

Performance bounds for stable matchings*

Juan F. Escobar[†]

Martin Valdevenito[‡]

February 5, 2024

Abstract

A common method to allocate scarce resources is to apply the deferred acceptance algorithm proposed by Gale and Shapley (1962) and obtain a stable matching. However, a stable matching need not be Pareto efficient and may assign many students to their worst or second-worst schools. We provide tight upper and lower bounds for the fraction of students assigned to their top schools and the fraction of students that can be Pareto-improved in a stable matching of a large market. Our results can be used to characterize the inefficiencies of different priority rules used in school choice applications, including distance-based and random priorities.

1 Introduction

The allocation of resources—such as school seats, college admissions, and residency programs—often occurs without monetary transfers. In practice, a common method for making these allocations is through the deferred acceptance (DA) algorithm proposed by Gale and Shapley (1962). This method yields a stable matching and provides a clear way to explain potentially uneven assignments to participants. As the literature has noted, there are fundamental tensions between efficiency and stability in matching

*Both authors acknowledge financial support from the Institute for Research in Market Imperfections and Public Policy, MIPP, ICS130002. Escobar acknowledges support from Anillo Information and Computation in Market Design ANID ACT 210005. All data used in this work is publicly available from the Chilean Ministry of Education.

[†]Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile. E-mail: juanesco-bar@uchile.cl.

[‡]Department of Economics, University of Zurich. E-mail: martin.valdevenito@econ.uzh.ch.

models. In particular, a stable matching need not be Pareto-efficient. Moreover, as Kesten (2010) shows in the school choice context, for any fixed supply of school seats, it is possible to construct the demand so that the stable matching assigns each student to her worst or second-worst school. The inefficiency of a stable matching is not just a theoretical possibility but also a key consideration in practical applications of matching theory (Abdulkadiroğlu et al., 2009).

This paper provides new performance estimates for stable matchings. We explore a large market model in which a continuum of students applies to a finite number of schools (Azevedo and Leshno, 2016; Abdulkadiroğlu et al., 2015). Students have preferences over schools, while schools have priorities over students. To capture the interplay between students' preferences and schools' priorities, each student has a type. These types govern students' preferences and also determine their scores within each school, subsequently influencing the schools' priorities over students. Introducing types into our matching model is a flexible way to allow for correlation between students' preferences and scores. Our model encompasses a variety of priority criteria used in applied school choice, including multiple and single tie breaking (Abdulkadiroğlu et al., 2009).¹ It also accommodates models in which students are geographically differentiated, and a student's preference is partly determined by the distances between the student and schools.

Our first main result, Theorem 1, provides tight upper and lower bounds for the fraction of students assigned to their top schools. Behind these bounds is the idea that the performance of a stable matching depends on how students can congest and get admission to schools they do not consider top choices. Notably, our bounds apply to a general school choice model that allows for asymmetries and arbitrary correlations between preferences and scores. The bounds provide a convenient solution for tackling examples and models that would otherwise pose significant challenges in analysis.

To establish our bounds, we write down market clearing conditions that characterize stable matchings (Azevedo and Leshno, 2016; Abdulkadiroğlu et al., 2015). The solutions to these equations are hard to solve in closed-form. We thus explore relaxed market clearing conditions that are used to obtain lower and upper bounds for the solutions to the original market clearing conditions. These bounds are then used to estimate the fraction of students assigned to their top schools.

We apply our methods to the problem of priority design in school choice. In districts employing

¹In some school choice applications, priorities are randomly determined. Under single tie breaking, each student obtains a unique random score that determines her priorities in all schools. Under multiple tie breaking, each student obtains a different random number for each school.

the deferred acceptance algorithm, oftentimes authorities design the priority criteria used by schools to increase students’ satisfaction (Abdulkadiroğlu et al., 2009). Consequently, an important literature has explored the role of different random tie breaking rules on the effectiveness of the deferred acceptance algorithm (Abdulkadiroğlu et al., 2009; Ashlagi and Nikzad, 2020; Arnosti, 2022; Allman et al., 2022). While several cities in Europe and the US employ proximity to schools to determine priorities (Dur et al., 2018; Çelebi and Flynn, 2021; Burgess et al., 2023), not much is known about how this policy choice impacts the performance of the deferred accepted algorithm.

We evaluate distance-based priorities in a general spatial model of school choice. Stable matchings under distance-based priorities are determined by how students value proximity, and also by the capacity and geographical distribution of schools. Under distance-based scores, when students significantly value proximity, students’ preferences and schools’ priorities are compatible: a student that likes a school also has a high score in the school. Naturally, in this case the resulting stable matching will be Pareto efficient and place many students into their top schools.

In contrast, when students’ preferences for proximity are not strong, distance-based priorities may result in important efficiency losses. Indeed, Theorem 2 shows that multiple tie breaking may result in more students assigned to their top schools and fewer students that can be Pareto improved than distance-based priorities. This happens even when students value proximity and, as a result, there is positive correlation between preferences and priorities.

Theorem 2 may appear counter-intuitive. After all, several papers have shown that, compared to single tie breaking, multiple tie breaking results in a relatively low number of students assigned to their top schools and a high number of Pareto improving pairs (Abdulkadiroğlu et al., 2009; Ashlagi and Nikzad, 2020; Arnosti, 2022; Allman et al., 2022). Moreover, distance-based priorities create consistency between preferences and priorities which –as we show in the text– is a force towards efficient stable matchings. We observe that in markets in which students care not only about proximity to schools but also about other aspects –such as scores in standardized tests, extracurricular activities, etc– the consistency between preferences and priorities is positive but weak. In these markets, under distance-based priorities a student may be stuck at a school that is not her top but just happens to live nearby. This force leaves relatively few students assigned to top schools.

Our theoretical results show that introducing proximity as a priority criterion has ambiguous effects on some important performance measures. We confirm our findings by simulating a market with a finite

number of students and seats.

The school choice literature has shown that even the student optimal stable matching need not be Pareto efficient (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003). Several papers derive conditions under which a stable matching is Pareto efficient.² Notably, Ergin (2002) introduces a class of school priorities such that, regardless of students’ preferences, the stable matching is efficient.³ See also Ehlers and Erdil (2010), Salonen and Salonen (2018), Reny (2021), Pakzad-Hurson (2023). In contrast, our main goal is to provide performance estimates for stable matchings. In many practical applications of the deferred accepted algorithm, efficiency will not be achieved and therefore understanding the magnitude of inefficiencies may be a useful step in the design of matching markets.

The idea that school priorities can be designed to impact the performance of the deferred acceptance algorithm is not new. Abdulkadiroğlu et al. (2009), Ashlagi and Nikzad (2020), Arnosti (2022), Shi (2022), Allman et al. (2022) notice that when schools solve indifferences by using random lotteries, the correlation between the scores of a student in different schools is important for efficiency. As the literature shows –and we confirm in Subsection 4.2– single tie breaking (under which the correlation between scores is perfect) results in a more efficient matching than multiple tie breaking (under which the correlation between scores is 0). We make two contributions to this literature. First, we notice that priority criteria such that a high score in a school implies low scores in other schools make efficiency hard to achieve; see the discussion of distance-based priorities following Theorem 2 part b. In this sense, priority criteria that result in no correlation between scores (such as multiple tie breaking) may produce a more efficient matching than criteria resulting in a negative correlation (such as distance-based priorities). Second, we observe that the correlation between students’ preferences and scores is also important to evaluate the efficiency of a stable matching.

Our paper also connects to research about distance-based priorities in school choice. Dur et al. (2018) explore how different precedence orders implementing walk-zone reserves impact the fraction of reserve-group students assigned to each school. More closely related, Çelebi and Flynn (2021) show that in a large market model, the optimal coarsening of scores is attained by splitting agents into at most three indifference classes. They also explore a model in which scores are determined by distance and show that the optimal number of zones depends on the diversity goals of the planner. Our

²There is also an important set of papers proposing alternative algorithms and solutions, including Shapley and Scarf (1974), Kesten (2010), Che and Tercieux (2018), Ehlers and Morrill (2020), Cantillon et al. (2022), and Reny (2022).

³Ergin (2002) introduces acyclical priorities. In school choice applications, priorities derived under single tie breaking are acyclical.

focus is different in that we explore alternative performance measures and our insights highlight how the correlation between preferences and priorities determine the effectiveness of the deferred acceptance algorithm. We thus see our analysis as complementary to Çelebi and Flynn’s (2021).

Finally, our work connects to the literature employing large market models to analyze market design problems (Azevedo and Leshno, 2016; Abdulkadiroğlu et al., 2015; Ashlagi and Nikzad, 2020). We provide a method to bound cutoffs in large market models to derive new insights for the design of matching markets.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents our estimates for the fraction of students assigned to their top schools. Section 4 applies our bounds to random priorities and distance-based priorities. Section 5 shows simulations for discrete economies. Section 6 presents concluding remarks. All proofs are in the Appendix.

2 Model

2.1 Environment

There is a finite set of schools $C = \{1, \dots, N\}$. There is a continuum S of students to be matched to schools. Each student s has a strict preference ordering \succ^s over $C \cup \{\emptyset\}$, where \emptyset is the outcome if s is unassigned. A student s has a score vector $e^s = (e_c^s)_{c=1}^N$. School c has capacity k_c . A school c prefers student s to student s' iff $e_c^s > e_c^{s'}$. We simplify exposition and assume that all schools and all students are acceptable.

