What matters in tie-breaking rules? How competition guides design PRELIMINARY DRAFT

Itai Ashlagi Afshin Nikzad*

August 2015

Abstract

School districts that adopt the Deferred Acceptance (DA) mechanism to assign students to schools face the tradeoff between fairness and efficiency when selecting how to break ties among equivalent students. We analyze a model with with random generated preferences for students and compare two mechanisms differing by their tie-breaking rules: DA with one single lottery (DA-STB) and DA with a separate lottery for each school (DA-MTB). We identify that the balance between supply and demand in the market is a prominent factor when selecting a tie-breaking rule. When there is a surplus of seats, we show that neither random assignments under these mechanisms stochastically dominates each other, and, the variance of student's assignments is larger under DA-STB. However, we show that there is essentially no tradeoff between fairness and efficiency when there is a shortage of seats: not only that DA-STB (almost) stochastically dominates DA-MTB, it also results in a smaller variance in student's rankings. We further find that under DA-MTB many pairs of students would benefit from directly exchanging assignments ex post when there is a shortage of seats, while only few such pairs exist when there is a surplus of seats. Our findings suggest that it is more desirable that "popular" schools use a single lottery over a separate lottery in order to break ties, while in other schools there is a real tradeoff.

1 Introduction

A growing number of cities that use a centralized assignment mechanism to assign students to schools adopted the Deferred Acceptance (DA) algorithm by Gale and Shapley (1962).[1] Their

*Ashlagi: Sloan, MIT, iashlagi@mit.edu, Nikzad: Economics Department, MS&E Department; Stanford. nikzad@stanford.edu. We thank Parag Pathak and Alvin Roth for helpful discussions. Ashlagi acknowledges the research support of the National Science Foundation grant SES-1254768.

[1]See Abdulkadiroğlu et al. (2009) and Abdulkadiroğlu et al. (2005) for redesigns of school choice mechanisms in NYC and Boston.

algorithm is designed to find a stable matching in a two-sided matching market, where participants on each side of the market have preferences over participants on the other side. However, while schools are often not strategic, schools' preferences are generated artificially in order to break ties among equivalent students (Abdulkadiroğlu and Sönmez (2003)).

Two tie-breaking rules have been considered by school districts: single and multiple. Under the multiple tie-breaking rule (MTB), each school independently selects a random order over students for breaking ties, and under the single tie-breaking rule (STB) all schools use the same order over students, selected uniformly at random. The question of which tie-breaking rule to pair with the deferred acceptance mechanism was raised by Abdulkadiroğlu and Sönmez (2003), who suggested that deferred acceptance paired with MTB (DA-MTB) may result in unnecessary inefficiency. While both DA-STB and DA-MTB mechanisms make it safe for students to reveal their true preferences, they naturally produce very different assignments.

Abdulkadiroğlu et al. (2009) and De Haan et al. (2015) find, using empirical data and simulations, that neither random assignments under DA-STB or DA-MTB stochastically dominate each other. They further find that the former assigns more students to their top choices but also more students obtain low ranked choices or even remain unassigned. Their findings have influenced policy makers: NYC adopted DA-STB as policy makers leaned towards more students getting their top choices, while, recently, Amsterdam adopted DA-MTB stating equity to be the key reason (as De Haan et al. (2015) write: "it spreads the 'pain' of a shortage of places at some popular schools, more equally".) In this paper we argue that these previous studies, which provide evidence in a high level, leave much room for taste. We provide novel insights based on market characteristics that alleviate much of the ambiguity, and hopefully would guide more informed decisions.

We study the impact of tie-breaking rules on students' assignments, and identify that competition, or the balance between supply in demand in the market, plays a crucial role in the design of such rules. (One reason to study markets with a shortage of seats is to restrict attention to the sub-market with popular schools.) To address this, we look at randomly generated school choice problems, where students' preferences are drawn uniformly at random and study the rank distribution of students' assignments under DA-STB and DA-MTB. Motivated by practical concerns, we study both efficiency and fairness of these mechanisms. To address efficiency we adopt the notion of stochastic dominance and for fairness we adopt the notion of "expected variation in students' ranks" (which also corresponds to the variance of a student's rank in our setting).

Our main contribution is showing that in an overdemanded market there is essentially no tradeoff between selecting efficiency and fairness, as DA-STB provides better outcomes with respect to both measures. However, in an underdemanded market, we find that MTB is more fair according to our measure of fairness (expected variation in students' ranks is significantly larger under DA-STB than under DA-MTB), and that neither random assignment is significantly more efficient than the other; in particular, neither stochastically dominates the other and both attain the the same average rank for students.

Interestingly, after the first assignment in Amsterdam has been completed, parents filed a lawsuit since students who wished to exchange seats were banned from doing so. This story and work by De Haan et al. (2015) motivates studying how common are *Pareto improving pairs*, which are pairs of students that wish to exchange seats after the assignment occurs.[2] Yet again we find the imbalance in the market to play a crucial role: when there is a shortage of seats, DA-MTB generates many Pareto improving pairs but when there is a surplus of seats, very few Pareto improving pairs exist. Observe that under DA-STB, the assignment is Pareto optimal.[3]

Formally, we consider a school choice problem with n students and m schools each of which has a single seat. For each student we independently draw a complete preference list over schools uniformly at random.[4] Consider first the case with $n+1$ schools. We find that (i) neither random assignment stochastically dominates the other, (ii) the ratio between the expected variation in students' rank under DA-MTB and that under DA-STB tends to 0, and (iii) the number of Pareto improving pairs under DA-MTB converges to 0. Next, we consider the case with $n-1$ schools. We find that (i) with high probability the random assignment under DA-STB (almost) stochastically dominates the random assignment under DA-MTB, (ii) the ratio between the expected variation in students' rank under DA-MTB and the one under DA-STB tends to infinity, and (iii) the number of Pareto improving pairs under DA-MTB converges to infinity and, for every student s, with probability converging to 1, there exist (many) other students with whom she forms a Pareto improving pair. We further explore how our insights extend to larger imbalances in the market and provide simulation results to test their robustness.

Our results complement findings by Ashlagi et al. (2013) who analyze how students rank on average: with $n+1$ schools, on average students' rank is the same under both mechanisms, but with $n-1$ schools, on average students rank significantly better under DA-STB.

Our findings imply that in a random market with a shortage of seats, using DA-STB results in better outcomes according to the properties we considered. Thus, in contrast Amsterdam's

[2]There may be Pareto improving cycles with more than two students. However, we limit ourselves to pairs of students since we believe that for a given student it is considerably simpler to find another student who is interested in exchanging seats, than to identify an indirect exchange through a cycle with at least two other students.

[3]DA-STB is equivalent to the Random Serial Dictatorship, under which each students chooses seats one by one in random order.

[4]Assuming unit capacities and complete preference lists simplify our analysis, however, we will discuss the robustness of our findings to these assumptions.

reasoning, our results suggest that, especially for popular schools, one should adopt a single tie-breaking rule. In a market with a surplus of seats, there is more ambiguity as both mechanisms lead to the same average rankings, but each mechanism has a different advantage. DA-MTB is more equitable than DA-STB, but more students receive their top choices under the latter (Ashlagi et al. (2015) and Arnosti (2015)). Thus our results suggest that it is more desirable that "popular" schools break ties using a single lottery over a separate lottery, while in non-popular schools there there is a real tradeoff between using the same lottery or a separate lottery.

To gain some intuition it is useful to consider a random market with just two schools, each of which has the large and identical capacity.[5] If there is a surplus of seats both mechanisms will generate very similar outcomes since almost every student will get his first choice.[6] Suppose that there are not enough seats for all students. A student who is rejected from her first choice is likely to have a bad draw (due to the large capacity) in that school, and is therefore less likely to be accepted at her second choice under DA-STB than under DA-MTB. Therefore, students who are temporarily assigned to their top choice are also more likely to maintain their seat under DA-STB than under DA-MTB. Therefore, with a shortage of seats, not only DA-STB stochastically dominates DA-MTB, but also many students get their second choice under DA-MTB, implying that there will be many Pareto improving pairs. We obtain our results in a model with unit capacities, which can be viewed as stronger results since with unit capacities much less information is revealed about the students who are rejected from their top choices under DA-STB. We provide further intuition through the paper (and see Section 7 for further discussion).

1.1 Related Work

Using NYC data and simulations, Abdulkadiroğlu et al. (2009) explored the rank distribution of single and multiple-tie breaking rules. They find that there is no stochastic dominance among the two random assignments, that more students receive top choices under DA-STB than under DA-MTB, and that the rank distributions under these mechanisms have a single crossing point. De Haan et al. (2015) conduct a similar study in the context of Amsterdam and finding similar patterns. Ashlagi et al. (2015) explain the differences regarding top choices in a model similar to ours. Independently to this paper, Arnosti (2015) explains the single crossing point pattern using a cardinal model. In his model, students' preference lists are bounded, which is essentially equivalent to a market with a large surplus of seats. Our results alleviate a lot of the ambiguity for

[5] It may be convenient to think about a continuum model with a student mass of 2 and a capacity of mass 1 in each school as in Azevedo and Leshno (2010).

[6] If the capacity in each school is q, only one school will be overdemanded by approximately $\sqrt{q}$ students and all rejected students will be assigned to their second choice, which is underdemanded.

selecting a tie-breaking rule, by distinguishing between overdemanded and underdemanded markets and showing that in overdemanded markets there is essentially no tradefoff.

Pathak and Sethuraman (2011) and Carroll (2014) extend results by Abdulkadiroğlu and Sönmez (1998) to show that under the Top Trading Cycles mechanism (Shapley and Scarf (1974)), there is no difference between single tie-breaking (equivalently, Random Serial Dictatorship) and multiple tie-breaking (top trading cycles with random endowments). The tradeoff between incentives and efficiency when preferences contain indifferences has led to different approaches, among which are the stable improvement cycles of Erdil and Ergin (2008), the efficiency-adjusted DA of Kesten (2011) , the choice-augmented DA of Abdulkadiroglu et al. (2008), and the circuit tie-breaker by Che and Tercieux (2014).

Finally, our work contributes to a growing literature that investigates properties assignments in random two-sided matching markets. Pittel (1989) studied the mens' average rank in a balanced stable marriage model with random preferences and found a very large gap between the men and woman optimal stable matchings. Our approach is motivated by Ashlagi et al. (2013) who show that when the market is unbalanced, agents on the short side are matched, on average, to at most their $\ln n$-th rank under any stable matching. This paper adds to this literature by studying the variance of a student's rank and the frequency of Pareto improving pairs, as well as proving concentration results for previously studied random variables (such as the the number of proposals made or received by a fixed agent).

2 The Model

In a school choice problem there are n students, each of whom can be assigned to one seat at one of m schools. Denote the set of students by S and the set of schools by C. Each student s has a strict preference ranking over all schools. Let the rank of a school c for student s be the number of schools that s weakly prefers to c. Thus the most preferred school for s has rank 1. A *random school choice problem* is generated by drawing a complete preference list for each student over schools independently and uniformly at random.

Each school $c \in C$ has a priority ranking over students that is used to break ties between students in the same priority class. We assume each school has a single priority class containing all students, and thus each school uses a single ordering to break ties between students. We also refer to a school's priority ranking over students by its preference list.

Our model assumes a a single priority class and unit capacities; see the Conclusion (Section 7) for further discussion regarding these assumptions.

A *matching* of students to schools assigns each student to at most one seat and at most one student to each school. Stability is often used as a solution concept in two-sided matching markets. A matching is said to be *unstable* if there is a student s and a school c such that s prefers to be assigned to c over his current assignment, and c either has a vacant seat or a student with lower priority than s. A matching is said to be *stable* if it is not unstable.

Gale and Shapley (1962) showed that a stable matching always exists and, under strict preferences, there is one matching that is weakly preferred by all students, called the student-optimal stable matching. They further proposed the student-proposing deferred acceptance algorithm that computes this matching with respect to students' revealed preferences and school priority rankings.

2.1 Tie-breaking rules

We consider two common tie-breaking rules that school districts use to determine priority rankings of schools over students. Under a *multiple tie-breaking rule* (MTB) each school independently selects a priority ranking over all students. Under a *single tie-breaking rule* (STB) all schools use the same priority order, which is selected uniformly at random. One way of implementing STB is to assign to each student a lottery number drawn independently and uniformly at random from $[0, 1]$. Similarly, MTB can be implemented in a similar fashion by using a different draw (and thus different lottery numbers) for each school. Naturally, MTB and STB lead to different outcomes.

2.2 Notations

We will use the following definitions. Denote by $\mathcal{A} = S \cup C$ the set of schools and students. We often refer to a school or a student by an *agent*. Consider a matching γ. Let $\gamma(x)$ be the agent to which x is matched to and for any subset of agents $A \in \mathcal{A}$, let $\gamma(A)$ be the set of agents matched to agents in A. Therefore, $\gamma(C)$ is the set of students who are assigned under γ.

For any student s, $\gamma^{\#}(s)$ denote the rank of school $\gamma(s)$ for s, and use similar notions for schools. Denote the average rank of students who are assigned under γ by $\mathcal{Ar}(\gamma) = \frac{1}{\gamma(C)} \cdot \sum_{s \in \gamma(C)} \gamma^{\#}(s)$. When it is clear from the context we will simply write r_s for $\gamma^{\#}(s)$, and r for $\mathcal{Ar}(\gamma)$.

We denote by μ_π and η_π the student-optimal and the school-optimal stable matching for a preference profile π, respectively. Finally, given students' preferences, let μ_{STB} and μ_{MTB} the random variables that denote the student-optimal stable matchings under STB and MTB, respectively.

Suppose $f, g : \mathbb{Z}_+ \to \mathbb{R}_+$. We use the notation $g = o(f)$ when $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$, $g = O(f)$ when $f \neq o(g)$, $g = \Theta(f)$ when $f = O(g)$ and $g = O(f)$, and finally, $g = \Omega(f)$ when $f = O(g)$.

2.3 Preliminary findings

The next result compares the average rankings of students in a random school choice problem under each of the tie-breaking rules: with a shortage of seats, the average rank for students is significantly better under the STB rule than under the MTB rule, but, when there is surplus of seats the average rankings under these tie-breaking rules are the same. The following fact follows from Ashlagi et al. (2013) and Knuth (1995).

Fact. *Consider a sequence of random school choice problems with n students and m schools.*

1. *If $n \leq m$,*
$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{STB}})\right]}{\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{MTB}})\right]} = 1.$$
2. *If $n > m$,*
$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{STB}})\right]}{\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{MTB}})\right]} = 0.$$

The first part follows from Ashlagi et al. (2013) who show that with a surplus of k seats, $\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{STB}})\right] = \mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{MTB}})\right] \approx (n+k)\ln\left(\frac{n+k}{k}\right)$.[7] The second part follows directly from Knuth (1995) and Ashlagi et al. (2013). Indeed Knuth (1995) shows that when $n = m$, $\mathcal{A}r(\mu_{\mathsf{STB}}) \approx \ln n$, implying that when $n > m$, $\mathcal{A}r(\mu_{\mathsf{STB}}) \approx \ln m$. Ashlagi et al. (2013) show that that when there is a shortage of seats, the average rank is at most $(1+\epsilon)\frac{n}{\log n}$.

the above fact suggests that the imbalance in the market maybe a crucial factor in selecting a tie-breaking rule, but only provides insights on how students rank on average under STB and MTB. In the next three sections we pursue this direction by analyzing properties of the random assignments under these tie-breaking rules distinguishing between overdemanded and underdemanded markets.

First, we investigate whether there is stochastic dominance between the assignments in Section 3 (and thus revisit previous empirical and simulation studies by Abdulkadiroğlu et al. (2009), De Haan et al. (2015)). Second, we compare the random assignments by analyzing the variance of a student's rank in Section 4. Third, we study the frequency of Pareto improving pairs (pair of students who wish to switch schools), under MTB in Section 5. Finally, in Section 6 we provide simulation results.

