate uncertainty, let us first specify the procedure we take to prepare ingredients for our the procedure. The main inputs for the SPDA algorithm are the capacities of programs, students' preferences and programs' preferences (priority scores). First we use adjusted

programs' capacities as described in Appendix D. Next, we use students' submitted ROL in the data as students' preferences.

School preferences are a bit more involved then the other two inputs. As described in Appendix D, NYC public high schools have coarse priority rules. First for eligibility criteria and admission priority groups which are publicly available before the admission process starts, we use the information listed on the High School Directory.

Next, we estimate the priority ranks for *Screened*, *Audition* and the screened part of *Educational option* programs which actively rank students. While there is information on priority rank of students in the data set provided by NYC DOE, it is limited only to students who ranked that program. Furthermore, how each program rank its applicants is not a public information, and each program has its individual assortment of criteria. For our purposes to calculate the probability of each program being feasible to a student regardless of she ranked it or not, we need to construct priority rank for all students at each program that actively ranks students. To do so, we assume that there exists a program specific latent variable v_{ij} for student i at actively ranking program j which determines the priority ranks:

$$v_{ij} = \beta_j X_i + \varepsilon_{ij} \quad \text{and} \quad i \succ_j i' \text{ if and only if } v_{ij} > v_{i'j},$$

where X_i is a vector of student characteristics including 7th Standardized Math and ELA scores, middle school Math, Social Studies, English, Science GPA, days absent and days late, and ε_{ij} is independent and identically distributed as extreme value type I conditional on X_i. We form a log-likelihood by considering all possible pairs of applicants to each program j, and estimate via MLE separately for each program. That is for $\mathcal{I}_j$, the set of applicants to program j,

$$\hat{\beta}_j = \arg\max_{\beta_j} l(\beta_j) \equiv \sum_{i>i' : i,i' \in \mathcal{I}_j} \log\Big(\frac{\exp(v_{ij})1\{i \succ_j i'\} + \exp(v_{i'j})1\{i' \succ_j i\}}{\exp(v_{ij}) + \exp(v_{i'j})}\Big)$$

With estimates $\hat{\beta}_j$, we predict $\hat{v}_{ij} = \hat{\beta}_j X_i$ for all students (not limited to applicants) and reconstruct priority ranks based on $\hat{v}_{ij}$.

E.2 Simulation of Uncertainties

As described in Section 3, the first step in Transitive Extension of Preferences from Stability is to trace out all the uncertainties present in the economy. Now we describe how we simulate uncertainties present in the matching environment. After reconstructing priority scores for each student at each program, we simulate $B_L = 10,000$ lotteries from a uniform distribution to break

ties at non-actively ranking programs and run the SPDA 10,000 times. This procedure would give us a empirical distribution of cutoffs of all programs $P = (P_1, \cdots, P_C)$ where the cutoff of a program is defined to be the lowest priority score of admitted students if the seats are filled, and zero if the seats are not filled. Note that this step simulates the uncertainty rising from small-sample uncertainties i.e., the uncertainty rising from other students' type which may be viewed random from the point of view of a single student.

Next, we draw L number of lotteries ($L = 5,000$) from a uniform distribution in order to account for the fact that a student's own ex-post score is uncertain due to the random tie-breaking rules. For each lottery draw, we use the cutoff distribution simulated above to figure out the set of feasible schools and its probability for each student in each realization of uncertainties. Together with the submitted ROLs, we can also compute the assigned program in each realization of uncertainties.

F Analysis of the NYC High School Choice Data

F.1 Preference Estimates

Tables F.4–F.7 present preference estimates for each covariate cell. We report the mean and standard deviation of the posterior distribution as the point estimate and the standard error. We iterate through the MCMC 1 million times and discard the first 0.9 million to ensure mixing. We calculate the Potential Scale Reduction Factor (PSRF) using the draws that we keep following (Gelman and Rubin, 1992). For those that did not converge, we additionally iterate 1 million times and keep the last 0.1 million. The resulting PSRFs for all parameters for all cells are below 1.1 which ensures convergence.