Students have types $i \in I$. We endow $I \subseteq \mathbb{R}^L$ with a measure ν so that $\int \nu(di) = 1$ and assume that ν is absolutely continuous. Preferences and scores are determined by types. Concretely, for each i there is a distribution F_i over the finite set of preferences over schools, with $\sum_{\succ} F_i(\succ) = 1$ and $F_i(\succ) \geq 0$, so $F_i(\succ)$ is the fraction of type i students having preference \succ . Additionally, a type i student has a score $e_c^s = e_c(i) \in [0, 1]$. We assume that the probability of a tie in a school is 0 so that for all c and all $x \in [0, 1]$, $\nu(\{i \in I \mid e_c(i) = x\}) = 0$. Conditional on a type i , the realizations of students’ preferences and scores are independent. We thus assume that any correlation between the preferences of a student s and her score in schools is determined by the type i of the student $s \in S$. Since a student’s type determines scores, a student s can be characterized by her type i and preferences \succ . We denote by $\bar{\nu}$

the measure induced by ν and $(F_i)_{i \in I}$ over the set of students.⁴

Let $F_i^{(k)}(c)$ be the fraction of type i students that put school c in the k -th position:

$$F_i^{(k)}(c) = \sum_{\succ \text{ such that } c_1 \succ \dots c_{k-1} \succ c} F_i(\succ)$$

and $\bar{F}_i(c)$ be the fraction of type i students listing school c :

$$\bar{F}_i(c) = \sum_{k=1}^N F_i^{(k)}(c).$$

Denote the set of schools listed by type i students by $\text{supp}(i) = \{c \mid \bar{F}_i(c) > 0\}$. We also abuse notation and for $c \in C$ we denote

$$\text{supp}(c) = \{\hat{c} \in C \setminus \{c\} \mid \exists i \in I: c \in \text{supp}(i), \hat{c} \in \text{supp}(i)\}$$

the set of schools that are listed by types that also list c . We assume that the all schools are popular in the sense that for all c ,

$$F^{(1)}(c) := \int F_i^{(1)}(c) \nu(di) > k_c.$$

Our analysis can be extended to the case in which this inequality holds for some but not all schools, but we simplify exposition by imposing the inequality in all schools. We also assume that $\bar{F}(c) > F^{(1)}(c)$ for all c so that each school has a nontrivial mass of students that demand it but not in the top position.

2.2 Examples

We now discuss how prominent school choice models can be cast as special cases of our model.

Example 1 (Horizontal differentiation and distance-based priorities). $I \subset \mathbb{R}_+^2$ models a city and a student's type is her location $i \in I$ in the city. Schools are located and spread across the city. Let $d(i, c) \in [0, 1]$ be a distance between a student located in i and school c .⁵ Similar to Abdulkadiroğlu et al. (2017), the utility that a student located in i derives from attending school c is in part determined by

⁴Given any subset of students $S' \subseteq S$, $\bar{\nu}(S') = \int \sum_{\succ \text{ such that } (i, \succ) \in S'} F_i(\succ) \nu(di)$.

⁵The distance function can be arbitrary. The only relevant property is that it satisfies the triangle inequality.

$d(i, c)$. For example, one can generate the utility that a type i student derives from school c as

$$u_{s,c} = -d(i, c) + \epsilon_{i,c}$$

where $\epsilon_i = (\epsilon_{i,c})_{c=1}^N$ is a shock vector and has a distribution H_i .⁶ In this case, we can construct the distribution over the finite set of preferences as:

$$F_i(\succ) = \text{Prob}[u_{s,c_1} \geq u_{s,c_2} \geq \dots \geq u_{s,c_N}]$$

where $c_1 \succ \dots \succ c_N$.

Schools can rank students using a variety of criteria (including random tie breaking, discussed below). Under distance-based priorities, the score that a student type i has in school c is given by $e_c^s = 1 - d(i, c)$.

Several papers compare single to multiple tie breaking in school choice problems (Abdulkadiroğlu et al., 2009; Ashlagi and Nikzad, 2020; Arnosti, 2022; Allman et al., 2022). Our model also accommodates these priorities.

Example 2 (Random tie breakings). Take a school choice model in which students have no types and students' preferences are given by a distribution $F(\succ)$. Given the set of schools, scores at each school are randomly determined in $[0, 1]$. Our general model can accommodate these random priorities as follows.

Let $I = [0, 1]^N$ be the set of types and ν be N independent uniform distributions over $[0, 1]$. The c -component of a student type $i \in [0, 1]^N$ determines the score that student i has in school c , that is, $e_c(i) = i_c$. In this case, our model becomes a school choice problem in which students are ranked according to multiple tie breaking (MTB) (Abdulkadiroğlu et al., 2009).

The model can also accommodate the case of single tie breaking (STB). When $I = [0, 1]$, and ν is the uniform distribution on $[0, 1]$. A type i student has score i at each school. Our model becomes a school choice problem in which students are ranked according to a single lottery (Abdulkadiroğlu et al., 2009).

2.3 Stable matchings

A matching is a function $\mu : S \cup C \rightarrow (C \cup \{\emptyset\}) \cup 2^S$ such that:

⁶Using this formulation, we can model fixed effects and also interaction effects other than distance (Abdulkadiroğlu et al., 2017).

- i. For all $s \in S$, $\mu(s) \in C \cup \{\emptyset\}$;
- ii. For all $c \in C$, $\mu(c) \subseteq S$ with $\bar{\nu}(\{s | \mu(s) = c\}) \leq k_c$;
- iii. For all $c \in C$ and all $s \in S$, $\mu(s) = c$ iff $s \in \mu(c)$.
- iv. For all c , $\{s \in S \mid c \succ_s \mu(s)\}$ is open.

The first condition says that each student is assigned to a school, the second condition says that each school is assigned to a measure of students that does not exceed its capacity, the third condition says that a student is assigned to a school iff the school is assigned to that student. The fourth condition is technical and eliminates redundant matchings that differ in a measure 0 of students (Azevedo and Leshno, 2016).

A matching μ is stable if for all $s \in S$, $\mu(s) \succ_s \emptyset$, and for all $c \in C$ such that $c \succ_s \mu(s)$, the following conditions hold: (i) $|\{s | \mu(s) = c\}| = k_c$; and (ii) $e_c^s < e_c^{s'}$ for all s' with $\mu(s') = c$. Intuitively, a matching is stable if there is no pair (s, c) that can block the matching (Gale and Shapley, 1962). Stability is an important desideratum in matching theory and its many applications (Roth, 1982; Abdulkadiroğlu et al., 2009).

To characterize stability, we follow Abdulkadiroğlu et al. (2015) and Azevedo and Leshno (2016) and find stable matchings as solutions to a supply and demand system of equations. Given cutoffs $p = (p_c)_{c=1}^N$, a student s can get admission to c if $e_c(s) \geq p_c$. A student's demand is given by her favorite school among those she can get admission given p . We thus define $D_c(p)$ as the measure of students that demand school c as a function of cutoffs p . A stable matching can be found by means of market-clearing cutoffs $p = (p_c)_{c=1}^N$ that solve

$$D_c(p) = k_c \quad \forall c \tag{2.1}$$

Given market-clearing cutoffs, a stable matching is built by assigning each student to her most preferred school among those where her score exceeds the cutoff.

While the system of equations (2.1) is neat and simple to interpret, it can be solved in closed-form solutions only for special cases. When we can find a closed-form solution to (2.1), it is simple to calculate statistics for the resulting stable matching. However, solving the model in closed-form is unfeasible even for relative simple models.⁷

⁷The system of equations (2.1) is non-linear in p . Under multiple tie breaking, each equation in (2.1) is polynomial of degree N .

3 Students assigned to their top schools

This Section states and discusses our bounds for the measure of students assigned to their top schools. We then provide some examples and sketch some of the arguments in the proof.

For a given matching, let $R^{(1)}(c)$ be the mass of students assigned to school c that put c as their top school. Obviously, $0 \leq R^{(1)}(c) \leq k_c$. $R^{(1)}(c)$ is an important metric usually employed by policy makers to evaluate the effectiveness of a matching (Abdulkadiroğlu et al., 2009). In the next Section, we discuss other performance measures.

For each school c , we compute the demands

$$\Lambda_c^1(x) = \int_{e_c(i) \geq x} F_i^{(1)}(c) \nu(di) \quad \text{and} \quad \bar{\Lambda}_c(x) = \int_{e_c(i) \geq x} \bar{F}_i(c) \nu(di)$$

for all $x \in [0, 1]$. Let $\phi_c \in [0, 1]$ and $\Phi_c \in [0, 1]$ be defined by the equations

$$\phi_c = \max \{x \in [0, 1] \mid \Lambda_c^1(x) = k_c\} \quad (3.1)$$

$$\Phi_c = \min \{x \in [0, 1] \mid \bar{\Lambda}_c(x) = k_c\}. \quad (3.2)$$

In contrast to the cutoff p_c that clears the market for school c in a stable matching, cutoffs ϕ_c and Φ_c are entirely determined by the local demand for school c : while ϕ_c is determined by the mass of students that demand c first $(F_i^{(1)}(c))_i$, Φ_c is determined by the mass of students that list c in any position $(\bar{F}_i(c))_i$.

The following result provides estimates for $R^{(1)}(c)$ in a stable matching.