3 Stochastic dominance

We investigate in this section whether there is a stochastic dominance relation between the assignments under DA-STB and DA-MTB. Our main finding is that when there is a shortage of seats,

[7] Although they show the result for $k \geq 0$, similar arguments hold for the balanced case.

the rank distribution under DA-STB (almost) stochastically dominates the one under DA-MTB. However, there is no such stochastic dominance between DA-STB and DA-MTB when there is a surplus of seats.

Before we state our results, we formalize the definitions. The *rank distribution of a subset of students S* is a function $\mathcal{R} : [1, m] \to [0, n]$ where $\mathcal{R}(i)$ denotes the number of students in S who are assigned to their (top) i-th choice in their preference list. Fix a constant $\epsilon > 0$. We say that a rank distribution $\mathcal{R}$ *stochastically dominates* rank distribution $\mathcal{R}'$ if, for any integer $i \in [1, m]$, $\sum_{j=1}^{i} \mathcal{R}(j) \geq \sum_{j=1}^{i} \mathcal{R}'(j)$. Furthermore, a rank distribution $\mathcal{R}$ *almost stochastically dominates* a rank distribution $\mathcal{R}'$ if, for any integer $i \in [1, m]$, either $\sum_{j=1}^{i} \mathcal{R}(j) \geq \sum_{j=1}^{i} \mathcal{R}'(j)$ or $\sum_{j=i}^{m} \mathcal{R}(j) \leq (\ln n)^{1+\epsilon}$. [8]

An intuitive way to think about *almost stochastic dominance* is the following. First remove the bottom $(\ln n)^{1+\epsilon}$ students (students who are assigned to their lowest preferences) from $\mathcal{R}'$; let $\overline{\mathcal{R}'}$ denote the resulting rank distribution. Then, $\mathcal{R}$ stochastically dominates $\overline{\mathcal{R}'}$ if and only if $\mathcal{R}$ almost stochastically dominates $\mathcal{R}'$. Similarly, let $\overline{\mathcal{R}}$ denote the rank distribution resulting from removing the bottom $(\ln n)^{1+\epsilon}$ students from $\mathcal{R}$; then, $\overline{\mathcal{R}}$ stochastically dominates $\overline{\mathcal{R}'}$ if and only if $\mathcal{R}$ almost stochastically dominates $\mathcal{R}'$.

Finally we define for any given set of students S with fixed known preferences the random variables $\mathcal{R}_{\mathsf{MTB}}(S), \mathcal{R}_{\mathsf{STB}}(S)$ as follows. Let $\mathcal{R}_{\mathsf{MTB}}(S)$ and $\mathcal{R}_{\mathsf{STB}}(S)$ be the rank distributions corresponding to a random draw of a MTB and STB, respectively. When S is clearly known form the context, we often denote these rank distributions by $\mathcal{R}_{\mathsf{MTB}}$ and $\mathcal{R}_{\mathsf{STB}}$. We are now ready to state our main result.

Theorem 3.1. *Consider a sequence of random school choice problems with n students and m schools with $n = m + 1$. Then, with high probability , $\mathcal{R}_{\mathsf{STB}}$ almost stochastically dominates $\mathcal{R}_{\mathsf{MTB}}$.*[9]

We give here the main idea for the proof, which is given in Section A. While in the introduction we gave intuition for markets with large capacities, the proof for Theorem 3.1 is much more involved as almost no information is revealed about a student who is rejected from her top choices. The main idea is to show that a much smaller fraction of students are assigned to one of their top choices under MTB than under STB. In particular, we show that there exist a rank r for which, whp, all students but at most $(\ln n)^{1+\epsilon}$ of them are assigned to a rank better than r under DA-STB, whereas at most $0.4n + o(n)$ students are assigned to a rank better than r under DA-MTB. In addition, it is not hard to see that about $0.5n$ students receive their top choice under DA-STB, with high

[8]Our findings hold for any (arbitrary small) constant $\epsilon > 0$.

[9]For a sequence of events $\{E_n\}_{n\geq 0}$, we say this sequence occurs *with high probability* (whp) if $\lim_{n\to\infty} \mathbb{P}[E_n] = 1$.

probability. These two facts together directly imply the theorem. It is worth pointing out that the analysis involves concentration results that complement the results of Ashlagi et al. (2013) on the expected rank of students.

Next, we show that almost stochastic dominance does not occur when there are more seats than students.

Theorem 3.2. *Consider a sequence of random school choice problems with n students and m schools with $n = m - 1$. Then, with very high probability, $\mathcal{R}_{\mathsf{STB}}$ does not almost stochastically dominate $\mathcal{R}_{\mathsf{MTB}}$.*[10] *Furthermore, wvhp $\mathcal{R}_{\mathsf{STB}}$ does not stochastically dominate $\mathcal{R}_{\mathsf{MTB}}[k]$ for any $k = o(n/ln^2 n)$, where $\mathcal{R}_{\mathsf{MTB}}[k]$ is the rank distribution resulting from the removal of the bottom k students from $\mathcal{R}_{\mathsf{MTB}}$.*

Theorem 3.2 suggests, that the lack of competition makes the decision between tie-breaking more ambiguous. Arnosti (2015) independently shows that there is no stochastic dominance between DA-STB and DA-MTB in a random market where there is an essentially a large surplus of seats. In fact he is able to show that the cumulative rank distribution has a single crossing point.

Our main result is that in a random market with a shortage of seats, DA-STB is more efficient than DA-MTB. While Theorem 3.1 shows almost stochastic dominance, we conjecture that $\mathcal{R}_{\mathsf{STB}}$ stochastically dominates $\mathcal{R}_{\mathsf{MTB}}$. Our conjecture is confirmed by the computational experiments that we present in Section 6.

4 Fairness

In this section we compare the random assignments under DA-STB and DA-MTB from an equity persecutive. For this purpose we adopt two measures of fairness, which are shown to be equivalent. Consider a matching μ. Define the *social inequity* of matching μ to be

$$\mathcal{S}i(\mu) = \frac{1}{|\mu(C)|} \cdot \sum_{s \in \mu(C)} (\mathcal{A}r(\mu) - \mu^{\#}(s))^2.$$

Another natural measure is the variation in the ranks of the schools that a given student could get assigned to. Define the *expected variation in the rank of student s* to be

$$\mathsf{Var}\,[r_s] = \mathbb{E}_{\{\pi(c):c\in C\}} \left[(\mathcal{A}r(\mu_\pi) - \mu_\pi^{\#}(s))^2 \middle| \mu_\pi(s) \neq \emptyset \right],$$

[10] For a sequence of events $\{E_n\}_{n\geq 0}$, we say the sequence occurs *with very high probability* (wvhp) if $\lim_{n\to\infty} \frac{1-\mathbb{P}[E_n]}{\exp\left(-(\ln n)^{0.4}\right)} = 1$.

where the expectation is taken taken over schools' preferences (which are generated by either the STB or MTB rule).

The next lemma shows that, when students' preference are iid, the expected social inequity is equal to the expected variation in the rank of a given student. The proof appears in Section B.1.

Lemma 4.1. *For any student* $s \in S$

$$\mathbb{E}_{\{\pi(s'):s'\in S, s'\neq s\}}\left[\mathit{Var}[r_s]\right] = \mathbb{E}_{\pi}\left[\mathcal{S}i(\mu_\pi)\right],$$

where expectation on the LHS is taken over all students' preferences except s, and expectation on the RHS is taken over all students' and schools' preferences with schools' preferences generated by either the STB or the MTB rule.

The next theorem shows that the imbalance in the market determines whether DA-MTB or DA-STB results in a larger social inequity. Intuitively DA-MTB seems to be more equitable than DA-STB. However, we find, surprisingly, that when there is a shortage of seats it is significantly less equitable than DA-STB.

Theorem 4.2. *Consider a sequence of random school choice problems with n students and m schools.*

1. *If* $n = m$ *or* $n = m - 1$*, then* $\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{MTB}})]} = \infty$.
2. *If* $n = m + 1$*, then* $\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{MTB}})]} = 0$.[11]

Theorem 4.2 follows directly from the next result, which quantifies the social inequities in our model.

Lemma 4.3. *Consider a sequence of random school choice problems with n students and m schools.*

1. *If* $n = m + 1$*, the expected social inequity under MTB is large is of order* $\Omega(\frac{n^2}{\ln^2 n})$ *and the expected social inequity under STB is* $\Theta(n)$*.*

2. *If* $n = m$*, the expected social inequity under MTB is* $O(\ln^4 n)$*, and the expected social inequity under STB is* $\Theta(n)$*.*

3. *If* $n = m - 1$*, the expected social inequity under MTB is* $O(\ln^2 n)$ *and the expected social inequity under STB is* $\Theta(n)$*.*

[11] Expectations are taken over students' preferences and the tie-breaking rules.

The proof for Lemma 4.3 is involved and is given in Section B. We provide here a proof sketch for the second part, which is rather simple.

Indeed assume $n = m$. Pittel (1989) shows that high probability no student does very bad: $\max_{s \in S} \mu^{\#}_{\mathsf{MTB}}(s) \leq 3 \ln^2 n$. Therefore, with high probability

$$\frac{1}{n} \cdot \sum_{s \in S} (\mathcal{A}r(\mu_{\mathsf{MTB}}) - \mu^{\#}_{\mathsf{MTB}}(s))^2 \leq 9 \ln^4 n,$$

implying that the expected social inequity under MTB is $O(\ln^4 n)$.

Next we compute a lower bound on the expected social inequity in STB. With probability at least 1/2, the student with the lowest priority number in STB gets assigned to a school that she has ranked on lower half of her preference list. So, for any student $s \in S$ we can write:

$$\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] = \mathbb{E}\left[\mathsf{Var}\left[r_{\underline{s}}\right]\right] \geq \frac{1}{n} \cdot \left(\mathcal{A}r(\mu^{\#}_{\mathsf{STB}}(s)) - n\right)^2.$$

Moreover Knuth (1995) showed that $\mathcal{A}r(\mu^{\#}_{\mathsf{STB}}(s)) = \Theta(\ln n)$. Plugging this into the above inequality implies $\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] \geq \Omega(n)$. Consequently, in this regime, the expected social inequity is much higher under STB than under MTB. Lemma B.3 shows that the bound for $\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right]$ is tight, i.e. $\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] = \Theta(n)$.

4.1 Robustness to large imbalances and short preference lists

Theorem 4.2 is robust to imbalances larger than 1. When $n > m + 1$, then the expected social inequality for STB remains the same as when $n = m + 1$, however, the expected social inequality for MTB remains as least as much as $n = m + 1$ (this is a direct implication from the proof of Lemma 4.3, part 1). Thus, part 1 of Theorem 2 always holds as long as $n > m$.

When $n < m$, as the surplus of seats grows large, one should expect that the social inequity decrease under STB as more students will match well. Next, we show that social inequity remains larger under STB than under MTB even when the surplus of seats is of the order of the number of students. (See Section B.3 for the proof)

Theorem 4.4. *Suppose $m = n + \lambda n$ for some positive λ. Then, $\lim_{n \to \infty} \frac{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{MTB}})]} > 1$, where the expectations are taken over preferences and the tiebreaking rules.*

We quantify the ratio between social inequities for different values of λ in our computational experiments in Section 6. In another set of experiments in this section, we show that the gap between social inequities is persistent even when the preference lists are significantly short (see

Section 6.3).

5 Pareto improving pairs

It is well-known that the student-optimal deferred acceptance mechanism paired with the STB rule is Pareto efficient.[12] The student-optimal deferred acceptance mechanism paired with the MTB can result in inefficiencies due to the different lottery numbers a student receives in different schools. In the section we explore Pareto improvements by pairs of students. The reason for concentrating on Pareto improving pairs and not longer cycles is that is rather easier for a student to identify a direct exchange with another student than an indirect one through a longer cycle.

Consider a matching γ. A pair of students $s, s' \in S$ is called a *pareto improving pair in* γ if $\gamma(s') \succ_s \gamma(s)$ and $\gamma(s) \succ_{s'} \gamma(s')$. Note that this definition ignores the lottery numbers of students in the schools. Further, define $\ddot{\gamma}(s)$ to be the number of Pareto improving pairs in γ which contain s.

We show that deferred acceptance paired with MTB generate many Pareto improving pairs when there is shortage of seats, and very few Pareto improving pairs when there is a surplus of seats.

Theorem 5.1. *Consider a sequence of random school choice problems with n students and m schools, let $\mu = \mu_{\mathsf{MTB}}$, and fix a student s.*

1. *If $n > m$,*

$$\lim_{n\to\infty} \mathbb{P}\left[\ddot{\mu}(s) \geq 1\right] \to 1, \qquad \lim_{n\to\infty} \mathbb{E}\left[\ddot{\mu}(s)\right] \to \infty.$$

2. *If $n < m$,*

$$\lim_{n\to\infty} \mathbb{P}\left[\ddot{\mu}(s) \geq 1\right] \to 0, \qquad \lim_{n\to\infty} \mathbb{E}\left[\ddot{\mu}(s)\right] \to 0.$$

For the overdemanded case, the proof of Theorem C.1 provides a stronger result, by showing that each student is contained in many Pareto improving pairs, whp. To give some intuition, we give a simple back of the envelope proof for the theorem, which ignores the dependencies of the involved random variables in the DA. The full proof becomes intricate since it involves handling these correlations; it is presented in the Section C.

[12]Observe that this mechanism is equivalent to the Random Serial dictatorship mechanism, under which students are randomly ordered, and each student in turn is assigned to her top choice from schools that are yet to be assigned.

Some intuition follows from Ashlagi et al. (2013) and Pittel (1989). When students are on the long side, they match to a low ranked school on their list, implying that each such student could have many potential candidates to form an improving pair with. When students, however, are on the short side, they match to very good choices, leaving few potential candidates to form an improving pair with: Students are roughly assigned, on average, to their $\frac{n}{\log n}$-th choice and to their $\log n$-th choice when they are on the short and on the long side, respectively. If a student s is assigned to her average rank, z, then there are about z students that can potentially form a Pareto improving pair with s. Note that for each other student s',

$$\mathbb{P}\left[s \succ_{s'} \mu_\pi(s')\right] = \frac{z}{n},$$

if we assume that $\pi(s')$ is selected independently uniformly at random after the matches are made, conditioned on having $\mu_\pi(s')$ in the position z of $\pi(s')$. (this, of course, is an assumption made for simplification and does not hold in general). Therefore, the chance that s cannot find a Pareto improving pair is roughly $(1 - \frac{z}{n})^z$, which converges to 1 when $z \geq \frac{n}{\log n}$, but converges to 0 when $z \leq \log n$. Similar intuition follows for the number of improving pairs. The proof, which builds on this intuition, is much more involved since one needs to deal with correlations.

6 Computational experiments

This section presents simulations results that complement our theoretical results. We first consider markets with complete preference lists for students and and varying capacities for schools. After that, we consider markets with short preference lists, and finally, tiered markets where a subset of of schools are preferred by all students over the rest of schools.