Table F.4: Preference Estimates: Cell 1–2

	Cell 1								Cell 2							
	WTT		TEPS^{all}		TEPS^{top}		Selected		WTT		TEPS^{all}		TEPS^{top}		Selected	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
School FE 4	-0.26	0.05	-0.45	0.08	-0.86	0.11	-0.55	0.08	0.11	0.05	-0.12	0.08	-0.38	0.13	-0.12	0.08
School FE 5	-0.65	0.08	-1.05	0.12	-1.45	0.17	-1.18	0.13	-0.45	0.09	-1.3	0.15	-1.55	0.21	-1.3	0.15
School FE 6	-0.24	0.08	-0.48	0.13	-0.64	0.18	-0.53	0.14	0.18	0.09	-0.44	0.15	-0.61	0.22	-0.44	0.15
School FE 7	0	0.09	-0.7	0.15	-0.87	0.21	-0.73	0.16	-0.4	0.1	-1.43	0.18	-1.58	0.25	-1.43	0.18
School FE 8	-0.29	0.06	-0.39	0.09	-0.47	0.12	-0.43	0.09	-0.2	0.06	-0.48	0.09	-0.58	0.13	-0.48	0.09
School FE 9	-0.03	0.08	-0.27	0.13	-0.59	0.18	-0.39	0.14	-0.69	0.11	-1.41	0.21	-1.51	0.3	-1.41	0.21
$\mathbb{1}$(STEM)																
Main Effect	0.22	0.06	0.33	0.11	0.44	0.16	0.3	0.13	-0.09	0.07	0.07	0.14	-0.26	0.25	0.07	0.14
×ELA Score	-0.22	0.05	-0.19	0.08	-0.14	0.1	-0.14	0.08	-0.06	0.06	-0.13	0.1	-0.2	0.15	-0.13	0.1
×Math Score	0.27	0.06	0.24	0.09	0.29	0.12	0.21	0.1	0.16	0.06	0.08	0.1	0.4	0.16	0.08	0.1
% High Perf. (ELA)																
Main Effect	-0.13	0.15	-0.86	0.23	-1.07	0.34	-0.73	0.26	0.16	0.16	-0.37	0.25	-0.48	0.35	-0.37	0.25
×ELA Score	0.32	0.07	0.4	0.15	0.49	0.2	0.38	0.16	0.42	0.08	0.35	0.18	0.56	0.25	0.35	0.18
% High Perf. (Math)																
Main Effect	0.46	0.2	1.49	0.33	1.58	0.51	1.36	0.36	0.7	0.22	2.31	0.4	2.96	0.6	2.31	0.4
×Math Score	0.91	0.08	0.71	0.17	0.84	0.25	0.77	0.19	1	0.09	1.12	0.21	0.7	0.27	1.12	0.21
% FRPL (program)																
Main Effect	1.07	0.18	-0.59	0.35	-0.23	0.52	-0.64	0.39	0.08	0.2	-2	0.37	-1.89	0.52	-2	0.37
×Med Income	-0.06	0.15	0.17	0.28	0.17	0.4	0.32	0.31	0.06	0.15	-0.03	0.3	-0.23	0.42	-0.03	0.3
×Avg Score	-0.62	0.14	-0.22	0.28	-0.03	0.44	-0.18	0.31	0.04	0.14	0.19	0.32	0.04	0.46	0.19	0.32
% FRPL (school)																
×Med Income	0.08	0.22	0.35	0.35	0.44	0.48	0.25	0.39	0.58	0.24	0.53	0.4	1.05	0.56	0.53	0.4
×Avg Score	0.39	0.18	-0.42	0.32	-0.83	0.45	-0.38	0.34	-0.43	0.2	-0.92	0.39	-1.03	0.51	-0.92	0.39
9th Grade Size (100s)	0.16	0.01	0.19	0.02	0.23	0.02	0.19	0.02	0.17	0.01	0.25	0.02	0.29	0.03	0.25	0.02
% Asian	0.68	0.31	0.75	0.51	0.64	0.71	0.75	0.55	-2.27	0.35	-3.84	0.63	-3.5	0.89	-3.84	0.63
% Black	-1.25	0.21	-0.62	0.33	-0.49	0.46	-0.6	0.36	-0.81	0.23	1.17	0.41	1.13	0.62	1.17	0.41
% Hispanic	-0.75	0.21	-0.66	0.33	-0.71	0.45	-0.51	0.36	-1.05	0.22	0.25	0.38	0.68	0.53	0.25	0.38
1(Nearest School)	0.28	0.04	0.26	0.07	0.26	0.09	0.27	0.07	0.2	0.05	0.26	0.07	0.24	0.11	0.26	0.07
Distance	-0.26	0.01	-0.3	0.02	-0.3	0.02	-0.3	0.02	-0.34	0.01	-0.35	0.02	-0.42	0.04	-0.35	0.02
σ^2_{STEM}	0.86	0.07	0.75	0.11	0.73	0.15	0.85	0.15	0.99	0.08	1.06	0.17	1.53	0.34	1.06	0.17