Theorem 1. *For any stable matching and all $c = 1, \dots, N$:*

$$R^{(1)}(c) \geq k_c - \bar{\eta}_c \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \quad (3.3)$$

and

$$R^{(1)}(c) \leq \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) + \eta_c \left(\int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) \geq \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \right) \quad (3.4)$$

where

$$\bar{\eta}_c = \min \left\{ 1, \frac{\inf_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \Lambda_c^1(x)}{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (\Lambda_c^1(x) + \int_{e_c(i) \geq x, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di))} \right\}$$

and

$$\eta_c = \frac{\inf_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \Lambda_c^1(x)}{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \bar{\Lambda}_c(x)}.$$

Theorem 1 provides estimates for the measure of students assigned to their top schools. To apply the bounds, one computes cutoffs ϕ_c and Φ_c that are entirely determined by each school c supply and demand. The real numbers η_c and $\bar{\eta}_c$ adjust for the fact that we do not employ stable matching cutoffs p but relaxed cutoffs ϕ_c and Φ_c . As we show below and in Section 4, Theorem 1 can be easily applied to several examples and models.

The idea behind bounds (3.3) and (3.4) is that the measure of students assigned to their top schools depends on how students can congest schools they do not rank top. A student that does not rank a school c at the top may still congest it depending on her scores in c and other schools $\hat{c} \neq c$. Thus, the measure of students assigned to their top schools critically depends on how types determine preferences for each school and scores across schools.

The first bound in the Theorem, inequality (3.3), provides a condition under which a high fraction of students assigned to school c will rank it as their top school. Most students will be assigned to their top school in c when (i) students that rank c apply to c first ($\bar{F}_i(c) \approx F_i^{(1)}(c)$), or (ii) students that rank c second, third, etc have a low score in c (that is, $\int_{e_c(i) \geq \phi_c} \bar{F}_i(c) - F_i^{(1)}(c) \nu(di) \approx 0$), or, more generally, (iii) most students that rank c second, third, etc and have a high score in c are also likely to get admission in some other school (that is, $\int_{e_c(i) \geq \phi_c} \bar{F}_i(c) - F_i^{(1)}(c) \nu(di) \approx \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) \geq \Phi_{\hat{c}} \text{ all } \hat{c} \in \text{supp}(c)} \bar{F}_i(c) - F_i^{(1)}(c) \nu(di)$). More generally, to evaluate inequality (4.7), we compute the measure of the set of students that rank c second, third, etc, and have a high score in c and a low score in some other school \hat{c} . When this measure is low, most students that get admission to c will naturally rank c top.

The second bound in the Theorem, inequality (3.4), provides a condition under which a low fraction of students assigned to school c will rank it as their top school. The Theorem shows that in a stable matching, few students assigned to school c will rank it as their top school when (i) most students that rank c top are unlikely to have sufficiently high scores (that is, $\sup \{F_i^{(1)}(c)/\bar{F}_i(c) \mid e_c(i) \geq \Phi_c\} \approx 0$), and (ii) most students that rank c second, third, etc and have a high score in c are unlikely to get admission

in some other school (that is, $\int_{e_c(i) \geq \phi_c} \bar{F}_i(c) - F_i^{(1)}(c) \nu(di) \approx \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) \leq \phi_{\hat{c}} \text{ all } \hat{c}} \bar{F}_i(c) - F_i^{(1)}(c) \nu(di)$). If (i) were not satisfied, then we could secure a non-negligible mass of students for whom c is the top choice and are sure to be assigned. Condition (ii) ensures that students for whom c is listed but is not top are admitted to c in a stable matching.

We now illustrate the bounds. The following example shows a stable matching that results in all students assigned to their top schools. This happens even when the preferences of both sides of the market do not conform: in our example, some students rank a school top, but that school does not rank those students highly. The example thus shows that system-wide effects may favor the efficiency of stable matchings even when preferences and priorities do not conform.⁸ The example below also shows that the lower bound (3.3) is tight.

Example 3. Suppose that $N = 2$ and $I = [0, 1]$. Each school has capacity $k = \frac{1}{4}$. Students $i \leq 1/4$ are elite students, with outstanding academic performance. For $i \leq 1/4$, scores are given by $e_{c_1}(i) = e_{c_2}(i) = 1 - i$. School c_1 (resp. c_2) is located in 0 (resp. 1) and students $i > 1/4$ are ranked according to distance. Concretely, for $i > 1/4$ $e_{c_1}(i) = 1 - i$ while $e_{c_2}(i) = i - 1/4$. For each i , a fraction $\alpha(i)$ (resp. $1 - \alpha(i)$) of students rank school c_1 first (resp. school c_2 first). Assume that $\alpha(i) = 1$ for $i \leq 1/2$ and $\alpha(i) = 0$ for $i > 1/2$.

It is simple to see that $\phi_{c_1} = 3/4$, $\Phi_{c_1} = 3/4$, $\phi_{c_2} = 1/2$, $\Phi_{c_2} = 3/4$. We can then compute

$$\int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di)$$

and note that for each $c = c_1, c_2$:

$$\int_{e_{c_1}(i) \geq 3/4, e_{c_2}(i) < 3/4} (1 - \alpha(i)) di = 0 \quad \text{and} \quad \int_{e_{c_2}(i) \geq 1/2, e_{c_1}(i) < 3/4} \alpha(i) di = 0$$

Using Theorem 1, $R^{(1)}(c_1) = R^{(1)}(c_2) = k$. Note that the matching is also Pareto-efficient as all students are assigned to their top schools.⁹

The next example shows that the upper bound (3.4) is tight.

⁸Erdil and Ergin (2008) show simulations in which the preferences of both sides of the market conform and, as a result, the stable matching is efficient. In those simulations, priorities are given by multiple tie breaking and walk zones. As distance becomes more important for students (in their model, that is captured by $\beta \rightarrow 1$), the efficiency loss in the stable matching goes to 0 since in the limit both sides of the market have perfectly conforming preferences. See also Salonen and Salonen (2018) for theoretical results. See also our Theorem 2 part a.

⁹We will provide results about the measure of students that can be Pareto-improved in Section 4.

Example 4. Suppose that $N = 2$ and $I = [0, 1]$. Each school has capacity $k = \frac{1}{4}$. Students i live in position i with preferences given by $F_i^{(1)}(c) = 1/2$ and $\bar{F}_i^{(1)}(c) = 1$ for each c . Schools c_1 and c_2 are located at the extremes of the interval. Priorities are distance-based so the scores of agent i are given by $e_{c_1}(i) = 1 - i$ and $e_{c_2}(i) = i$. It is simple to see that $\phi_c = 1/2$ and $\Phi_c = 3/4$. Thus

$$R^{(1)}(c) \leq \int_{1-i \geq 1-k} \frac{1}{2} di + \frac{1}{2} \int_{1-i \geq 1/2, i \geq 1/2} \frac{1}{2} di = \frac{k}{2}.$$

Thus bound is tight since in the unique stable matching, the cutoff equals $p_c = 3/4$, and thus in each school only half of the students assigned to the school rank the school first.

We close this Section by discussing the main ideas behind the proof of Theorem 1. Since ϕ_c solves a market-clearing condition for a demand Λ_c^1 that is below the total demand D_c , we deduce that $\phi_c \leq p_c$ for any cutoff vector p from a stable matching. Analogously, $p_c \leq \Phi_c$. See Lemma 1 in the Appendix for details.

Cutoffs ϕ_c and Φ_c are important in that they provide bounds for cutoffs p characterizing stable matchings. More subtly, ϕ_c and Φ_c are informative about the measure of students assigned to their top schools. Indeed, when $p_c = \phi_c$, then the number of students assigned to their top schools in c equals k_c .¹⁰

$$R^{(1)}(c) = \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) = \int_{e_c(i) \geq \phi_c} F_i^{(1)}(c) \nu(di) = k_c.$$

Similarly, we note that

$$R^{(1)}(c) = \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) = \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) + \int_{p_c \leq e_c(i) \leq \Phi_c} F_i^{(1)}(c) \nu(di).$$

It follows that $R^{(1)}(c) = 0$ when $\Phi_c = p_c$ and $F_i^{(1)}(c) = 0$ for all $e_c(i) \geq \Phi_c$. In general, however, $\phi_c < p_c < \Phi_c$. The key technical observation is that we can bound $p_c - \phi_c$ and $\Phi_c - p_c$ by using several market-clearing conditions. The proof (which is presented in the Appendix) thus bounds the distance between the solutions to different non-linear market clearing equations to derive estimates for the measure of students assigned to their top schools.

¹⁰Moreover, $R^{(1)}(c) = k_c$ iff $\phi_c = p_c$.

4 Priorities in school choice

In districts employing the deferred acceptance algorithm, oftentimes authorities design the priority criteria employed by schools. For example, in cities such as Boston and Copenhagen, students get priority based on how close they live to a school. In New Haven, students get higher priority in schools in which they have siblings. In many cities, schools use random priorities. While in New York City a student gets a random score that applies to all schools, in Chile each student gets a different random score for each school. All these priority decisions impact the final assignment and performance measures. This Section applies our bounds to distance-based priorities and random tie breaking.

Given any matching μ , a positive measure set of students $S' \subseteq S$ can be Pareto-improved if there exists a matching $\bar{\mu}$ such that for almost all $s \in S$, $\bar{\mu}(s) \succeq_s \mu(s)$ with strict preferences for $s \in S'$. When the matching $\bar{\mu}$ is such that $\bar{\mu}(c) = \mu(c)$ for all $c \in C \setminus \{c', c''\}$, with $c' \neq c''$, we say that S' is part of Pareto-improving pairs. Define

$$P = \bar{\nu} \left(\bigcup_{S' \text{ can be Pareto-improved}} S' \right) \quad \text{and} \quad P^2 = \bar{\nu} \left(\bigcup_{S' \text{ is part of Pareto-improving pairs}} S' \right)$$

When $P = 0$, the measure of students that can be Pareto-improved is 0 and thus the matching is Pareto-efficient. More generally, P provides the measure of all students who could envision a Pareto-improvement of the proposed matching μ and thus P is a metric of the efficiency of the matching.¹¹

The set-up for this Section is the model of horizontal differentiation presented in Example 1. We fix the demand and the capacity of each school and compute the bounds from Theorem 1 for distance-based priorities and random priorities. While we explore random priorities within the context of the model of horizontal differentiation, some of our bounds apply to the general random tie breaking model presented in Example 2.