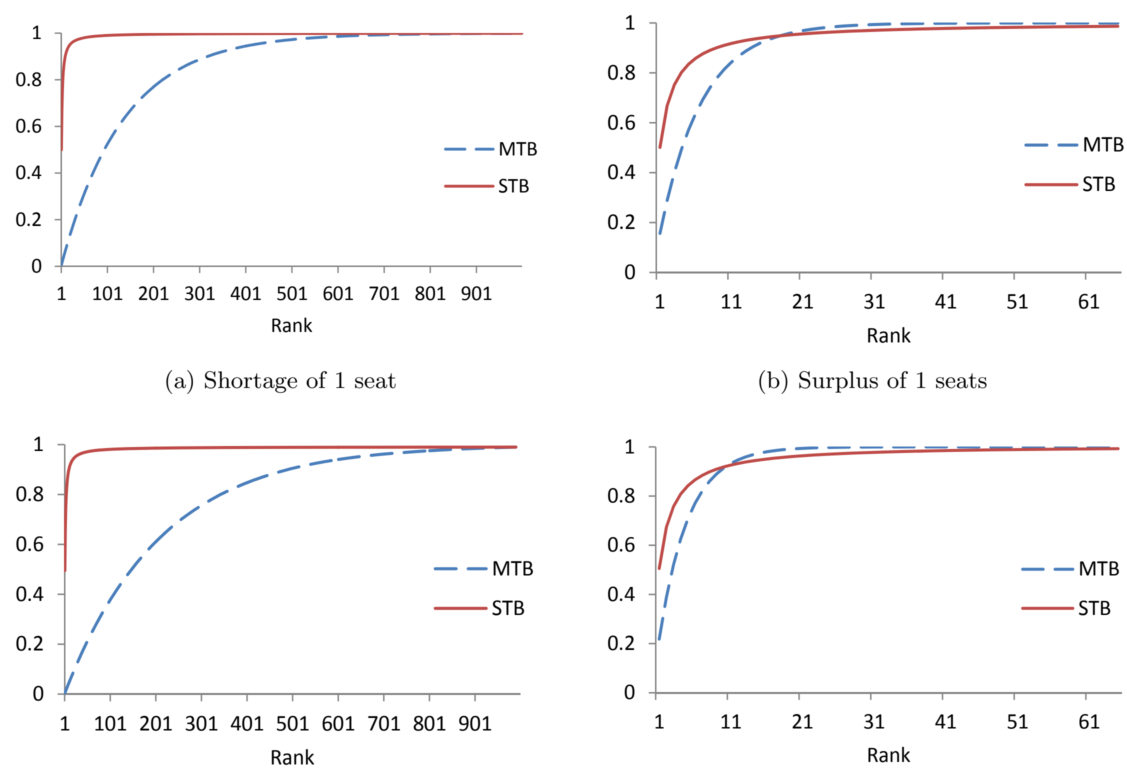
6.1 Numerical results for our model

The first computational experiments illustrates the effect of the imbalance in the market on the students' rank distributions under DA-STB and DA-MTB and the relationship between the two. For each instance that we consider,[13] 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

[13] An instance contains the information regarding market characteristics (size, capacities, list length), and the choice of tiebreaking rule.

at random. Then, we compute the student optimal stable matching. The plots and the tables that we present here are generated by taking average over several (between 100 to 1000) samples [14] for each instance.

Figure 1 shows the cumulative rank distribution under each tiebreaking rule in a market with 1000 students. We consider instances with a small imbalance of either 1 or 10 seats, i.e. four different instances with 1000 ± 1 and 1000 ± 10 seats. Each school has one seat (capacity 1). Observe that when there is a shortage of seats (left), the rank distribution under STB stochastically dominates the rank distribution under MTB. When there is a surplus of seats (right), there is no stochastic dominance.



(a) Shortage of 1 seat

(b) Surplus of 1 seats

(c) Shortage of 10 seats

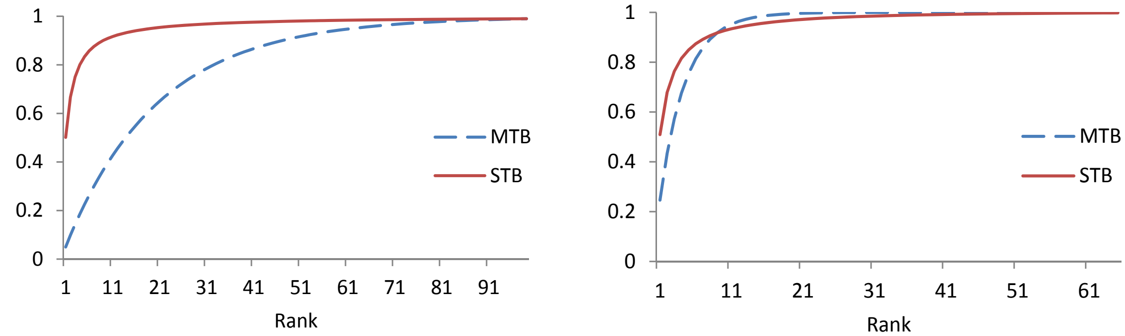
(d) Surplus of 10 seats

Figure 1: The cumulative rank distribution under MTB and STB in random market with 1000 students. Panels 1a and 1c plot the distributions in markets with a shortage of 1 and 10 seats, respectively. Panels 1b and 1d plot the distributions in markets with a surplus of 1 and 10 seats, respectively. The dashed and solid lines indicates the rank distribution under MTB and STB, respectively.

Figure 2 illustrates similar findings for a market with only 100 students, unit capacities, and a

[14] The number of samples were chosen large enough so that increasing this number changes the numerical results insignificantly.

shortage or surplus of a single seat.



(a) Shortage of seats (b) Surplus of seats

Figure 2: The rank distribution under MTB and STB in random market with 100 students with a shortage (left) and surplus (right) of one seat. The dashed and solid lines indicates the rank distribution under MTB and STB, respectively.

Table 1 reports the expected average rank and expected social inequity (or the variance of a student's rank) for markets with varying imbalances where each school has one seat. Observe that the variance of the rank is larger under MTB (than under STB) when there is a shortage of seats and that the variance increases significantly in this case as the shortage grows from 1 to 10. Furthermore, notice that the variance of the rank is smaller under MTB when there is a surplus of seats.

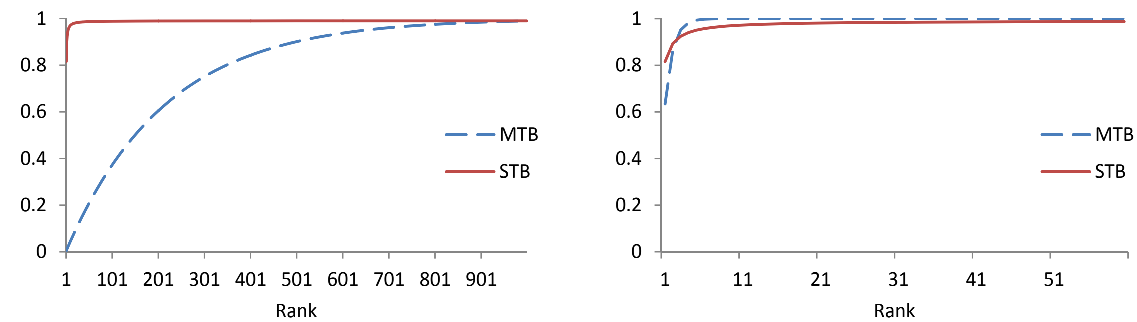
m \ $n-m$		-10	-1	1	10
100	$\mathcal{Ar}(\mu_{\text{STB}})/\mathcal{Ar}(\mu_{\text{MTB}})$	2.52/2.54	3.78/4.1	4.14/29.5	4.23/19.79
	$\mathcal{Si}(\mu_{\text{STB}})/\mathcal{Si}(\mu_{\text{MTB}})$	9.47/3.87	49.8/12.6	69.6/516.9	78.2/322.9
1000	$\mathcal{Ar}(\mu_{\text{STB}})/\mathcal{Ar}(\mu_{\text{MTB}})$	4.53/4.59	6/6.46	4.14/203.5	6.5/136.8
	$\mathcal{Si}(\mu_{\text{STB}})/\mathcal{Si}(\mu_{\text{MTB}})$	144.4/16.51	628.9/35.7	69.6/35780	947/18300

Table 1: Average rank and social inequity under under STB and MTB in the student optimal stable matching for different markets. A student's most preferred rank is 1 and larger rank indicates a less preferred school.

6.2 Robustness to large imbalances and capacities

This section presents simulation results to examine the effect of different imbalances as well as capacities on the random assignments under MTB and STB. We find that for all markets with a shortage of seats, DA-STB stochastically dominates DA-MTB.

Figure 3 shows the rank distribution under each tiebreaking rule in markets with 10000 students. Each school has 10 seats, and there is a total imbalance of 100 seats.



(a) Shortage of seats (b) Surplus of seats

Figure 3: The rank distribution under MTB and STB in random market with 1000 schools where each school has 10 seats and in total there is a shortage (left) or surplus (right) of 100 seats. The dashed and solid lines indicates the rank distribution under MTB and STB, respectively.

Table 3 reports the expected average rank and social inequity for eight markets with imbalances 1 or 100 and school capacities 5 or 10. All schools have the same capacity in each instance; we denote this capacity by q.

m (q) \ $n - qm$		-100	-1	1	100
1000 (5)	$\mathcal{Ar}(\mu_{\mathsf{STB}})/\mathcal{Ar}(\mu_{\mathsf{MTB}})$	1.77/1.77	2.74/2.94	2.86/112	2.86/234.9
	$\mathcal{Si}(\mu_{\mathsf{STB}})/\mathcal{Si}(\mu_{\mathsf{MTB}})$	7.36/1.37	213.6/5.8	280/12429	289.2/44348
1000 (10)	$\mathcal{Ar}(\mu_{\mathsf{STB}})/\mathcal{Ar}(\mu_{\mathsf{MTB}})$	1.57/1.57	2.15/2.25	2.19/104.2	2.19/206.8
	$\mathcal{Si}(\mu_{\mathsf{STB}})/\mathcal{Si}(\mu_{\mathsf{MTB}})$	6.29/0.9	134.7/2.844	166.7/10851	36773/167.5

Table 2: Average rank and social inequity under under STB and MTB in the student optimal stable matching for different markets. A student's most preferred rank is 1 and larger rank indicates a less preferred school.

Figure 4 shows the ratio between $\mathcal{Si}(\mu_{\mathsf{STB}})$ to $\mathcal{Si}(\mu_{\mathsf{MTB}})$ in a market with 10000 students, unit capacities, and the surplust of seats varying from 100 to 1000. Observe that the ratio decreases as the surplus grows because the larger the surplus, the more students will receive their top choices.

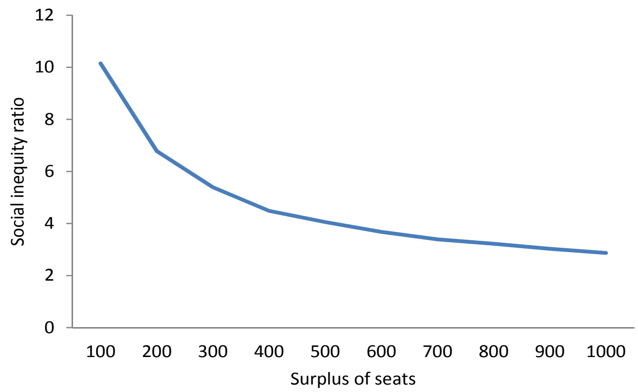


Figure 4: The ratio between $\mathcal{S}i(\mu_{\mathsf{STB}})$ to $\mathcal{S}i(\mu_{\mathsf{MTB}})$ in a random market with 10000 students, unit capacities, and a surplus of seats. (the x-axis denotes the surplus of schools)

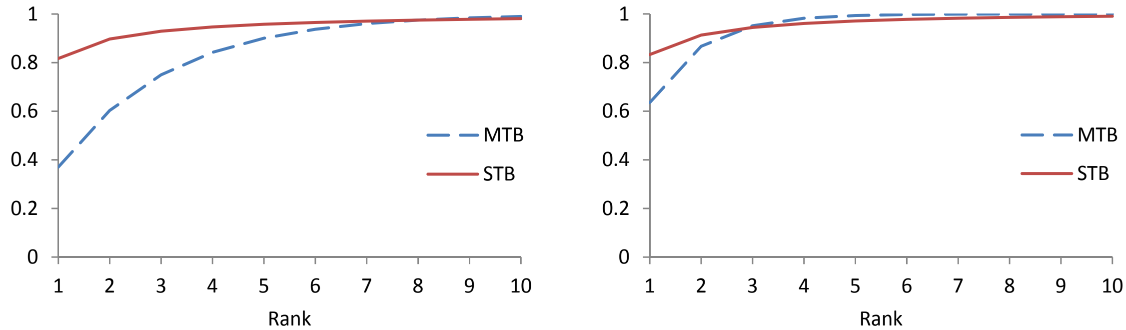
6.3 Short preference lists

In this section, we present simulations to illustrate the effect of shortening the students' preference lists on our results.

Figure 5 presents the rank distribution in random market with 1000 schools, each with capacity with of 10. In addition there are either 10100 or 9900 students, each of which ranks independently uniformly at random 10 schools. (Note that we considered the same instance with complete preference lists in Section 6.2, Table 3). When there is a shortage of seats and the preference lists are complete, our simulations showed that the rank distribution under STB stochastically dominates the rank distribution under MTB; when the preference lists are short, stochastic dominance, not strictly, but "almost" holds.

Shortening the list reduces competition among students (see Ashlagi et al. (2015)), which impacts the market balance, i.e. whether students are "effectively" on the long side or the short side of the market. Therefore, whether there is a surplus or shortage in the market, as the preference lists become shorter, the crossing point of the rank distributions moves to the left (if the crossing happens at all).[15] In overdemanded markets, shortlists and large capacities act as two forces pushing in opposite directions: the former slows down the competition and the latter speeds it up: When the capacities are large in an overdemanded market, MTB creates a much harsher competition relative to when the capacities are small, i.e. rejection chains become longer. In the other hand, under STB, a rejection gives much information about the rejected student's priority number, and thus, that student is less likely to initiate rejection chains. Consequently, as the capacities increase, the crossing point moves to the right (if crossing happens at all).

[15] The extreme case is when the list length is 1, where both distributions become identical.



(a) Shortage of seats (b) Surplus of seats

Figure 5: The rank distribution under MTB and STB in random market with 1000 schools each with 10 seats and in total there is a shortage (left) or surplus (right) of 100 seats. Each student ranks 10 schools. The dashed and solid lines indicates the rank distribution under MTB and STB, respectively.

m (q) \ $n - qm$		-100	100
1000 (10)	$\mathcal{A}r(\mu_{\mathsf{STB}})/\mathcal{A}r(\mu_{\mathsf{MTB}})$	1.36/1.57	1.4/2.6
	$\mathcal{S}i(\mu_{\mathsf{STB}})/\mathcal{S}i(\mu_{\mathsf{MTB}})$	1.24/0.89	1.44/3.59

Table 3: Average rank and social inequity under under STB and MTB in the student optimal stable matching for different markets. A student's most preferred rank is 1 and larger rank indicates a less preferred school.

6.4 Comparison to a hybrid tiebreaking rule

This section provides simulation results for two different tiered markets where some schools are considered as *top* schools and others are considered as *bottom* schools. In these markets every student prefers every top school to every bottom school and the preferences within a tier are drawn independently uniformly at random. Motivated by our findings, we compare three tiebreaking rules: (i) STB, (ii) MTB, and (iii) HTB (Hybrid Tie-Breaking rule), which is tiebreaking rule that uses a single preference list for all of the top schools and an independently drawn preference list for each bottom school.

Example: unit capacity Figure 6 shows the rank distribution under the three tiebreaking rules in a market with 1000 students and 1000 schools, each with unit capacity. We consider 100 schools to be the top schools. Notice that up to rank 100, the STB and HTB graphs agree with each other and are above the one under MTB. Conditioning on being above the 100 rank, MTB and HTB agree with other; there is no stochastic dominance in this range.

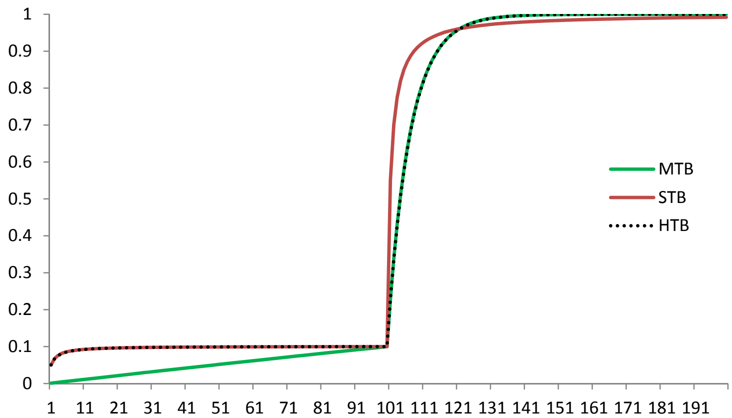


Figure 6: Students' rank distribution under STB, MTB and HTB. The market consists of $n = m = 1000$ and 100 top schools.