Notes: We report the full estimates of the parameters in (1). All % variables are within $[0, 1]$. For conciseness, we report only the estimates based on WTT, TEPS^{all}, TEPS^{top}, and the selected estimates obtained by following the procedure described in Section 3.5. We report the mean and standard deviation of the posterior distribution as the point estimate and the standard error.

Table F.5: Preference Estimates: Cell 3–4

	Cell 3								Cell 4							
	WTT		TEPS^{all}		TEPS^{top}		Selected		WTT		TEPS^{all}		TEPS^{top}		Selected	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
School FE 4	-0.35	0.06	-0.44	0.08	-0.55	0.13	-0.49	0.1	-0.01	0.06	-0.12	0.08	-0.42	0.14	-0.39	0.13
School FE 5	-0.53	0.09	-1.06	0.13	-1.18	0.19	-1.21	0.16	-0.42	0.09	-1.17	0.14	-1.36	0.21	-1.38	0.2
School FE 6	-0.5	0.09	-0.61	0.14	-0.72	0.2	-0.8	0.17	-0.08	0.09	-0.72	0.15	-0.86	0.22	-0.95	0.22
School FE 7	0.18	0.11	-0.47	0.17	-0.43	0.24	-0.51	0.21	-0.27	0.11	-1.03	0.2	-1.3	0.29	-1.19	0.27
School FE 8	-0.34	0.07	-0.42	0.1	-0.38	0.14	-0.44	0.12	-0.15	0.06	-0.54	0.1	-0.58	0.15	-0.62	0.15
School FE 9	-0.42	0.09	-0.57	0.14	-0.53	0.2	-0.64	0.17	-0.76	0.11	-1.63	0.21	-1.7	0.31	-1.79	0.3
𝟙(STEM)																
Main Effect	0.19	0.09	0.27	0.14	-0.02	0.28	-0.07	0.24	-0.26	0.1	-0.29	0.2	-0.54	0.35	-0.41	0.32
×ELA Score	-0.04	0.07	0.01	0.09	0.09	0.15	0.13	0.13	-0.02	0.08	-0.01	0.14	-0.22	0.19	-0.1	0.18
×Math Score	0.13	0.07	0.01	0.1	0.07	0.16	-0.05	0.13	0.1	0.08	-0.04	0.13	0.38	0.2	0.12	0.18
% High Perf. (ELA)																
Main Effect	0.21	0.17	-0.26	0.26	-0.2	0.36	0	0.3	0.13	0.17	-0.2	0.26	-0.47	0.38	-0.26	0.35
×ELA Score	0.43	0.09	0.69	0.17	0.76	0.24	0.72	0.21	0.31	0.09	0.14	0.19	0.41	0.27	0.3	0.25
% High Perf. (Math)																
Main Effect	-0.34	0.24	0.65	0.37	0.65	0.56	0.69	0.44	0.8	0.23	2.27	0.41	3.12	0.63	2.95	0.58
×Math Score	0.79	0.09	0.73	0.18	0.97	0.27	0.92	0.23	0.59	0.09	0.9	0.21	0.23	0.29	0.43	0.28
% FRPL (program)																
Main Effect	1.17	0.21	-0.39	0.35	-1.13	0.48	-0.66	0.43	0.29	0.2	-0.94	0.38	-1.01	0.55	-0.52	0.52
×Med Income	0	0.16	0.14	0.27	0.17	0.39	0.1	0.32	0.03	0.14	-0.04	0.29	0.04	0.4	0.03	0.4
×Avg Score	-0.25	0.15	0.08	0.27	0.08	0.37	0.19	0.31	-0.15	0.15	0.05	0.33	-0.23	0.49	-0.5	0.48
% FRPL (school)																
×Med Income	-0.4	0.27	-0.38	0.37	-0.32	0.52	-0.29	0.43	0.06	0.24	0.17	0.36	0.27	0.53	-0.03	0.57
×Avg Score	0.29	0.21	-0.3	0.31	-0.06	0.43	-0.37	0.38	0.12	0.21	0.01	0.4	-0.27	0.59	0.24	0.56
9th Grade Size (100s)	0.14	0.01	0.19	0.02	0.25	0.03	0.23	0.03	0.15	0.01	0.23	0.02	0.29	0.04	0.28	0.03
% Asian	2.3	0.38	1.7	0.6	2.41	0.87	1.88	0.72	-1.31	0.38	-3.27	0.67	-2.99	0.95	-3.36	0.9
% Black	-1.29	0.26	-0.59	0.38	0.08	0.59	0.03	0.47	-0.48	0.26	1.18	0.43	1.85	0.64	1.66	0.62
% Hispanic	-0.7	0.25	-0.08	0.38	0.73	0.54	0.46	0.45	-0.51	0.25	0.45	0.4	0.71	0.56	0.91	0.55
1(Nearest School)	-0.01	0.06	0.06	0.08	0.01	0.11	0.03	0.09	0.26	0.06	0.22	0.08	0.22	0.12	0.3	0.11
Distance	-0.31	0.02	-0.33	0.02	-0.35	0.03	-0.35	0.03	-0.26	0.01	-0.29	0.02	-0.3	0.03	-0.28	0.03
σ^2_{STEM}	1.17	0.12	0.97	0.17	1.28	0.36	1.42	0.33	1.21	0.14	1.53	0.3	1.66	0.47	1.7	0.47

