# Assigning more students to their top choices: A comparison of tie-breaking rules

Itai Ashlagi Afshin Nikzad Assaf Romm *

**Abstract**

School districts that implement stable matchings face various decisions that affect students' assignments to schools. We study the properties of the rank distribution of students with random preferences when schools use different tie-breaking rules to rank equivalent students. Under a single tie-breaking rule, where all schools use the same ranking, a constant fraction of students are assigned to one of their top choices. In contrast, under a multiple tie-breaking rule, where each school independently ranks students, a vanishing fraction of students are matched with one of their top choices. However, if students are allowed to submit only relatively short preference lists under a multiple tie-breaking rule, a constant fraction of students will be matched with one of their top choices, while only a "small" fraction of students will remain unmatched.

**Keywords.** school choice; tie-breaking rule; deferred acceptance; stable matching

*Ashlagi: MS&E department, Stanford University, iashlagi@stanford.edu. Nikzad: Department of Economics, UC Berkeley, nikzad@berkeley.edu. Romm: Department of Economics, Hebrew University of Jerusalem, and Microsoft Research, Israel, assaf.romm@mail.huji.ac.il. We thank Sergiu Hart, Avinatan Hassidim, Déborah Marciano, Parag Pathak, and Alvin Roth for helpful discussions. We also thank the editors and two referees in GEB for very useful comments. Ashlagi acknowledges the research support of National Science Foundation grant SES-1254768. Romm acknowledges the research support of the Maurice Falk Institute, Israel Science Foundation grant 1780/16, and the U.S.-Israel Binational Science Foundation grant.

# 1 Introduction

Many school districts have adopted student assignment mechanisms based on the student-proposing deferred acceptance (DA) algorithm (Gale and Shapley, 1962).[1] This algorithm assigns students to schools based on students' preferences and schools' priorities, which are used to ration seats in over-demanded schools. However, it is often the case that large groups of students receive the same priority, and the manner in which ties are resolved has consequences for students' assignments (Abdulkadiroğlu and Sönmez, 2003; Erdil and Ergin, 2008).

Policymakers frequently consider two natural tie-breaking rules. The first is the *multiple tie-breaking rule* (MTB), under which every school independently assigns each applicant a random (lottery) number that is used to resolve ties. The other is the *single tie-breaking rule* (STB), which assigns each student a single random number that is used by all schools to resolve ties. MTB seems fairer to many observers, as students with bad draws at some schools may still have good chances at other schools, but it may be unnecessarily inefficient (Abdulkadiroğlu and Sönmez, 2003).

Numerical comparisons between MTB and STB using data from New York City (Abdulkadiroğlu et al., 2009) and the city of Amsterdam (De Haan et al., 2015) reveal similar patterns. STB assigns more students to one of their top choices, but MTB assigns fewer students to their lower-rank choices and leaves fewer students unassigned. These cities made different decisions. Amsterdam first adopted MTB for equity reasons.[2][3] In New York City policymakers also initially leaned toward MTB, but as Abdulkadiroğlu et al. (2009) write: "the greater number of students obtaining one of their top choices[...]convinced New York City to employ a single tiebreaker in their assignment system." We note that MTB has also been adopted recently in school districts in Chile.[4]

In this paper we are mainly concerned with the fraction of students who are assigned to one of their top choices. This measure is often reported by the media and administrators. For instance: "just 66% of pupils in the greater London region received an offer from their

[1] See also Abdulkadiroğlu and Sönmez (2003). Examples include the city of Boston (Abdulkadiroğlu et al., 2005) and New York City (Abdulkadiroğlu et al., 2009).

[2] See https://www.nemokennislink.nl/publicaties/schoolstrijd-in-amsterdam/ (Schoolstrijd in Amsterdam)(Arnout Jaspers, Kennislink, July 1, 2015).

[3] Following a lawsuit by families who wished to switch assignments, the city of Amsterdam moved to STB in the second year of operation.

[4] From personal communication.

first preference school," [5] and, "city officials said less than 60 percent of students were offered a seat at their top choice, but 85 percent were offered a seat at one of their three school choices."[6] We take here a first analytical approach to comparing STB with MTB with respect to this measure, which explains the qualitative insights described above.

We consider a stylized two-sided matching model with students and schools, in which students have randomly drawn preferences over all schools, and schools have fixed capacities. To keep things simple, all students in our model belong to a single priority class, meaning that they all have a priori equal access to every school. We assume students are assigned using the student-proposing DA, and ties are resolved using either STB or MTB. We begin by analyzing the model assuming students submit their complete preference lists and later we discuss the implications of limiting the length of students' preference lists (as is done in several school districts).

We are interested in analyzing how many students are likely to be assigned to one of their "top" choices under STB and MTB as the numbers of students and schools grow large. Unless otherwise specified, by a student's top choice we refer to one of her $k$ most preferred schools for some constant $k$ that is independent of the number of schools. This notion is meant to capture the case in which students receive one of their very top choices in a large market. We note that this definition, which requires a top choice to be held constant as the market grows large, may seem too restrictive. However, with $n$ students and $n$ schools, both STB and MTB result in students being assigned, on average, to one of their top $\Theta(\ln n)$ choices as the market grows large (Knuth, 1976; Pittel, 1989)).[7] Therefore, it is interesting to consider a top choice to be a quantity that is significantly smaller than $\Theta(\ln n)$.[8]

We show that under STB, a constant fraction of students are assigned to their first choice. In sharp contrast, we show that under MTB students are *unlikely* to be assigned to one of their top choices. Formally, for any constant $k$ and as the market grows large, with high probability only a vanishing fraction of students are matched to one of their $k$ most preferred schools.[9] The reason that STB assigns many students to a top choice is that the mechanism

[5] "One-third of London families miss out on first choice of secondary school" (Richard Adams and Emily Hawkins, *The Guardian*, March 1, 2018).

[6] "This is how the D.C. school lottery is supposed to work" (Alejandra Matos, *The Washington Post*, May 17, 2017).

[7] Under MTB, with high probability all students are assigned to one of their top $O(\ln^2 n)$ choices.

[8] The significance of our definition of top choices is also demonstrated in the context of a simple cardinal utility model in Section E.1, where a student's utility from being assigned to her $i$-th choice is $1/i$.

[9] We make a similar observation in Section E.1 in the context of (simple) cardinal utilities.

is equivalent to students choosing schools in random order (random serial dictatorship). The following provides some intuition for the contrast between STB and MTB. Let $s$ be an arbitrary student, who has been temporarily assigned to her top choice by the end of the first round of DA. Since $s$ was assigned to her top choice, say, school $c$, she is likely to have a "good" lottery number at $c$. Consider future contenders for a seat at school $c$ who were rejected by other schools. Under STB, these students are likely to have "worse" lottery numbers than $s$, due to their being already rejected elsewhere. Under MTB, however, being rejected from a different school reveals no information about her (new) lottery number at $c$. These renewed lottery numbers generate long sequences of rejections, causing many students to be assigned to lower choices.

One way to increase the number of students who are assigned to a top choice is bounding the length of the preference list each student can submit, thus reducing competition.[10] The drawback of shortened preference lists is that more students may remain unassigned. Since few students are assigned to top choices under MTB with long preferences lists, we ask how does shortening preference lists affect students' ranks. We show that, under MTB, the number of unassigned students vanishes rapidly as the length of the preference list increases (e.g., we prove that the rate is exponential when there is a linear surplus or shortage of seats). This implies that under MTB, one can shorten preference lists such that a constant fraction of students are assigned to a top choice without leaving too many students unassigned.

Policymakers often find the MTB rule attractive because of its apparent fairness (as discussed earlier). Our results suggest that policymakers who find MTB more attractive may wish to consider shortening the preference lists as a way to improve the assignments of students to top choices. This gives a new argument for bounding the length of preference lists, especially under MTB.[11]

Restricting the length of preference lists is indeed practiced in several cities (e.g., 12 schools in NYC, 5 schools in Denver, 10 in Amsterdam). We remark that for the case of short preference lists, empirical (Abdulkadiroğlu et al., 2009; De Haan et al., 2015) and theoretical (Arnosti, 2015) papers find that STB assigns more students to top choices and leaves more students unassigned than MTB does.

[10]For some intuition, note that letting students submit only 1 school maximizes the number of students assigned to their first choice under both tie-breaking rules.

[11]Shortening preference lists may increase students' welfare through equilibrium behavior (Abdulkadiroğlu et al., 2011). This is not the case in our model, which assumes that preferences are drawn uniformly at random and capacities are identical (and thus agents have no reason to behave strategically).

From a technical standpoint, one novelty in the proofs is the explicit use of (nontrivial) capacities. Most papers that study random matching markets assume either explicitly or implicitly that the market contains large imbalances (Arnosti, 2015; Ashlagi et al., 2014; Kojima and Pathak, 2009), which leads to an analysis similar to that of one-to-one matching markets.[12] Our method allows us to extend the arguments to capacities that are not held constant as the market grows large. The main idea is to simplify the random process that corresponds to the deferred acceptance algorithm by coupling it with a simpler random process.

## 1.1 Related work

Several papers study the trade-offs between STB and MTB. Abdulkadiroğlu et al. (2009) and De Haan et al. (2015) run numerical experiments using NYC and Amsterdam school choice data, respectively, and find that STB results in many more students being assigned to one of their top choices and, furthermore, that students' rank distributions under STB and MTB cross each other once.

Independent of this work, Arnosti (2015) studies a cardinal utility model, in which the length of preference lists is held constant. He finds that the rank distributions under MTB and STB cross each other once (STB assigns more students to top choices than MTB but also assigns more students to lower choices than MTB does). He quantifies the number of students that remain unassigned under both tie-breaking rules. Our work differs in two ways. First, Arnosti's model assumes the length of preference lists is held constant as the number of schools grows large, whereas we analyze markets in which the length of preference lists grows large. (Our finding that only a vanishing fraction of students receive one of their top choices under MTB holds even when the length of students' preference lists grows at a slow rate as the market grows large.) Second, Arnosti's model assumes that the imbalance between seats and students is of the order of the number of agents, whereas our model allows also for a small additive imbalance.

Abdulkadiroğlu et al. (2015) analyze a model with a continuum of students and a finite number of schools, each with a mass of seats, in order to compare the two tie-breaking rules. They establish the existence of a unique set of cutoffs that clear the market under each

[12]This is not to say that capacities do not significantly affect these papers' approaches and analyses (see, for example, Kojima and Pathak, 2009).

tie-breaking rule and find that STB is ordinally efficient.[13]

The trade-off between incentives and efficiency when preferences contain indifferences has led to several novel tie-breaking approaches, among which are stable improvement cycles (Erdil and Ergin, 2008), efficiency-adjusted DA (Kesten, 2011), and choice-augmented DA (Abdulkadiroğlu et al., 2015). Another closely related paper is Che and Tercieux (2015), who propose a DA-like mechanism based on a "circuit breaker." Their algorithm does not allow students who are preferred by a certain school to push out other students who rank that school much higher than the preferred students do, thus improving the rank distribution at the expense of a few students' priorities. Their approach, which can be viewed as a dynamic truncation of preference lists (and is reminiscent of the Chinese Parallel, described by Chen and Kesten, 2013), aims to limit the inefficiency resulting from a lot of competition.

The effect of requiring students to submit short preference lists has been described and studied by Haeringer and Klijn (2009) and experimentally tested by Calsamiglia et al. (2010). However, the idea of using truncation and dropping strategies dates back even further; see, for example, Roth and Rothblum (1999) and Kojima and Pathak (2009).[14]

We are not the first to study properties of the rank distribution in two-sided random matching markets. Pittel (1989, 1992) studied the men's average rank in a balanced stable marriage model with random preferences and showed that under men-optimal stable matching, the men's average rank is approximately $\ln n$ and, with high probability, all men are assigned to one of their top $ln^2n$ preferred women. Ashlagi et al. (2016) showed that if there are fewer men than women, men are matched on average with one of their top $\ln n$ choices under any stable matching. Our main result here, which shows that few students obtain one of their top preferences under MTB, is not implied by Ashlagi et al. (2016). In followup work Ashlagi and Nikzad (2015) show that, under complete preferences lists, when there is a shortage of seats, the rank distribution under STB almost stochastically dominates the rank distribution under MTB.

Finally, a few studies compare tie-breaking rules under the top trading cycles (TTC) algorithm, which find Pareto efficient assignments. Both Pathak and Sethuraman (2011) and Carroll (2014) extend the results of Abdulkadiroğlu and Sönmez (1998) and show that there essentially no difference between a single tie-breaking rule (i.e., random serial dictatorship)

[13]See also Che and Kojima (2010), Liu and Pycia (2012), and Ashlagi and Shi (2014), who calculate cutoffs under STB in markets with a continuum of students.