We list down the expected average rank and social inequity under the three tiebreaking rules below.

$$\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{STB}})\right] \approx 96.23 \qquad \mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{MTB}})\right] \approx 101.48 \qquad \mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{HTB}})\right] \approx 96.97$$

$$\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] \approx 1752.81 \qquad \mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{MTB}})\right] \approx 422.40 \qquad \mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{HTB}})\right] \approx 1005.34$$

Example: large capacity Figure 7 shows the rank distribution under the three tiebreaking rules in a market with 1000 students, 26 schools each with capacity 50. We consider 5 schools to be the top schools. Observe the same patterns as in the previous example.

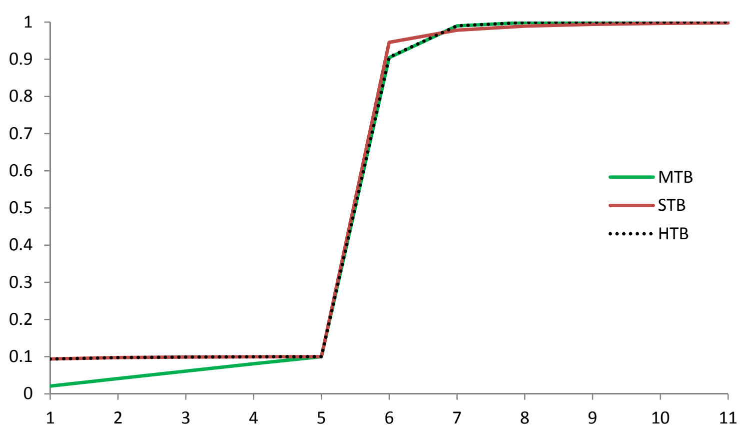


Figure 7: Students' rank distribution under STB, MTB and HTB. The market consists of 1000 students, 26 schools each with 50 seats and 5 top schools.

We list down the expected average rank and social inequity under the three tiebreaking rules below.

$$\mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{STB}})\right] \approx 5.61 \qquad \mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{MTB}})\right] \approx 5.80 \qquad \mathbb{E}\left[\mathcal{A}r(\mu_{\mathsf{HTB}})\right] \approx 5.61$$

$$\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] \approx 2.60 \qquad \mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{MTB}})\right] \approx 1.21 \qquad \mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{HTB}})\right] \approx 2.39$$

7 Discussion

The deferred acceptance algorithm has been increasingly adopted by school districts and how they break ties between students can have a significant impact on students' assignments. Often the decision is between two simple tie-breaking rules: a single lottery for all schools or a separate lottery for each school, where the former is Pareto optimal but the latter is considered to be more fair.

This paper compares these mechanisms in a market where students' preferences are drawn uniformly at random and finds that balance between supply and demand is a prominent factor in the design of tie-breaking rules. Consistent with intuition and empirical evidence, we find that when there is a surplus of seats, neither random assignment stochastically dominates the other and having separate lotteries leads to lower rank variance. Most importantly, we find that with a shortage of seats (or an overdemanded market), there is essentially no tradeoff: not only DA-STB (almost) stochastically dominates DA-MTB, it is also also more fair. So, while Amsterdam selected DA-MTB in order to spread fairness across popular schools, our results suggest that especially in these schools, a single lottery is preferable. In fact, the more competition there is over popular schools (or the larger the shortage of seats), the more inefficiencies are caused by the independently chosen lottery numbers at these schools.

We next discuss some potential extensions to our model.

Large capacities. Our model assumes that each school has a single seat. First, we believe that our main results do not (even) quantitatively change if every school has a finite and constant capacity, i.e. our quantifications remain in the same order of magnitude when the capacity is a constant larger than one.

Roughly speaking, when there is a shortage of seats, the smaller the capacities, the more difficult it is to establish our main results. Let us reconsider the simple example from the introduction of a random market with (large) n students and 2 schools, each with a $0.4n$ seats. Students rejected from their top choice are likely to have a lottery number in $[0.4, 1]$ and tentatively accepted students are likely to have a lottery number in $[0, 0.4]$. Therefore, while under a single lottery all assigned students are likely to obtain their top choice, under separate lotteries many rejected students will be able to draw a better lottery number and get their second choice. So roughly in this setting DA-STB stochastically dominates DA-MTB, and also, there are many Pareto improving pairs under DA-MTB. In addition, the variance is smaller under a single lottery than under a separate lottery. Observe that when there is a linear surplus of seats (linear in n), separate and single lotteries are likely to lead to very similar assignments.

Therefore in overdemanded markets, the larger the capacities are (relative to the market size), the harsher the competition becomes under MTB, and rejection chains become longer. Under STB, however, larger capacities make it less likely that a rejected student will trigger a rejection chain. Consequently, with larger capacities, the the crossing point between the rank distributions under STB and MTB moves further to the "right" (if crossing happens at all). Our model assumes complete preference lists, and it is worth noting, as demonstrated through simulations, that shortening preference lists decreases competition in contrast to the effect of increasing capacities.

Priority classes. Our model assumes a single priority class, and of course, having multiple priority classes affects students' assignments. However, this, interestingly, becomes another reason for studying markets with small capacities. Indeed, schools often fill in many seats with students that belong to top priority classes in these schools (for example top schools are likely to assign many students due to their proximity or sibling priorities). Thus, competition is essentially over the remaining seats. More generally, we believe that adding multiple priority classes to the model would not impact our qualitative insights.

Finally, it is important to remark that in practice, determining which schools are popular (or overdemanded) may not be a straightforward task, yet, a starting point would be using previous students rankings as well as public data regarding school qualities. One of the benefits of this approach compared to exploiting the currently submitted preference lists (i.e. the input to the mechanism) is that it maintains strategy-proofness.

References

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

Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *The American Economic Review*, 93(3):729–747, 2003. ISSN 00028282.

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 Abdulkadiroglu, Yeon-Koo Che, and Yosuke Yasuda. Expanding" choice" in school choice. 2008.

Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *The American Economic Review*, pages 1954–1978, 2009.

Barry C Arnold, Narayanaswamy Balakrishnan, and Haikady Navada Nagaraja. *A first course in order statistics*, volume 54. Siam, 1992.

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

Itai Ashlagi, Yashodhan Kanoria, and Jacob D Leshno. Unbalanced random matching markets. In *ACM Conference on Electronic Commerce*, pages 27–28, 2013.

Itai Ashlagi, Afshin Nikzad, and Assaf I Romm. Assigning more students to their top choices: A tiebreaking rule comparison. *Available at SSRN 2585367*, 2015.

Eduardo Azevedo and Jacob D. Leshno. The college admissions problem with a continuum of students. Mimeo, 2010.

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

Yeon-Koo Che and Olivier Tercieux. Efficiency and stability in large matching markets. Mimeo, 2014.

Monique De Haan, Pieter A Gautier, Hessel Oosterbeek, Bas Van der Klaauw, et al. The performance of school assignment mechanisms in practice. Technical report, CEPR Discussion Papers, 2015.

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

D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962. ISSN 00029890.

Onur Kesten. On two kinds of manipulation for school choice problems. *Economic Theory*, pages 1–17, 2011. ISSN 0938-2259. 10.1007/s00199-011-0618-6.

Donald Ervin Knuth. *An exact analysis of stable allocation.* J. Algorithms 20 (1996), no. 2, 431–44, 1995.

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.

Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of mathematical economics*, 1(1):23–37, 1974.

A Proofs for Section 3

A.1 Proof of Theorem 3.1

We need the following lemmas before proving this theorem.

A.1.1 Computing $\mathcal{R}_{\mathsf{MTB}}$

Lemma A.1. *When $n = m + 1$, wvhp there at most $\frac{3n \ln n}{t}$ students who receive more than t proposals in the school-proposing DA.*

Proof. The proof is a direct consequence of the following result by Pittel (1989): When $n = m+1$, the school-proposing DA takes no more than $3n \ln n$ proposals, wvhp. □

Definition A.2. *Let $\bar{t} = 3\theta \ln m$, where $\theta > 1$ is a large constant that we set later.*

Proposition A.3. *At most n/θ students receive more than $\bar{t}$ offers in school-proposing DA wvhp. This is a direct consequence of A.1.*

Lemma A.4. *Suppose a student s receive t proposals in the school-proposing DA such that $1 \leq t \leq \bar{t}$. Then, for any constant $\alpha > 2$*

$$\mathbb{P}\left[\eta^{\#}(s) > \frac{m}{\alpha t}\right] \geq \mathsf{exp}\left(-\frac{2m}{\alpha(m-t)}\right)$$

Proof. By the principle of deferred decisions, we can assume that students rank proposals upon receiving them. Upon receiving each proposal, the student assigns a (yet unassigned) rank to the school who offers the proposal. The probability that the first school is ranked worse than $\frac{m}{\alpha t}$ is

$1 - \frac{m/\alpha t}{m}$. In general, the probability that i-th school who proposes to s gets ranked better than $1 - \frac{m/\alpha t}{m-i}$. Thus, we have

$$\mathbb{P}\left[\eta^{\#}(s) > \frac{m}{\alpha t}\right] = \prod_{i=1}^{t} 1 - \frac{1}{\alpha t(1 - i/m)}$$
$$\geq \exp\left(-\sum_{i=1}^{t} \frac{2}{\alpha t(1 - i/m)}\right) \geq \exp\left(-\frac{2m}{\alpha(m-t)}\right)$$

where in the first inequality we have used the fact that $1 - x \geq e^{-2x}$ for any $x < 1/2$. □

Lemma A.5. *For any constant $\alpha > 4$, $\mathcal{R}_{\mathsf{MTB}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \leq 0.4n + o(n)$, wvhp.*

Proof. To compute $\mathcal{R}_{\mathsf{MTB}}$, first we run the school-proposing DA and prove the lemma statement for the school-optimal matching. Then, using the fact that almost every student has the same match in the student-optimal matching Ashlagi et al. (2013), we establish the lemma statement (which holds for the student-optimal matching).

For any student s, let x_s be a binary random variable that is 1 iff $\eta^{\#}(s) > \frac{m}{\alpha \bar{t}}$. Also, let S' denote the subset of students who received at least one but no more than $\bar{t}$ offers. For any $s \in S'$, Lemma A.4 implies

$$\mathbb{P}\left[\eta^{\#}(s) > \frac{m}{\alpha \bar{t}}\right] \geq e^{-1/2},$$

since $\alpha > 4$. This means $\mathbb{E}\left[\sum_{s \in S'} x_s\right] \geq e^{-1/2} \cdot |S'| \geq 0.606(n-1)$. Now, applying a standard Chernoff concentration bound implies that wvhp $\sum_{s \in S'} x_s \geq 0.6n$. This fact, together with the fact that $|S \backslash S'| = o(n)$ (which holds by A.3, there are at most $0.4n + o(n)$ students s for whom $\eta^{\#}(s) \leq \frac{m}{\alpha \bar{t}}$.

It is straight-forward to imply a similar result for the student-optimal matching, μ. Note that the number of students who have different matches in μ and η is at most $n/\sqrt{\ln n}$, wvhp Ashlagi et al. (2013). Consequently, there are at most $0.4n + o(n) + n/\sqrt{\ln n}$ students s for whom $\mu^{\#}(s) \leq \frac{m}{\alpha \bar{t}}$. □

A.1.2 Computing $\mathcal{R}_{\mathsf{STB}}$

Lemma A.6. *Suppose student $s \in S$ has priority number $n - x$. Then, the probability that s is not assigned to one of her top i choices is at most $(1 - \frac{x}{n})^i$*

Proof. The probability that s is not assigned to his top choice is $1 - \frac{x}{n}$. The probability that s is not assigned to his second top choice is $(1 - \frac{t}{n})(1 - \frac{x}{n-1})$, which is at most $(1 - \frac{x}{n})^2$. Similarly, it is straightforward to see that the probability that s is not assigned to her i-th top choice is at most $(1 - \frac{x}{n})^i$. □

Lemma A.7. *A student s who has priority number $n-x$ is assigned to one her top $\frac{2n\ln(n)}{x}$ choices with probability at least $1-1/n^2$.*

Proof. The probability that s is not assigned to his top choice is $1-\frac{x}{n}$. The probability that s is not assigned to his second top choice is $(1-\frac{t}{n})(1-\frac{x}{n-1})$, which is at most $(1-\frac{x}{n})^2$. Similarly, it is straightforward to see that the probability that s is not assigned to her i-th top choice is at most $(1-\frac{x}{n})^i$, which is at most $e^{-\frac{xi}{n}}$. Setting $i=\frac{2n}{x}\ln(n)$ proves the claim. $\square$

Lemma A.8. *For any positive constant $\alpha > 1$, $\mathcal{R}_{\mathsf{STB}}\left(\lfloor\frac{m}{\alpha\bar{t}}\rfloor\right) \geq n - O(\ln n\cdot\ln\ln n)$*

Proof. Define $x=\frac{\alpha\bar{t}2n\ln n}{m}$. Let S' be the subset of students who have priority numbers better than $n-x$. First, we apply Lemma A.7 on each student in S'. Lemma A.7 implies that a student with priority number $n-x$ or better, gets assigned to one of her top $\frac{m}{\alpha\bar{t}}$ choices with probability at least $1-n^{-2}$. Taking a union bound over all students with priority number no worse than $n-x$, implies that at least $n-x$ students are assigned to one of their top $\frac{m}{\alpha\bar{t}}$ choices, with probability at least $1-1/n$. This means $\mathcal{R}_{\mathsf{STB}}\left(\frac{m}{\alpha\bar{t}}\right) \geq n-x = n-O(\ln^2 n)$ holds with probability at least $1-1/n$. To prove the sharper bound in the lemma statement, we need to take the students in $S\backslash S'$ into account.

Let $S''\subset S\backslash S'$ denote the subset of students who have priority number between $\beta\bar{t}\cdot\ln\ln n$ and $n-x$, where $\beta = 2\alpha^2\bar{t}/\ln n$. Lemma A.6 implies that for any $s\in S''$,

$$\mathbb{P}\left[\mu^{\#}(s) > \frac{m}{\alpha\bar{t}}\right] \leq \exp\left(-\frac{\beta}{\alpha}\cdot\ln\ln n\right).$$

Having $\beta = 2\alpha^2\bar{t}/\ln n$ implies

$$\mathbb{P}\left[\mu^{\#}(s) > \frac{m}{\alpha\bar{t}}\right] \leq (\ln n)^{-\frac{2\alpha\bar{t}}{\ln n}}.$$

Now, we use the above bound to write a union bound over all $s\in S''$:

$$\mathbb{P}\left[\max_{s\in S''}\mu^{\#}(s) > \frac{m}{\alpha\bar{t}}\right] \leq |S''|\cdot(\ln n)^{-\frac{2\alpha\bar{t}}{\ln n}} \leq O(1/\ln^4 n),$$

where in the last inequality we have used the fact that $x=\frac{\alpha\bar{t}2n\ln n}{m} = O(\ln^2 n)$.[16]

Taking a union bound over the students in S', S'' implies that

$$\mathbb{P}\left[\max_{s\in S'\cup S''}\mu^{\#}(s) > \frac{m}{\alpha\bar{t}}\right] \leq 1/n + O(1/\ln^4 n).$$

[16]The convergence rate $O(1/\ln^4 n)$ can be easily improved to $O(1/n)$ in the expense of changing $\ln n\cdot\ln\ln n$ to $(\ln n)^{1+\epsilon}$ in the lemma statement. Note that we already proved this fact for $\epsilon = 1$ in the current proof.