Notes: We report the full estimates of the parameters in (1). All % variables are within $[0, 1]$. For conciseness, we report only the estimates based on WTT, TEPS^{all}, TEPS^{top}, and the selected estimates obtained by following the procedure described in Section 3.5. We report the mean and standard deviation of the posterior distribution as the point estimate and the standard error.

Table F.6: Preference Estimates: Cell 5–6

	Cell 5								Cell 6							
	WTT		TEPS^{all}		TEPS^{top}		Selected		WTT		TEPS^{all}		TEPS^{top}		Selected	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
School FE 4	-0.29	0.1	-0.5	0.12	-0.5	0.24	-0.64	0.13	-0.13	0.11	-0.23	0.15	-0.41	0.25	-0.32	0.16
School FE 5	-0.8	0.12	-1.47	0.19	-1.72	0.33	-1.61	0.21	-0.58	0.14	-1.19	0.24	-0.98	0.34	-1.2	0.25
School FE 6	-0.45	0.12	-0.83	0.19	-0.54	0.33	-0.85	0.2	-0.07	0.14	-0.56	0.23	-0.22	0.37	-0.54	0.24
School FE 7	0.22	0.18	-0.62	0.29	-1.37	0.55	-0.74	0.31	0.14	0.2	-0.55	0.37	-1.16	0.53	-0.62	0.39
School FE 8	-0.3	0.1	-0.44	0.15	-0.33	0.26	-0.48	0.16	-0.07	0.11	-0.48	0.19	-0.7	0.29	-0.56	0.2
School FE 9	-0.3	0.12	-0.64	0.17	-0.78	0.33	-0.8	0.19	-0.85	0.16	-1.22	0.27	-0.88	0.39	-1.14	0.27
𝟙(STEM)																
Main Effect	0.28	0.11	0.46	0.18	0.02	0.43	0.41	0.22	-0.15	0.16	-0.48	0.34	-0.85	0.54	-0.45	0.37
×ELA Score	-0.02	0.07	0.07	0.09	0.17	0.2	0.08	0.1	-0.17	0.12	-0.43	0.21	-0.4	0.33	-0.36	0.22
×Math Score	0	0.07	-0.08	0.1	0.11	0.23	-0.09	0.11	0.14	0.11	0.38	0.23	0.64	0.43	0.3	0.24
% High Perf. (ELA)																
Main Effect	0.4	0.25	-0.12	0.37	-0.64	0.6	-0.14	0.4	0.74	0.3	0.24	0.48	-0.09	0.66	0.38	0.51
×ELA Score	0.52	0.12	0.61	0.23	1.19	0.42	0.7	0.25	0.37	0.15	0.33	0.31	0.38	0.41	0.29	0.31
% High Perf. (Math)																