[14]See also Coles and Shorrer (2014), who study truncation in random markets, and Gonczarowski (2014) for an algebraic approach.

and a multiple tie-breaking rule (TTC with random endowments).

## 2 The model

In a school choice problem there is a set of $n$ students, each of whom can be assigned to one seat at one of $m$ schools. Denote the set of students by $\mathcal{S}$ and the set of schools by $\mathcal{C}$, each of which has a fixed capacity.

Every student has a strict preference ranking over all schools. Students prefer being assigned to any school over remaining unassigned (that is, all schools are acceptable). The *rank* of a school $c$ in a preference list of student $s$ is the number of schools that $s$ weakly prefers to $c$ (so $s$'s most preferred school has rank 1). Unless otherwise specified, students' preferences over schools are drawn independently and uniformly at random. Each school $c \in \mathcal{C}$ has a strict priority ranking over students that is used to break ties between students.

A *matching* is a mapping from students to schools such that each student is assigned at most one school and the number of students assigned to any given school is at most the capacity of that school. A matching is *unstable* if there is a student $s$ and a school $c$ such that $s$ prefers to be assigned to $c$ over her current assignment and $c$ either has a vacant seat or an assigned student with a lower priority ranking than $s$. A matching is *stable* if it is not unstable.

We consider two common tie-breaking rules that school districts use to determine schools' priority rankings over students. Under the *multiple tie-breaking rule* (MTB) each school independently selects a priority ranking over all students uniformly at random. Under the *single tie-breaking rule* (STB) all schools use the same priority ranking order, which is selected a priori uniformly at random.[15] We study the rank distribution of students in the assignments generated by the student-proposing deferred acceptance (DA) algorithm when schools' priority rankings are generated by each of these tie-breaking rules.[16] For brevity we refer to the assignments as the outcomes under MTB and STB.

Our model considers school choice problems with many (small) schools, each of which has the the same capacity $\overline{q}$, which is a constant independent of $n$ and $m$. We further denote by $q = \overline{q}m$ the total number of seats.

[15] One way to implement STB is to assign each student a lottery number drawn independently and uniformly at random from $[0, 1]$. MTB can be implemented similarly by using different draws (and thus different lottery numbers) for each school.

[16] See Shapley (1955); Abdulkadiroğlu and Sönmez (2003) for the DA algorithm.

## 2.1 Asymptotic notations

This paper uses asymptotic notations, which we define here for clarity.

Given two nonnegative functions $f, g : \mathbb{N} \to \mathbb{R}_+$, we say that $f(n) = O(g(n))$ if there exists $n_0$ and $M > 0$ such that $f(n) \leq Mg(n)$ for all $n \geq n_0$. We say that $f(n) = o(g(n))$ if for any $\epsilon > 0$, there exists $n_0$ such that $f(n) \leq \epsilon g(n)$ for all $n \geq n_0$. We say that $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$.

Given a sequence of events $\{E_n\}$, we say that the sequence occurs *with high probability (w.h.p.)* if $\lim_{n\to\infty} \frac{1-\mathbb{P}[E_n]}{n^{-\lambda}} = 0$, for some constant $\lambda > 0$. We also say that this sequence occurs *with very high probability (w.v.h.p.)* if $\lim_{n\to\infty} \frac{1-\mathbb{P}[E_n]}{e^{-n^{\lambda}}} = 0$, for some constant $\lambda > 0$.

# 3 Students' assignment to top choices

**Main Theorem.** *Suppose $|n - q| \leq d$ for some constant $d \geq 0$. Then, for any constant $k$, as $n$ approaches infinity, the expected fraction of students who are matched to one of their top $k$ choices under the student-optimal stable matching approaches zero under MTB, but approaches a positive constant under STB.*

We discuss the robustness of the theorem in the following subsection.

The main theorem implies that MTB is not an attractive tiebreaker when it comes to assigning students to top choices. Nevertheless, some policymakers find MTB attractive because of its apparent fairness (as discussed in the introduction). One natural way to increase the number of students assigned to top choices under MTB is imposing a limit, $k$, on the number of schools that students can rank (so assigned students obtain one of their top $k$ choices). The drawback is leaving more students unassigned. In Section 3.2, we show that the additional number of unassigned students vanishes rapidly as their preferences lists grow longer. In Section 5.2 we will further observe that MTB leaves fewer students unassigned than STB does when preference lists are short.

## 3.1 Discussion of the main theorem

The proof for the main theorem has two parts: our claim about MTB, which is proved by Proposition 3.1, and our claim about STB, which is proved by Proposition 3.2. Denote by $R_k(\pi)$ the expected fraction of students who get one of their top $k$ choices under student-optimal stable matching when the tie-breaking rule $\pi$ is used.

**Proposition 3.1.** *Suppose $|n - q| \leq d$ for some constant $d \geq 0$. Then, for any constant $k$, $R_k(MTB)$ approaches $0$ as $n$ approaches infinity.*

Proposition 3.1 shows that very few students receive one of their top choices under MTB as the market grows large even when students are on the short side of the market. We note that the proposition is true even when $d = n^b$, where $b$ is a constant less than 1.[17] Moreover, the proposition remains true even when $k = k(n)$ is not held constant but grows sufficiently slow.

**Proposition 3.2.** *Let $t = \min\{n, q\}$ be the total number of assigned students under STB. Then: (i) the expected number of students assigned to their first choice under STB is at least $t/2$. Moreover, (ii) the total number of students assigned to their first choice under STB is w.h.p. at least $(1 - \varepsilon)(t/2)$ for any $\varepsilon > 0$.*[18]

It follows directly from Proposition 3.2 that at least half the assigned students are assigned to their top choice in expectation.

A few comments are appropriate:

1. The main theorem also holds true when preference lists are short, as long as they grow at a nonzero rate. That is, preference lists are of length $f(m)$ where $f$ approaches infinity as $m$ approaches infinity.

2. Ashlagi et al. (2016) compute the average rank of students in a one-to-one large random matching market. They find that when students are on the short side ($n \leq q$), they obtain on average roughly their $\ln n$-ranked and when students are on the long side ($n > q$), they obtain on average at least their $\frac{n}{\ln n}$-ranked school. Neither result implies Proposition 3.1, which shows that a vanishing fraction of students obtain one of their top $k$ choices for any constant $k$ ($k$ is independent of $n$). Hence, our result is especially surprising when students are on the short side of the market. Moreover, our analysis extends to many-to-one markets with constant capacities.

3. Proposition 3.1 applies when there is up to a small surplus of seats. When the surplus of seats is sufficiently large, many students will obtain one of their top choices. Ashlagi et al. (2016) analyze a one-to-one matching market, in which the number of students

[17] The same proof goes through, with minor modifications (merely different choices of constants).
[18] For the definition of w.h.p. see Section 2.1.

is a fraction of the number of seats ($n = \lambda qm$ for some $\lambda < 1$ and $\overline{q} = 1$) and find that students will obtain, on average, a constant rank (so a constant fraction of students will be assigned to one of their top choices).

A few more words are appropriate regarding the assumption that preferences are drawn uniformly at random. Our analysis still holds when there is small alignment in students' preferences, allowing for schools to differ slightly in their popularity, resulting in the same qualitative insights. As students' preferences become more (positively) correlated, the rank distributions under STB and MTB become more similar; at the extreme of fully aligned preferences, the rank distributions are identical.

The proof of Proposition 3.1, given in Appendix A, is based on the following steps. Fix some arbitrary school $c$. Denote by $X_c$ the number of students assigned to $c$, for whom $c$ is a top $k$ choice. By symmetry and linearity of expectation, the expected fraction of students who get one of their top $k$ choices under MTB is $R_k(MTB) = m\frac{\mathbb{E}[X_c]}{n}$. This equality says the fraction of students who get one of their top $k$ choices in expectation (over students' and schools' preferences) equals the ratio between the number of schools and students multiplied by the number of students who get one of their top choices at any fixed school $c$. We show that school $c$ receives many more applications than the number of students for whom $c$ is a top $k$ choice. Formally, $c$ receives at least $\Omega(\ln n)$ proposals with high probability. Moreover, with high probability the number of students for whom $c$ is a top $k$ choice, denoted by $\psi(k)$, is very small (sub-logarithmic in $n$). Therefore, since each student who applies to $c$ is admitted with a probability of at most $O(\overline{q}/\ln n)$, $\mathbb{E}[X_c] \leq O\left(\frac{\overline{q}\,\psi(k)}{\ln n}\right)$, implying that $\mathbb{E}[X_c]$ approaches 0 as $n$ approaches infinity.

The proof for Proposition 3.2 essentially shows that when preferences are drawn independently and uniformly at random, then in the random serial dictatorship mechanism, a constant fraction of the assigned students are likely to obtain their top choice.

## 3.2 MTB with short preference lists

School districts often limit the number of schools students can rank. When students do not apply to all schools they find acceptable, competition is reduced, which is a way to resolve the issues that our main theorem exposes. Of course, shortening the lists may also increase the number of unassigned students. The next theorem quantifies this trade-off and shows that the number of unassigned students resulting from shortening the lists vanishes rapidly as the length of the preference list increases. (In particular, the rate is exponential when there is a linear surplus or shortage of seats.)

**Theorem 3.3.** *Let $U_k$ denote the number of unassigned students given that students submit only their $k$ most preferred schools under the MTB rule. Also, let $\underline{q} = \min(q, n)$.*

1. *If $n = (1+\varepsilon)q$ for some constant $\varepsilon$ ($\varepsilon$ can be negative), then there exists a random variable $\Delta$ such that $U_k \le \max(\varepsilon q, 0) + \Delta$, where*
   (a) $\mathbb{E}[\Delta] \le e^{-|\varepsilon|k}\underline{q}$,
   (b) *$\Delta$ is not much larger than $\mu = e^{-|\varepsilon|k}\underline{q}$, in the following sense (Chernoff bounds):*

$$\mathbb{P}[\Delta > \mu(1+\delta)] \le e^{\frac{-\delta^2\mu}{3}} \qquad \forall\, 0 < \delta < 1$$
$$\mathbb{P}[\Delta > \delta\mu] \le \left(\frac{e^{\delta-1}}{\delta^\delta}\right)^\mu \qquad \forall\, \delta > 1.$$

2. *If $n = q + d$ with $|d| = o(q)$ ($d$ can be negative), then there exists a random variable $\Delta$ such that $U_k < \max(d, 0) + \Delta$, where:*
   (a) $\mathbb{E}[\Delta] \le 2\underline{q}/k$ *for all* $k \le m$.
   (b) *w.v.h.p.* $\Delta \le (2+\varepsilon')\underline{q}/k$, *for all* $\varepsilon, \varepsilon' > 0$ *and* $k \le q^{1/3-\varepsilon}$.

When preference lists are long, very few students are assigned to one of their top choices under MTB. Theorem 3.3 suggests that shortening the lists provides a remedy. Furthermore, when there is an excess of seats, the average rank of assigned students under MTB is essentially no better than $\frac{n}{\ln n}$, as discussed in the next section. Shortening the lists can then significantly improve the average rank of assigned students, while ensuring that the number of assigned students does not decrease significantly or does not decrease at all, as shown in the following corollary.

**Corollary 3.4.** *Suppose $n = (1+\varepsilon)q$, and let $k = \frac{1+\delta}{\varepsilon} \cdot \ln q$, where $\varepsilon$ and $\delta$ are positive constants. Then w.h.p. $\Delta = 0$, where $\Delta$ is defined as in Theorem 3.3.*

*Proof.* By Theorem 3.3 we have $\mathbb{E}[\Delta] \le e^{-(1+\delta)\ln q} q = q^{-\delta}$. So, $\mathbb{P}[\Delta > 1] \le q^{-\delta}$. □

Observe that when preference lists are short, DA is not strategyproof. However, since preferences are drawn uniformly at random in our model, it is safe to abstract away from strategic decisions, and in equilibrium each student ranks her top $k$ choices. In Section 5.2 we provide further simulation results that quantify the number of unassigned students when students' preferences are drawn from a multinomial logit discrete choice model.

# 4 Other properties of the rank distributions

Although this paper is mainly concerned with the assignment of students to one of their top-ranked schools, we also briefly discuss another measure for the quality of students' matches: the average rank of assigned students. Here, we quantify the average rank when preference lists are unrestricted. Under MTB, the average rank is essentially computed in Ashlagi et al. (2016), which we briefly summarize here. Denote by $Avg(\pi)$ the matched students' average ranking of schools under the student-optimal stable matching when the tie-breaking rule $\pi$ is used.