4.1 Distance-based priorities

Under distance-based priorities, school c ranks students according to $e_c^s = 1 - d(i, c)$. In this Subsection, we argue that the fraction of students assigned to their top schools depends on several factors, including how much students value proximity, and the capacity and geographical dispersion of schools.

¹¹If S' and S'' can be Pareto-improved, it does not follow that $S' \cup S''$ can be Pareto-improved.

To derive a lower bound for $R_{DB}^{(1)}(c)$, it is useful to consider the set of all students that can get admission to c given cutoff ϕ_c^{DB} but are rejected by some school \hat{c} given $\Phi_{\hat{c}}^{DB}$:

$$H(c) = \left\{ i \mid d(i, c) \leq 1 - \phi_c^{DB} \text{ and } d(i, \hat{c}) > 1 - \Phi_{\hat{c}}^{DB} \text{ some } \hat{c} \in \text{supp}(c) \right\}.$$

$H(c)$ estimates the set of all the students that could get admission to c but would be rejected by some school \hat{c} . Theorem 1 can be used to deduce:

$$R_{DB}^{(1)}(c) \geq k_c - \nu(H(c)) \sup_{d(i, c) \leq 1 - \phi_c^{DB}} (\bar{F}_i(c) - F_i^{(1)}(c)) \quad (4.1)$$

This bound shows two forces that make $R_{DB}^{(1)}(c)$ close to k_c .

A. Consistent preferences and priorities. When all students living within distance $1 - \phi_c^{DB}$ of school c list c at the top,¹² then $R_{DB}^{(1)}(c) = k_c$. In this case, preferences and priorities are consistent in the sense that students that have a high score in c (in other words, that live close to c) also rank school c at the top. When preferences and priorities are consistent, all students will be assigned to their top school and the matching will be efficient. The observation that consistent preferences and priorities favor efficiency is not new and is discussed by Salonen and Salonen (2018), Echenique et al. (2020), Cantillon et al. (2022).

B. Clustered schools. When the measure $\nu(H(c))$ is small, then $R_{DB}^{(1)}(c)$ is close to k_c . When $\nu(H(c))$ is close to 0, then most students that have a score high enough for school c also have score high enough in other schools \hat{c} . In this case, schools are clustered and distance-based priorities result in a subset of students who are close to all schools and can get admission anywhere. As a result, many of those students get accepted to the school they like the most.¹³

The following example shows that under distance-based priorities, all students can be assigned to their top schools because, even when preferences and priorities are not consistent, because there is a set of students that live close to all schools and thus each of those students get accepted to her top school.

¹²That is, when for all i such that $d(i, c) \leq 1 - \phi_c^{DB}$, $\bar{F}_i(c) = F_i^{(1)}(c)$.

¹³This intuition is similar to the idea that under single-tie breaking, many students are assigned to their top schools (Allman et al., 2022).

Example 5 (Clustered schools). Suppose that $N = 2$ and $I = [0, 1]$. Each school has capacity $k < 1/2$. A fraction $\alpha(i)$ (resp. $1 - \alpha(i)$) of students rank school c_1 first (resp. school c_2 first) and we assume that $\alpha(i) = 1 - \alpha(1 - i)$ for all $i < 1/2$. Both schools are located at $1/2$. In the unique stable matching, $p_{c_1} = p_{c_2} = k$. Students in $\tilde{I} = [\frac{1}{2} - p, \frac{1}{2} + p]$ could get accepted to both schools and thus

$$R_{DB}^{(1)}(c_1) = R^{(1)}(c_2) = k.$$

C. Strong competition and weak preferences for location. We now derive an upper bound for $R_{DB}^{(1)}(c)$ and show how competition places upper bounds on the number of students assigned to their top schools. Consider the set of all students that could get admission to c and \hat{c} given cutoffs ϕ_c^{DB} and $\phi_{\hat{c}}^{DB}$:

$$A(c, \hat{c}) = \left\{ i \mid d(i, c) \leq 1 - \phi_c^{DB} \right\} \cap \left\{ i \mid d(i, \hat{c}) \leq 1 - \phi_{\hat{c}}^{DB} \right\}. \quad (4.2)$$

Note that if $d(i, c) \leq 1 - \phi_c^{DB}$, by the triangle inequality, $d(i, \hat{c}) \geq d(c, \hat{c}) - d(i, c) \geq d(c, \hat{c}) - 1 + \phi_c$. So, the set in equation (4.2) is empty whenever

$$2 \leq d(c, \hat{c}) + \phi_c^{DB} + \phi_{\hat{c}}^{DB} \quad (4.3)$$

for all $c \neq \hat{c}$. The triangle inequality used to derive this condition captures an important intuition about congestion under distance-based priorities: When cutoffs in schools are high, having a score high enough for some schools implies that the scores in other schools are below the cutoffs. This means that under distance-based priorities, students located near a school will have limited chances to attend other schools which, as shown below, makes efficiency much harder.

Condition (4.3) holds for all schools provided that for all c

1. The function $x \in [0, 1] \mapsto \int_{d(i, c) < x} F_i^{(1)}(c) \nu(di)$ has strictly positive derivative at $x = 0$;
2. $d(c, \hat{c}) > 0$ for all $\hat{c} \neq c$; and
3. k_c is small enough;

The first condition says that each school has demand arbitrarily close to it. It is relatively simple to show that under the first condition, $\phi_c^{DB} \rightarrow 1$ as $k_c \rightarrow 0$. Since $d(c, \hat{c}) > 0$, it follows that (4.3) holds when all capacities $(k_c)_{c=1}^N$ are small enough.

Under (4.3), it is simple to apply Theorem 1 to obtain:

$$R_{DB}^{(1)}(c) \leq k_c \sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \frac{F_i^{(1)}(c)}{\bar{F}_i(c)} \quad (4.4)$$

When capacities are low, under distance-based priorities some students get assigned to a nearby school that is not their top choice. This puts an upper bound on the fraction of students assigned to their most preferred schools.¹⁴ Under distance-based priorities, each student may get admission to only one school, which may or may not be her top school.

The next example shows that under (4.3), it is entirely possible that an arbitrarily small fraction of students are assigned to their top schools.

Example 6. Suppose that $N = 2$ and $I = [0, 1]$. Each school has capacity $k < 1/2$. A fraction $\alpha(i)$ (resp. $1 - \alpha(i)$) of students rank school c_1 first (resp. school c_2 first) and we assume that $\alpha(i) = 1 - \alpha(1 - i)$ for all $i < 1/2$. Assume that $\alpha(i)$ is increasing in i with $\alpha(i) > 0$ for all $i \in [0, 1]$. This means that students tend to value schools that are farther away. Schools rank using distance-based priorities. Under

$$\int_{i \leq 1/2} \alpha(i) > k. \quad (4.5)$$

it follows that $\phi_{DB}(c) > 1/2$ and (4.3) holds. Since $\bar{\Lambda}(x) = 1 - x$, it is simple to see that $\Phi = 1 - k$. Clearly,

$$R_{DB}^{(1)}(c) \leq k \sup_{i \leq k} \alpha(i) = k\alpha(k)$$

It follows that for any $\epsilon > 0$, there exists an increasing function α and $k < 1/2$ such that (4.5) holds and $R_{DB}^{(1)}(x) < \epsilon$ for all c .¹⁵

4.2 Random priorities

This Subsection applies our bounds to the widely studied model of school choice with random priorities (Abdulkadiroğlu et al., 2009; Ashlagi and Nikzad, 2020; Arnosti, 2022; Allman et al., 2022).

¹⁴Clearly, the bound is non-trivial only when some of the students living close to c list the school not in the top.

¹⁵Take $\alpha(i) \geq \epsilon i$ for all $i \in [0, 1/2]$ and $k < \epsilon/2$. Then, $R^{(1)} \leq k\alpha(k) < \epsilon/2 < \epsilon$.

It is simple to see that under single or multiple tie breaking, cutoffs are identical and given by

$$\phi_c^{RP} = 1 - \frac{k_c}{F^{(1)}(c)} \quad \text{and} \quad \Phi_c^{RP} = 1 - \frac{k_c}{\bar{F}(c)} \quad (4.6)$$

Using Theorem 1, we deduce that under multiple tie breaking, for each school c ¹⁶

$$\begin{aligned} & k_c \left(\frac{F^{(1)}(c)}{F^{(1)}(c) + (\bar{F}(c) - F^{(1)}(c)) \left(1 - \prod_{\hat{c} \in \text{supp}(c)} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})}\right)} \right) \\ & \leq R_{MTB}^{(1)}(c) \leq k_c \left(1 - \frac{\bar{F}(c) - F^{(1)}(c)}{\bar{F}(c)} \prod_{\hat{c} \in \text{supp}(c)} \left(1 - \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right) \right) \end{aligned} \quad (4.7)$$

Since $\bar{F}(c) > F^{(1)}(c)$, both the upper and the lower bound for $R_{MTB}^{(1)}(c)$ are informative. Under multiple tie breaking, some students will be assigned to school c even when c is not their top school, but there will always be some students assigned to their top schools.