Main Effect	-0.75	0.34	0.85	0.5	1.56	0.88	0.87	0.55	-0.36	0.39	1.48	0.69	2.91	0.96	1.38	0.74
×Math Score	0.51	0.13	0.61	0.25	0.44	0.47	0.57	0.27	0.35	0.14	0.69	0.37	0.45	0.51	0.81	0.39
% FRPL (program)																
Main Effect	0.78	0.3	-1.02	0.51	-2.79	0.81	-1.12	0.56	0.6	0.35	0.19	0.68	-0.18	0.95	0.09	0.73
×Med Income	-0.14	0.18	-0.18	0.33	-0.82	0.56	-0.16	0.36	-0.09	0.24	-0.36	0.48	-0.69	0.69	-0.56	0.5
×Avg Score	-0.05	0.21	0.46	0.36	1.02	0.54	0.42	0.39	-0.51	0.26	-0.86	0.56	-1.46	0.88	-0.59	0.6
% FRPL (school)																
×Med Income	0.44	0.26	0.12	0.44	1.23	0.75	0.18	0.46	-0.54	0.37	-0.21	0.59	1.24	0.93	0.09	0.64
×Avg Score	-0.07	0.31	-0.66	0.46	-1.14	0.75	-0.44	0.48	0.84	0.36	1.42	0.65	1.98	0.96	1.16	0.66
9th Grade Size (100s)	0.1	0.02	0.16	0.04	0.33	0.07	0.17	0.04	0.09	0.03	0.18	0.04	0.33	0.07	0.19	0.05
% Asian	2.34	0.51	1.28	0.75	0.45	1.34	1.1	0.82	-0.41	0.63	-2.23	1.08	-3.27	1.6	-1.81	1.16
% Black	-0.96	0.34	0.11	0.47	0.67	0.84	-0.06	0.52	-0.46	0.41	0.12	0.64	0.35	0.96	0.16	0.67
% Hispanic	-0.24	0.32	0.68	0.48	1.84	0.82	0.84	0.54	-0.98	0.39	-0.32	0.65	-0.16	0.88	-0.22	0.68
1(Nearest School)	-0.12	0.08	-0.22	0.11	-0.07	0.18	-0.15	0.12	0.13	0.09	-0.01	0.15	-0.02	0.21	-0.01	0.16
Distance	-0.3	0.02	-0.29	0.03	-0.35	0.05	-0.31	0.03	-0.23	0.02	-0.25	0.04	-0.3	0.06	-0.25	0.04
σ^2_{STEM}	1.18	0.18	0.86	0.21	2.17	0.88	1.03	0.29	1.1	0.2	1.31	0.42	1.39	0.58	1.3	0.45

Notes: We report the full estimates of the parameters in (1). All % variables are within $[0, 1]$. For conciseness, we report only the estimates based on WTT, TEPS^{all}, TEPS^{top}, and the selected estimates obtained by following the procedure described in Section 3.5. We report the mean and standard deviation of the posterior distribution as the point estimate and the standard error.

