# What matters in school choice tie-breaking?
# How competition guides design

Itai Ashlagi          Afshin Nikzad*

## Abstract

Many school districts apply the student-proposing deferred acceptance algorithm after ties among students are resolved exogenously. This paper compares two common tie-breaking rules: one in which all schools use a common lottery, and one in which each school uses a separate independent lottery. We identify the balance between supply and demand as the determining factor in this comparison.

We analyze a two-sided assignment model with random preferences in over-demanded and under-demanded markets. In a market with a surplus of seats, a common lottery is less equitable, and there are efficiency tradeoffs between the two tie-breaking rules. However, a common lottery is always preferable when there is a shortage of seats in the sense of stochastic dominance of the rank distribution. The theory suggests that popular schools should use a common lottery to resolve ties. We run numerical experiments with New York City school choice data after partitioning the market into popular and non-popular schools. The experiments support our findings.

**JEL codes: D47, C78**

# 1 Introduction

Centralized assignment mechanisms that offer students a seat in a school have been adopted in many school districts around the world, including New York City, Boston, Denver, New Orleans, Amsterdam, and Santiago. Many of these districts use versions of the deferred acceptance (DA) algorithm (Gale and Shapley, 1962; Abdulkadiroğlu and Sönmez, 2003) to assign students to schools after eliciting families' preferences over schools. Schools, however, have much less detailed priorities for students, and many students fall into the same bin. So over-demanded schools must resolve ties when they have more applicants of a given priority than they can accommodate.

The question of how to break ties initially came up in school choice designs in New York City in 2003 (Abdulkadiroğlu et al., 2009) and later in other districts including Amsterdam in 2015 (De Haan et al., 2015).[1] Two natural tie-breaking rules were considered: a *single tie-breaking rule* (STB), under which each student receives a single lottery number to be used for tie-breaking by all schools, and a *multiple tie-breaking rule* (MTB), under which every school independently assigns to each applicant a random number that is used to break ties.[2]

Should a single lottery be conducted, or multiple independent lotteries? The literature does not provide a crisp answer to this question. A separate lottery for each school seems fairer, as students with bad draws at some schools may still have good draws at other schools, but at the same time may lead to unnecessary inefficiency (Abdulkadiroğlu and Sönmez, 2003).

Abdulkadiroğlu et al. (2009) and De Haan et al. (2015) find similar patterns in the data, which verify an intuitive trade-off. STB assigns more students to their top choices than MTB does, but MTB assigns fewer students to their lower-rank choices and leaves fewer students unassigned.[3] These numerical findings have led to different choices in practice: NYC adopted STB,[4] whereas Amsterdam initially adopted MTB, citing "fairness" as a major reason (Kennislink, 2015).[5][6]

The main message of this paper is that the trade-off between the two tie-breaking rules does not spread through the entire market. Loosely speaking, we find that the trade-offs between the tie-breaking rules disappear when restricting attention to the assignments to "popular" schools; within the set of popular schools a single lottery is preferable to independent lotteries. Our findings remove much of the earlier ambiguity and suggest that at least popular schools should use the same tie-breaker.

To understand better how tie-breaking rules affect students' assignments, we consider a stylized model that partitions schools into two tiers, *popular* and *non-popular*, based on student demand.

---

[1] The manner in which ties are resolved in matching markets impacts the assignment of students to their (top) choices (Abdulkadiroğlu et al., 2009).

[2] It is worth noting that under both STB and MTB, the student-proposing DA mechanism remains strategyproof, since ties are resolved independently of students' preferences.

[3] The difference in the assignments to the top choice is approximately 4% in favor of STB.

[4] In fact Abdulkadiroğlu et al. (2009) document that prior to the numerical experiments, NYC policymakers favored MTB due to its fairness.

[5] After the first year of using MTB (2014), however, Amsterdam switched to STB, following a lawsuit by two families who were interested in switching their assigned schools.

[6] Chile initiated a school choice system in 2017 and adopted the MTB rule.

All students prefer any popular school to any non-popular school. Within each tier, students rank schools uniformly at random, independently. Importantly, there are not enough seats in popular schools for all students, but there are enough seats overall.[7] This simple model will also help to explain the empirical observations when we revisit the NYC data.