Our results can also be used to obtain bounds when priorities are derived using single tie breaking. It is immediate to see that

$$R_{STB}^{(1)}(c) \geq R_{MTB}^{(1)}(c)$$

whenever

$$\frac{k_c}{F^{(1)}(c)} \leq \frac{\min\{\frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \mid \hat{c} \neq c\}}{1 - \prod_{\hat{c} \neq c} \left(1 - \frac{k_{\hat{c}}}{\bar{F}(\hat{c})}\right)}.$$

This bound says that when school c is sufficiently popular (that is, $\frac{k_c}{F^{(1)}(c)}$ is small enough), more students are assigned to c in the top position under single tie breaking than under multiple tie breaking. See Abdulkadiroğlu et al. (2009), Allman et al. (2022), Ashlagi and Nikzad (2020), and Arnosti (2022) for similar results.¹⁷

¹⁶For multiple tie breaking, it is possible to derive bounds that do not use Theorem 1. By definition,

$$R^{(1)}(c) = F^{(1)}(c)(1 - p_c) \in [k_c(1 - \Phi_c), k_c(1 - p_c)] = \left[k_c \frac{F^{(1)}(c)}{\bar{F}(s)}, k_c\right].$$

The bounds given in (4.7) are strictly sharper than these simple bounds.

¹⁷Our bound restricts the popularity of the school c but puts no restriction on the demand for schools, in contrast to previous results.

4.3 Comparing priorities

We now compare distance-based priorities and multiple tie breaking. We evaluate these priority criteria using the fraction of students assigned to their top schools and the fraction of students that can be Pareto-improved.

Theorem 2. *a. Suppose that for all c and all $d(i, c) \leq 1 - \phi_c^{DB}$, $\bar{F}_i(c) = F_i^{(1)}(c)$. Then, for all c*

$$R_{DB}^{(1)}(c) = k_c \quad \text{and} \quad P_{DB}^2 = P_{DB} = 0.$$

In particular, no alternative priority criterion can result in more students assigned to top schools than distance-based priorities.

b. Assume condition (4.3) and that for all c ,

$$\sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \left(\frac{F_i^{(1)}(c)}{\bar{F}_i(c)} - \frac{F^{(1)}(c)}{\bar{F}(c)} \right) \leq \frac{\left(1 - \frac{F^{(1)}(c)}{\bar{F}(c)}\right) \left(\prod_{\hat{c} \in \text{supp}(c)} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})}\right)}{1 + \left(\frac{\bar{F}(c)}{F^{(1)}(c)} - 1\right) \left(1 - \prod_{\hat{c} \in \text{supp}(c)} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})}\right)} \quad (4.8)$$

Then, for all c

$$R_{DB}^{(1)}(c) < R_{MTB}^{(1)}(c).$$

If we additionally assume that $\mathbb{P}[c \succ c' \mid i] > 0$ for all $c \neq c'$ and all $i \in I$, then

$$P_{MTB}^2 = P_{MTB} < P_{DB} = P_{DB}^2.$$

The first part shows that when students value distance strongly, then all students are assigned to their top schools under distance-based priorities. To see how the sufficient conditions can be satisfied, fix I , ν , the set of schools C , the distance function $d(i, c)$, the distribution $\bar{F}_i(c)$, and the capacities k_c . For $c \in C$, compute Φ_c^{DB} and assume that capacities are low enough so that $\{i \mid e_c(i) \geq \Phi_c\} \cap \{i \mid e_{\hat{c}}(i) \geq \Phi_{\hat{c}}\} = \emptyset$ for all $c \neq \hat{c}$. Now, construct F_i such that for all $i \in I_c$, type i students rank c first. This implies that $\bar{F}_i(c) = F_i^{(1)}(c)$ for all $i \in I_c$ and therefore $\phi_c^{DB} = \Phi_c^{DB}$ and

$$\sup_{d(i,c) \leq 1 - \phi_c^{DB}} \bar{F}_i(c) - F_i^{(1)}(c) = 0.$$

The second part of the result provides conditions under which multiple tie breaking assigns more

students to their top schools than distance-based priorities. Note that when types do not determine preferences, that is $F_i(\succ) = F(\succ)$ for all $i \in I$, then the left-hand side of (4.8) equals 0 and thus condition (4.8) holds. More generally, condition (4.8) captures the idea that types have only a mild impact on preferences so that the ratio $\frac{F_i^{(1)}(c)}{\bar{F}_i(c)}$ stays relatively flat as a function of i and close to $\frac{F^{(1)}(c)}{\bar{F}(c)}$.¹⁸ Behind this result is the idea that when preferences for nearby schools are weak and competition is strong, distance-based priorities assign some students to schools just because they live nearby even when those schools are not ranked top by them, while under multiple tie breaking those students still have a chance to get accepted in their top schools.

Theorem 2 also compares the fractions of students that can be Pareto improved. It is relatively simple to prove that for any matching μ , $P_2 \leq P$ and

$$\sum_{c=1}^N R^{(1)}(c) + P \leq \sum_{c=1}^N k_c. \quad (4.9)$$

We then prove that, under the conditions of the Theorem, these inequalities bind and therefore the fraction of students assigned to top schools and the fraction of students that can be Pareto-improved add up to the total capacity of schools. See Appendix B.1.

Inequality (4.9) can be strict in some important cases. For example, under single tie breaking the matching is Pareto-efficient, but it is possible that not all students are assigned to their top schools. Example 7 in Appendix B.1 shows that under distance-based priorities, inequality (4.9) can be strict when condition (4.3) does not hold.

5 Simulations

We now verify our results in a simulated economy with a finite number of students and a finite number of schools. We construct an economy with 10,000 students and 100 schools with 50 seats each. Schools are located at the integer points of a 10×10 grid and students are randomly located within this grid.

¹⁸Note that both conditions (4.8) and (4.3) restrict F and k . They simultaneously hold when types have a limited impact on preferences so that for all c

$$\sup_i \frac{F_i^{(1)}(c)}{\bar{F}_i(c)} \leq \frac{F^{(1)}(c)}{F^{(1)}(c) + (\bar{F}(c) - F^{(1)}(c)) \left(1 - \prod_{\hat{c} \in \text{supp}(c)} \frac{k_{\hat{c}}}{F^{(1)}(\hat{c})}\right)}$$

and, given preferences, k_c is small enough so that (4.3).

Each student is surrounded by four schools, and those are the only schools a student applies to. The utility that a student s gets if she is assigned to one of the schools c that surround her is

$$u(s, c) = \alpha(1 - d(s, c)) + (1 - \alpha)\epsilon_{s,c}$$

where $d(s, c)$ is the Euclidean distance between student s and school c , $\epsilon_{s,c} \sim U[0, 1]$ are i.i.d idiosyncratic noise terms. Parameter α captures the relevance of school proximity in students' preferences. When α is close to 1, proximity to schools fully determines a student's preferences. Conversely, when α is close to 0, students' preferences are random and proximity to schools plays no role.

We compare three priority criteria (MTB, STB, DB) in terms of the fraction of students assigned to their top schools. As suggested by Theorem 2, these comparisons will crucially depend on α . When proximity is important (α close to 1), Theorem 2 (a) shows that DB results in more students assigned to their top schools than SB. More subtly, when proximity is moderately important (α small), Theorem 2 (b) shows that MTB assigns more students to their top schools than DB. Figure 1 shows the expected fraction of students assigned to top schools for MTB, STB, and DB for different values of $\alpha \in [0, 1]$. This exercise confirms our theoretical predictions in the finite economy model.¹⁹

¹⁹Figure 1 also shows that STB results in more students assigned to top schools than MTB priorities, as in Abdulkadiroğlu et al. (2009).

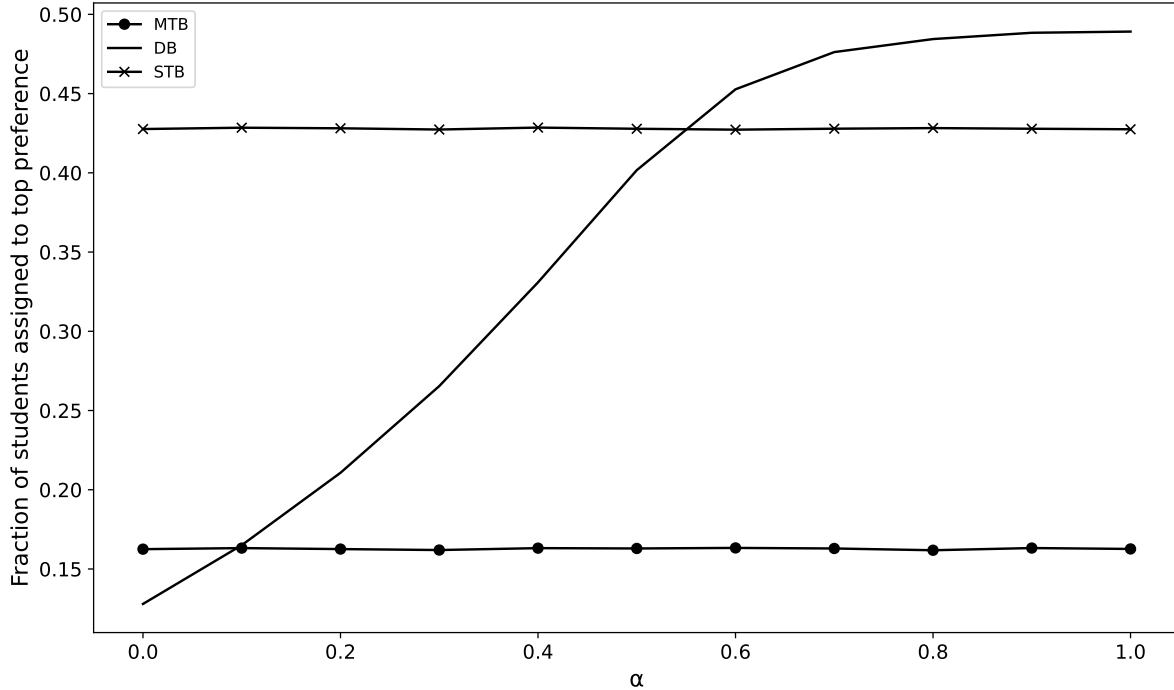


Figure 1: Comparison of priority criteria for different preferences.