Table F.7: Preference Estimates: Cell 7–8

	Cell 7								Cell 8							
	WTT		TEPS^{all}		TEPS^{top}		Selected		WTT		TEPS^{all}		TEPS^{top}		Selected	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
School FE 4	-0.11	0.06	-0.25	0.08	-0.28	0.13	-0.31	0.09	0.32	0.06	0.13	0.08	0.16	0.14	0.06	0.12
School FE 5	-0.5	0.07	-0.98	0.1	-1.03	0.17	-1.03	0.12	-0.22	0.07	-0.85	0.11	-0.74	0.17	-0.79	0.16
School FE 6	-0.17	0.08	-0.44	0.12	-0.42	0.19	-0.48	0.14	0.09	0.08	-0.51	0.12	-0.43	0.19	-0.48	0.18
School FE 7	0.66	0.1	0.24	0.16	-0.33	0.28	0.08	0.18	0.06	0.11	-0.6	0.19	-1.1	0.37	-1.04	0.35
School FE 8	0.19	0.06	0.07	0.09	0.03	0.15	0.08	0.1	0.17	0.06	-0.17	0.1	0.04	0.17	-0.04	0.15
School FE 9	-0.04	0.07	-0.24	0.1	-0.25	0.17	-0.33	0.12	-0.48	0.08	-0.89	0.12	-0.73	0.2	-0.79	0.18
𝟙(STEM)																
Main Effect	0.18	0.07	-0.19	0.18	-0.99	0.55	-0.15	0.2	-0.39	0.1	-0.92	0.29	-0.93	0.55	-0.7	0.41
×ELA Score	-0.04	0.04	-0.08	0.07	-0.09	0.12	-0.03	0.07	-0.03	0.06	0.02	0.1	-0.1	0.16	-0.09	0.14
×Math Score	-0.01	0.04	-0.1	0.07	-0.11	0.14	-0.16	0.08	-0.07	0.06	-0.27	0.11	-0.1	0.19	-0.2	0.17
% High Perf. (ELA)																
Main Effect	0.31	0.15	0.23	0.23	0.39	0.34	0.38	0.26	0.33	0.14	0.17	0.22	-0.28	0.31	-0.08	0.29
×ELA Score	0.44	0.07	0.57	0.15	0.66	0.2	0.54	0.15	0.4	0.08	0.5	0.17	0.27	0.23	0.34	0.21
% High Perf. (Math)																
Main Effect	-0.57	0.19	0.18	0.3	0.61	0.47	0.24	0.33	0.62	0.19	1.34	0.34	1.68	0.48	1.43	0.45
×Math Score	0.3	0.07	0.41	0.15	0.49	0.22	0.54	0.16	0.66	0.08	1.05	0.2	1.08	0.28	1.25	0.28
% FRPL (program)																
Main Effect	1.29	0.18	0.5	0.34	0.18	0.47	0.53	0.36	0.76	0.18	0.07	0.34	-0.48	0.49	-0.3	0.49
×Med Income	-0.14	0.1	-0.3	0.18	-0.67	0.3	-0.46	0.19	0.07	0.1	0.15	0.2	0.03	0.3	-0.01	0.29
×Avg Score	0.01	0.12	0.12	0.21	0.17	0.31	0.24	0.23	-0.12	0.13	-0.26	0.27	-0.5	0.38	-0.26	0.35
% FRPL (school)																
×Med Income	0.27	0.15	0.26	0.23	0.89	0.42	0.37	0.27	-0.5	0.16	-0.83	0.25	-0.71	0.42	-0.67	0.41
×Avg Score	0.22	0.18	0.13	0.25	0.45	0.39	0.25	0.28	0.71	0.18	1.02	0.32	0.94	0.49	0.89	0.44
9th Grade Size (100s)	0.01	0.01	0.05	0.02	0.16	0.03	0.08	0.02	0.04	0.02	0.11	0.03	0.24	0.04	0.22	0.04
% Asian	1.24	0.28	0.78	0.45	0.2	0.73	0.59	0.5	-1.78	0.31	-2.92	0.52	-1.76	0.76	-1.97	0.73
% Black	-1.01	0.2	-0.78	0.28	-0.42	0.46	-0.65	0.31	0.19	0.2	0.71	0.31	1.74	0.48	1.49	0.44
% Hispanic	-0.58	0.17	-0.17	0.26	0.27	0.41	0.14	0.3	-0.28	0.18	0.23	0.28	0.5	0.39	0.37	0.37
1(Nearest School)	0.13	0.05	0.12	0.06	0.15	0.1	0.12	0.07	0.14	0.05	0.15	0.07	0.29	0.1	0.28	0.1
Distance	-0.19	0.01	-0.22	0.02	-0.25	0.03	-0.23	0.02	-0.14	0.01	-0.15	0.02	-0.14	0.03	-0.12	0.02
σ^2_{STEM}	1.14	0.11	1.43	0.25	2.75	1.03	1.43	0.29	1.13	0.13	1.6	0.36	1.68	0.69	1.42	0.49

Notes: We report the full estimates of the parameters in (1). All % variables are within $[0, 1]$. For conciseness, we report only the estimates based on WTT, TEPS^{all}, TEPS^{top}, and the selected estimates obtained by following the procedure described in Section 3.5. We report the mean and standard deviation of the posterior distribution as the point estimate and the standard error.