**Proposition 4.1** (Ashlagi et al. (2016))**.** *Suppose* $\overline{q} = 1$*,* $n = q + d$ *for some (possibly negative) constant* $d$*, and fix any* $\epsilon > 0$*. If* $d > 0$*, the probability that* $Avg(MTB)$ *is at least* $(1-\epsilon)\frac{n}{\ln n}$ *converges to* $1$ *as* $n$ *grows large. If* $d \leq 0$*, the probability that* $Avg(MTB)$ *is at most* $(1+\epsilon)\ln n$ *converges to* $1$ *as* $n$ *grows large.*

Therefore, shortening the lists can significantly improve the average rank of assigned students under MTB, while ensuring that the number of assigned students does not decrease significantly or does not decrease at all, as shown in Corollary 3.4. The effect is more pronounced with a shortage of seats, because in that case students are, on average, assigned to much worse choices than they are when there is a surplus of seats.

The following proposition shows that the average rank of students under STB is not sensitive to market imbalance.

**Proposition 4.2.** $Avg(STB)$ *is* $O(\ln \min\{n, q\})$*.*

The intuition for the result follows from observing that the DA algorithm under STB is equivalent to the random serial dictatorship algorithm, in which students pick in random order their favorite school among schools that still have vacancies. When it is the turn of the $(i+1)$-th student to choose a school, the rank she will obtain is, roughly speaking, a geometric random variable with a success probability of at least $1 - i/m$ and summing up the expectations of these random variables gives the result.

One can observe from Propositions 3.1, 3.2, 4.1, and 4.2 that STB significantly outperforms MTB in assigning students to top choices. Moreover, when there is a shortage of schools, STB also outperforms MTB with respect to the average rank.

# 5 Simulations

This section presents simulation results that complement our theoretical results. We first run simulations for the case of complete preference lists to illustrate the predictions of the main theorem. Then we run a simulation for the case of short preference lists to illustrate the effects predicted in Theorem 3.3.

## 5.1 Complete preference lists

The first computational experiment illustrates the effect of the market size on the cumulative rank distributions under STB and MTB. For each set of parameters considered, we sample realizations by drawing complete preference lists uniformly at random and independently for each student. In addition, under MTB, for each market realization we draw a complete order over students for each school, independently and uniformly at random. Under STB, for each market realization we draw a single order over students uniformly at random. Then, we compute the student-optimal stable matching. The statistics presented are generated by taking averages over at least 1000 samples for each set of parameters.

Figure 1 plots the (average) students' cumulative rank distribution under MTB (left panel) and under STB (right panel) in three different balanced markets and where each school has unit capacity. The left panel (under MTB) shows that as the number of students (and seats) grows large, the percentage of students who get their top choices under MTB becomes smaller. Note that the difference between the percentage of those students who get their top choices when $n = 10^4$ and when $n = 10^3$ is roughly similar to the difference between $n = 10^5$ and $n = 10^4$; this follows from the logarithmic rate of decay as shown in the proof of the main theorem. Under MTB, as the number of students increases, the rank distribution becomes flatter. The corresponding graphs for STB almost coincide with each other for all simulated markets.[19]

[19]Under STB, the larger the capacity in each school, the "better" the rank distribution is, since more students are assigned to their top choices.

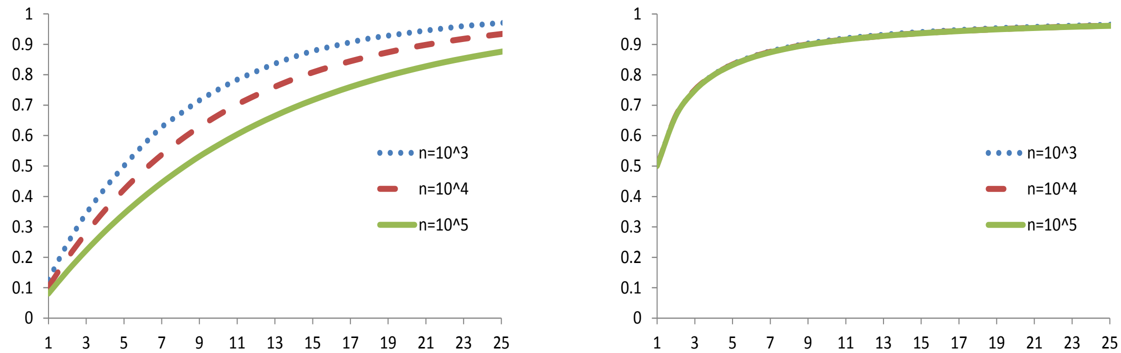


Figure 1: Students' cumulative rank distribution in three different balanced markets ($n = 10^3, 10^4, 10^5$) in which each school has capacity $\overline{q} = 1$. The left panel plots the distributions under MTB, and the right panel plots the distributions under STB. The x-axis represents the rank and the y-axis represents the fraction of students that are assigned to one of their top x ranked schools.

## 5.2 Short preference lists

Next we simulate the effect of limiting the length of preference lists on the rank distribution under MTB. Figure 2 plots the (average) cumulative rank distributions under MTB for markets, in which schools have relatively large capacities. The markets simulated consist of $m = 100$ schools, each with capacity $\overline{q} = 100$, and $n = 100^2 + 100$ students. When students submit preference lists that are 50 schools long, the fraction of students assigned to their first choice is just about 0.09 (with complete preference lists this fraction is 0.045). In contrast, recall that under STB almost half of the students are assigned to their first choice. However, as lists become shorter, the rank distribution under MTB (for assigned students) improves, whereas the number of unassigned students does not notably increase.

### 5.2.1 Number of unassigned students under MTB and STB

Table 1 reports the expected percentage of unassigned students in simulated markets with 1000 schools, where each school has 10 seats, an imbalance of 100 students, and varying lengths of preference lists. Observe that the number of unassigned students under MTB rapidly decreases with the length of the list (as predicted by Theorem 3.3). A similar effect is observed under STB. Also observe that MTB leaves fewer students unassigned than STB does, which is aligned with the experiments in New York City (Abdulkadiroğlu et al., 2009) and the city of Amsterdam (De Haan et al., 2015). This observation is consistent with the

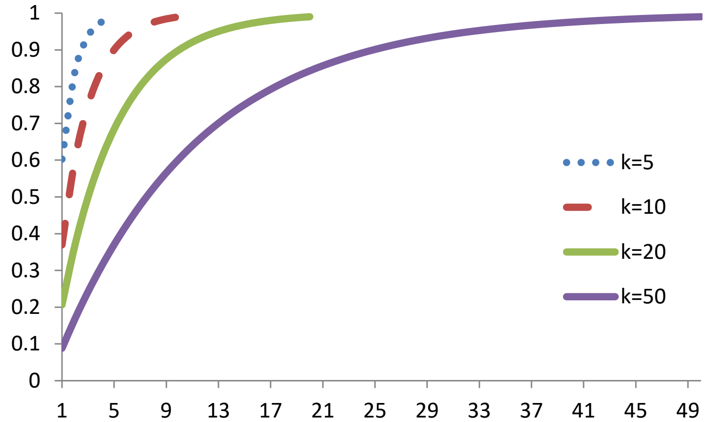


Figure 2: Students' cumulative rank distribution under MTB with different lengths of preference lists (5, 10, 20, and 50). There are $m = 100$ schools, each school has capacity $\overline{q} = 100$, and $n = 100^2 + 100$. The x-axis represents the rank and the y-axis represents the fraction of students that are assigned to one of their top x ranked schools.

trade-off that we discussed earlier: STB assigns more students to their top choices, but also more students to their lower choices.

| $n - qm$ | STB/MTB | 5 | 8 | 11 | 14 | 20 |
|---|---|---|---|---|---|---|
| +100 | STB | 0.032 | 0.02 | 0.018 | 0.015 | 0.012 |
| | MTB | 0.014 | 0.01 | 0.01 | 0.01 | 0.01 |
| -100 | STB | 0.023 | 0.013 | 0.008 | 0.005 | 0.003 |
| | MTB | 0.004 | 0.0002 | 0 | 0 | 0 |

Table 1: Percentage of unassigned students under STB and MTB in student-optimal stable matching for different markets. The first two rows report results for markets with 10,100 students and the last two rows report results for markets with 9,900 students. Column headings (from the third column onward) represent the length of preference lists.

We next explore the effect of correlation in students' preferences on the number of unassigned students. We simulated markets with 100 schools, each with a capacity of 100, and 10,000 students. Students' preference rankings are drawn independently from a multinomial logit discrete choice model: we assume each school $c$ has a weight $w_c = e^{-\frac{c}{m}\cdot\lambda}$ representing its "popularity" and students rank schools on their list proportionally to these weights. That is, the preference list of a given student is drawn independently as follows: her first choice is drawn independently from the distribution in which school $c \in \mathcal{C}$ is assigned probability $\frac{w_c}{\sum_{c' \in \mathcal{C}} w_{c'}}$. After her first choice is drawn, say, $c_1$, her second choice is drawn from the distribution in which each school $c \in \mathcal{C} \setminus \{c_1\}$ is assigned probability $\frac{w_c}{\sum_{c' \in \mathcal{C}\setminus\{c_1\}} w_{c'}}$, and so forth.

In this model, the larger $\lambda$ is, the more students' preferences are aligned (setting $\lambda = 0$ results in the uniform random distribution). Figure 3 reports the percentage of unassigned students for different choices of $\lambda$ (degree of correlation) and varying lengths of preference lists under both STB and MTB. Observe that a larger $\lambda$ results in a larger fraction of unassigned students. Observe that MTB leaves fewer students unassigned than STB.

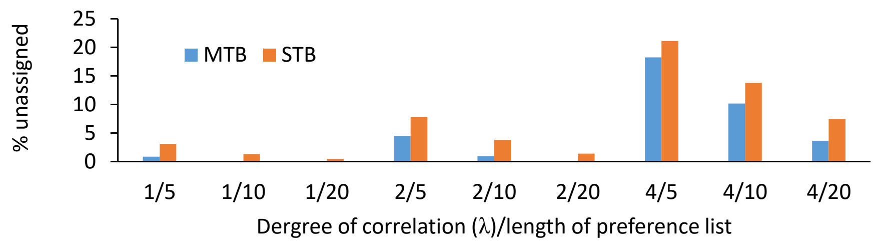


Figure 3: Percentage of students remaining unassigned under MTB and STB with varying lengths of preference lists and varying demand over schools.

# 6 Discussion

In their seminal paper comparing school choice mechanisms, Abdulkadiroğlu and Sönmez write:

> Using a single tie-breaking lottery might be a better idea [...]. In this case, any inefficiency will be necessarily caused by a fundamental policy consideration and not by an unlucky lottery draw. In other words, the tie-breaking will not result in additional efficiency loss if it is carried out through a single lottery (while that is likely to happen if the tie-breaking is independently carried out across different schools). (Abdulkadiroğlu and Sönmez, 2003, footnote 14).

MTB, however, is often considered fairer than STB by school choice practitioners, since a bad draw for a student under STB is irreversible. For example, in a balanced market it follows from Pittel (1992) that with high probability every student is assigned to at most her $O\left(\ln^2 n\right)$ rank under MTB. However, under STB the last assigned student is assigned to her $\frac{n}{2}$-th choice or worse, with probability at least half (because in a balanced market, the last student has a unique choice which is uniformly distributed). In general there is a "tipping rank" at which the rank distributions under STB and MTB cross: STB assigns more students cumulatively to the ranks before the tipping rank, whereas MTB assigns more students (cumulatively) to the ranks after the tipping rank. In particular, when preference lists are short, previous findings and our simulations confirm that MTB leaves fewer students unassigned than STB. Arnosti (2015) further establishes that the cumulative rank distributions cross each other once for the case of short preference lists.

This paper analyzes how well STB and MTB assign students to one of their top choices in random markets. When preferences lists are unbounded, MTB assigns almost no student to a top choice, in sharp contrast to STB, which assigns a constant fraction of students to a top choice.

Some school districts including New York City and Denver allow students to rank only a limited number of schools. This boosts the rankings of assigned students. The drawback is that imposing bounds on preference lists potentially leaves more students unassigned, resulting in an excessive administrative burden.[20] Our results provide a rationale for shortening lists in addition to some guidance for practitioners (Theorem 3.3). By limiting (appropriately) the length of preference lists under MTB, the social planner can bound the fraction of students who get assigned to a school they do not like, whereas allowing long lists may grant only a very small number of students a top choice. This observation also resonates with the logic behind the "circuit-breaker" mechanism of Che and Tercieux (2015). Indeed, offering a small number of choices and then later administratively assigning unmatched students can be viewed as a rudimentary way to implement a two-stage mechanism. We emphasize that our results for short preferences lists are obtained in a stylized model with uncorrelated preferences; in practice shortening preference lists is a challenging task due to the correlation in students' preferences, which also results in strategic concerns.