Consequently, $\mathcal{R}_{\mathsf{STB}}\left(\lfloor\frac{m}{\alpha\bar{t}}\rfloor\right) \geq n - |S\backslash(S' \cup S'')|$ holds whp. To finish the proof, just note that $|S\backslash(S' \cup S'')| = \beta\bar{t} \cdot \ln\ln n = O(\ln n \cdot \ln\ln n)$. □

Lemma A.9. *Wvhp, half of students are assigned to their top choice in STB.*

Proof. □

Now, we are ready to prove Theorem 3.1.

Proof of Theorem 3.1. Lemma A.5 says that $\mathcal{R}_{\mathsf{MTB}}\left(\lfloor\frac{m}{\alpha\bar{t}}\rfloor\right) \leq 0.4n + o(n)$ wvhp. This and Lemma A.9 together imply that $\mathcal{R}_{\mathsf{MTB}}\left(\lfloor\frac{m}{\alpha\bar{t}}\rfloor\right) < \mathcal{R}_{\mathsf{STB}}(1)$ wvhp. On the other hand, Lemma A.8 says that $\mathcal{R}_{\mathsf{STB}}\left(\lfloor\frac{m}{\alpha\bar{t}}\rfloor\right) \geq n - (\ln n)^{1+\epsilon}$ with high probability. The two latter facts, by definition, imply that $\mathcal{R}_{\mathsf{STB}}$ almost stochastically dominates $\mathcal{R}_{\mathsf{MTB}}$. □

A.2 Proof of Theorem 3.2

Lemma A.10. *When $n = m - 1$, at least $\frac{n(1-\epsilon)}{16\ln^2 n}$ students are not assigned to one of their top $3\ln^2 n$ choices in STB, wvhp, for any $\epsilon > 0$.*

Proof. Let $x = 3\ln^2 n$ and $t = \frac{n}{4\ln^2 n}$. Also, let X_s be a binary random variable which is 1 iff student s is not assigned to one of her top x choices. By the principle of deferred decisions, we can assume that $\{X_s\}_{s\in S}$ are independent random variables.

Applying Lemma A.11 implies that any student with priority number below $n-t$ is assigned to one of her top x choices with probability at most $3/4$; in other words, it implies $\mathbb{P}[X_s = 1] \geq 1/4$. Now, let S_t denote the set of students with lowest t priority numbers in STB. A standard application of Chernoff bound implies that $\sum_{s\in S_t} X_s \geq |S_t|(1-\epsilon)/4$, wvhp, for any $\epsilon > 0$. □

Lemma A.11. *A student with priority number $n-t$ in STB is assigned to one of her top x choices with probability at most $\frac{tx}{n-t+1}$.*

Proof. The probability that s is not assigned to her top choice is $1 - \frac{t}{n}$. The probability that s is not assigned to her top two choices is $(1 - \frac{t}{n})(1 - \frac{t}{n-1})$. Similarly, the probability that s is not assigned to her top i choices is $\Pi_{j=1}^{i}(1 - \frac{t}{n-j+1})$. To complete the proof, it is enough to see that $\Pi_{j=1}^{x}(1 - \frac{t}{n-i+1}) \geq 1 - \frac{tx}{n-x+1}$. □

B Proofs for Section B.1

B.1 Equivalence of social inequity and variance

Proof of Lemma 4.1. Let $q = \min\{m, n\}$ be the number of assigned students (which is the same for all stable matchings). Then,

$$\begin{aligned}
\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &= \mathbb{E}_\pi\left[\frac{1}{|\mu_\pi(C)|}\cdot\sum_{t\in\mu_\pi(C)}\left(\mathcal{A}r(\mu_\pi)-\mu_\pi^{\#}(t)\right)^2\right] \\
&= \mathbb{E}_\pi\left[\frac{1}{q}\cdot\sum_{t\in\mu_\pi(C)}\mathcal{A}r(\mu_\pi)^2+\mu_\pi^{\#}(t)^2-2\mathcal{A}r(\mu_\pi)\mu_\pi^{\#}(t)\right] \\
&= \sum_{t\in S}\mathbb{P}_\pi\left[t\in\mu_\pi(C)\right]\cdot\mathbb{E}_\pi\left[\frac{1}{q}\cdot\mathcal{A}r(\mu_\pi)^2+\mu_\pi^{\#}(t)^2-2\mathcal{A}r(\mu_\pi)\mu_\pi^{\#}(t)\middle| t\in\mu_\pi(C)\right] \\
&= \sum_{t\in S}\frac{q}{n}\cdot\mathbb{E}_\pi\left[\frac{1}{q}\cdot\mathcal{A}r(\mu_\pi)^2+\mu_\pi^{\#}(t)^2-2\mathcal{A}r(\mu_\pi)\mu_\pi^{\#}(t)\middle| t\in\mu_\pi(C)\right] \\
&= \frac{1}{n}\cdot\sum_{t\in S}\mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2+\mu_\pi^{\#}(t)^2-2\mathcal{A}r(\mu_\pi)\mu_\pi^{\#}(t)\middle| t\in\mu_\pi(C)\right] && (1)\\
&= \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2+\mu_\pi^{\#}(s)^2-2\mathcal{A}r(\mu_\pi)\mu_\pi^{\#}(s)\middle| s\in\mu_\pi(C)\right] && (2)\\
&= \mathbb{E}_{\pi(s)}\mathbb{E}_{\{\pi(s'):s'\in S,s'\neq s\}}\left[\mathsf{Var}\left[r_s\right]\right] && (3)\\
&= \mathbb{E}_{\{\pi(s'):s'\in S,s'\neq s\}}\left[\mathsf{Var}\left[r_s\right]\right]. && (4)
\end{aligned}$$

In the above inequalities, (2) holds because the term inside the expectation in (1) is equal for all students by symmetry. Also, (4) holds since the inner expectation in (3) is equal for all preference profiles of s by symmetry. □

B.2 Proof of Lemma 4.3

B.2.1 Proof of Lemma 4.3 - Part 1

The proof for Part 1 of Lemma 4.3 is directly implied by Lemmas B.1 and B.3.

Lemma B.1. *When $n = m + 1$, Social inequity in MTB is large is of order of $O(\frac{n^2}{\ln^2 n})$.*

Proof. The proof has two steps. In Step 1, we show that if we run the school-proposing DA, then the variance of the rank of each student is high. In Step 2, we show that even when we move from the school-optimal matching to the student-optimal matching, the variance remains high. The intuition behind Step 2 is, roughly, that only $o(n)$ of the students would have different matched under the school-optimal matching and the student-optimal matching.

Step 1. Since the expected social inequity and the expected deviation in the rank of a fixed student are equal by Lemma 4.1, in this step we use the latter notion. We will switch to the former notion in Step 2. We are interested in providing a lower bound on $\mathbb{E}\left[(r_s - r)^2\right]$, where r_s is a random variable denoting the rank for student s and $r = \mathcal{A}r(\eta)$ (note that r is also equal to the average rank of s, conditioned on being assigned). Since $\mathbb{E}\left[(r_s - r)^2\right] = \mathbb{E}\left[r_s^2\right] - r^2$, we can instead provide a lower bound on the RHS of the equality.

Fix an arbitrary small constant $\epsilon > 0$. Let E_s denote the event in which student s receives at most $(1+\epsilon)\ln n$ proposals. Then

$$\mathbb{E}\left[r_s^2\right] \geq \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}\left[E_s\right]) \cdot (0). \tag{5}$$

To provide a lower bound on the RHS of (5), we provide a lower bund on $\mathbb{E}\left[(r_s - r)^2 \middle| E_s\right]$. If student s receives d_s proposals in school-proposing DA, then it chooses the best out of these d_s proposals, which means its rank is the first order statistic among the proposals that she had received. In Proposition B.6, we calculate $\mathbb{E}\left[r_s^2 \middle| d_s\right]$ (which is the expected rank squared for s conditioned on receiving d_s proposals).

Using Proposition B.6 and (5) together we can write

$$\begin{aligned}
\mathbb{E}\left[r_s^2 | E_s\right] &\geq \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}\left[E_s\right]) \cdot (0) \\
&\geq (1 - o(1)) \cdot \frac{3n^2}{2\ln^2 n} + o(1) \cdot (0),
\end{aligned} \tag{6}$$

where (6) holds by Lemma B.2, which shows event E_s happens whp.

It is known that, $r \in [\frac{(1-\delta)n}{\ln n}, \frac{(1+\delta)n}{\ln n}]$ for any constant $\delta > 0$ and large enough n (see Ashlagi et al. (2013)). This fact, and (6) together imply that

$$\mathbb{E}\left[(r_s - r)^2\right] = \mathbb{E}\left[r_s^2\right] - r^2 \geq (1 - o(1)) \cdot (3/2 - (1+\delta)^2) \cdot \frac{n^2}{\ln^2 n} = \Theta(\frac{n^2}{\ln^2 n}).$$

This finishes Step 1.

Step 2. In this step, instead of working with the notion of expected deviation in the rank of a fixed student, we switch to its equivalent notion, expected social inequity. Step 1 and Lemma 4.1 together imply that $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]$ is $O(\frac{n^2}{\ln^2 n})$. In this step, we show that moving from the school-optimal matching to the student-optimal matching does not change the social inequity much in expectation, and as the result, we would prove that $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right]$ is also $O(\frac{n^2}{\ln^2 n})$. This is done as follows.

$$\begin{aligned}
m \cdot \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] = \mathbb{E}_\pi\Bigg[&\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)^2 + \mathcal{A}r(\mu_\pi)^2 - 2\mu_\pi^{\#}(s)\mathcal{A}r(\mu_\pi) \\
&- \sum_{s\in\eta_\pi(C)} \eta_\pi^{\#}(s)^2 + \mathcal{A}r(\eta_\pi)^2 - 2\eta_\pi^{\#}(s)\mathcal{A}r(\eta_\pi)\Bigg] \\
&= m \cdot \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right] \\
&+ \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)^2 - \eta_\pi^{\#}(s)^2\right] \\
&- 2\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)\mathcal{A}r(\mu_\pi) - \sum_{s\in\eta_\pi(C)} \eta_\pi^{\#}(s)\mathcal{A}r(\eta_\pi)\right]. \qquad (7)
\end{aligned}$$

We can rewrite the above inequality by simplifying (7) as

$$\begin{aligned}
&2\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)\mathcal{A}r(\mu_\pi) - \sum_{s\in\eta_\pi(C)} \eta_\pi^{\#}(s)\mathcal{A}r(\eta_\pi)\right] \\
=&2m \cdot \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right],
\end{aligned}$$

which together with the previous equation implies

$$\begin{aligned}
m \cdot \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] = & \qquad (8) \\
&- m \cdot \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right] \qquad (9) \\
&+ \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)^2 - \eta_\pi^{\#}(s)^2\right] \qquad (10)
\end{aligned}$$

To prove the lemma, we provide lower bounds for (9) and (10). When we move from the school-optimal matching to the student-optimal matching, each student gets assigned to a school at least as good as before. Let $\Delta_\pi(s) = \eta_\pi^{\#}(s) - \mu_\pi^{\#}(s)$, and $\Delta_\pi = \sum_{s\in\mu_\pi(C)} \Delta_\pi(s)$.

An Lower Bound for (9)**.** First, we note that $m\left(\mathcal{A}r(\eta_\pi) - \mathcal{A}r(\mu_\pi)\right) = \Delta_\pi$ is "small" wvhp. This is a direct consequence of Theorem 5 of Ashlagi et al. (2013); they show that there exist constants

$n_0, \delta > 0$ such that for $n > n_0$, we have

$$\mathbb{P}_{\pi \sim \Pi}\left[\Delta_\pi \geq \delta n \ln n\right] < \mathsf{exp}\big\{-(\ln n)^{0.4}\big\}. \tag{11}$$

According to this bound, we have that

$$m \cdot (\mathcal{Ar}(\mu_\pi)^2 - \mathcal{Ar}(\eta_\pi)^2) = m \cdot \big((\mathcal{Ar}(\eta_\pi) - \Delta_\pi/m)^2 - \mathcal{Ar}(\eta_\pi)^2\big) = \Delta_\pi^2/m - 2\Delta_\pi \mathcal{Ar}(\eta_\pi).$$

By taking expectation from both sides of the above equation, we can write

$$\begin{aligned} m \cdot \mathbb{E}_\pi\left[\mathcal{Ar}(\mu_\pi)^2 - \mathcal{Ar}(\eta_\pi)^2\right] &= \mathbb{E}_\pi\left[\Delta_\pi^2/m - 2\Delta_\pi \mathcal{Ar}(\eta_\pi)\right] \\ &\leq (\overline{\delta} n \ln n)^2/m, \end{aligned} \tag{12}$$

where the last inequality holds by (11), for any constant $\overline{\delta} > \delta$ and sufficiently large n. This implies a lower bound of $-(\overline{\delta} n \ln n)^2/m$ for (9).

A Lower Bound for (10)**.** First, we rewrite (10) as follows.

$$\begin{aligned} \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^\#(s)^2 - \eta_\pi^\#(s)^2\right] &= \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} (\eta_\pi^\#(s) - \Delta_\pi(s))^2 - \eta_\pi^\#(s)^2\right] \\ &\geq -2\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \eta_\pi^\#(s)\Delta_\pi(s)\right]. \end{aligned} \tag{13}$$

We proceed by providing a lower bound on (13). First, we use the Cauchy-Schwarz inequality to write

$$\begin{aligned} \sum_{s\in\eta_\pi(C)} \eta_\pi^\#(s)\Delta_\pi(s) &\leq \left(\sum_{s\in\eta_\pi(C)} (\eta_\pi^\#(s))^2 \cdot \sum_{s\in\eta_\pi(C)} (\Delta_\pi(s))^2\right)^{1/2} \\ &\leq m^{3/2} \cdot \left(\sum_{s\in\eta_\pi(C)} (\Delta_\pi(s))^2\right)^{1/2} \end{aligned}$$

Taking expectation from both sides of the above inequality implies

$$\mathbb{E}_\pi\left[\sum_{s\in\eta_\pi(C)} \eta_\pi^\#(s)\Delta_\pi(s)\right] \leq m^{3/2} \cdot \mathbb{E}_\pi\left[\left(\sum_{s\in\eta_\pi(C)} (\Delta_\pi(s))^2\right)^{1/2}\right].$$

Using (11), we can rewrite the above upper bound:

$$\mathbb{E}_\pi\left[\sum_{s\in\eta_\pi(C)}\eta_\pi^{\#}(s)\Delta_\pi(s)\right]\leq m^{3/2}\cdot n(\overline{\delta}\ln n)^{1/2},$$

which holds for any constant $\overline{\delta}>\delta$. This upper bound can be directly translated into the lower bound $-2m^{3/2}\cdot n(\overline{\delta}\ln n)^{1/2}$ for (10).

Using the lower bounds that we provided for (9) and (10), we can rewrite equation (8) as follows:

$$m\cdot\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)-\mathcal{S}i(\eta_\pi)\right]\geq-(\overline{\delta}n\ln n)^2/m-2m^{3/2}\cdot n(\overline{\delta}\ln n)^{1/2}.$$

In the other hand, In Step 1 we established that $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]\geq O(n^2/\ln^2 n)$. The two latter inequalities together imply that

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right]=\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]+\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)-\mathcal{S}i(\eta_\pi)\right]\geq O(n^2/\ln^2 n).$$

The claim is proved. □

Lemma B.2. *Suppose $n=m+1$, and fix a student s. Then, under MTB, in the school-proposing DA, the number of offers received by s is whp at most $(1+\epsilon)\ln n$ for any constant $\epsilon>0$.*

Proof. The proof idea is defining another stochastic process that we denote by $\mathcal{B}$. Process $\mathcal{B}$ is defined by a sequence of binary random variables $X_1,\ldots,X_k$, where $k=(1-\delta)n\ln n$ for some arbitrary small constant $\delta>0$. Each random variable in this sequence is 1 with probability $\frac{\ln^3 n}{n}$, and is 0 otherwise. For convenience, we also refer to these random variables by *coins*, and the process that determines the value of a random variable by *coin-flip*.