F.2 Additional Figures

F.2.1 Program Characteristics by Feasibility Status

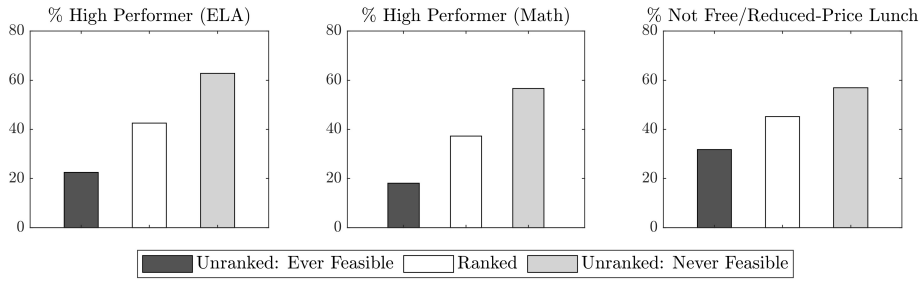


Figure F.4: Characteristics of Ranked and Unranked Programs by Feasibility Status

Notes: For each student, we classify the programs into three types—ever-feasible-unranked, never-feasible-unranked, and ranked. Ranked programs are those included in the student's ROL, and Stages 1 and 2 of the TEPS procedure determine feasibility of each unranked program. We use the fraction of high-performing students (measured by ELA and math scores), and the fraction of those who are not eligible for free/reduced-price lunch in each program. The figure reports the average across all students for each type of programs.

F.2.2 Commuting Distances

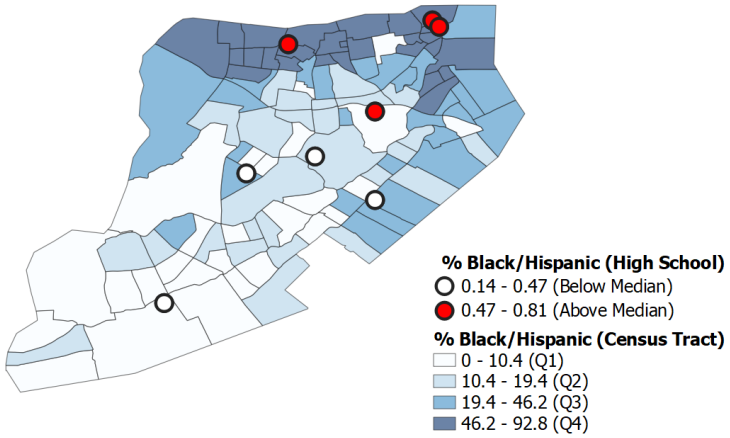


Figure F.5: Percentage of Black or Hispanic: Census Tracts and High Schools

Source: 2017 American Community Survey: 5-Year Data, US Census TIGER/Line Shapefiles.

These seemingly drastic counterfactual policies seem to have small effects on desegregation in Figure 7, which may be explained by the fact that SI has a high degree of residential segregation (Figure F.5). Many Black/Hispanic students live far away from schools with few Black/Hispanic students and low fraction of FRPL students or schools with more high performers. As students have

a distaste for commuting (see Appendix Tables F.4-F.7), the racial gap remains even when every student has an equal opportunity to attend any school.

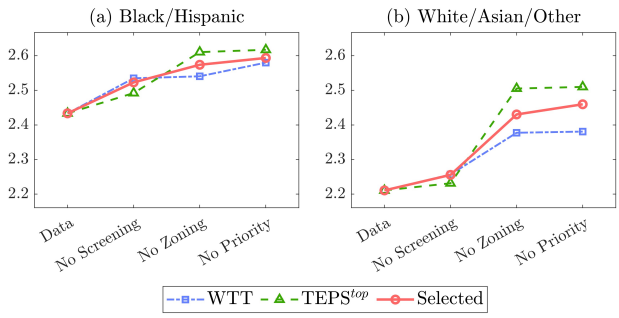


Figure F.6: Distance to Assigned Programs

Nevertheless, there is still hope: Students do commute more under counterfactual regimes, leading to some desegregation effects. Figure F.6 reports the average predicted distance to the assigned program under each counterfactual policy by the estimation method. Recall that all three counterfactual policies alleviate constraints (either academic or geographical) that students face in school choice. As those constraints are lifted, students more actively exercise choice options by applying and being assigned to schools outside of their neighborhoods, which is manifested in the increased commuting distance compared to the current regime. This pattern is stronger for non-Black/Hispanic students as shown by the larger increase in commuting distance. Again, WTT predicts a smaller increase in the commuting distance, which contributes to smaller predicted effects of counterfactual policies in Figure 7.