The tie-breaking rule may affect the resulting rank distribution significantly. Information about this distribution is not only arguably important for guiding design, but also valuable for students who go through the (probably) costly task of ranking schools.[21] This paper contributes to the study of tie-breaking rules by applying methods from a growing line of research on two-sided matching markets with random preferences that aims to understand

[20] Abdulkadiroğlu et al. (2009) provide empirical evidence that many students in New York City remain unmatched after the first round.

[21] Abdulkadiroğlu et al. (2009) document that a significant fraction of students rank very few schools.

outcomes in typical markets. While clearly the quantitative results should be treated with caution, our qualitative findings help to explain the impact of STB and MTB on student assignments.

# References


Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

Atila Abdulkadiroğlu, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. The boston public school match. *American Economic Review*, 95(2):368–371, 2005.

Atila Abdulkadiroğlu, Parag A. Pathak, and Alvin E. Roth. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the New York City high school match. *American Economic Review*, 99(5):1954–1978, 2009.

Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. Resolving conflicting preferences in school choice: The "Boston mechanism" reconsidered. *American Economic Review*, 101 (1):399–410, 2011.

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

Nick Arnosti. Short lists in centralized clearinghouses. *Available at SSRN*, 2015.

Itai Ashlagi and Afshin Nikzad. What matters in tie-breaking rules? how competition guides design. *Unpublished working paper*, 2015.

Itai Ashlagi and Peng Shi. Improving community cohesion in school choice via correlated-lottery implementation. *Operations Research*, 62(6):1247–1264, 2014.

Itai Ashlagi, Mark Braverman, and Avinatan Hassidim. Stability in large matching markets with complementarities. *Operations Research*, 62(4):713–732, 2014.

Itai Ashlagi, Yashodhan Kanoria, and Jacob D. Leshno. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy*, 2016. Forthcoming.

Caterina Calsamiglia, Guillaume Haeringer, and Flip Klijn. Constrained school choice: An experimental study. *American Economic Review*, 100(4):1860–1874, 2010.

Gabriel Carroll. A general equivalence theorem for allocation of indivisible objects. *Journal of Mathematical Economics*, 51:163–177, 2014.

Yeon-Koo Che and Fuhito Kojima. Asymptotic equivalence of probabilistic serial and random priority mechanisms. *Econometrica*, 78(5):1625–1672, 2010.

Yeon-Koo Che and Olivier Tercieux. Efficiency and stability in large matching markets. Mimeo, 2015.

Yan Chen and Onur Kesten. From Boston to Chinese parallel to deferred acceptance: Theory and experiments on a family of school choice mechanisms. Discussion Paper, Social Science Research Center Berlin (WZB), Research Area Markets and Politics, Research Unit Market Behavior, 2013.

Peter Coles and Ran Shorrer. Optimal truncation in matching markets. *Games and Economic Behavior*, 87:591–615, 2014.

Monique De Haan, Pieter A. Gautier, Hessel Oosterbeek, and Bas Van der Klaauw. The performance of school assignment mechanisms in practice. CEPR Discussion Paper No. DP10656, 2015.

Aytek Erdil and Haluk Ergin. What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, 98(3):669–689, 2008.

David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

Yannai Gonczarowski. Manipulation of stable matchings using minimal blacklists. In *Proceedings of the Fifteenth ACM conference on Economics and Computation*, page 449. ACM, 2014.

Guillaume Haeringer and Flip Klijn. Constrained school choice. *Journal of Economic Theory*, 144(5):1921–1947, 2009.

Onur Kesten. On two kinds of manipulation for school choice problems. *Economic Theory*, pages 1–17, 2011.

Donald Ervin Knuth. *Mariages stables et leurs relations avec d'autres problèmes combinatoires*. Presses de l'Université de Montréal, 1976.

Fuhito Kojima and Parag A. Pathak. Incentives and stability in large two-sided matching markets. *American Economic Review*, 99(3):608–27, 2009. doi: 10.1257/aer.99.3.608.

Qingmin Liu and Marek Pycia. Ordinal efficiency, fairness, and incentives in large markets. Working paper, 2012.

Parag A. Pathak and Jay Sethuraman. Lotteries in student assignment: An equivalence result. *Theoretical Economics*, 6(1):1–17, 2011.

Boris Pittel. The average number of stable matchings. *SIAM Journal on Discrete Mathematics*, 2(4):530–549, 1989.

Boris Pittel. On likely solutions of a stable marriage problem. *The Annals of Applied Probability*, 2(2):358–401, 1992.

Alvin E. Roth and Uriel G. Rothblum. Truncation strategies in matching markets-in search of advice for participants. *Econometrica*, 67(1):21–43, 1999.

Lloyd S Shapley. *Markets as cooperative games*. Rand Corporation, 1955.

# A Analysis

We fix some notation and then proceed to the proof. We use a version of the deferred acceptance (DA) algorithm that accepts an infinite stream of integers, representing a source of randomness from which students' preferences are drawn. We use square brackets to denote a range, i.e., $[n] = \{1, \ldots, n\}$. Let $S$ be a sequence of integers in $[m]$, such that each integer in $[m]$ appears an infinite number of times in $S$. Let $S[j]$ denote the prefix of $S$ up to the $j$th place, and let $S_h$ denote the $h$th integer in $S$. The student-proposing DA algorithm with $k$ rounds based on $S$ (denoted by $\mathrm{DA}\,(k)$) works as described in Algorithm 1 by letting all proposers in a specific round draw schools one by one, while skipping any school they have already proposed to. Note that when a prefix of size $j$ is read, it is possible that fewer than $j$ proposals were made, because some integers were read by students who already proposed to those schools. The regular DA algorithm (with no limit on the number of rounds) is denoted by DA ($=$ DA $(\infty)$).

---

**Algorithm 1:** DA $(k)$ based on $S$

---

```
Input: k, S, schools' preferences
Output: Matching μ
P ← [n]
h ← 0
∀c ∈ C : μ[c] ← ∅
for round ← 1 to k do
    for s in P do
        if s already proposed to all schools then
            continue
        end
        while s already proposed to S_h do
            h ← h + 1
        end
        μ[S_h] ← μ[S_h] ∪ {s}
        h ← h + 1
    end
    P ← ∅
    for c in C do
        T ← top q̄ students in μ[c] according to c's preference
        P ← P ∪ (μ[c] \ T)
        μ[c] ← T
    end
    sort P
end
```

---

We almost always omit the dependency of DA $(k)$ and DA on the schools' preferences, assuming they are drawn at random as described above. Whenever we refer to DA $(k)$ or DA without specifying a stream $S$, we are referring to the operation of the DA algorithm on a random stream in which every integer is drawn uniformly at random from $[m]$. At the same time we ignore the (measure 0) event of having a stream in which some integer appears only finitely many times.

Finally, we say that a probability mass function (PMF) $P : \mathbb{Z}_+ \to \mathbb{R}_+$ stochastically dominates PMF $Q : \mathbb{Z}_+ \to \mathbb{R}_+$ if for any $i \in \mathbb{Z}_+$ we have: $\sum_{j=0}^{i} P(j) \geq \sum_{j=0}^{i} Q(j)$.

## A.1 Proof of Proposition 3.1

Let us first provide a proof sketch.

*Proof Sketch.* For each school $c$, we define a random variable $X_c$ that takes values in $\{0, 1, \ldots, \overline{q}\}$ and denotes the number of students who are admitted to school $c$ for whom $c$ is a top $k$ choice. Note that $R_k = \mathbb{E}\left[\sum_{c \in \mathcal{C}} X_c\right]/n$. By linearity of expectation and symmetry, we have $R_k = \mathbb{E}[X_c] \cdot \frac{m}{n}$ for any fixed $c \in \mathcal{C}$. Since $\frac{m}{n} \leq 1$, the theorem is proved if we show that $\mathbb{E}[X_c]$ approaches 0 as $n$ approaches infinity.

To prove the latter fact, first we show that school $c$ receives at least $\Omega(\ln n)$ proposals w.h.p. Second, we show that w.h.p., the number of students who have listed $c$ in one of their top $k$ positions, denoted by $\psi(k)$, is sub-logarithmic in $n$. These two facts imply that $\mathbb{E}[X_c] \leq O\left(\frac{\overline{q}\,\psi(k)}{\ln n}\right)$, because each of the students who applied to $c$ and have listed $c$ in one of their top $k$ positions would be admitted with probability at most $O(\overline{q}/\ln n)$. So, each of the students contributes at most $O(\overline{q}/\ln n)$ to $\mathbb{E}[X_c]$. Since there are at most $\psi(k)$ such students, it follows that $\mathbb{E}[X_c] \leq O\left(\frac{\overline{q}\,\psi(k)}{\ln n}\right)$. This proves the promised claim: $\mathbb{E}[X_c]$ approaches 0 as $n$ approaches infinity. □

**Remark 1.** *Given that $k$ is a constant, Proposition 3.1 still holds when $\overline{q} = o(\ln n)$. In fact, the same proof works; here, we verify this by following the same proof sketch. Note that $\mathbb{E}[X_c] \approx \frac{\overline{q}\,\psi(k)}{\ln(n/\overline{q})}$. So, $\mathbb{E}[X_c]$ approaches 0 when $\overline{q} = o(\ln(n/\overline{q}))$, or, equivalently, when $\overline{q} = o(\ln n)$.*

**Remark 2.** *Given that $\overline{q}$ is a constant, Proposition 3.1 holds for any $k = o(\ln n)$. To see why, it is enough to follow the proof for Proposition 3.1 and note that when $k = o(\ln n)$, the right-hand side of concentration bound* (1) *still approaches* 0 *as $n$ approaches infinity, for a suitable choice of $\theta > 1$.*

Before proceeding to the proof, we introduce two parameters, $r$ and $\delta$, that are frequently used in our analysis. During the analysis, we typically find it helpful to run DA for only $r$ rounds. We define $r = 4m^{1/2}/\overline{q}$. Another frequently used parameter in our analysis is $\delta$, which is set to $3/4$ in this proof.

Next, we state two lemmas that are required to prove Proposition 3.1. In the first lemma, we show that w.h.p., most of the schools are not empty by the end of DA $(r)$.

**Lemma A.1.** *The probability of having more than $m^\delta$ empty schools by the end of DA $(r)$ is at most $r \cdot e^{-\frac{\overline{q} m^{2\delta-1}}{16}}$.*

We defer the proof for Lemma A.1 to Section B. Next, we show that w.h.p. the total number of proposals sent in DA $(r)$ is at least $\Omega(m \ln m)$; this is a consequence of the next lemma.

**Lemma A.2.** *Let $l = m(\ln m - t)$ for some $t > 0$. Then, with probability at least $1 - \frac{m^\delta+1}{e^t} - r \cdot e^{-\frac{\overline{q}m^{2\delta-1}}{16}}$, at least $l$ proposals are sent in DA $(r)$.*

A suitable choice of $t$, e.g., $t = 4/5 \ln m$, implies that w.h.p. the total number of proposals sent in DA $(r)$ is at least $\Omega(m \ln m)$. Lemma A.2 is proved in Section B.

We are now ready to prove Proposition 3.1.

*Proof of Proposition 3.1.* We follow the proof sketch. Recall that we define a random variable $X_c$ for each school $c$; $X_c$ is the number of students who are admitted to $c$, and $c$ is among their top $k$ choices. Notice that $R_k = \mathbb{E}\left[\sum_{c\in\mathcal{C}} X_c\right]/n$ and therefore $R_k = \mathbb{E}[X_c] \cdot \frac{m}{n}$ for any fixed $c \in \mathcal{C}$. So the theorem is proved if we show that $\mathbb{E}[X_c]$ approaches 0 as $n$ approaches infinity.

We now formally prove the above fact in three steps. In Step 1, we show that school $c$ receives at least $\Omega(\ln n)$ proposals w.h.p. In Step 2, we show that, w.h.p., the number of students who have listed $c$ in one of their top $k$ positions is a sub-logarithmic function of $m$. Step 3 completes proof using what we proved in the first two steps.

**Step 1** We use Lemma A.2 with $t = 4/5 \ln m$. This implies DA $(r)$ makes at least $m \ln m/5$ proposals with probability at least

$$1 - \frac{m^\delta + 1}{e^t} - r \cdot e^{-\frac{\overline{q}m^{2\delta-1}}{16}},$$

which is at least $1 - 2m^{-1/20}$ for large enough $m$.[22] Now, given school $c$, we show that this school receives at least $O(\ln m)$ proposals. We just showed that at least $m \ln m/5$ proposals will be sent w.h.p. Thus, DA $(r)$ reads a prefix of $S$ with length at least $j = m \ln m/5$. To complete Step 1, we show that w.h.p. school $c$ appears at least $\Omega(\ln m)$ times in $S[j]$. Using this fact we then show that at least $\Omega(\ln m)$ different students propose to $c$.