Define $X=\sum_{i=1}^k X_k$. The goal is to show that X is a good upper bound on the number of proposals that are received by s. The high-level idea is based on two facts: First, the number of total proposals is stochastically dominated by the coupon-collector problem, and so is wvhp at most k. Second, by Pittel (1989), we know that wvhp, each school makes at most $3\ln^2 n$ proposals, and so, each proposal is made to s with probability at most $\frac{1}{n-3\ln^2 n}$. Consequently, the number of proposals made to s cannot be more than $\frac{k(1+\delta')}{n-3\ln^2 n}$ whp, for any constant $\delta>0$. (The latter fact is a direct consequence of the Chernoff bound which is applicable since the coin flips are independent).

To formalize this argument, we have to define a new random process $(\text{DA},\mathcal{B})$, which is the coupling of the random processes DA, $\mathcal{B}$. This process would have to components, one for each of the original random process. Each component, if considered independently, behaves identical to its corresponding original process. It is straight-forward to define a simple coupling in which in almost all sample paths (i.e. whp), the number of coin flips. The coupling is defined naturally by flipping a coin for each proposal and deciding whether to make a proposal to s based on the outcome of the coin-flip. We omit the details. $\square$

Lemma B.3. *Suppose $|n - m| = 1$. Then, under STB, the expected social inequity is $\Theta(n)$.*

Proof. First, we compute a lower bound on the expected social inequity in STB. With probability at least $1/2$, the student with the lowest priority number in STB gets assigned to a school that she has ranked on lower half of her preference list. So, for any student $s \in S$ we can write:

$$\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] = \mathbb{E}\left[\mathsf{Var}\left[r_{\underline{s}}\right]\right] \geq \frac{1}{n} \cdot \left(\mathcal{A}r(\mu^{\#}_{\mathsf{STB}}(s)) - n\right)^2.$$

It is proved by Knuth (1995) that $\mathcal{A}r(\mu^{\#}_{\mathsf{STB}}(s)) = \Theta(\ln n)$. Plugging this into the above inequality implies $\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right] \geq \Omega(n)$. In the other hand, using Lemma B.4 we have

$$\begin{aligned}\mathcal{S}i(\mu_{\mathsf{STB}}) &= \mathbb{E}_{\pi}\left[(\mathcal{A}r(\mu_{\pi}) - \mu^{\#}_{\pi}(s))^2 \middle| \mu_{\pi}(s) \neq \emptyset\right] \\ &= \mathbb{E}_{\pi}\left[\mu^{\#}_{\pi}(s)^2 \middle| \mu_{\pi}(s) \neq \emptyset\right] - \mathbb{E}_{\pi}\left[\mathcal{A}r(\mu_{\pi})\right]^2 \\ &\leq \mathbb{E}_{\pi}\left[\mu^{\#}_{\pi}(s)^2 \middle| \mu_{\pi}(s) \neq \emptyset\right] = O(n)\end{aligned}$$

which completes the proof. $\square$

Lemma B.4. *Suppose $|n - m| = 1$. Then, under STB, for any student s,*

$$\mathbb{E}_{\pi}\left[\mu^{\#}_{\pi}(s)^2 \middle| \mu_{\pi}(s) \neq \emptyset\right] = O(n).$$

Proof. We prove this assuming that $m \geq n$. The proof for $m < n$ is identical to the proof for $m = n$: To see why, suppose $n = m$, and note that the expected social inequity does not change when one more student is added.

Let $t = \sqrt{n} \ln n$ and let p_s be the "priority number" of s in the corresponding random serial dictatorship. We consider two cases: either $p_s \leq n - t$ or not. See that:

$$\begin{aligned}\mathbb{E}_{\pi}\left[\mu^{\#}_{\pi}(s)^2 \middle| \mu_{\pi}(s) \neq \emptyset\right] &= \mathbb{P}\left[p_s \leq n - t\right] \cdot \mathbb{E}\left[\mu^{\#}_{\pi}(s)^2 \middle| p_s \leq n - t\right] \\ &\quad + \mathbb{P}\left[n - t < p_s\right] \cdot \mathbb{E}\left[\mu^{\#}_{\pi}(s)^2 \middle| n - t < p_s\right]. \end{aligned} \tag{14}$$

We provide an upper bound for each of the terms in the right-hand side of (14).

By Lemma B.5, we have:

$$\mathbb{E}\left[\mu_\pi^\#(s)^2 \middle| p_s \leq n-t\right] \leq (1-\frac{1}{n}) \cdot (n\ln(n)/t)^2 + \frac{1}{n} \cdot (n^2) \leq 2n,$$

which implies

$$\mathbb{P}\left[p_s \leq n-t\right] \cdot \mathbb{E}\left[\mu_\pi^\#(s)^2 \middle| p_s \leq n-t\right] \leq 2n. \tag{15}$$

Also, we have that:

$$\begin{aligned}\mathbb{P}\left[n-t < p_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| n-t < p_s\right] &\leq \frac{t}{n} \cdot \sum_{i=1}^{t} \frac{1}{t} \mathbb{E}\left[r_s^2 \middle| p_s = n-i+1\right] \\ &\leq \frac{1}{n} \cdot \sum_{i=1}^{t} 2(n/i)^2. &(16)\\ &\leq n \cdot \frac{\pi^2}{3}. &(17)\end{aligned}$$

where (16) holds since for a geometric random variable X with mean p we have $\mathbb{E}\left[X\right] = \frac{2-p}{p^2}$.

Finally, putting (15) and (17) together implies

$$\mathbb{E}_\pi\left[\mu_\pi^\#(s)^2 \middle| \mu_\pi(s) \neq \emptyset\right] \leq n(2 + \frac{\pi^2}{3})$$

□

Lemma B.5. *Suppose $n \leq m$. Then, a student s with priority number $n-t$ is assigned to one of her top $\frac{n\ln(n)}{t}$ choices with probability at least $1 - 1/n$.*

Proof. The probability that s is not assigned to his top choice is $1 - \frac{t}{n}$. The probability that s is not assigned to his second top choice is $(1-\frac{t}{n})(1-\frac{t}{n-1})$, which is at most $(1-\frac{t}{n})^2$. Similarly, it is straightforward to see that the probability that s is not assigned to her i-th top choice is at most $(1-\frac{t}{n})^i$, which is at most $e^{-\frac{ti}{n}}$. Setting $i = \frac{n}{t}\ln(n)$ proves the claim. □

Proposition B.6. *Suppose $d \leq n$, and define the random variable $X = \min\{X_1, \ldots, X_d\}$, where $X_1, \ldots, X_d$ respectively represent the first d elements of a permutation over $[n]$ that is chosen uniformly at random. Then, $\mathbb{E}\left[X^2\right] = \frac{d(n+1)(n-d)}{(d+1)^2(d+2)} + \frac{(n-d)^2}{(d+1)^2}$.*

Proof. It is known that $\mathbb{E}\left[X\right] = \frac{n-d}{d+1}$ and $\mathsf{Var}[X] = \frac{d(n+1)(n-d)}{(d+1)^2(d+2)}$ (see Arnold et al. (1992), Page 55). Plugging these equations into $\mathsf{Var}\left[X\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$ proves the claim. □

B.2.2 Proof of Lemma 4.3 - Part 2

We presented the proof earlier in Section 4. For completeness, we repeat it here.

Pittel (1989) shows that wvhp, $\max_{s\in S} \mu^{\#}_{\mathsf{MTB}}(s) \leq 3\ln^2 n$. Therefore, wvhp

$$\frac{1}{n} \cdot \sum_{s\in S} (\mathcal{Ar}(\mu_{\mathsf{MTB}}) - \mu^{\#}_{\mathsf{MTB}}(s))^2 \leq 9\ln^4 n.$$

This implies that the expected social inequity under MTB is $O(\ln^4 n)$. On the other hand, Lemma B.3 implies that the expected social inequity under STB is $\Theta(n)$.

B.2.3 Proof of Lemma 4.3 - Part 3

First, note that as a consequence of Part 2, we have already proved a weaker version of Part 3: If $n = m - 1$, the expected social inequity under MTB is still $O(\ln^4 n)$, by the same analysis for $n = m$. In the other hand, the expected social inequity under STB is $\Theta(n)$, by Lemma B.3. This gap is large enough that Theorem 4.2 still holds, even with this weaker version of Part 3.

Here, we prove the gap is even larger, by showing how the bound on the expected social inequity under MTB can be improved to $O(\ln^2 n)$. The proof follows the same steps as the proof of Lemma B.8, where we provide an upper bound on $\mathbb{E}\left[\mathcal{Si}(\mu_{\mathsf{MTB}})\right]$ when the imbalance is linear. During the proof, we will also use Lemma C.7, which is proved independently in Section C.

The proof is done in 2 Steps (which are the same steps as in Lemma B.8). In Step 1, we show that that the variance of the rank of student s in the student-proposing DA is approximately equal to the variance of its rank in the school-proposing DA. Then, in Step 2, we provide an upper bound on the variance of rank in the school-proposing DA. Steps 1,2 then together will prove the claim.

Step 1 is identical to the Step 1 of Lemma B.8. It remains to prove Step 2. Suppose we are running the school-proposing DA. First, see that

$$\begin{aligned}
\mathbb{E}_\pi\left[\mathcal{Si}(\eta_{\mathsf{MTB}})\right] &= \mathbb{E}_\pi\left[(\mathcal{Ar}(\eta_\pi) - \eta^{\#}_\pi(s))^2 \middle| \eta_\pi(s) \neq \emptyset\right] \\
&= \mathbb{E}_\pi\left[\eta^{\#}_\pi(s)^2 \middle| \eta_\pi(s) \neq \emptyset\right] - \mathbb{E}_\pi\left[\mathcal{Ar}(\eta_\pi)\right]^2 \\
&\leq \mathbb{E}_\pi\left[\eta^{\#}_\pi(s)^2 \middle| \eta_\pi(s) \neq \emptyset\right]
\end{aligned}$$

For notational simplicity, let r_s denote the rank of student s. Note that since s is always assigned, then $r_s \in [m]$. We can write the above bound as

$$\mathbb{E}_\pi\left[\mathcal{Si}(\eta_{\mathsf{MTB}})\right] \leq \mathbb{E}\left[r_s^2\right]. \tag{18}$$

Next, we provide an upper bound on $\mathbb{E}\left[r_s^2\right]$. Fix an arbitrary small constant $\epsilon > 0$. Let E_s denote the event in which student s receives at least $\kappa = \frac{(1-\epsilon)n}{2\ln n}$ proposals. Lemma C.7 (which is proved independently in Section C) shows that E_s happens wvhp. Consequently,

$$\mathbb{E}\left[r_s^2 | E_s\right] \lesssim \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] \leq O(\ln^2 n), \tag{19}$$

where we have used by Proposition B.6 to bound $\mathbb{E}\left[r_s^2 \middle| E_s\right]$.

Now we are ready to finish the proof of Part 3. See that (18) and (19) together imply that

$$\mathbb{E}_\pi\left[\mathcal{S}i(\eta_{\mathsf{MTB}})\right] \leq O(\ln^2 n).$$

By Step 1, and by the above bound, we have

$$\begin{aligned}\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &\leq \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] + \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] \\ &\leq o(1) + \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] \approx O(\ln^2 n).\end{aligned}$$

B.3 Proof of Theorem 4.4

Instead of Proving Theorem 4.4, we first prove a weaker version of it (Theorem B.7). In the end of this section, we will explain how the proof can be adapted (with a more careful analysis) for Theorem 4.4.

Theorem B.7. *Suppose $m = n + \lambda n$ for some positive $\lambda \leq 0.01$. Then, $\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{MTB}})]} > 1$, where the expectations are taken over preferences and the tiebreaking rules.*

Proof of Theorem B.7. The proof is directly implied by Lemmas B.8 and B.12 below.

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right]}{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{MTB}})\right]} \geq \frac{\frac{2(1+\lambda)}{\lambda} - (1+\lambda)\ln(1+1/\lambda) - (1+\lambda)^2 \ln(1+\frac{1}{\lambda})^2}{7k^2 - 6k + 1}.$$

where $k = (1+\lambda)\ln(1+1/\lambda)$. For $\lambda \leq 0.01$, RHS of the above inequality is strictly greater than one. ☐

Lemma B.8. *Suppose $m = n + \lambda n$. Then, under MTB we have*

$$\lim_{n\to\infty} \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] \leq 7k^2 - 6k + 1,$$

where $k = (1+\lambda)\ln(1+1/\lambda)$.

Proof. We use Lemma 4.1, by which the expected social inequity and the expected deviation in the rank of a fixed student are equal. So, to prove the lemma, we fix an student s and show that

$$\lim_{n\to\infty} \mathbb{E}_{\{\pi(s'):s'\in S, s'\neq s\}} [\mathsf{Var}[r_s]] \leq 7k^2 - 6k + 1. \tag{20}$$

We prove (20) in 2 Steps. In Step 1, we show that that the variance of the rank of student s in the student-proposing DA is approximately equal to the variance of its rank in the school-proposing DA. Then, in Steps 2, we provide an upper bound $7k^2 - 6k + 1$ on the variance of rank in the school-proposing DA. Steps 1,2 then together will imply that (20) holds.

To prove the lemma, it remains to prove each of the steps separately. After that, we will conclude by putting these steps together.

Step 1. First, we rewrite the following equality from the proof of Lemma B.1.

$$m \cdot \mathbb{E}_\pi [\mathcal{Si}(\mu_\pi) - \mathcal{Si}(\eta_\pi)] = \tag{21}$$

$$- m \cdot \mathbb{E}_\pi \left[\mathcal{Ar}(\mu_\pi)^2 - \mathcal{Ar}(\eta_\pi)^2\right] \tag{22}$$

$$+ \mathbb{E}_\pi \left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)^2 - \eta_\pi^{\#}(s)^2\right]. \tag{23}$$

To complete Step 1, we need to provide upper bounds for (22) and (23).

An upper bound for (22) We will use the following relation between average ranks, provided by Theorem 3 of Ashlagi et al. (2013): wvhp we have

$$\mathcal{Ar}(\eta_\pi) \leq \mathcal{Ar}(\mu_\pi)(1 + o(1)).$$

Consequently, $m \cdot o(1) \cdot \mathbb{E}_\pi [\mathcal{Ar}(\mu_\pi)]$ is a valid upper bound for (22).

An upper bound for (23) 0 is a valid upper bound since, by the definition of μ, η, we always have $\mu_\pi^{\#}(s) \leq \eta_\pi^{\#}(s)$.

Plugging the provided upper bounds into (21) implies

$$\mathbb{E}_\pi [\mathcal{Si}(\mu_\pi) - \mathcal{Si}(\eta_\pi)] \leq o(1) \cdot \mathbb{E}_\pi [\mathcal{Ar}(\mu_\pi)].$$

When there are linearly more seats, $\mathbb{E}_\pi[\mathcal{A}r(\mu_\pi)] = O(1)$. This implies

$$\mathbb{E}_\pi[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)] \leq o(1), \tag{24}$$

which concludes Step 1.

Step 2. This Step is similar to Step 1 in the proof of Lemma B.1.

Since the expected social inequity and the expected deviation in the rank of a fixed student are equal by Lemma 4.1, in this step we use the latter notion. We will switch to the former notion in Step 2. We are interested in providing an upper bound on $\mathbb{E}\left[(r_s - r)^2\right]$, where r_s is a random variable denoting the rank for student s and $r = \mathcal{A}r(\eta)$ (note that r is also equal to the average rank of s, conditioned on being assigned). Since $\mathbb{E}\left[(r_s - r)^2\right] = \mathbb{E}\left[r_s^2\right] - r^2$, we can instead provide an upper bound on the RHS of the equality.