We compare the impact of STB and MTB on students' assignments under the student-proposing DA using measures of efficiency and equity.[8] The first measure is the *rank distribution* of an assignment, which counts for each $r$ the number of students who are assigned to their $r$-th choice. We ask whether, and when, one rank distribution *rank-wise dominates* the other, which means that it assigns more students to their first choice, more students to their first or second choice, and so on. The second measure is the *variance* of rank distribution. Intuitively, the larger the variance, the larger the dispersion of students' ranks, which we interpret as a larger level of inequity. (We note that, in our theoretical model, every school is indifferent over all students.)

We find that the balance between demand and supply is the determining factor for both measures:

- In popular schools: the rank distribution of assigned students under STB (almost) rank-wise dominates the rank distribution under MTB. Moreover, the variance of rank distribution is lower under STB than under MTB.

- In non-popular schools: Neither rank distribution (of the assignments under STB and MTB) rank-wise dominates the other. Moreover, the variance of the rank distribution is higher under STB than under MTB.

These results imply that within the set of popular schools there is essentially no trade-off between our notions of efficiency and equity, since a single lottery generates better assignments than separate lotteries with respect to both measures. This contrasts with the intuition that MTB is fairer. For non-popular schools the decision remains ambiguous and consistent with the intuition; separate lotteries are more equitable than a single lottery, and more students are assigned to their higher choices under a single lottery whereas fewer students are assigned to their lower choices under MTB.[9] These findings also suggest the design of a hybrid tie-breaking rule that uses a single lottery in popular schools and separate lotteries in non-popular schools. In the stylized model with popular and non-popular schools, the rank distribution under the hybrid rule rank-wise dominates the rank distribution under the multiple tie-breaking rule.

Some intuition for why the trade-off vanishes in popular schools is the following. Throughout the DA algorithm[10] unassigned students apply to their favorite school they have not yet applied to,

---

[7]As explained in Section 2.1, since all students prefer the popular schools, one can equivalently compare the tie-breaking rules in a (simpler) single tiered random market with either a shortage or an excess of seats.

[8]Note that when schools have a single priority class, no assignment can result in a Pareto improvement for students over STB, because DA with STB is equivalent to a serial dictatorship mechanism, in which students select in sequential order their seats according to the tie-breaking order (Appendix 2.1). This is not necessarily the case with multiple priority classes. It is noteworthy, however, that any gains derived from finding an "optimal" stable assignment for students would require the use of a non-strategyproof mechanism (Abdulkadiroğlu et al., 2009).

[9]See also Ashlagi et al. (2019) and Arnosti (2015).

[10]The algorithm is described in detail in Section 2.

while schools tentatively admit students with the highest priority and reject the other. Consider a rejected student who applies to her next choice. Under MTB she is assigned a new lottery number, but under STB her lottery number remains the same. Therefore, she is more likely to get tentatively accepted under MTB than under STB. This implies that she is more likely to cause the rejection of another student under MTB. In popular schools, these chains of rejections are much longer under MTB, resulting in the rank-wise dominance relation.

We examine the predictions of our stylized model in the school choice data from New York City public high school assignments during the 2007–2008 school year. In the main assignment round, students submitted rank-ordered lists of at most 12 programs, and the deferred acceptance algorithm was used to assign students. First, consistent with the findings of Abdulkadiroğlu et al. (2009), neither the rank distribution under STB nor the rank distribution under MTB rank-wise dominates the other. Next we separate the market into popular and non-popular schools using a simple heuristic. We define the popularity of a school as the ratio between the number of students that rank the school as their first choice and the capacity of the school.[11] The set of popular schools is then defined as the set of schools whose popularity is above a certain threshold, which we call the *popularity threshold*. When restricting attention to students who are assigned to popular schools, we find that STB rank-wise dominates MTB. Unlike in the stylized model, there is no clear separation between popular and non-popular schools in the NYC data. However, rank-wise dominance holds when the popularity threshold is at least 1.[12] For the variance of rank distribution as well, we find qualitatively similar results in the data as in the stylized model. Finally, motivated by a lawsuit in Amsterdam filed by two families who requested to exchange seats, we test the vulnerability of MTB and STB to the existence of such pairs.[13] The experiments reveal that there are no such pairs under STB, while under MTB there are "many" such pairs in popular schools and "few" in non-popular schools. (See Sections 5 and G.2 for details.)

Many of our assumptions in the stylized model do not hold in the data (e.g., the market is not perfectly tiered, schools have different capacities, schools assign students to various priority classes prior to breaking ties), hinting at the robustness of the predictions of our model. That said, a limitation of our theoretical results is that they rely on the assumption that schools have unit capacities. (Extending the results to the case of larger capacities seems to be nontrivial.)