6 Concluding remarks

This paper provides performance estimates for stable matchings. Stable matchings are hard to analyze because comparative statics results and closed-form formulas are typically unfeasible. Quantifying the performance of stable matching is a useful step to understand some tradeoffs in the design of matching markets and inform policy decisions.

Our paper makes two main contributions. First, we provide tight lower and upper bounds for the measure of students assigned to top schools. These bounds apply to a general matching model and can be easily used in applications. Second, we examine the impact of distance-based priorities in school choice. While several school districts employ proximity as a priority criterion, little is known about how this policy decision affects the outputs of the deferred acceptance algorithm. We show that when students highly value proximity, efficient outcomes are achievable. However, under weaker proximity preferences, even multiple tie breaking may assign more students to their top choices than distance-

based priorities. Future research could sharpen our bounds. It would also be interesting to estimate other performance measures, including different diversity metrics.

Appendix

A Proof of Theorem 1

Define $P_{k,c}$ as the set of all orderings \succ such that school c lies in position k . Given cutoffs $p \in [0, 1]^N$, the demand for school c can be written as

$$D_c(p) = \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \sum_{k=2}^N \sum_{\succ \in P_{k,c}} \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \succ c} F_i(\succ) \nu(di)$$

The demand is built in the following way. Fix a student type i that has school c as its k -th preference. For each one of these student types, a mass $F_i(\succ)$ reveals preference ordering $\succ \in P_{k,c}$. However, only a fraction of $F_i(\succ)$ effectively demands school c . These are students rejected at all $k-1$ schools preferred over c according to \succ ($e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \succ c$) and accepted at school c ($e_c(i) \geq p_c$). Then, adding up over all possible ranking positions k , all preference orderings $\succ \in P_{k,c}$ with positive measure ($F_i(\succ) > 0$), and aggregating over all student types $i \in I$, we get the total demand for school c .

The following result is useful to derive our efficiency bounds.

Lemma 1. *Let p be a market-clearing cutoff vector characterizing a stable matching. Then, for all c , $\phi_c \leq p_c \leq \Phi_c$.*

Proof. For any $x \in [0, 1]^N$, $\Lambda_c^1(x_c) \leq D_c(x) \leq \bar{\Lambda}(x_c)$ which are all decreasing in x_c . Then, fix any p_{-c} and let ϕ_c, p_c, Φ_c be solutions to $\Lambda_c^1(\phi_c) = k_c$, $D_c(p_c, p_{-c}) = k_c$ and $\bar{\Lambda}_c(\Phi_c) = k_c$ respectively, it is true that $\phi_c \leq p_c \leq \Phi_c$. \square

A.1 Upper Bound

Let p be a cutoff vector for a stable matching. Then

$$\begin{aligned}
k_c &= D_c(p) \\
&= \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \sum_{k=2}^N \sum_{\succ \in P_{k,c}} \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \succ c} F_i(\succ) \nu(di) \\
&\geq \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \sum_{k=2}^N \sum_{\succ \in P_{k,c}} \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} F_i(\succ) \nu(di) \\
&= \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} \sum_{k=2}^N \sum_{\succ \in P_{k,c}} F_i(\succ) \nu(di) \\
&= \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} \sum_{k=2}^N F_i^{(k)}(c) \nu(di) \\
&= \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) + \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \\
&:= \Lambda(p_c).
\end{aligned}$$

To see the inequality above, note that for any $k = 2, \dots, N$, $\succ \in P_{k,c}$, and $\hat{c} \succ c$, it follows that $\hat{c} \in \text{supp}(i) \setminus \{c\}$. Thus, for any $k = 2, \dots, N$ and $\succ \in P_{k,c}$

$$\left\{ i \in I \mid e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \succ c \right\} \supseteq \left\{ i \in I \mid e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\} \right\}$$

and therefore

$$\int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \succ c} F_i(\succ) \nu(di) \geq \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} F_i(\succ) \nu(di)$$

Since $\bar{\Lambda}(\Phi_c) = k_c \geq \Lambda(p_c)$

$$0 \leq \bar{\Lambda}(\Phi_c) - \Lambda(p_c) = \bar{\Lambda}(p_c) + \int_{p_c}^{\Phi_c} \bar{\Lambda}'(s) ds - \Lambda(p_c) \leq (\Phi_c - p_c) \sup_{x \in [\phi_c, \Phi_c]} \bar{\Lambda}' + \bar{\Lambda}(p_c) - \Lambda(p_c)$$

therefore

$$\begin{aligned}
& (\Phi_c - p_c) \left(-\sup \bar{\Lambda}' \right) \\
& \leq \bar{\Lambda}(p_c) - \Lambda(p_c) \\
& = \int_{e_c(i) \geq p_c} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) - \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) < p_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \\
& = \int_{e_c(i) \geq p_c, e_{\hat{c}}(i) \geq p_{\hat{c}} \text{ for some } \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \\
& \leq \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) \geq \phi_{\hat{c}} \text{ for some } \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \\
& = \int_{e_c(i) \geq \phi_c} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) - \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \phi_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di).
\end{aligned}$$

Since $\sup \bar{\Lambda}' < 0$, it follows that

$$\begin{aligned}
& (\Phi_c - p_c) \\
& \leq \frac{1}{(-\sup \bar{\Lambda}')} \left(\int_{e_c(i) \geq \phi_c} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) - \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \phi_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \right)
\end{aligned}$$

Finally,

$$\begin{aligned}
R^{(1)}(c) & = \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) \\
& = \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) + \int_{p_c \leq e_c(i) \leq \Phi_c} F_i^{(1)}(c) \nu(di) \\
& \leq \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) \\
& + \frac{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (-\int_{e_c(i) \geq x} F_i^{(1)}(c) \nu(di))}{(-\sup \bar{\Lambda}')} \left(\int_{e_c(i) \geq \phi_c} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) - \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \phi_{\hat{c}} \forall \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \right)
\end{aligned}$$

where the inequality follows since

$$\int_{p_c \leq e_c(i) \leq \Phi_c} F_i^{(1)}(c) \nu(di) = \Lambda^1(p_c) - \Lambda^1(\Phi_c) \leq (\Phi_c - p_c) \sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (-\Lambda^1(x)).$$

It follows that

$$\begin{aligned}
R^{(1)}(c) &\leq \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) + \eta_c \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) > \phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(i) \setminus \{c\}} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di) \\
&\leq \int_{e_c(i) \geq \Phi_c} F_i^{(1)}(c) \nu(di) + \eta_c \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) > \phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di)
\end{aligned}$$

A.2 Lower Bound

Let p be the cutoff vector for a stable matching. Define

$$\hat{\Lambda}_c(x) = \Lambda_c^1(x) + \int_{e_c(i) \geq x, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di)$$

and note that

$$\hat{\Lambda}_c(p_c) \geq D_c(p) = k_c$$

Since $\Lambda_c^1(\phi_c) = k_c$,

$$\Lambda_c^1(\phi_c) \leq \hat{\Lambda}_c(p_c) \leq \hat{\Lambda}_c(\phi_c) + (p_c - \phi_c) \sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \hat{\Lambda}_c(x)$$

Rearranging terms,

$$p_c - \phi_c \leq \frac{-1}{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \hat{\Lambda}_c(x)} \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di)$$

Now,

$$\begin{aligned}
R^1(c) &= \int_{e_c(i) \geq p_c} F_i^{(1)}(c) \nu(di) \\
&= \int_{e_c(i) \geq \phi_c} F_i^{(1)}(c) \nu(di) - \int_{\phi_c \leq e_c(i) \leq p_c} F_i^{(1)}(c) \nu(di) \\
&\geq k_c - (p_c - \phi_c) \sup_{x \in [\phi_c, \Phi_c]} -\frac{d}{dx} \Lambda_c^1(x) \\
&\geq k_c - \frac{\sup_{x \in [\phi_c, \Phi_c]} -\frac{d}{dx} \Lambda_c^1(x)}{-\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \hat{\Lambda}_c(x)} \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^{(1)}(c)) \nu(di)
\end{aligned}$$

Note that

$$\begin{aligned}
R^1(c) &\geq k_c - \sum_{k=2}^N \sum_{\succ \in P_{k,c}} \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_c \quad \forall \hat{c} \succ c} F_i(\succ) \nu(di) \\
&\geq k_c - \sum_{k=2}^N \sum_{\succ \in P_{k,c}} \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_c \quad \text{some } \hat{c} \succ c} F_i(\succ) \nu(di) \\
&= k_c - \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_c \quad \text{some } \hat{c} \succ c} (\bar{F}_i(c) - F_i^1(c)) \nu(di),
\end{aligned}$$

Setting

$$\bar{\eta}_c = \min \left\{ 1, \frac{\sup_{x \in [\phi_c, \Phi_c]} -\frac{d}{dx} \Lambda_c^1(x)}{-\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} \hat{\Lambda}_c(x)} \right\}$$

it follows that

$$R^1(c) \geq k_c - \bar{\eta}_c \int_{e_c(i) \geq \phi_c, e_{\hat{c}}(i) < \Phi_c \quad \text{some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^1(c)) \nu(di)$$

B Proof of Theorem 2

We apply Theorem 1 for each priority scheme.