Let $E(c)$ be the event in which $c$ appears at least $\ln m/10$ times in $S[j]$. To complete Step 1, we first prove that $E(c)$ holds w.h.p.; this is done by a standard application of Chernoff bounds. For each index $h$, let $Y_h$ be a binary random variable that is 1 iff $S_h = c$. Let

[22]E.g., $m > 5$ works if $\overline{q} \geq 16$; $m > 10^5$ suffices for $\overline{q} = 1$. We have not tried to optimize this constant.

$\mu = \mathbb{E}\left[\sum_{h=1}^{j} Y_h\right]$; note that since $Y_h$ is a Bernoulli random variable with mean $1/m$, we have $\mu = \ln m/5$. By Chernoff bounds we have

$$\mathbb{P}\left[\sum_{h=1}^{j} Y_h < \mu(1-\epsilon)\right] \leq e^{\frac{-\epsilon^2\mu}{2}} = m^{-1/40},$$

for $\epsilon = 1/2$. So, the probability that $c$ appears less than $\ln m/10$ times in $S[j]$ is at most $m^{-1/40}$.

**Proposition A.3.** *For any school $c$, $E(c)$ holds with probability at least $1 - m^{-1/40}$.*

Assuming that $c$ appears at least $\ln m/10$ times in $S[j]$, we show that $\Omega(\ln m)$ of these appearances correspond to proposals made by disjoint students. This will complete the proof of Step 1. To this end, define $E$ to be the event in which each student makes at most 4 (possibly redundant) proposals in any round of DA $(r)$. We show that $E$ holds w.h.p. See that by Lemma B.3, the probability that each student makes more than 4 offers in each round is at most $(r/m)^4 = (4/\overline{q})^4 m^{-2}$. A union bound over all students and all rounds implies that $E$ holds with probability at least $1 - 4(4/\overline{q})^4 m^{-1/2}$.

**Proposition A.4.** *$E$ holds with probability at least $1 - 4(4/\overline{q})^4 m^{-1/2}$.*

By Propositions A.3 and A.4, $E(c) \wedge E$ holds with probability at least $1 - m^{-1/40} - 4(4/\overline{q})^4 m^{-1/2}$. To complete the proof of Step 1, we simply note that when $E(c) \wedge E$ holds, $c$ must have received proposals from at least $\ln m/40$ disjoint students.

**Step 2** In this step, we will show that w.h.p., the number of students who have listed $c$ in one of their top $k$ positions is a sub-logarithmic function of $m$. For any student $s$, let $Z_s$ be a binary random variable that is 1 iff $s$ lists the school $c$ in one of her top $k$ positions. Let $\mu = \mathbb{E}\left[\sum_{s \in \mathcal{S}} Z_s\right]$. Note that since the preferences are uniform, it follows that $\mathbb{P}[Z_s = 1] = k/n$, which means $\mu = k$. We prove that, w.h.p., $\sum_{s \in \mathcal{S}} Z_s$ is not much larger than its mean. This is done by applying the following version of Chernoff bounds:

$$\mathbb{P}\left[\sum_{s \in \mathcal{S}} Z_s > \theta\mu\right] < \left(\frac{e^{\theta-1}}{\theta^\theta}\right)^\mu, \tag{1}$$

which holds for any $\theta > 1$. Let the right-hand side of (1) be denoted by $f(\theta)$ and observe that by setting $\theta = \sqrt{\ln m}$, $f(\theta)$ approaches 0 as $n$ approaches infinity.[23]

**Step 3** Let $E^*$ be the event in which school $c$ receives at least $\ln m/40$ proposals during $\mathrm{DA}\,(r)$ and at most $\theta k$ students list $c$ among their top $k$ choices. As a consequence of Steps 1 and 2, we know that $E^*$ holds w.h.p., i.e., with probability at least $1 - m^{-1/40} - 4(4/\overline{q})^4 m^{-1/2} - f(\theta)$. Notice that

$$\mathbb{E}\left[X_c | E^*\right] \leq \theta k \cdot \frac{\overline{q}}{\ln m/40} = 40k\overline{q}/\sqrt{\ln m}.$$

Using this, we can write

$$\begin{aligned}\mathbb{E}\left[X_c\right] \leq &\left(1 - m^{-1/40} - 4(4/\overline{q})^4 m^{-1/2} - f(\theta)\right) \cdot \left(40k\overline{q}/\sqrt{\ln m}\right) \\ &+ \left(m^{-1/40} + 4(4/\overline{q})^4 m^{-1/2} + f(\theta)\right) \cdot k. \end{aligned} \tag{2}$$

Now, observe that the right-hand side of (2) approaches 0 as $m$ approaches infinity. This completes the proof. □

## A.2 Proof of Theorem 3.3

Before proving Theorem 3.3, we extend the definition of $\mathrm{DA}\,(k)$ by specifying which "seat" each student takes if accepted to a school. Give each seat a unique label, and let $\mathcal{L}$ denote the set of labels of all $q$ available seats. In each round of $\mathrm{DA}\,(k)$, when a student $s$ proposes to a school with empty seats, one of the empty seats is chosen uniformly at random (among all the empty seats in that school) and is assigned to that student. This is the only change that we make to $\mathrm{DA}\,(k)$; as this change is not structural, we still denote this process by $\mathrm{DA}\,(k)$.

We prove Theorem 3.3 for the case in which there is a shortage of seats. The proof for the case in which there is an excess of seats is similar and is therefore omitted.

*Proof for Part 1 of Theorem 3.3.* The main idea is to define a much simpler process, $\mathrm{DA}''\,(k)$, which, roughly speaking, leaves more students unassigned at the end of round $k$. We next analyze $\mathrm{DA}''\,(k)$ and use the number of unassigned students in it as an upper bound on

[23]In fact, the proof works for any non-constant function that grows slower than $\ln m$, i.e., $\theta = o(\ln m)$.

the number of unassigned students in $\mathrm{DA}(k)$. More precisely, suppose that $P : \mathbb{Z}_+ \to \mathbb{R}_+$ and $P'' : \mathbb{Z}_+ \to \mathbb{R}_+$ respectively denote the PMFs of *the number of unassigned students* in $\mathrm{DA}(k)$ and $\mathrm{DA}''(k)$. Then, we show that $P$ stochastically dominates $P''$. Because of this stochastic domination, we need only to prove the bounds (stated in the theorem statement) for $\mathrm{DA}''(k)$; the identical bounds will hold for $\mathrm{DA}(k)$.

We define $\mathrm{DA}''(k)$ in two steps. In Step 1, we convert $\mathrm{DA}(k)$ to $\mathrm{DA}'(k)$ (a random process slightly more complicated than $\mathrm{DA}''(k)$). This conversion is done so that $P$ stochastically dominates $P'$, the PMF of the number of unassigned students in $\mathrm{DA}'(k)$. In Step 2, we convert $\mathrm{DA}'(k)$ to $\mathrm{DA}''(k)$ while ensuring that $P'$ stochastically dominates $P''$.

We start by defining $\mathrm{DA}'(k)$. This process involves $k$ rounds. In each round, $\varepsilon q$ new students show up and, one by one, each proposes to a school picked uniformly at random. When a student $s$ proposes to a school that has at least one empty seat, then one of the empty seats in that school is chosen uniformly at random and is assigned to $s$. Otherwise, if the school is full, $s$ is rejected in that round. Lemma C.1 shows that $P'$ is stochastically dominated by $P$.

We emphasize that the main difference between $\mathrm{DA}(k)$ and $\mathrm{DA}'(k)$ is that in $\mathrm{DA}'(k)$, a school accepts new students if and only if it is not full (and so a student that applies to a unfilled school will surely maintain her seat). In $\mathrm{DA}(k)$, however, a new applicant can result in a different student being rejected from an already full school. This creates "rejection chains" in $\mathrm{DA}(k)$ which are not present in $\mathrm{DA}'(k)$, which give rise to the stochastic dominance relation.

Next, we convert $\mathrm{DA}'(k)$ to $\mathrm{DA}''(k)$ while ensuring that $P''$ is stochastically dominated by $P'$. We start by defining $\mathrm{DA}''(k)$: this process involves $k$ rounds. In each round, $\varepsilon q$ new students show up and, one by one, they propose to a school picked uniformly at random. When a student $s$ proposes to a school, she picks one of the seats uniformly at random among all the available $\overline{q}$ seats; her proposal in that round is accepted iff the seat is empty. Lemma C.2 shows that $P''$ is stochastically dominated by $P'$.

To prove the theorem, we first prove the bounds given in the theorem statement for $\mathrm{DA}''(k)$ (instead for $\mathrm{DA}(k)$). Stochastic dominance then implies that the bounds hold for $\mathrm{DA}(k)$ as well. More formally, suppose that $U''_k$ denotes the number of unassigned students at the end of $\mathrm{DA}''(k)$. We then show that there exists a random variable $\Delta''$ such that $U''_k \leq \varepsilon q + \Delta''$, where $\mathbb{E}[\Delta''] \leq e^{-\varepsilon k} q$ and, moreover, w.h.p. $\Delta''$ is not much larger than

$\mu = e^{-\varepsilon k}q$, in the following sense (Chernoff concentration bounds):

$$\mathbb{P}\left[\Delta'' > \mu(1+\delta)\right] \leq e^{\frac{-\delta^2\mu}{3}} \qquad \forall\, 0 < \delta < 1,$$
$$\mathbb{P}\left[\Delta'' > \delta\mu\right] \leq \left(\frac{e^{\delta-1}}{\delta^\delta}\right)^\mu \qquad \forall\, \delta > 1.$$

This will prove the theorem.

Given the above definition for $\mathrm{DA}''(k)$, it is straightforward to verify that, in each round, each seat will receive a proposal with probability $\frac{1}{\bar{q}m} = 1/q$. We use this fact to prove the theorem. For any seat $l \in \mathcal{L}$, let $X_l$ be a binary random variable that is 1 iff seat $l$ is empty at the end of $\mathrm{DA}''(k)$. By considering applications sent out by $\varepsilon qk$ students (arriving over $k$ rounds) we have:

$$\mathbb{E}\left[X_l\right] \leq \left[\left(1 - \frac{1}{q}\right)^{\varepsilon q}\right]^k \leq e^{-\varepsilon k}. \tag{3}$$

Let $\Delta'' = \sum_{l\in\mathcal{L}} X_l$; since $\mathbb{E}\left[\Delta''\right] = \sum_{l\in\mathcal{L}} \mathbb{E}\left[X_l\right]$, we have $\mathbb{E}\left[\Delta''\right] \leq e^{-\varepsilon k}q$, which is the promised claim. For the concentration result, just note that the variables $X_l$ are negatively correlated, which means Chernoff bounds are applicable. □

*Proof of Part 2 of Theorem 3.3.* We assume that $d = 0$; almost the same proof works for the general case when $d = o(q)$. Thus, we present the proof assuming that $n = q$.

**Proof of Part (a)** We run the process from round 1 to round $k$. We prove that in each round $j$, the expected number of unassigned students is at most $\theta n/j$, where $\theta \geq 1$ is a constant that we will set later. The proof is by induction. The induction base is $j = 1$, which holds trivially. Suppose that the induction hypothesis holds for $j = i$; for the induction step we need to prove that $\mathbb{E}\left[U_{i+1}\right] \leq \theta n/(i+1)$.