The effect of imbalance in two-sided marriage markets has been studied by Ashlagi et al. (2017), who compare the average rank of agents when preferences on both sides of the market are random. Here the concern is with students' *rank distribution* (rather than just average rank), which is relevant for practical engineering. From a technical perspective, the "approximate core-uniqueness" result

---

[11]When students' preferences are drawn from a multinomial logit model, this notion of popularity is an unbiased estimator for the weight of a school in that model normalized by its capacity (Appendix F).

[12]This threshold can be understood through our motivating stylized example with two tiers of schools (discussed earlier); schools in the popular tier have, in expectation, popularity larger than 1, whereas schools in the non-popular tier have, in expectation, popularity smaller than 1. In our experiments with NYC data, a popularity threshold of 1 results in rank-wise dominance while for some thresholds below 1 rank-wise dominance fails to hold (e.g., when the threshold is zero (Abdulkadiroğlu et al., 2009)).

[13]We intentionally focus on such short *Pareto-improving* cycles, which in comparison to longer cycles, are arguably easier for families to identify.

of Ashlagi et al. (2017) serves us as a first step in some of the proofs, which allows us to study the school-proposing DA instead of the student-proposing DA. The rest of the proofs are independent of their work. The proofs involve analyzing the stochastic processes that correspond to the school- and student-proposing DA.

## 1.1 Related Work

Closely related are papers that investigate the trade-offs between STB and MTB that were observed in Abdulkadiroğlu et al. (2009) and De Haan et al. (2015). Ashlagi et al. (2019) explain why STB assigns many more students to top choices than MTB does in a model with random preferences (even in a slightly under-demanded market). Independent of this work, Arnosti (2015) explains the single crossing point pattern using a cardinal utility model. His model, which assumes students' preference lists are short, is essentially equivalent to analyzing a market with a large surplus of seats. This paper takes a novel approach, which distinguishes between over-demanded and under-demanded schools, that explains the source of these trade-offs both theoretically and empirically.

This paper complements results by Ashlagi et al. (2017), who analyze the *average* student rank in unbalanced two-sided random markets. Their results, together with those of Knuth (1995), imply that the *average rank* of students is significantly better under STB than under MTB in an over-demanded market (one that has more students than seats), but these average ranks are essentially the same in a market with surplus seats. These papers limit attention to students' average rank and do not study the rank distributions.

Also related are papers that study economic properties in large random matching markets (Immorlica and Mahdian, 2005; Kojima and Pathak, 2009). Closest to our paper are studies on agents' ranks under DA (Pittel, 1989; Ashlagi et al., 2017) and on inefficiency under DA (Lee and Yariv, 2014; Che and Tercieux, 2019). Che and Tercieux (2019) find that in an over-demanded market, with high probability the assignment under MTB is not Pareto efficient. While the STB assignment is Pareto efficient in this model, it does not imply that it rank-wise dominates the MTB assignment.

The trade-off between incentives and efficiency when preferences or priorities contain indifferences has led to papers suggesting several novel tie-breaking 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 Abdulkadiroğlu et al. (2015), and the circuit tiebreaker by Che and Tercieux (2019).

Several papers study tie-breakings under the top trading cycles algorithm, which finds Pareto-efficient outcomes. Pathak and Sethuraman (2011) and Carroll (2014) extend results by Abdulkadiroğlu and Sönmez (1998) to show that under the top trading cycles algorithm (Shapley and Scarf (1974)), there is no difference between a single lottery (equivalently, random serial dictatorship) and multiple independent lotteries (top trading cycles with random endowments). Che and Tercieux (2018) show that all Pareto-efficient mechanisms (and not only top trading cycles) are asymptotically payoff-equivalent under certain assumptions.

4

# 2 Setup

A school choice market contains $n$ students and $m$ schools. We denote the set of students by $S$ and the set of schools by $C$. Each school $c \in C$ has a capacity of $q_c$ seats.

Each student $s$ has a *preference order* $\succ_s$, which is a strict linear order[14] over the set $C$. Student $s$ *prefers* school $c$ to $c'$ if $c \succ_s c'$. Student $s$ *weakly prefers* school $c$ to $c'$ if $c \succ_s c'$ or $c = c'$. We use the terms *preference order* and *preference list* interchangeably.