Multiple tie breaking: Recall that in this setting $I = [0, 1]^N$ and ν are N independent uniform distributions. First, we can specify $\Lambda_c^1(x)$ and $\bar{\Lambda}_c(x)$:

$$\Lambda_c^1(x) = \int_{e_c(i) \geq x} F_i^1(c) \nu(di) = \int \left[\int_x^1 F_i^1(c) du \right] \nu(di) = \int F_i^1(c) \nu(di) \int_x^1 du = F^1(c)(1 - x)$$

$$\bar{\Lambda}_c(x) = \int_{e_c(i) \geq x} \bar{F}_i(c) \nu(di) = \int \left[\int_x^1 \bar{F}_i(c) du \right] \nu(di) = \int \bar{F}_i(c) \nu(di) \int_x^1 du = \bar{F}_i(c)(1 - x)$$

where the first equality obviates the $N - 1$ integrals of measure 1. Therefore

$$\phi_c = 1 - \frac{k_c}{F^1(c)} \quad \Phi_c = 1 - \frac{k_c}{\bar{F}(c)}$$

Similarly,

$$\begin{aligned}
\Lambda_c(x) &= \Lambda_c^1(x) + \int_{e_c(i) \geq x, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ for some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^1(c)) \nu(di) \\
&= F^1(c)(1-x) + \int_{e_c(i) \geq x} (\bar{F}_i(c) - F_i^1(c)) \nu(di) - \int_{e_c(i) \geq x, e_{\hat{c}}(i) \geq \Phi_{\hat{c}} \forall \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^1(c)) \nu(di) \\
&= F^1(c)(1-x) + \int \left[\int_x^1 (\bar{F}_i(c) - F_i^1(c)) du \right] \nu(di) - \int \left[\int_x^1 \int_{\Phi_{\hat{c}}, \forall \hat{c} \neq c}^1 (\bar{F}_i(c) - F_i^1(c)) du \right] \nu(di) \\
&= F^1(c)(1-x) + (\bar{F}(c) - F^1(c))(1-x) - (\bar{F}(c) - F^1(c))(1-x) \prod_{\hat{c} \neq c} (1 - \Phi_{\hat{c}}) \\
&= F^1(c)(1-x) + (\bar{F}(c) - F^1(c))(1-x) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]
\end{aligned}$$

Having calculated, $\Lambda_c^1(x), \bar{\Lambda}_c(x), \Lambda_c(x)$, we have that

$$\begin{aligned}
\bar{\eta}_c &= \min \left\{ 1, \frac{\inf_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (\Lambda_c^1(x))}{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (\Lambda_c(x))} \right\} \\
&= \min \left\{ 1, \frac{\inf_{x \in [\phi_c, \Phi_c]} -F^1(c)}{\sup_{x \in [\phi_c, \Phi_c]} -F^1(c) - (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} \right\} \\
&= \frac{F^1(c)}{F^1(c) + (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]}
\end{aligned}$$

and

$$\begin{aligned}
\eta_c &= \min \frac{\inf_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (\Lambda_c^1(x))}{\sup_{x \in [\phi_c, \Phi_c]} \frac{d}{dx} (\bar{\Lambda}_c(x))} = \frac{\inf_{x \in [\phi_c, \Phi_c]} -F^1(c)}{\sup_{x \in [\phi_c, \Phi_c]} -\bar{F}(c)} \\
&= \frac{F^1(c)}{\bar{F}(c)}
\end{aligned}$$

Using the same logic to measure sets correctly, we can now calculate our bounds

$$\begin{aligned}
R^{(1)}(c) &\geq k_c - \frac{F^1(c)}{F^1(c) + (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} \frac{k_c}{F^1(c)} (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right] \\
&= k_c \left(\frac{F^1(c)}{F^1(c) + (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} \right)
\end{aligned}$$

and

$$\begin{aligned}
R^{(1)}(c) &\leq F^1(c) \frac{k_c}{\bar{F}(c)} + \frac{F^1(c)}{\bar{F}(c)} (\bar{F}(c) - F^1(c)) \frac{k_c}{F^1(c)} \left[1 - \prod_{\hat{c} \neq c} \left(1 - \frac{k_{\hat{c}}}{F^1(\hat{c})} \right) \right] \\
&= \frac{k_c}{\bar{F}(c)} \left[F^1(c) + \frac{\bar{F}(c) - F^1(c)}{\bar{F}(c)} \left[1 - \prod_{\hat{c} \neq c} \left(1 - \frac{k_{\hat{c}}}{F^1(\hat{c})} \right) \right] \right] \\
&= k_c \left[1 - \frac{\bar{F}(c) - F^1(c)}{\bar{F}(c)} \prod_{\hat{c} \neq c} \left(1 - \frac{k_{\hat{c}}}{F^1(\hat{c})} \right) \right]
\end{aligned}$$

Distance-based: In this setting we compute ϕ_c^{DB} . Then, we take our bounds from Theorem 1 and continue to bound

$$\begin{aligned}
R^{(1)}(c) &\geq k_c - \bar{\eta}_c \int_{e_c(i) \geq \phi_c^{DB}, e_{\hat{c}}(i) < \Phi_{\hat{c}} \text{ for some } \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^1(c)) \nu(di) \\
&\geq k_c - 1 \left[\int_{e_c(i) \geq \phi_c^{DB}} (\bar{F}_i(c) - F_i^1(c)) \nu(di) - \int_{e_c(i) \geq \phi_c^{DB}, e_{\hat{c}}(i) > \Phi_{\hat{c}} \forall \hat{c} \in \text{supp}(c)} (\bar{F}_i(c) - F_i^1(c)) \nu(di) \right] \\
&\geq k_c - \int_{e_c(i) \geq \phi_c^{DB}} (\bar{F}_i(c) - F_i^1(c)) \nu(di) \\
&= k_c - \int_{e_c(i) \geq \phi_c^{DB}} \frac{\bar{F}_i(c) - F_i^1(c)}{F^1(c)} F^1(c) \nu(di) \\
&\geq k_c - k_c \sup_{i: e_c(i) \geq \phi_c^{DB}} \frac{\bar{F}_i(c) - F_i^1(c)}{F^1(c)} \\
&= k_c \left(1 - \sup_{i: d(i, c) \leq 1 - \phi_c^{DB}} \frac{\bar{F}_i(c) - F_i^1(c)}{F^1(c)} \right)
\end{aligned}$$

and

$$\begin{aligned}
R^{(1)}(c) &\leq \int_{e_c(i) \geq \Phi_c^{DB}} F_i^1(c) \nu(di) + 0 \\
&= \int_{e_c(i) \geq \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} \bar{F}_i(c) \nu(di) \\
&\leq k_c \sup_{i: e_c(i) \geq \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} \\
&= k_c \sup_{i: d(i, c) \leq 1 - \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)}
\end{aligned}$$

Now that we have our bounds, we need to construct a condition such that the lower bound for multiple tie breaking is larger than the upper bound for distance-based priorities:

$$\begin{aligned}
\sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} &\leq k_c \left(\frac{F^1(c)}{F^1(c) + (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} \right) \\
\sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} - \frac{F^1(c)}{\bar{F}(c)} &\leq \frac{F^1(c)}{F^1(c) + (\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} - \frac{F^1(c)}{\bar{F}(c)} \\
\sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} - \frac{F^1(c)}{\bar{F}(c)} &\leq \frac{F^1(c)\bar{F}(c) - F^1(c)F^1(c) - F^1(c)(\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]}{\bar{F}(c)F^1(c) + \bar{F}(c)(\bar{F}(c) - F^1(c)) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]} \\
\sup_{d(i,c) \leq 1 - \Phi_c^{DB}} \frac{F_i^1(c)}{\bar{F}_i(c)} - \frac{F^1(c)}{\bar{F}(c)} &\leq \frac{\left(1 - \frac{F^1(c)}{\bar{F}(c)} \right) \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})}}{1 + \left(\frac{\bar{F}(c)}{F^1(c)} - 1 \right) \left[1 - \prod_{\hat{c} \neq c} \frac{k_{\hat{c}}}{\bar{F}(\hat{c})} \right]}
\end{aligned}$$

which proves the result.