Notice that each empty seat remains empty (at the end of round $i+1$) with probability at most $\left(1 - \frac{1}{\bar{q}m}\right)^{U_i}$; since the number of empty seats equals the number of unassigned students, we then have:

$$\begin{aligned}\mathbb{E}\left[U_{i+1}|U_i\right] &\leq U_i\left(1 - \frac{1}{n}\right)^{U_i} \\ &\leq U_i e^{-U_i/n} \leq U_i\left(1 - \frac{U_i}{2n}\right),\end{aligned}$$

where the last inequality is due to $e^{-x} \leq 1 - x/2$, which holds for all $0 \leq x \leq 1$. Then, we take another expectation from both sides of the above inequality to infer that

$$\mathbb{E}\left[U_{i+1}\right] \leq \mathbb{E}\left[U_i - \frac{U_i^2}{2n}\right].$$

Now we can use linearity of expectation and then Jensen's inequality to write

$$\mathbb{E}\left[U_{i+1}\right] \leq \mathbb{E}\left[U_i\right] - \mathbb{E}\left[\frac{U_i^2}{2n}\right] \leq \mathbb{E}\left[U_i\right] - \frac{\mathbb{E}\left[U_i\right]^2}{2n}.$$

We are almost done. First note that if we have $\mathbb{E}\left[U_i\right] = \theta n/i$, then we get

$$\mathbb{E}\left[U_{i+1}\right] \leq \mathbb{E}\left[U_i\right] - \frac{\mathbb{E}\left[U_i\right]^2}{2n} \leq \theta n/i - \frac{(\theta n/i)^2}{2n}.$$

Since for any $\theta \geq 2i/(i+1)$, we have

$$\theta n/i - \frac{(\theta n/i)^2}{2n} \leq \theta n/(i+1),$$

and so by setting $\theta = 2$ we always get

$$\mathbb{E}\left[U_{i+1}\right] \leq \theta n/(i+1).$$

This proves the induction step if $\mathbb{E}\left[U_i\right] = \theta n/i$. On the other hand, note that the function $g(x) = x - x^2/(2n)$ is an increasing function of $n$ for all $x < n$. So, even when $\mathbb{E}\left[U_i\right] < \theta n/i$, we have

$$\begin{aligned}\mathbb{E}\left[U_{i+1}\right] &\leq \mathbb{E}\left[U_i\right] - \frac{\mathbb{E}\left[U_i\right]^2}{2n} \\ &\leq \theta n/i - \frac{(\theta n/i)^2}{2n} \leq \theta n/(i+1),\end{aligned}$$

as long as we have $\theta n/i \leq n$, which holds for all $i \geq 2$. This completes the induction step. Thus, we have shown that $\mathbb{E}\left[U_k\right] \leq 2n/k$ for all $k \leq n$. This proves the theorem.

**Proof of Part (b)** We prove that, w.h.p. in any round $j$, we have $U_j \leq \theta n/j$, where $\theta = 2 + \varepsilon'$. The induction base is $j = 1$, which holds trivially. Suppose the induction hypothesis holds for $j = i$; for the induction step, we prove that w.h.p. we have $U_{i+1} \leq \theta n/(i+1)$.

Then a union bound over all steps ensures that w.h.p. every step holds.

Before proving the induction step, notice that we can safely assume that $U_i \geq n^{2/3+\varepsilon}$. Otherwise, we have

$$\theta n/(i+1) \geq n^{2/3+\varepsilon} > U_i \geq U_{i+1},$$

which would prove the induction step. Now, suppose $E$ denotes the set of empty seats at the beginning of round $i$. For any $l \in E$, let $X_l$ be a binary random variable that is 1 iff $l$ is empty at the end of round $i$. Note that $U_{i+1} = \sum_{l \in E} X_l$. We prove that w.h.p. $U_{i+1}$ is not too large, in the following sense.

**Claim A.5.** *Let $\zeta = U_i(1-1/n)^{U_i}$, then for any positive $\delta < 1$ we have:*

$$\mathbb{P}\left[U_{i+1} > \zeta(1+\delta)\right] \leq e^{-\frac{\delta^2 n^{2/3+\varepsilon}}{6}}. \tag{4}$$

We note that this is not a direct corollary of Chernoff concentration bounds, since the random variables $\{X_l\}_{l \in E}$ are neither independent nor negatively correlated. We prove (4) as follows.

*Proof of Claim A.5.* We define a random process $\mathcal{B}$, which corresponds to round $i$ of DA $(k)$. $\mathcal{B}$ is in fact a simple "balls and bins process," defined as follows. In $\mathcal{B}$, students propose to the same school as in (round $i$ of) DA $(k)$; however, when a student $s$ proposes to a school $c$, student $s$ picks one of the $\overline{q}$ seats in $c$ uniformly at random. Then, $s$ is accepted at $c$ iff the seat she picks is empty.

Let $U'_{i+1}$ denote the number of empty seats at the end of $\mathcal{B}$ and suppose that $P, P'$ respectively denote the PMFs of $U_{i+1}, U'_{i+1}$. The proof is done in two steps. First, we prove that $P$ stochastically dominates $P'$. Then, we prove the promised bound for the random variable $U'_{i+1} \sim P'$ (instead of $U_{i+1} \sim P$). The stochastic domination property then implies that the bound holds for $U_{i+1}$ as well.

To prove stochastic domination, we use a simple coupling argument as follows. We start running round $i$ of DA $(k)$ and define $\mathcal{B}$ based on the evolution of DA $(k)$. Unassigned students in $\mathcal{B}$ submit proposals in the same order as (in round $i$ of) DA $(k)$. Suppose that in DA $(k)$, it is the turn of an unassigned student $s$, who proposes to seat $l_s$ at school $c$. Let $E(c), E'(c)$ denote the set of empty seats in $c$ in the processes DA $(k)$, $\mathcal{B}$, respectively. We use the variable $l'_s$ to denote the seat to which $s$ proposes in $\mathcal{B}$, and define it as follows:

1. If $|E(c)| = \overline{q}$, then $l'_s = l_s$.

2. If $|E(c)| < \overline{q}$, then with probability $|E(c)|/\overline{q}$, let $l'_s = l_s$, and with probability $1-|E(c)|/\overline{q}$, let $l'_s = l'$, where $l'$ is a seat picked uniformly at random from the set of full seats in $c$.

It is straightforward to see that in any sample path we have $U_{i+1} \le U'_{i+1}$; i.e., the coupled process ($\mathcal{B}$) will have more unassigned students than in DA$\,(k)$. This holds simply because $\mathcal{B}$ never allocates a seat that was not allocated in round $i$ of DA$\,(k)$. In other words, our coupling guarantees that $E(c') \subseteq E'(c')$ always holds during the process, for any school $c' \in C$. Since $U_{i+1} \le U'_{i+1}$ in any sample path, $P$ stochastically dominates $P'$. Consequently, to prove the claim, we just need to show that

$$\mathbb{P}\left[U'_{i+1} > \zeta(1+\delta)\right] \le e^{-\frac{\delta^2 n^{2/3+\varepsilon}}{6}}.$$

To this end, let $X'_l$ be a binary random variable that is 1 iff seat $l$ is still empty at the end of $\mathcal{B}$. Note that $U'_{i+1} = \sum_{l\in E} X'_l$. Let $\mu' = \mathbb{E}\left[U'_{i+1}\right]$. Since the random variables $\{X'_l\}_{l\in E}$ are negatively correlated (which holds by the definition of $\mathcal{B}$), we have

$$\mathbb{P}\left[U'_{i+1} > \mu'(1+\delta)\right] \le e^{-\frac{\delta^2\mu'}{3}}$$

for any $\delta > 0$. Because of stochastic dominance, we have

$$\mathbb{P}\left[U_{i+1} > \mu'(1+\delta)\right] \le e^{-\frac{\delta^2\mu'}{3}}. \tag{5}$$

To complete the proof, we compute an upper bound $\mu'_U$ and a lower bound $\mu'_L$ for $\mu'$. We then plug these values into (5) to get

$$\mathbb{P}\left[U_{i+1} > \mu'_U(1+\delta)\right] \le e^{-\frac{\delta^2\mu'_L}{3}}, \tag{6}$$

which proves our claim.

First, we compute $\mu'_U$. Fix a seat $l$ and an unassigned student $s$; notice that in $\mathcal{B}$, this seat receives a proposal from $s$ with probability at least $1/(\overline{q}m) = 1/n$. Consequently, $\mathbb{P}\left[X'_l = 1\right] \le (1-1/n)^{U_i}$, which implies that $\mu' \le U_i(1-1/n)^{U_i}$. Thus, we set

$$\mu'_U = U_i(1-1/n)^{U_i}.$$

To find $\mu'_L$, notice that in $\mathcal{B}$, $s$ proposes to $l$ with probability at most $\frac{1}{\overline{q}(m-i)}$. So we have

$$\begin{aligned}\mathbb{P}\left[X'_l = 1\right] \geq \left(1 - \frac{1}{\overline{q}(m-i)}\right)^{U_i} &\geq 1 - \frac{U_i}{n - \overline{q}i} \\ &\geq 1/2, \end{aligned} \tag{7}$$

where (7) holds with very high probability by Lemma B.4. So, we can set $\mu'_L$ to be any number not larger than $U_i/2$. Now, recall that we assumed $U_i \geq n^{2/3+\varepsilon}$. So, we can safely set $\mu'_L = n^{2/3+\varepsilon}/2$.

Note that $\zeta = \mu'_R$, and plug the values for $\mu'_L, \mu'_R$ into (6); this completes the proof:

$$\mathbb{P}\left[U_{i+1} > \zeta(1+\delta)\right] \leq e^{-\frac{\delta^2 n^{2/3+\varepsilon}}{6}}.$$

□

We use Claim A.5 to prove the induction step; this is done by finding the $\delta$ for which $\zeta(1+\delta) \leq \theta n/(i+1)$; after finding such $\delta$, we plug it into (4) and complete the proof. From the latter inequality, we should have:

$$\delta \leq \frac{\theta n}{\zeta(i+1)} - 1. \tag{8}$$

So, to find the right value for $\delta$, we provide an upper bound on $\zeta$ and plug it into the right-hand side of (8). This is done as follows.

$$\zeta = \left(1 - \frac{1}{n}\right)^{U_i} \leq U_i e^{-U_i/n} \leq U_i \left(1 - \frac{U_i}{2n}\right),$$

where the last inequality is due to $e^{-x} \leq 1 - x/2$, which holds for all $0 \leq x \leq 1$. Using the induction hypothesis, we can rewrite the above inequality as:

$$\zeta \leq \theta n/i \left(1 - \frac{\theta n/i}{2n}\right), \tag{9}$$

where in writing (9) we have used the fact that $U_i\left(1 - \frac{U_i}{2n}\right)$ is an increasing function of $U_i$ for all $U_i < n$.

By plugging (9) into (8) we get $\delta \leq \varepsilon'/i$. Since $i \leq n^{1/3-\varepsilon}$, we can set $\delta = \varepsilon' n^{\varepsilon-1/3}$. Now, we are ready to finish the proof using (4). Recall that our choice of $\delta$ guarantees that

$\zeta(1+\delta) \leq \theta n/(i+1)$. So, we can use (4) to write

$$\mathbb{P}\left[U_{i+1} > \theta n/(i+1)\right] \leq e^{-\frac{\delta^2 n^{2/3+\varepsilon}}{6}}.$$

Since we have $\delta = \varepsilon' n^{\varepsilon - 1/3}$, we can rewrite the above bound as

$$\mathbb{P}\left[U_{i+1} > \theta n/(i+1)\right] \leq e^{\frac{-(\varepsilon')^2 n^{3\varepsilon}}{6}}. \tag{10}$$

To complete the proof, we just need a union bound over all rounds: (10) holds for all $i$ with probability at least $1 - ke^{-(\varepsilon')^2 n^{3\varepsilon}/6}$, i.e., with very high probability. □ □

# B Proofs of lemmas used in Proposition 3.1

*Proof of Lemma A.1.* The proof contains two main steps. Fix any round of DA $(r)$ and suppose that there are more than $m^\delta$ empty schools at the beginning of this round. This also means that there should be at least $\overline{q}m^\delta$ unassigned students. In the first step of the proof, we will prove that w.h.p., at least $O(\sqrt{m})$ of the unassigned students get assigned by the end of this round. Then, in the second step, we will use a union bound over all $r$ rounds to show that w.h.p., we have at most $m^\delta$ empty schools at the end of round $r$.

First, we show that if there is a subset $E$ of empty schools with size at least $m^\delta$ at the beginning of a round, then w.h.p., $O(\sqrt{m})$ students must get assigned to these schools by the end of this round. For each $c \in E$, let $X_c$ be a binary random variable that is set to 0 if school $c$ is still empty at the end of this round and is set to 1 otherwise. Each unassigned student then proposes to school $c$ with probability at least $1/m$, and the probability that $c$ receives no proposals by the end of this round is at most $(1-1/m)^{\overline{q}m^\delta}$. Since this quantity is at most

$$(1-1/m)^{\overline{q}m^\delta} \leq e^{-\overline{q}m^{\delta-1}} \leq \frac{1}{1+\overline{q}m^{\delta-1}} \leq 1 - \frac{\overline{q}m^{\delta-1}}{2},$$

we have $X_c = 1$ with probability at least $\overline{q}m^{\delta-1}/2$. This means that $\mathbb{E}\left[\sum_{c\in E} X_c\right] \geq \overline{q}m^{2\delta-1}/2$. We show that w.h.p., the sum (inside the expectation) is not too small relative to its mean.