Each school $c$ has a *priority order* $\succeq_c$, which is a weak order (complete and transitive) over the set of students $S$. We denote by $\succ_c$ and $\sim_c$ the asymmetric and symmetric parts of $\succeq_c$, respectively. If $s \sim_c s'$, $s$ and $s'$ have the same priority at school $c$, and we say that they belong to the same *priority class at c*. A school priority order $\succeq_c$ partitions students into priority classes. A priority order of $c$ is *strict* if it is a strict linear order.

**Definition 2.1.** *A school choice market is* simple *if every school has a capacity of 1 seat and a single priority class containing all students.*

An *assignment* is a function $\mu : S \cup C \to C \cup 2^S \cup \{\varnothing\}$ such that for every student $s \in S$, $\mu(s) \in C \cup \{\varnothing\}$, for every school $c \in C$, $\mu(c) \in 2^S$ and $|\mu(c)| \leq q_c$, and furthermore, $\mu(s) = c$ if and only if $s \in \mu(c)$. In other words, $\mu$ assigns each student to at most one school and the number of students assigned to each school is less than or equal to the capacity of that school. If $\mu(s) = \varnothing$, then we say that $s$ is *unassigned*.

An assignment $\mu$ is *stable* if for every student-school pair $(s, c)$ such that $c \succ_s \mu(s)$, $|\mu(c)| = q_c$ and $s' \succeq_c s$ for all students $s'$ with $\mu(s') = c$.

A mechanism is a function that takes the preference orders of students and the priority orders of schools as its input and outputs an assignment. We focus on a class of mechanisms that are based on the *student-proposing deferred acceptance* (DA) algorithm (Gale and Shapley, 1962). The algorithm requires to receive as input a strict priority order from every school and a preference list from every student, and begins with all students being unassigned. It then assigns students in a sequence of rounds as follows. In the beginning of each round, every unassigned student applies to her most preferred school that she has not already applied to and is *tentatively* assigned to that school. Then, every school that has more tentatively assigned students than seats considers all of these students, and rejects students with the lowest priority, one at the time, until the remaining number of tentatively assigned students equals its capacity. The algorithm proceeds to the next round and terminates when every student is assigned to a school or when every unassigned student has applied to all schools.

The DA algorithm receives strict priority orders over students from the schools, whereas we are interested in school choice markets in which schools have weak priority orders. In such markets, DA can be adopted after breaking ties between students with equal priority at each school.

---

[14]Recall that a linear order is a transitive and complete preference relation (also called a weak order). A linear order $R$ is *strict* if there exist no two distinct elements $x, y$ in its domain such that both $xRy$ and $yRx$ hold.

A *tie-breaker at school c* is a strict priority order over students, denoted by $\succ^\rho$, which is used to transform the priority order of school $c$, $\succeq_c$, to a strict priority order, $\succ_c^\rho$, as follows. For any two students $s, s'$, $s \succ_c^\rho s'$ if

$$(s \succ_c s') \quad \text{or} \quad (s \sim_c s' \quad \text{and} \quad s \succ^\rho s').$$

Observe that the tie-breaker of a school with a single priority class is identical to the strict priority order of that school after ties are resolved. So the assignment in such markets is determined by the students' realized preferences and the schools' tie-breakers.

We next define two rules for generating tie-breakers that are commonly used in practice. In the *multiple tie-breaking* (MTB) rule, each school $c$ draws a tie-breaker independently and uniformly at random from the set of all strict linear orders over students. In the *single tie-breaking* (STB) rule, one strict linear order is drawn uniformly at random from the set of all strict linear orders over students, and all schools use that order as their tie-breaker.

Define DA-MTB to be the following assignment algorithm: it receives students' preference lists and schools' weak priority orders as its input, uses the multiple tie-breaking rule to transform the schools' weak priority orders to strict priority orders, and then outputs the assignment generated by the student-proposing DA algorithm. DA-STB is defined similarly, with the difference that it transforms schools' priority orders to strict orders using the single tie-breaking rule.

The theoretical results focus on comparing the assignments generated by DA-MTB and DA-STB in school choice markets with randomly generated preferences, which we will describe next.

## 2.1 Random simple markets

Our theory will focus on a class of simple school choice markets with randomly generated preferences; the preference order of every student is drawn independently and uniformly at random from the set of all strict preference orders over the schools.