B.1 Students assigned to top schools and Pareto efficiency

Lemma 2. *a. For any matching μ , $P_2 \leq P$ and*

$$\sum_{c=1}^N R^{(1)}(c) + P \leq \sum_{c=1}^N k_c. \tag{B.1}$$

- b. Assume that $\mathbb{P}[c \succ c' \mid i] > 0$ for all $c \neq c'$ and all $i \in I$. When priorities are built from multiple tie breaking and μ is stable, then $P_2 = P$ and (B.1) holds with equality.*
- c. Assume that $\mathbb{P}[c \succ c' \mid i] > 0$ for all $c \neq c'$ and all $i \in I$. Assume priorities are distance-based, and that conditions (4.8) and (4.3) hold. Then $P_2 = P$ and (B.1) holds with equality.*

Proof of Lemma 2. Let μ be any matching and take S^1 as the set of all students assigned to their top schools. Clearly, no subset of S^1 can be Pareto improved. Let

$$S^P = \bigcup_{S'' \text{ can be Pareto improved}} S''$$

and note that $S^P \subseteq \{s \mid s \text{ is assigned by } \mu\}$. It follows that

$$S^1 \cup S^P \subseteq \{s \mid s \text{ is assigned}\} \text{ and } S^1 \cap S^P = \emptyset.$$

As a result,

$$\bar{\nu}(S^1) + \bar{\nu}(S^P) \leq \bar{\nu}(\{s \text{ is assigned}\}).$$

Since $\bar{\nu}(S^1) = \sum_c R^{(1)}(c)$, $\bar{\nu}(S^P) = P$ and $\bar{\nu}(\{s \text{ is assigned}\}) \leq \sum_c k_c$, we deduce that

$$\sum_c R^{(1)}(c) + P \leq \sum_c k_c.$$

Take now a stable matching under multiple tie breaking. Consider any student s who is assigned to a school that is not her top choice. We will argue that there is a positive measure set S' , that contains s , such that S' is part of Pareto-improving pairs. Let $c = \mu(s)$ and consider a school \hat{c} and a set of students \hat{S} assigned to c such that \hat{S} has positive measure and contains s , and all students in \hat{S} prefer \hat{c} over c . Consider the set of all students who prefer c over \hat{s} but only have scores to get admission to \hat{c} :

$$\bar{S} = \{s \in S \mid c \succ_s \hat{c}, \quad i_{\hat{c}} \geq p_{\hat{c}}, \quad p_{c'} > i_{c'} \forall c' \neq \hat{c}\}.$$

Clearly,

$$\bar{\nu}(\bar{S}) = \left(\int \mathbb{P}[c \succ \hat{c} \mid i] \nu(di) \right) (1 - p_{\hat{c}}) \prod_{c' \neq \hat{c}} p_{c'} > 0$$

Without loss, assume that $\bar{\nu}(\bar{S}) = \bar{\nu}(\hat{S})$.²⁰ Construct the matching $\bar{\mu}$ by $\bar{\mu}(c') = \mu(c')$ for all $c' \in C \setminus \{c, \hat{c}\}$ and

$$\bar{\mu}(c) = (\mu(c) \cup \bar{S} \setminus \hat{S}) \quad \text{and} \quad \bar{\mu}(\hat{c}) = (\mu(\hat{c}) \cup \hat{S} \setminus \bar{S}).$$

It follows that $\bar{\mu}$ is a matching and $S' = \bar{S} \cup \hat{S}$ is part of Pareto-improving pairs. As a result,

$$\{s \text{ is assigned to a school that is not her top}\} \leq \bigcup_{\bar{S} \text{ is part of Pareto-improving pairs}} \bar{S}$$

and since

$$\bar{\nu}(\{s \text{ is assigned to a school that is not her top}\}) = \sum_c (k_c - R^{(1)}(c))$$

²⁰If not, scale down the set with the largest measure so that the measures coincide.

it follows that

$$\sum_{c=1}^N (k_c - R^{(1)}(c)) \leq \bar{\nu} \left(\bigcup_{\tilde{S} \text{ is part of Pareto-improving pairs}} \tilde{S} \right) = P_2.$$

We deduce that under multiple-tie breaking, $P = P_2$ and $\sum_{c=1}^N (k_c - R^{(1)}(c)) = P_2$. The proof for distance-based priorities is analogous. \square

The following example shows that under distance-based priorities, inequality (4.9) can be strict when condition (4.3) does not hold.

Example 7. Suppose that $N = 2$ and $I = [0, 1]$. School c_1 has capacity k_1 and is located in $3/4$, while school c_2 has capacity k_2 and is located in 1 . Students find both schools acceptable and for each i , a fraction $1/2$ of students prefer c_1 over c_2 .

Under distance-based priorities, we characterize a stable matching such that all students with score above the cutoff p_1 also have score above p_2 for school c_2 :

$$\int_{|i-3/4| < 1-p_1} \frac{1}{2} di = k_1 \quad \text{and} \quad \int_{i > p_2} di - \int_{|i-3/4| < 1-p_1} \frac{1}{2} di = k_2.$$

The first condition is the market clearing condition for school c_1 : the demand for school 1 is given by half of the student living within distance p_1 of the schools. The second condition is the market clearing condition for school c_2 : the demand for school 2 is given by all students living with distance p_2 of schools 2 minus the fraction of students that get admission to school 1. We can solve for the cutoffs:

$$(1 - p_1) = k_1 \quad 1 - p_2 = k_1 + k_2$$

with $3/4 - k_1 > 1 - (k_1 + k_2)$ (so that students that can be accepted to c_1 can also be accepted to c_2). For $k_2 > 1/4$ and $k_1 < k_2$,

$$R_{DB}^{(1)}(c_1) = k_1 \quad \text{and} \quad R_{DB}^{(1)}(c_2) = \frac{(k_2 - k_1)}{2}.$$

The matching is Pareto-efficient and $P = 0$, but

$$\sum_c R^{(1)}(c) + P < \sum_c k_c.$$

References

- Abdulkadiroğlu, A., Agarwal, N., and Pathak, P. A. (2017). The welfare effects of coordinated assignment: Evidence from the new york city high school match. American Economic Review, 107(12):3635–3689.
- Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifference: Redesigning the nyc high school match. American Economic Review, 99(5):1954–78.
- Abdulkadiroğlu, A., Agarwal, N., and Pathak, P. A. (2017). The welfare effects of coordinated assignment: Evidence from the new york city high school match. American Economic Review, 107:3635–3689.
- Abdulkadiroğlu, A., Che, Y.-K., and Yasuda, Y. (2015). Expanding ”choice” in school choice. American Economic Journal: Microeconomics, 7:1–42.
- Abdulkadiroğlu, A., Pathak, P., and Roth, A. (2009). Strategy-proofness versus efficiency in matching with indifference: Redesigning the nyc high school match. American Economic Review, 99:1974–1978.
- Abdulkadiroğlu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. American Economic Review, 93:729–747.
- Allman, M., Ashlagi, I., and Nikzad, A. (2022). On rank dominance of tie-breaking rules. Theoretical Economics.
- Arnosti, N. (2022). Lottery design for school choice. Management Science.
- Ashlagi, I. and Nikzad, A. (2020). What matters in school choice tie-breaking? how competition guides design. Journal of Economic Theory, 190:105–120.
- Azevedo, E. and Leshno, J. (2016). A supply and demand framework for two-sided matching markets. Journal of Political Economy, 124.
- Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: student placement. Journal of Economic theory, 84(1):73–94.

- Burgess, S., Cantillon, E., Cavallo, M., Greaves, E., and Zhang, M. (2023). School admissions in england: The rules schools choose on which pupils to admit. Technical report, ULB–Universite Libre de Bruxelles.
- Cantillon, E., Chen, L., and Pereyra, J. S. (2022). Respecting priorities versus respecting preferences in school choice: When is there a trade-off? arXiv preprint arXiv:2212.02881.
- Che, Y.-K. and Tercieux, O. (2018). Efficiency and stability in large matching markets. Journal of Political Economy, 127:0–12.
- Dur, U., Kominers, S. D., Pathak, P. A., and Sönmez, T. (2018). Reserve design: Unintended consequences and the demise of boston’s walk zones. Journal of Political Economy, 126(6):2457–2479.
- Echenique, F., Gonzalez, R., Wilson, A., and Yariv, L. (2020). Top of the batch: Interviews and the match. American Economic Review: Insights.
- Ehlers, L. and Erdil, A. (2010). Efficient assignment respecting priorities. Journal of Economic Theory, 145(3):1269–1282.
- Ehlers, L. and Morrill, T. (2020). (il)legal assignments in school choice. The Review of Economic Studies, 87(4):1837–1875.
- Erdil, A. and Ergin, V. (2008). What’s the matter with tie-breaking? improving efficiency in school choice. American Economic Review, 98:669–689.
- Ergin, H. I. (2002). Efficient Resource Allocation on the Basis of Priorities. Econometrica, 70(6):2489–2497.
- Gale, D. and Shapley, L. (1962). College admissions and the stability of marriage. The American Mathematical Monthly, 69:9–15.
- Kesten, O. (2010). School choice with consent. The Quarterly Journal of Economics, 125(3):1297–1348.
- Pakzad-Hurson, B. (2023). Stable and efficient resource allocation with contracts. American Economic Journal: Microeconomics, 15(2):627–659.
- Reny, P. J. (2021). A simple sufficient condition for a unique and student-efficient stable matching in the college admissions problem. Economic Theory Bulletin, 9(1):7–9.

- Reny, P. J. (2022). Efficient matching in the school choice problem. American Economic Review, 112(6):2025–43.
- Roth, A. E. (1982). The economics of matching: Stability and incentives. Mathematics of Operations Research, 7:617–628.
- Salonen, H. and Salonen, M. A. (2018). Mutually best matches. Mathematical Social Sciences, 91:42–50.
- Shapley, L. and Scarf, H. (1974). On cores and indivisibility. Journal of Mathematical Economics, 1:23–37.
- Shi, P. (2022). Optimal priority-based allocation mechanisms. Management Science, 68(1):171–188.
- Çelebi, O. and Flynn, J. P. (2021). Priority Design in Centralized Matching Markets. The Review of Economic Studies, 89(3):1245–1277.