To prove the latter fact, we use a Chernoff bound on the set of variables $\{X_c\}_{c\in E}$. A straightforward calculation shows that these variables are negatively correlated, and so Chernoff bounds are applicable. Let $\mu = \bar{q}m^{2\delta-1}/2$. By Chernoff bounds we have

$$\mathbb{P}\left[\sum_{c\in E} X_c \leq \mu(1-\epsilon)\right] \leq e^{\frac{-\epsilon^2\mu}{2}}.$$

Setting $\epsilon = 1/2$ implies that

$$\mathbb{P}\left[\sum_{c\in E} X_c \leq \mu/2\right] \leq e^{-\frac{\bar{q}m^{2\delta-1}}{16}}.$$

So far, we have shown that if $|E| \geq m^\delta$ at the beginning of a round, then, with probability at least $1 - e^{-\frac{\bar{q}m^{2\delta-1}}{16}}$, at least $\bar{q}m^{2\delta-1}/4$ of the schools in $E$ will not be empty at the end of that round. Using a union bound over all the $r$ rounds implies that, with probability $1 - r\cdot e^{-\frac{\bar{q}m^{2\delta-1}}{16}}$, we have at most $\max\{m^\delta, m - r\cdot\bar{q}m^{2\delta-1}/4\}$ empty schools by the end of round $r$. Observing that $r\cdot\bar{q}m^{2\delta-1}/4 \geq m$ proves the lemma. □

The next proposition will be used in the Proof of Lemma A.2.

**Proposition B.1.** *Suppose that $X$ is a geometric random variable with mean $1/p$. Then for any $\theta > 0$ we have $\mathbb{E}\left[e^{-\theta X}\right] = \frac{pe^{-\theta}}{1-(1-p)e^{-\theta}}$.*

*Proof.*

$$\begin{aligned}\mathbb{E}\left[e^{-\theta X}\right] &= \sum_{i=1}^{\infty} p(1-p)^{i-1}e^{-i\theta} \\ &= \frac{p}{1-p}\cdot\sum_{i=1}^{\infty}\left((1-p)e^{-\theta}\right)^i = \frac{pe^{-\theta}}{1-(1-p)e^{-\theta}}\end{aligned}$$

□

*Proof of Lemma A.2.* Let $E_1$ be the event that $\mathrm{DA}\,(r)$ observes at least $m - m^\delta$ different schools in the sequence $S$. By Lemma A.1, $E_1$ happens w.h.p. Also, let $E_2$ be the event that $S[l]$ does not contain $m - m^\delta$ different schools. The next claim shows that $E_2$ happens w.h.p.

**Claim B.2.** *Let $l < m(\ln m - t)$ for some $t > 0$. Then, with probability at least $1 - \frac{m^\delta+1}{e^t}$, $S[l]$ contains fewer than $m - m^\delta$ different schools.*

*Proof.* Let $\zeta = m - m^\delta$. For any $i \geq 1$, let $X_i$ be a variable that denotes the smallest integer $j$ such that $S[j]$ contains $i$ different schools. Define $X_0 = 0$. Note that since $S$ is a random variable, so is $X_i$. Also, define $Z_i = X_i - X_{i-1}$ for all positive $i$. It is straightforward to see that $Z_i$ is a geometric random variable with mean $\frac{m}{m-i+1}$. We provide a (Chernoff-type) concentration bound such that the random variable $Z = \sum_{i=1}^{\zeta} Z_i$ is highly concentrated around its mean. To do so, first notice that

$$\mathbb{P}\left[Z < \beta\right] = \mathbb{P}\left[e^{-\theta Z} > e^{-\theta\beta}\right] \leq \mathbb{E}\left[e^{\theta\beta - \theta Z}\right]. \tag{11}$$

Now we use the independence of $Z_i$'s to rewrite (the right-hand side of) (11):

$$\begin{aligned}
\mathbb{P}\left[Z < \beta\right] \leq \mathbb{E}\left[e^{\theta\beta - \theta Z}\right] &= e^{\theta\beta} \cdot \prod_{i=1}^{\zeta} \mathbb{E}\left[e^{-\theta Z_i}\right] \\
&= e^{\theta\beta} \cdot \prod_{i=1}^{\zeta} \frac{\frac{m-i+1}{m} \cdot e^{-\theta}}{1 - (1 - \frac{m-i+1}{m}) \cdot e^{-\theta}}, \qquad (12)
\end{aligned}$$

where (12) is due to Proposition B.1. Then, we choose $\theta = 1/m$ and use the fact that $e^{1/m} \geq 1 + 1/m$ to bound the right-hand side of (12):

$$\begin{aligned}
\mathbb{P}\left[Z < \beta\right] &\leq e^{\theta\beta} \cdot \prod_{i=1}^{\zeta} \frac{\frac{m-i+1}{m} \cdot e^{-\theta}}{1 - (1 - \frac{m-i+1}{m}) \cdot e^{-\theta}} \\
&\leq e^{\theta\beta} \cdot \prod_{i=1}^{\zeta} \frac{\frac{m-i+1}{m}}{1 + 1/m - (1 - \frac{m-i+1}{m})} \\
&= e^{\theta\beta} \cdot \prod_{i=1}^{\zeta} \frac{m-i+1}{m-i+2} = e^{\theta\beta} \cdot \frac{m-\zeta+1}{m+1}. \qquad (13)
\end{aligned}$$

Plugging $\beta = m(\ln m - t)$ into (13) implies that

$$\mathbb{P}\left[X_\zeta < m(\ln m - t)\right] \leq e^{-t}(m - \zeta + 1) = \frac{m^\delta + 1}{e^t}.$$

Then, recall that $l < m(\ln m - t)$. Thus, this bound says the probability of seeing $\zeta$ different schools in $S[l]$ is at most $\frac{m^\delta+1}{e^t}$. The lemma is proved. □

A union bound over the probabilistic bounds provided by Lemma A.1 and Claim B.2 imply that $E_1 \wedge E_2$ happens with probability at least $1 - \frac{m^\delta+1}{e^t} - r \cdot e^{-\frac{\bar{q}m^{2\delta-1}}{16}}$. When $E_1 \wedge E_2$

is true, DA $(r)$ reads a prefix of length at least $l$ from $S$, which means it makes at least $l$ proposals. □

**Lemma B.3.** *Fix a student $s$. The probability that $s$ makes more than 4 proposals in any round of DA $(r)$ is at most $(r/m)^4$.*

*Proof.* Suppose that we are in round $t$. Since there are at most $r$ rounds, $s$ has made at most $r$ proposals so far. The probability that $s$ makes 4 redundant proposals is then at most $(r/m)^4$. □

**Lemma B.4.** *Suppose that we are running DA $(k)$ with $n = q$. Denote the expected number of unassigned students at the end of round 1 by $U_1$. Then, the following holds:*

1. $\mathbb{E}[U_1] \leq q/e$.
2. *For any positive $\delta < 1$, $U_1$ is not larger than $(1+\delta)q/e$ with very high probability.*

*Proof.* We prove the lemma for when $\overline{q} = 1$. It is straightforward to use a coupling argument (similar to the coupling in the proof for Claim A.5) and show that the same bounds hold for $\overline{q} > 1$. For each school $c \in \mathcal{C}$, let $X_c$ be a binary random variable that is 1 iff $c$ has received no proposals by the end of round 1. Notice that

$$\mathbb{P}[X_c = 1] = (1 - 1/n)^n \leq e^{-1}. \tag{14}$$

So, we have $\mathbb{E}[U_1] = \sum_{c \in \mathcal{C}} \mathbb{E}[X_c] \leq n/e$, which proves Part 1. To prove Part 2, we use the negative correlation of random variables $\{X_c\}_{c \in \mathcal{C}}$ to apply Chernoff concentration bounds. To this end, first we need to give a lower bound for $\mathbb{E}[U_1]$. See that

$$\mathbb{P}[X_c = 1] = (1 - 1/n)^n \geq e^{-1.01},$$

which means that $\mathbb{E}[U_1] \geq ne^{-1.01}$. Using this lower bound and the upper bound given by (14) we can write the following bound, which completes the proof:

$$\mathbb{P}\left[U_1 > ne^{-1.01}(1+\delta)\right] \leq e^{\frac{-\delta^2 n}{3e}}, \qquad \forall\, 0 < \delta < 1.$$

□

# C Couplings required for the proof of Theorem 3.3

**Lemma C.1.** *Let $P : \mathbb{Z}_+ \to \mathbb{R}_+$ and $P' : \mathbb{Z}_+ \to \mathbb{R}_+$ respectively denote the PMF of unassigned students in DA $(k)$ and DA$'$ $(k)$. Then, $P$ stochastically dominates $P'$.*

*Proof.* To prove stochastic domination, we couple the random process $\mathrm{DA}'(k)$ with $\mathrm{DA}(k)$; i.e., we start running $\mathrm{DA}(k)$ and define another random process $\overline{\mathrm{DA}'}(k)$ based on the evolution of $\mathrm{DA}(k)$. We define this coupling so that the resulting process $\overline{\mathrm{DA}'}(k)$ becomes the same process as $\mathrm{DA}'(k)$.

Suppose that we are in round $i$ of $\mathrm{DA}(k)$; the coupling is then defined as follows. Let $q = \epsilon q$ and let $Q = \{s_1, \ldots, s_q\}$ denote the first $q$ unassigned students visited in $\mathrm{DA}(k)$. Also, let $Q' = \{s'_1, \ldots, s'_q\}$ be the $q$ new students arriving in round $i$ of $\overline{\mathrm{DA}'}(k)$. We define the proposals of the students in $Q'$ based on the proposals made by the students in $Q$. Suppose that $s_j$ has applied to the subset $H_j \subseteq \mathcal{C}$ of schools in previous rounds and is applying to (a new) school $c_j$ in this round. Suppose $E(c), E'(c)$ denote the set of empty seats in any school $c \in \mathcal{C}$ respectively in the processes $\mathrm{DA}(k), \overline{\mathrm{DA}'}(k)$. Our coupling would guarantee that $E(c) \subseteq E'(c)$ always holds during the process.

Moreover, suppose that $l_j$ denotes the seat (from $c_j$) that $s_j$ is assigned to; set $l_j = \emptyset$ if $s_j$ is rejected from $c_j$. We now define the proposal made by $s'_j$ in round $i$ of $\overline{\mathrm{DA}'}(k)$. Let $l'_j$ denote the seat for which $s'_j$ submits a proposal, and define it as follows:

1. With probability $|H_j|/m$, $s'_j$ applies to a school $c$ picked uniformly at random from $H_j$. If $E'(c) = \emptyset$, then let $l'_j = \emptyset$; otherwise, let $l'_j$ be an empty seat selected uniformly at random from $c$.

2. Otherwise (with probability $1 - |H_j|/m$), $s'_j$ applies to $c_j$. If $E'(c_j) = \emptyset$, then let $l'_j = \emptyset$. If $E'(c_j) \neq \emptyset$, then with probability $\frac{|E'(c_j) - E(c_j)|}{|E'(c_j)|}$, let $l'_j$ be a seat picked uniformly at random from $E'(c_j) - E(c_j)$; and with probability $1 - \frac{|E'(c_j) - E(c_j)|}{|E'(c_j)|}$, let $l'_j = l_j$.

Define $\overline{U'}_k$ as the number of empty seats at the end of $\overline{\mathrm{DA}'}(k)$. It is straightforward to see that in any sample path we have $U_k \leq \overline{U'}_k$; i.e., $\overline{\mathrm{DA}'}(k)$ will have more empty seats than $\mathrm{DA}(k)$. This holds simply because in any round $i$, $\overline{\mathrm{DA}'}(k)$ never allocates a seat that was not allocated in $\mathrm{DA}(k)$. Since $U_k \leq \overline{U'}_k$ in any sample path, then $P$ stochastically dominates $\overline{P'}$, the PMF of the number of unassigned seats at the end of $\overline{\mathrm{DA}'}(k)$. To complete the proof, it is enough to observe that $\overline{\mathrm{DA}'}(k)$ is the same random process as $\mathrm{DA}'(k)$, by definition. □

**Lemma C.2.** *Let $P' : \mathbb{Z}_+ \to \mathbb{R}_+$ and $P'' : \mathbb{Z}_+ \to \mathbb{R}_+$ respectively denote the PMF of unassigned students in $DA'(k)$ and $DA''(k)$. Then, $P'$ stochastically dominates $P''$.*

*Proof.* To prove stochastic dominance, we couple the random process $\mathrm{DA}''(k)$ with $\mathrm{DA}'(k)$; i.e., we start running $\mathrm{DA}'(k)$ and define another random process $\overline{\mathrm{DA}''}(k)$ based on the evolution of $\mathrm{DA}'(k)$. We define this coupling so that the resulting process $\overline{\mathrm{DA}''}(k)$ becomes the same process as $\mathrm{DA}''(k)$.