To motivate the study of such school choice markets, it is useful to consider the following two-tiered market with popular and non-popular schools. Each student prefers any popular school to any non-popular school and preferences within each tier are drawn independently and uniformly at random. Suppose that the number of students is larger than the number of seats available in popular schools, but smaller than the number of seats available in all schools. Moreover, there is a single priority class in each school, and ties are resolved using MTB or STB. Observe that the outcome of DA in the two-tiered market can be generated by first running DA while ignoring non-popular schools, and then running DA with the remaining unassigned students and the non-popular schools. Thus, to compare DA-MTB and DA-STB in the two-tiered market, one can study a single-tiered market with either a shortage or a surplus of seats.

Motivated by the above observation, we analyze simple school choice markets in which the students' preference orders are generated independently and uniformly at random. These assumptions will be removed in Section 4 where we conduct numerical experiments using NYC data.

## 2.2 Notions of comparison

First, we define two operators for comparing vectors of the same dimensionality and the same L1 norm.[15] Let $Q, R \in R_+^d$ be $d$-dimensional vectors both with an L1 norm of $l$. We write $Q \preceq R$ if for every integer $i \in [1, d]$, $\sum_{j=1}^{i} R(j) \geq \sum_{j=1}^{i} Q(j)$, where $Q(j)$ and $R(j)$ denote the $j$-th entry of $Q$ and $R$, respectively. Also, for any $\epsilon > 0$ we write $Q \preceq^\epsilon R$ if, for every integer $i \in [1, d]$, either $\sum_{j=1}^{i} R(j) \geq \sum_{j=1}^{i} Q(j)$ or $\sum_{j=i}^{d} R(j) \leq (\log l)^{1+\epsilon}$. We write $Q \not\preceq^\epsilon R$ if $Q \preceq^\epsilon R$ does not hold.

Consider a school choice market with $n$ students and $m$ schools. The *rank* of a school $c$ for student $s$ is the number of schools that $s$ weakly prefers to $c$. Thus the most preferred school for $s$ has rank 1. Consider an assignment $\mu$. We say that the *rank of a student $s$ in $\mu$ is $r$* if, in $\mu$, $s$ is assigned to a school that has rank $r$ for her. The *rank distribution* of assignment $\mu$ is a vector $R_\mu \in \mathbb{Z}_+^m$ where $R_\mu(i)$ denotes the number of students who have rank $i$ in $\mu$.

**Definition 2.2.** *Given two assignments $\mu, \eta$, we say that $R_\mu$ rank-wise dominates $R_\eta$ if $R_\eta \preceq R_\mu$.*

We also consider a slightly weaker notion than rank-wise dominance.

**Definition 2.3.** *We say that $R_\mu$ $\epsilon$-rank-wise-dominates $R_\eta$ when $R_\eta \preceq^\epsilon R_\mu$.*

An intuitive way to think about the relation $R_\eta \preceq^\epsilon R_\mu$ is the following. First, exclude from $\eta$ the $(\log \min\{n, m\})^{1+\epsilon}$ students with the worst (i.e., largest) ranks; let $\eta'$ denote the resulting assignment. Define $\mu'$ from $\mu$ in the same way. Then, $R_\eta \preceq^\epsilon R_\mu$ holds if and only if $R_{\eta'} \preceq R_{\mu'}$ holds.

For every student $s$, $\mu^\#(s)$ denotes the rank of school $\mu(s)$ for student $s$. When $s$ is unassigned, we define $\mu^\#(s) = \infty$. Let $\mathcal{A}r(\mu)$ denote the average rank over students that are assigned under $\mu$.

**Definition 2.4.** *The* social inequity *in assignment $\mu$ is*

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

This notion measures the dispersion of students' ranks from the average rank. When all schools have a single priority class and the preference orders of all students are independently and identically distributed (iid), one can interpret a lower expected level of "dispersion" as a more equitable treatment of students. We note that since all students have complete preference lists, the number of unassigned students is the same in all stable assignments. Hence, we omit this number from the above definition. When the students' preference orders are iid, the expected social inequity is equal to the variance of a student's rank conditional on the student being assigned (Lemma D.4).