Fix an arbitrary small constant $\epsilon > 0$. Let E_s denote the event in which student s receives at least $\kappa = \frac{(1-\epsilon)n}{2k}$ proposals, where recall that $k = (1+\lambda)\ln(1+1/\lambda)$. Then

$$\mathbb{E}\left[r_s^2\right] \leq \mathbb{P}[E_s] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}[E_s]) \cdot (n + \lambda n)^2. \tag{25}$$

We proceed by providing an upper bound on the RHS of (25). Lemma B.9 implies E_s happens wvhp, and so, we can ignore the second term in the RHS of (25) since it is a lower order term. We provide an upper bound on the first term in the RHS of (25), i.e. on $\mathbb{E}\left[r_s^2 \middle| E_s\right]$. If student s receives d_s proposals in school-proposing DA, then it chooses the best out of these d_s proposals, which means its rank is the first order statistic among the proposals that she had received. In Proposition B.6, we calculate $\mathbb{E}\left[r_s^2 \middle| d_s\right]$ (which is the expected rank squared for s conditioned on receiving d_s proposals).

Using Proposition B.6 and (25) together we can write

$$\begin{aligned}
\mathbb{E}\left[r_s^2 | E_s\right] &\lesssim \mathbb{P}[E_s] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] \\
&\lesssim (\frac{n}{\kappa} - 1)(\frac{2n}{\kappa} - 1) \qquad (26) \\
&= \left(\frac{2k-1}{1-\epsilon} - 1\right)\left(\frac{4k-1}{1-\epsilon} - 1\right). \qquad (27)
\end{aligned}$$

Now, (27) implies that

$$\lim_{n\to\infty} \mathbb{E}_\pi[\mathcal{S}i(\eta_\pi)] = \lim_{n\to\infty} \mathbb{E}_\pi\left[r_s^2 - r^2\right] = (2k-1)(4k-1) - k^2 = 7k^2 - 6k + 1. \tag{28}$$

This completes Step 2.

Now we are ready to finish the proof of the lemma. See that

$$\begin{aligned}\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &\leq \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)-\mathcal{S}i(\eta_\pi)\right]+\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]\\ &\leq o(1)+\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] \qquad (29)\\ &\approx 7k^2-6k+1, \qquad (30)\end{aligned}$$

where (29) holds by Step 1, and (30) holds by (28). ☐

Lemma B.9. *Suppose $m = n + \lambda n$. Then, for any positive constant ϵ, the number of proposals received by a fixed student in the school-proposing DA is wvhp at least $(1-\epsilon)\kappa$, where $\kappa = \frac{n}{2(1+\lambda)\ln(1+1/\lambda)}$.*

Proof. The proof idea is defining another stochastic process that we denote by $\mathcal{B}$. Process B is defined by a sequence of binary random variables $X_1, \ldots, X_k$. Each random variable in this sequence is 1 with probability $1/n$, and is 0 otherwise. For convenience, we also refer to these random variables by *coins*. We describe the process $\mathcal{B}$ in a high level and then define it formally. First,we set the number of coins (k) and then we start flipping them. Based on the outcome of each coin-flip, we might decrease the number of remaining coin-flips (by dismissing some of the coins). The process is finished when there are no coins left. We define the process formally below.

1. Fix a small constant $\delta > 0$.
2. Let $k = 2n\kappa(1-\delta)$.
3. Let $i = 1$.
4. While $i \leq k$ do
   (a) Flip coin i.
   (b) If the outcome is 0 then $i \leftarrow i+1$, otherwise $k \leftarrow k-n$.

Next, we would like to use the number of successful coin-flips, defined by $X = \sum_{i=1}^{k} X_i$, as a lower bound for the number of proposals made to s, which we denote by d_s. To this end, we couple the process $\mathcal{B}$ with the school-proposing DA, and denote the coupled process by $(\text{DA}, \mathcal{B})$. Our coupling has the property that in almost all of its sample paths (except for a negligible fraction), $X \leq d_s$. In other words, if we pick a sample path of $(\text{DA}, \mathcal{B})$ uniformly at random (from the space of all sample paths), then $X \leq d_s$ holds in that sample path wvhp.

Claim B.10. *In* $(DA, \mathcal{B})$*, wvhp we have* $d_s \geq X$*.*

We relegate the definition of our coupling and the proof of Claim B.10 to Section B.4. The rest of the proof is straight-forward. In Lemma B.11, we show that for any constant $\delta' > 0$,

$$X \geq (1 - \delta')(1 - \delta)\kappa$$

holds wvhp. This, and Claim B.10 together imply that $d_s \geq (1 - \epsilon)\kappa$ holds wvhp, for any constant $\epsilon > 0$. □

B.4 Remaining proofs from Section B.3

In this section, we complete the proof of Lemma B.9 by providing the proof for Claim B.10. As mentioned before, the proof involves defining a new process, $(\mathrm{DA}, \mathcal{B})$, which is in fact a coupling of the processes $\mathrm{DA}, \mathcal{B}$. First, we define the coupling formally, and after that we prove Claim B.10.

Definition of the Coupling

Recall that we fixed a student s, with the purpose of providing a lower bound on the number of proposals made to s during the DA algorithm. We define the process $(\mathrm{DA}, \mathcal{B})$ by running both of DA and $\mathcal{B}$ simultaneously. The results of coin-flips in $\mathcal{B}$ would be used to decide whether each proposal in DA is made to s or not.

Suppose we are running the school-proposing DA. Let S_c denote the set of students that c has proposed to them so far. In the coupled process, each school could have 3 possible states: *active*, *inactive*, and *idle*. In the beginning, all schools are active. We will see that as the process evolves, schools might change their state from active to inactive or idle and from inactive to idle.

In the coupled process, a coin-flip corresponds to a new proposal. If there are no coins left to flip (in $\mathcal{B}$), or no proposals left to make (in DA), then $(\mathrm{DA}, \mathcal{B})$ stops. Suppose it is the turn of a school c to make a new proposal. This will be done by considering the following cases:

1. If c is active, then use a coin-flip to decide whether c proposes to s in her next move. This is done as it follows: Flip one of the unflipped coins. If it is a successful flip (with probability $1/n$), then c will propose to s; make c idle, and dismiss n of the unflipped coins. Otherwise, if the coin-flip is not successful then: with probability $1 - \frac{1-1/|S \backslash S_c|}{1-1/n}$ propose to s and make c inactive, and with probability $\frac{1-1/|S \backslash S_c|}{1-1/n}$ propose to one of the students in $S \backslash (S_c \cup \{s\})$ uniformly at random (without changing the state of c).

2. If c is inactive, then flip one of the unflipped coins. If it is a successful flip, make c idle, and dismiss n of the unflipped coins; otherwise, do not change the state of c. Propose to one of the students in $S \backslash S_c$ uniformly at random.

3. If c is idle, then do not flip any coins. Propose to one of the students in $S \backslash S_c$ uniformly at random.

This completes the description of $(\mathrm{DA}, \mathcal{B})$.

Proof of Claim B.10. For any school c who has made a proposal to s, there is at most one successful coin-flip corresponding to c. This holds since

(i) A successful coin-flip that corresponds to school c happens when c is either active or inactive. In both of these cases, c must have made a proposal to s.

(ii) After a successful coin-flip that corresponds to school c, n coins are removed (which account for the next proposals from c). So, there will be no two successful coin-flips both of which correspond to a proposal from c to s.

Consequently, the number of successful coin-flips is no larger than the number of proposals made to s. □

Lemma B.11. *For any constant $\delta' > 0$, $X \geq (1 - \delta')(1 - \delta)\kappa$ holds wvhp.*

Proof. First, we show that wvhp $(\mathrm{DA}, \mathcal{B})$ terminates with no coins left. To see this, note that in $(\mathrm{DA}, \mathcal{B})$, the number of proposals that are made is at most equal to the number of flipped or dismissed coins. In the other hand, Ashlagi et al. (2013) show that wvhp the number of proposals made by the school-proposing DA is at least $2n\kappa(1 - \delta)$. Since $\mathcal{B}$ starts with $k = 2n\kappa(1 - \delta)$ coins, then, wvhp, $(\mathrm{DA}, \mathcal{B})$ ends when there are no coins left.

We are now ready to prove the lemma. Partition the set of k coins into two subsets with equal size, namely subsets A, B. Correspond the operation $k \leftarrow k - n$ (in the process $\mathcal{B}$) to the operation of removal of n coins from the subset B (as long as B is non-empty). One way of running $\mathcal{B}$ would be flipping the coins in A one by one and removing n coins from B whenever a coin-flip is successful. This will be continued until B is empty. Suppose X' denotes the number of successful coin-flips in this process. Since $X \geq X'$ in each sample path of the process, it is enough to prove the lemma statement for X' (instead of X). A standard application of Chernoff bound implies that $X' \geq \frac{|A|}{n} \cdot (1 - \delta')$ wvhp. This proves the lemma since $|A| \geq n\kappa(1 - \delta)$, by definition. □

Lemma B.12. *Suppose $m = n + \lambda n$. Then, under STB we have*

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] \geq \frac{2(1+\lambda)}{\lambda} - (1+\lambda)\ln(1+1/\lambda) - (1+\lambda)^2\ln(1+\frac{1}{\lambda})^2$$

proof sketch. Suppose students indexed with respect to their priority number in STB, i.e. the student with the highest priority number is indexed 1, and the student with the lowest priority number is indexed with n. Fix a student s. Using Lemma 4.1, we can write

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] = \mathsf{Var}\left[r_s\right] = \mathbb{E}\left[r_s^2\right] - \mathbb{E}\left[r_s\right]^2, \tag{31}$$

where r_s denotes the rank assigned to student s.

To provide a lower bound for (31), we provide a lower bound for $\mathbb{E}\left[r_s^2\right]$ and an upper bound for $\mathbb{E}\left[r_s\right]^2$.

Upper bound for $\mathbb{E}\left[r_s\right]^2$. By Proposition B.13, we have that

$$\mathbb{E}\left[r_s\right]^2 \approx (1+\lambda)^2\ln(1+\frac{1}{\lambda})^2.$$

Lower bound for $\mathbb{E}\left[r_s^2\right]$ First, see that

$$\mathbb{E}\left[r_s^2\right] = \frac{1}{n}\cdot\sum_{i=0}^{n-1}\mathbb{E}\left[r_s^2 \middle| s \text{ has priority } i+1\right].$$

Then, we use Proposition B.14 to calculate an upper bound on the RHS of the above inequality:

$$\begin{aligned}
\mathbb{E}\left[r_s^2\right] &= \frac{1}{n}\cdot\sum_{i=0}^{n-1}\mathbb{E}\left[r_s^2 \middle| s \text{ has priority } i+1\right] \\
&\gtrsim \frac{1}{n}\sum_{i=0}^{n-1}\frac{2}{(\frac{m-i}{m})^2} - \frac{1}{(\frac{m-i}{m})} \\
&\approx \frac{1}{n}\sum_{i=0}^{n-1}\frac{2}{(\frac{m-i}{m})^2} - (1+\lambda)\ln(1+1/\lambda).
\end{aligned}$$

Now, using the inequality $\frac{1}{x^2} \geq \frac{1}{x} - \frac{1}{x+1}$ we can write

$$\begin{aligned}
\mathbb{E}\left[r_s^2\right] &\gtrsim \frac{2m^2}{n} \cdot \sum_{i=0}^{n-1} \frac{1}{(m-i)^2} - (1+\lambda)\ln(1+1/\lambda). \\
&\geq \frac{2m^2}{n} \cdot \left(\frac{1}{\lambda n} - \frac{1}{(\lambda+1)n}\right) - (1+\lambda)\ln(1+1/\lambda). \\
&= \frac{2(1+\lambda)}{\lambda} - (1+\lambda)\ln(1+1/\lambda).
\end{aligned}$$

By combining the above bounds, we can provide the promised lower bound on (31).

$$\begin{aligned}
\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &= \mathbb{E}\left[r_s^2\right] - \mathbb{E}\left[r_s\right]^2 \\
&\gtrsim \frac{2(1+\lambda)}{\lambda} - (1+\lambda)\ln(1+1/\lambda) - (1+\lambda)^2\ln(1+\frac{1}{\lambda})^2
\end{aligned}$$

□

Proposition B.13. *Suppose $m = (1+\lambda)n$. Then,* $\mathbb{E}\left[r_s\right] \approx (1+\lambda)\ln(1+\frac{1}{\lambda})$.

Proof. This is implied by Proposition **??** and Ashlagi et al. (2013). □

Proposition B.14. *Suppose $m = (1+\lambda)n$. Then,* $\mathbb{E}\left[r_{k+1}^2\right] \geq \frac{2-p}{p} - O(\frac{\ln^5 m}{m})$, *where* $p = \frac{m-k}{m}$.

Proof. A straight-forward calculation gives

$$\mathbb{E}\left[r_{k+1}^2\right] = \sum_{j=0}^{k}(j+1)^2 \cdot (1 - \frac{k-j}{m-j}) \cdot \prod_{l=0}^{j-1} \frac{k-l}{m-l}. \tag{32}$$

Define $\bar{t} = \min\{k, 5\log_{1+\lambda}^n\}$. To provide a lower bound, we will only consider the first $\bar{t}$ summands in the above sum (the rest of the summands will be very small). Fix an arbitrary $t \leq \bar{t}$. We provide a lower bound for the summand corresponding to $j = t$. This summand contains the term $\prod_{l=0}^{t-1} \frac{k-l}{m-l}$, which is at least

$$\prod_{l=0}^{t-1} \frac{k-l}{m-l} \geq \prod_{l=0}^{t-1} \frac{k}{m} - \sum_{l=0}^{t-1} \left|\frac{k}{m} - \frac{k-l}{m-l}\right| \geq \prod_{l=0}^{t-1} \frac{k}{m} - \frac{\lambda t^2}{m-t} = (k/m)^t - \frac{\lambda \bar{t}^2}{2m}.$$

Now, using the above inequality, we can provide the following upper bound on (32):

$$\mathbb{E}\left[r_{k+1}^2\right] \geq \left(\sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1 - \frac{k}{m})(\frac{k}{m})^j\right) - \frac{\lambda \bar{t}^5}{2m}. \tag{33}$$

We are almost done. In the RHS of (33), we bound the first term from below by

$$\sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1-\frac{k}{m})(\frac{k}{m})^j \geq \frac{2-p}{p} - O(n^{-2}),$$

which holds because of the following well-known fact: $\mathbb{E}\left[Z^2\right] = \frac{2-q}{q}$ where Z is a geometric random variable with success probability q. Using the above bound, we can rewrite (33) as

$$\mathbb{E}\left[r_{k+1}^2\right] \geq \frac{2-p}{p} - O(\frac{\ln^5 m}{m}),$$

which completes the proof. □

B.5 Proof Sketch for Theorem 4.4

Finally, we describe how proof of Theorem B.7 can be adapted for Theorem 4.4. The main difference is in Lemma B.9. By proving a stronger version of Lemma B.9, the same proof would work for $\lambda > 0$.

We define the stronger version of Lemma B.9 simply by using, the variable $\kappa' = \frac{n}{0.5+(1+\lambda)\ln(1+1/\lambda)}$ instead of a variable κ. Replacing κ with κ' in the lemma statement would give the stronger version of the lemma.

The intuition on why the stronger version holds goes back to the definition of the coupling, $(\mathrm{DA}, \mathcal{B})$, which we defined for the proof. There, for each successful coin-flip (a proposal made to s), we removed n coins. However, instead of doing that, we could just remove a coin for each of the future proposals of the proposer. Everything else in the process remains identically the same, including the number of coins that will be flipped (this number will remain $2n\kappa(1-\delta)$).