Suppose that we are in round $i$ of $\mathrm{DA}'(k)$; the coupling is then defined as follows. Let $q = \epsilon q$ and let $Q' = \{s'_1, \ldots, s'_q\}$ be the $q$ new students arriving in round $i$ in $\mathrm{DA}'(k)$. Also, let $Q'' = \{s''_1, \ldots, s''_q\}$ be the $q$ new students arriving in round $i$ of $\overline{\mathrm{DA}''}(k)$. We define the proposals of the students in $Q''$ based on the proposals made by the students in $Q'$. Suppose that $s'_j$ applies to school $c_j$ in this round; $s''_j$ also applies to $c_j$. To complete the definition of our coupling, we still must define the seat assigned to $s''_j$. Suppose $l_j$ denotes the seat (from $c_j$) that $s_j$ is assigned to; set $l_j = \emptyset$ if $c_j$ is full (and so cannot accept $s_j$). The seat assigned to $s''_j$ in the process $\overline{\mathrm{DA}''}(k)$ is denoted by $l''_j$ and is defined as follows:

1. If $c_j$ is full in $\mathrm{DA}'(k)$, then let $l''_j$ be a seat picked uniformly at random from the set of seats in $c_j$.

2. Otherwise, let $E(c_j)$ denote the set of empty seats in school $c_j$ in the process $\mathrm{DA}'(k)$. With probability $1 - \frac{|E(c_j)|}{\bar{q}}$ let $l''_j$ be a seat picked uniformly at random from the set of full seats in $c_j$. Otherwise (with probability $\frac{|E(c_j)|}{\bar{q}}$), let $l''_j = l_j$.

Define $\overline{U''}_k$ to be the number of empty seats at the end of $\overline{\mathrm{DA}''}(k)$. It is straightforward to see that in any sample path we have $U'_k \leq \overline{U''}_k$; i.e., $\overline{\mathrm{DA}''}(k)$ will have more empty seats than $\mathrm{DA}'(k)$. This holds simply because in any round $i$, $\overline{\mathrm{DA}''}(k)$ never allocates a seat that was not allocated in $\mathrm{DA}'(k)$. Since $U'_k \leq \overline{U''}_k$ in any sample path, then $P'$ stochastically dominates $\overline{P''}$, the PMF of the number of unassigned seats at the end of $\overline{\mathrm{DA}''}(k)$. To complete the proof, it is enough to observe that $\overline{\mathrm{DA}''}(k)$ is the same random process as $\mathrm{DA}''(k)$, by definition. □

# D Proofs for Proposition 3.2 and Proposition 4.2

*Proof of Proposition 4.2.* Without loss of generality, assume that students are ordered from 1 to $n$. When it is the turn of student $i + 1$ to select a school, there are already $i$ assigned students. The probability of success at each attempt made by student $i+1$ is at least $\frac{m-i/\overline{q}}{m} = \frac{q-i}{q}$. So, the expected number of attempts made by student $i+1$ is upper-bounded by $\frac{q}{q-i}$. Consequently, the expected total number of attempts made is upper-bounded by

$$\sum_{i=1}^{n} \frac{q}{q-i} = O(q \cdot (\ln q - \ln \min\{q-n, 1\})).$$

Now consider the two cases of $q \geq n$ and $q < n$ separately. Plugging them into the above expression proves the lemma. $\square$

*Proof of Proposition 3.2.* Suppose that students are ordered from 1 (highest priority) to $n$ (lower priority) in the tiebreaking. It is easy to see that if instead of running DA, we run Random Serial Dictatorship (RSD) in this order (student 1 choosing first), we will get exactly the same outcome. Now, notice that in RSD, student $i$ gets her top choice with probability at least $\frac{m-(i-1)/\overline{q}}{m}$, which is at least $\frac{q-i+1}{q}$. Therefore, proving that $\sum_{i=1}^{t} \frac{q-i+1}{q} \geq t/2$ would prove the first claim. To see this, consider two cases: either $q \leq n$ or $q > n$. In the former case,

$$\sum_{i=1}^{t} \frac{q-i+1}{q} = (q+1) - \frac{q(q+1)}{2q} = \frac{q+1}{2} > t/2.$$

In the latter case,

$$\sum_{i=1}^{t} \frac{q-i+1}{q} = \frac{n(q+1)}{q} - \frac{n(n+1)}{2q} > n - \frac{n}{2} = t/2.$$

This proves the first claim.

The second claim can be proved using a Chernoff bound. For each student $s$, let $E_s$ denote the event in which student $s$ is assigned to her top choice. Note that if the events $\{E_s\}_{s \in S}$ were independent, we could apply Chernoff bound and finish the proof. Here, however, more work is needed since the events are not independent. To handle these correlations, we use a coupling technique that simplifies the stochastic process corresponding to RSD by coupling that process to a coinflipping process. The formal proof follows after a brief, informal proof sketch.

Before we formally proceed, we briefly describe a sketch of the formal argument. We will use the Principle of Deferred Decisions and suppose students' preferences are not fixed, rather, they are generated while RSD is run. When it is the turn of student $i$ to choose, she first generates her preference list, and then is assigned to her most favorite available (unfilled) school. Let $f_i$ be the probability that the student is assigned to her first choice. Note that $f_i \geq \frac{m-(i-1)/\overline{q}}{m}$, because there are at least $m-(i-1)/\overline{q}$ unfilled schools available when student $i$ is choosing, regardless (independent) of how many students are assigned to their top choices in the past. In the coupled process, flip a coin for each student $i$ with success probability $\frac{m-(i-1)/\overline{q}}{m}$. This success probability accounts for a portion $\frac{m-(i-1)/\overline{q}}{m}$ of the probability $f_i$. Independence of the coinflips corresponds to the inequality $f_i \geq \frac{m-(i-1)/\overline{q}}{m}$ holding regardless of the rank assignments of the previous students. Finally, since the coinflips are independent, a Chernoff bound gives the desired concentration result.

The formal proof defines an auxiliary process, $\mathcal{B}$, which simulates the original process, but uses independent coin flips for each student. We will then couple $\mathcal{B}$ with the original process to show that the number of students who get their top choice in the original process first-order stochastically dominates the number of "successful coin flips" in $\mathcal{B}$. Therefore, by applying Chernoff bound on the number of successful coin flips in $\mathcal{B}$, we would prove the second claim.

We define $\mathcal{B}$ as follows. Suppose we are running RSD and it is student $s$'s turn. There should be at least $m - \frac{i-1}{\overline{q}}$ schools which are not full yet. Pick a subset $C$ of such schools so that $|C| = \lceil m - \frac{i-1}{\overline{q}} \rceil$. Flip a coin, and with probability $\frac{|C|}{|\mathcal{C}|}$ (i.e., in case of a successful coin flip) let $s$'s first choice be one of the members of $C$ selected uniformly at random. Otherwise, in case of an unsuccessful coin flip, let $s$'s first choice be one of the schools in $\mathcal{C} \backslash C$, uniformly at random. By the definition of the coin flipping process, it is clear that the number of successful coin flips is at most the number of students who get their top choice in any sample path. Therefore, the number of students who get their top choice stochastically dominates the number of successful coin flips. So, we can prove the second claim in the lemma by applying Chernoff bound on the number of successful coin flips. $\square$

# E Examples

## E.1 An example with cardinal utilities

Consider the model in Section 2 with $n$ students and $m = n$ schools. In addition, suppose that the utility that of a student from being assigned to her $i$-th choice is $\frac{1}{i}$; if the student is unassigned her utility is 0. Let $u_n(\pi)$ denote the expected utility of a student when the tiebreaking rule $\pi$ is used. (The expectation is taken over students' preferences and the schools' priorities defined by the tiebreaking rule.) Let $u(\pi) = \lim_{n\to\infty} u_n(\pi)$.

Similarly, let $u_{n,k}(\pi)$ denote the expected utility of a student for when the preference lists of students are shortened to contain only $k$ schools (i.e. the preference list of each student has length $k$). Let $u_k(\pi) = \lim_{n\to\infty} u_{n,k}(\pi)$.

**Proposition E.1.** *For any $k > 0$, $u_k(MTB) > 0$. However, $u(MTB) = 0$.*

*Proof.* Fix $k > 0$. First, we show that $u_k(MTB) > 0$. When the length of the preference lists are $k$, the expected number of students that are assigned under MTB is at least $n(1 - (1 - 1/n)^n)$. (This is the expected number of seats that are filled in the first round of DA.) Observe that each of the assigned students receive a utility at least $1/k$, and therefore

$$u_k(MTB) \geq \lim_{n\to\infty} \frac{n(1-(1-1/n)^n)}{n} \cdot \frac{1}{k} \geq \frac{1-1/e}{k},$$

which is a positive constant.

Next, we show that $u(MTB) = 0$. Fix an arbitrary positive integer $i$. By the main theorem, under MTB only a vanishing fraction of students are assigned to one of their top $i$ choices when the preference lists are complete. Therefore,

$$u(MTB) \leq \lim_{n\to\infty} \frac{o(n) \cdot 1 + n/i}{n}$$

where the first summand in the numerator on the right-hand side is an upper bound on the utility of students who are assigned to one of their top $i$ choices and the second summand is an upper bound on the utility of the rest of the students. Note that the above inequality holds for any positive integer $i$, which implies that $u(MTB) = 0$. $\square$

As discussed in Section 3.1 our Main Theorem also holds true when preference lists are short, as long as they grow at a nonzero rate. Using this generalization we can prove a stronger version of the second part of Proposition E.1: Suppose the length of the preference lists, $k(n)$, is a growing function of $n$ (for example $\log n$). Then, it can be shown that $\lim_{n\to\infty} u_{n,k(n)}(MTB) = 0$. The proof follows the same steps as the proof in Proposition E.1.

For completeness, we also provide a counterpart of the above proposition for STB.

**Proposition E.2.** *For any $k > 0$, $u_k(STB) > 0$. Furthermore, $u(STB) > 0$.*

*Proof.* Under STB, so long as the length of preference lists are positive, at least a fraction $n/2$ of the students are assigned to their first choice, in expectation. To see why, note that running DA with STB is equivalent to running the Random Serial Dictatorship (RSD) mechanism. Under RSD, the $j$-th student who chooses her favorite school can choose her first choice with probability $\frac{n-j+1}{n}$. Therefore, the expected number of students who are assigned to their first choice is equal to $\sum_{j=1}^{n} \frac{n-j+1}{n}$, which is at least $n/2$. Note that this holds regardless of the length of students' preference lists. Therefore, $u_k(STB) > 1/2$ for any $k > 0$, and $u(STB) > 1/2$. $\square$

## E.2 MTB may assign more students to their top choice

STB typically assigns more students to their first choice than MTB does: this holds in our model where students' preferences are drawn uniformly at random; also, this pattern is observed in New York City (Abdulkadiroğlu et al., 2009) and the city of Amsterdam (De Haan et al., 2015).) This, however, is not always the case.

**Remark 3.** *The following example shows that the expected number of students assigned to their first choice can be larger under MTB than under STB. There are 5 schools with capacities $q_1 = 40$, $q_2 = 10$, $q_3 = 500$, $q_4 = 5000$, and $q_5 = 20000$. There are 4 types of students. There are 50 students of type 1 with preference $c_1 > c_2 > c_3 > c_4 > c_5$, 10 students of type 2 with preference $c_2 > c_5$, 500 students of type 3 with preference $c_3 > c_4 > c_5$, and 5000 students of type 4 with preferences $c_4 > c_5$. Computer simulations show that the expected fractions of students who obtain their first choice under STB and under MTB are 0.9951 and 0.9955, respectively.*[24]

*The main idea behind this example is the following. Run the student-proposing DA such that in the first round all students propose to their first choice, and after that students propose sequentially where the last rejected student is always the next one to propose.*[25] *We argue roughly that after the first round, more students are likely to be rejected under STB than under MTB. Observe that after the first round, all students of types 2 and 3 and 4 are tentatively accepted to their first choice and 10 students of type 1 are rejected from their first choice, $c_1$. These rejected students are less likely to be accepted to their second choice, $c_2$, under STB than under MTB. Furthermore, each student who is accepted to $c_2$ after the first*

[24] The averages are taken over $i$ iterations. For any large enough $i$ (e.g., for $i \geq 10^4$), the averages coincide with the two numbers mentioned above, regardless of the random seed initializer.

[25] It is well known that the order in which proposals are made does not affect the outcome.

*round causes a rejection of a different student from $c_2$ who is in turn accepted at school $c_5$. Finally, a student of type 1 who is accepted to $c_3$ after the first round is likely to trigger the rejection of two students (one student from $c_3$, who in turn is likely to cause the rejection of a student from $c_4$).*

*This example is carefully crafted, and the gap in favor of MTB is very small. We believe, however, that more students are assigned to their first choice under STB than under MTB under more general assumptions about preferences (for example, when preferences are generated from a symmetric multinomial logit model).*