In the new process, after each successful coin-flip (a proposal made to s by a school, namely, c), only x coins are removed where x is the number of future proposals of c. Since preferences of students are randomly generated, the expected value of x is about $\frac{n}{2k}$. Independence of preference lists allows us to use Chernoff bounds and imply that in the average, for each successful coin-flip only about $\frac{n}{2k}$ coins are removed wvhp (rather that n coins, which was the case in the original coupling). This will imply that the number of successful coin-flips isbe at least $\frac{\kappa}{nk(1+1/(2k))} \cdot (1-\delta) = \kappa'(1-\delta)$. Recall that the number of successful coin flips represent a lower bound on the number of proposals made to s. Consequently, at least $\kappa'(1-\delta)$ proposals are made to s, wvhp. This would prove the stronger version of Lemma B.9.

The rest of the proof is straight-forward. We follow the proof of Theorem 4.4 by rewriting (26) and (27) as follows.

$$
\begin{aligned}
\mathbb{E}\left[r_s^2|E_s\right] &\lesssim \mathbb{P}\left[E_s\right]\cdot\mathbb{E}\left[r_s^2|E_s\right] \\
&\lesssim (\frac{n}{\kappa'}-1)(\frac{2n}{\kappa'}-1) \qquad (34)\\
&= 2k^2-k \qquad (35)
\end{aligned}
$$

Now, (35) implies that

$$\lim_{n\to\infty}\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]=\lim_{n\to\infty}\mathbb{E}_\pi\left[r_s^2-r^2\right]\le 2k^2-k-k^2=k^2-k. \qquad (36)$$

(36) is an improved upper bound. On the other hand, recall that we had shown that

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right]\approx\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]$$

(in Step 1 of the proof of Lemma B.8). Consequently,

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right]}{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{MTB}})\right]}\ge\frac{\frac{2(1+\lambda)}{\lambda}-(1+\lambda)\ln(1+1/\lambda)-(1+\lambda)^2\ln(1+\frac{1}{\lambda})^2}{k^2-k.}.$$

where recall that $k=(1+\lambda)\ln(1+1/\lambda)$. The RHS of the above inequality is strictly greater than one for any constant $\lambda>0$.

C Improving pairs: proofs for Section 5

Theorem C.1. *Suppose $n=m+1$ and fix a student $s\in S$. Then, under MTB, we have*

$$\lim_{n\to\infty}\mathbb{P}\left[\ddot{\mu}(s)\ge\frac{n}{(\ln n)^{2+\epsilon}}\right]\to 1,$$

for any constant $\epsilon>0$.

We need the following definition and lemma before proving Theorem C.1.

Definition C.2. *For a fixed student s, we define a random variable $\Pi(s)$, which is a subset of preference profiles. We define $\Pi(s)$ by constructing it, this would implicitly define the corresponding support and probability mass function (PMF); we denote the PMF by $\mathcal{P}(s)$. We define $\Pi(s)$ by first defining a* partial preference profile $\hat{\pi}$*, as follows:*

1. *For all students $s' \neq s$, let $\hat{\pi}(s')$ be drawn independently uniformly at random.*

2. *Positions $\underline{r}, \dots, \overline{r}$ in $\hat{\pi}(s)$ are filled with schools $\underline{r}, \dots, \overline{r}$, respectively.*

$\Pi(s)$ contains the set of all preference profiles π who are consistent with $\hat{\pi}$. Given a realization $\Pi(s)$, let $\mathcal{U}(\Pi(s))$ denote the uniform distribution over the elements of $\Pi(s)$.

Lemma C.3. *Suppose $\Pi(s) \sim \mathcal{P}(s)$. Also, suppose π, π' are preference profiles that are drawn independently uniformly at random from $\Pi(s)$. Then, whp $\mu_\pi = \mu_{\pi'}$. (i.e., almost all student-optimal matchings in $\Pi(s)$ are identical, whp)*

Proof. π, π' are selected so that they are identical everywhere except on a fixed student, namely s; π, π' coincide on the interval $[n/(\ln n)^{1+\epsilon}, n/(\ln n)^{1-\epsilon}]$ of the preference list of s, but they are constructed independently (and uniformly at random) everywhere else in the preference list of s. (In other words, the schools listed in the interval $[n/(\ln n)^{1+\epsilon}, n/(\ln n)^{1-\epsilon}]$ of $\pi'(s)$ are the same as $\pi(s)$, but the rest of the schools in $\pi'(s)$ are shuffled randomly)

Using lemma C.5, we have

$$\begin{aligned}
&\mathbb{P}\left[\mu_\pi^{\#}(s) \notin [\underline{r}, \overline{r}] \bigvee \mu_{\pi'}^{\#}(s) \notin [\underline{r}, \overline{r}]\right] \\
\leq &\mathbb{P}\left[\mu_\pi^{\#}(s) \notin [\underline{r}, \overline{r}]\right] + \mathbb{P}\left[\mu_{\pi'}^{\#}(s) \notin [\underline{r}, \overline{r}]\right] = o(1). \qquad (37)
\end{aligned}$$

We have simply used a union bound in writing (37).

The preference list of each student $s' \neq s$ is the same in π, π'; also, whp, $\mu_\pi(s), \mu_{\pi'}(s)$ are both in the interval $[\underline{r}, \overline{r}]$ of the preference list of s. Since the preference lists $\pi(s), \pi'(s)$ are identical in this internval, then $\mu_\pi = \mu_{\pi'}$, whp.

□

Proof of Theorem C.1. For a preference profile π, define $B_\pi(s)$ to be the subset of students s' for which $\mu_\pi(s) \succ_{s'} \mu_\pi(s')$. Define $A_\pi(s)$ to be the subset of students s' for which $s' \in B_\pi(s)$, and moreover, $\mu_\pi(s') \succ_s \mu_\pi(s)$. The proof is done in two steps. In Step 1, we show that $|B_\pi(s)|$ is "large", whp. In Step 2, we show that $|A_\pi(s)|$ is "large", whp; this would prove the lemma.

Step 1. Consider an arbitrary school $c \in C$. We show that wvhp, there are "many" students who rank c above their match in the student-optimal matching. Then, taking a union bound over all schools $c \in C$ would show that wvhp, many students rank $\mu_\pi(s)$ above their current match, which means $|B_\pi(s)|$ is large. Instead of showing that many students rank c above their match in the student-optimal matching, we can equivalently show that c receives many proposals in the student-proposing DA. This is what Lemma C.7 proves.

We now formalize our proof sketch. By Lemma C.7, for any constant $\epsilon > 0$, each school receives at least $\frac{n(1-\epsilon)}{2\ln n}$ proposals wvhp, which also implies that all schools receive at least $\frac{n(1-\epsilon)}{2\ln n}$ proposals wvhp. Thus, $\mu_\pi(s)$ receives at least $\frac{n(1-\epsilon)}{2\ln n}$ proposals wvhp, which means for any constant $\epsilon > 0$, wvhp we have $|B_\pi(s)| > \frac{n(1-\epsilon)}{2\ln n}$. This completes Step 1.

Before going to Step 2, let us rephrase what we proved in Step 1. In Step 1, we showed that

$$\mathbb{P}_{\pi\sim\mathcal{P}}\left[|B_\pi(s)| > \frac{n(1-\epsilon)}{2\ln n}\right] \geq 1 - o(1), \tag{38}$$

where $\mathcal{P}$ denotes the uniform distribution over all preference profiles. Next, we write an alternative version of (38), which will be used later in Step 2.

Recall Definition C.2, by which $\Pi(s)$ is a random variable containing the set of all the possible placements of schools $[m]\backslash[\underline{r}, \overline{r}]$ in positions $[m]\backslash[\underline{r}, \overline{r}]$. Note that, without loss of generality, we can assume that schools listed on positions $\underline{r}, \ldots, \overline{r}$ of $\pi(s)$ are schools $\underline{r}, \ldots, \overline{r}$, respectively. Thus, we can rewrite (38) as

$$\mathbb{P}_{\Pi(s)\sim\mathcal{P}(s),\pi\sim\mathcal{U}(\Pi(s))}\left[|B_\pi(s)| > \frac{n(1-\epsilon)}{2\ln n}\right] \geq 1 - o(1). \tag{39}$$

Step 2 Lemma C.3 shows that, when $\Pi(s) \sim \mathcal{P}(s)$, almost all student-optimal matchings in $\Pi(s)$ (i.e. a fraction $1 - o(1)$ of them) are the same whp. Let μ denote this matching. Suppose that, for $\pi, \pi' \in \Pi(s)$, we have $\mu_\pi = \mu_{\pi'} = \mu$. Then, see that by the definition of $\Pi(s)$, we have $B_\pi(s) = B_{\pi'}(s)$. Thus, we let $B(s)$ denote $B_\pi(s)$ for any $\pi \in \Pi(s)$ for which $\mu_\pi = \mu$. Now, (39) implies that $|B(s)|$ is large, whp. This means, if $\pi \sim \mathcal{U}(\Pi(s))$, then, both of the events $\mu_\pi = \mu$ and $|B_\pi(s)| \geq \frac{n(1-\epsilon)}{2\ln n}$ hold whp. We use this fact to prove that $|A_\pi(s)|$ is large, whp. This would conclude Step 2.

Let $\pi \sim \mathcal{U}(\Pi(s))$. We will show that whp, a large number of schools in $B(s)$ have a rank better than $\underline{r}$ in $\pi(s)$. This would imply that $|A_\pi(s)|$ is large, whp. First, see that we can safely assume $\mu_\pi = \mu$ (and so $B_\pi(s) = B(s)$), since $\mu_\pi \neq \mu$ is a low-probability event (has probability $o(1)$) by Lemma C.3. This lets us safely assume that the event $\mu_\pi = \mu$ holds in the rest of the analysis.

Let $X(c)$ be a binary random variable which is 1 iff school c has a rank $\underline{r}$ or better in $\pi(s)$. Also, let $X = \sum_{c\in\mu(B(s))} X_c$. For any $c \in \mu(B(s))$, we have

$$\mathbb{P}\left[X_c = 1\right] \geq \frac{\underline{r}}{n} = \frac{1}{(\ln n)^{1+\epsilon}}.$$

Thus, $\mathbb{E}[X] \geq \frac{|B(s)|}{(\ln n)^{1+\epsilon}}$. A standard application of Chernoff bounds imply that for any $\delta > 0$, we have

$$\mathbb{P}[X < (1-\delta) \cdot \mathbb{E}[X]] \leq \mathsf{exp}\left(-\frac{\delta^2 \mathbb{E}[X]}{2}\right).$$

Thus, $|A_\pi(s)|$ is at least $\frac{(1-\delta)\cdot|B(s)|}{(\ln n)^{1+\epsilon}}$ whp. In Step 1, (39) shows that $|B(s)|$ is large whp. Consequently,

$$\mathbb{P}_{\Pi(s)\sim\mathcal{P}(s),\pi\sim\mathcal{U}(\Pi(s))}\left[|A_\pi(s)| \geq \frac{n(1-\epsilon)(1-\delta)}{2(\ln n)^{2+\epsilon}}\right] \geq 1 - o(1),$$

for any constants $\epsilon, \delta > 0$. This concludes Step 2 and completes the proof. □

For notational convenience in this section, we adopt the following definition.

Definition C.4. *Let $\underline{r}, \overline{r}$ respectively denote $n/(\ln n)^{1+\epsilon}, n/(\ln n)^{1-\epsilon}$.*

Lemma C.5. *Suppose $n = m + 1$ and fix a student $s \in S$. Then, for any constant $\epsilon > 0$ we have*

$$\mathbb{P}\left[\mu^{\#}(s) \notin \left[\frac{n}{(\ln n)^{1+\epsilon}}, \frac{n}{(\ln n)^{1-\epsilon}}\right]\right] = o(1).$$

Proof. Instead of proving the claim directly, we will show that

$$\mathbb{P}\left[\eta^{\#}(s) \notin [\underline{r}, \overline{r}]\right] = o(1). \tag{40}$$

Ashlagi et al. (2013) show that $\mathbb{P}[\mu(s) \neq \eta(s)] \leq \frac{\sqrt{\ln n}}{n}$, consequently, proving (40) would prove the lemma.

We use Lemma C.6 to prove (40). Let d denote the number of proposals received by s. Lemma C.6 implies that

$$\mathbb{P}[d < \underline{\alpha} \ln n] = o(1), \tag{41}$$

where $\underline{\alpha}$ is a positive constant. So, we can safely assume $d \geq \underline{\alpha} \ln n$. Let $X_1, \ldots, X_d$ be random variables that denote the utility of s from the j-th proposal she receives. Note that $\eta^{\#}(s) = \min\{X_1, \ldots, X_d\}$.

Since students preferences are chosen uniformly at random, we can write

$$\begin{aligned}\mathbb{P}\left[\eta^{\#}(s) \geq \overline{r}\right] &= \prod_{i=1}^{d} 1 - \frac{\overline{r}}{m-i+1} \\ &\leq \left(1 - \frac{\overline{r}}{m}\right)^d \leq e^{-\frac{d\overline{r}}{m}} \leq \mathsf{exp}(-\underline{\alpha}(\ln n)^\epsilon) = o(1). \end{aligned} \tag{42}$$

In the other hand, we have

$$\mathbb{P}\left[\eta^{\#}(s) \le \underline{r}\right] = 1 - \mathbb{P}\left[\eta^{\#}(s) > \underline{r}\right]$$
$$\le 1 - \prod_{i=1}^{d} 1 - \frac{\underline{r}}{m-i+1}$$
$$\le 1 - \left(1 - \frac{\underline{r}}{m-d}\right)^{d} \le 1 - \left(1 - \frac{d\underline{r}}{m-d}\right) \le O\left(\frac{\overline{\alpha}}{(\ln n)^{\epsilon}}\right) = o(1). \tag{43}$$

Taking a union bound over the bounds (41), (42), and (43) completes the proof.

□

Lemma C.6. *Suppose $n = m+1$. Fix an arbitrary small constant $\epsilon > 0$. Then, in the school-proposing DA, the number of proposals received by a fixed student in the school-proposing algorithm is whp at least $(1-\epsilon)\cdot\kappa$, where $\kappa = \frac{\ln n}{2}$.*

Proof. The same proof for Lemma B.9 works; we only need to use a new the definition for κ, which is stated in the lemmas statement. Everything else in the proof remains the same. □

Lemma C.7. *Suppose $n = m+1$. Then, for any positive constant ϵ, the number of proposals received by a fixed school in the student-proposing DA is wvhp at least $(1-\epsilon)\kappa$, where $\kappa = \frac{n}{2\ln n}$.*

Proof. The same proof for Lemma B.9 works; we only need to use a new the definition for κ, which is stated in the lemmas statement. Everything else in the proof remains the same. □

Theorem C.8. *Fix a student s. Under MTB, if $n < m$*

$$\lim_{n\to\infty} \mathbb{P}\left[\ddot{\mu}(s) \ge 1\right] \to 0.$$

Proof. Let $l = 3\ln^2 n$. Pittel (1989) proves that wvhp, every student is assigned to one of her top l choices. Let $L(s)$ denote the top l schools listed by student s. We show that for any student $s' \neq s$,

$$\mathbb{P}\left[|L(s)\cap L(s')| \ge 2\right] < O\left(\frac{\ln^4 n}{n^2}\right). \tag{44}$$

That is, the probability that (s,s') is a pareto improving pair is very low. If we prove (44), then the lemma is proved by a union bound over all $s' \neq s$; the union bound would imply that

$$\mathbb{P}\left[\ddot{\mu}(s) \ge 1\right] \le n \cdot O\left(\frac{\ln^4 n}{n^2}\right),$$

which proves the lemma.

To prove (44), first fix $L(s)$ and then start constructing $L(s')$ randomly. It is straight-forward to verify that

$$\begin{aligned}\mathbb{P}\left[|L(s) \cap L(s')| \geq 2\right] \leq \binom{l}{2} \cdot (l/m)^2 \leq l^4/m \\ = O\left(\frac{\ln^4 n}{n^2}\right).\end{aligned}$$

□