## 3 Main results

To present our main results, the following notations will be useful. A *preference profile* is a function $\varphi$ defined on $S$ such that for every student $s$, $\varphi(s)$ is the preference order of $s$. A *tie-breaking profile*

---

[15]Recall that the L1 norm of a vector is the sum of the absolute values of its entries.

is a function $\psi$ defined on $C$ such that for every school $c$, $\psi(c)$ gives the tie-breaker at $c$. Let $\Phi^{n,m}$ and $\Psi^{n,m}$ respectively denote the set of preference profiles and the set of tie-breaking profiles of all simple school choice markets with $n$ students and $m$ schools. Define $\Psi^{n,m}_{\text{MTB}} = \Psi^{n,m}$ and

$$\Psi^{n,m}_{\text{STB}} = \{\psi : \psi \in \Psi^{n,m} \text{ and } \psi(c) = \psi(c') \text{ for all } c, c' \in C\}.$$

A *market profile* is a tuple $\pi = (\varphi, \psi)$, where $\varphi$ is a preference profile and $\psi$ is a tie-breaking profile. For a simple school choice market with market profile $\pi$, the (deterministic) assignment generated by the student-proposing DA algorithm in that market is denoted by $\mu_\pi$. To simplify notation, we denote the rank distribution $R_{\mu_\pi}$ by $R_\pi$.

**Theorem 3.1.** *Consider a sequence of simple school choice markets indexed by $i = 1, 2, \ldots$, with $n_i$ students and $m_i$ schools in market $i$. In every market $i$, the students' preference profile $\varphi_i$ is drawn independently and uniformly at random from $\Phi^{n_i, m_i}$. In addition, $\psi_{i,\text{MTB}}$ and $\psi_{i,\text{STB}}$ are tie-breaking profiles that are drawn independently and uniformly at random from $\Psi^{n_i, m_i}_{\text{MTB}}$ and $\Psi^{n_i, m_i}_{\text{STB}}$, respectively. Define the market profiles $\pi_{i,\text{MTB}} = (\varphi_i, \psi_{i,\text{MTB}})$ and $\pi_{i,\text{STB}} = (\varphi_i, \psi_{i,\text{STB}})$.*

*(i) Suppose $m_i = i$ and $n_i = i + \lambda(i)$ for a function $\lambda : \mathbb{N} \to \mathbb{N}$. Then, for every constant $\epsilon > 0$,*

$$\lim_{i \to \infty} \mathbb{P}\left[R_{\pi_{i,\text{MTB}}} \preceq^\epsilon R_{\pi_{i,\text{STB}}}\right] = 1.$$

*Furthermore, if $\lambda(i) \leq i^\gamma$ for a positive constant $\gamma < \frac{3}{2}$ and every $i$, then*

$$\lim_{i \to \infty} \frac{\mathbb{E}_{\pi_{i,\text{MTB}}}\left[Si(\mu_{\pi_{i,\text{MTB}}})\right]}{\mathbb{E}_{\pi_{i,\text{STB}}}\left[Si(\mu_{\pi_{i,\text{STB}}})\right]} = \infty.$$

*(ii) Suppose $n_i = i$ and $m_i = i + \lambda(i)$ for a function $\lambda : \mathbb{N} \to \mathbb{N}$. Then, for every constant $\epsilon > 0$,*

$$\lim_{i \to \infty} \mathbb{P}\left[R_{\pi_{i,\text{MTB}}} \preceq^\epsilon R_{\pi_{i,\text{STB}}}\right] = 0, \quad \lim_{i \to \infty} \mathbb{P}\left[R_{\pi_{i,\text{STB}}} \preceq^\epsilon R_{\pi_{i,\text{MTB}}}\right] = 0$$

*and*

$$\lim_{i \to \infty} \frac{\mathbb{E}_{\pi_{i,\text{MTB}}}\left[Si(\mu_{\pi_{i,\text{MTB}}})\right]}{\mathbb{E}_{\pi_{i,\text{STB}}}\left[Si(\mu_{\pi_{i,\text{STB}}})\right]} = 0$$

*hold when there exists a positive constant $\gamma < 1$ such that $\lambda(i) \leq i^\gamma$ for all $i$.*

The first part of Theorem 3.1 shows that in over-demanded markets, STB performs better than MTB in the sense of $\epsilon$-rank-wise dominance and expected social inequity. Our proof does not directly use the condition that $\lambda(i) \leq i^\gamma$ for a positive constant $\gamma < \frac{3}{2}$. This condition is necessary for the concentration bounds on the number of proposal in DA, provided by Pittel (2019). He conjectures that this condition is not necessary for the concentration bounds to hold, and is merely a limitation of his proof approach.

The second part shows that in under-demanded markets, there can be a trade-off in terms of rank distributions, as the $\epsilon$-rank-wise dominance relation does not hold in either direction. Moreover, the

expected social inequity is lower under MTB, in contrast to the case of over-demanded markets. As opposed to the first part of the theorem, our proofs here directly rely on the condition that $\lambda(i) \leq i^{\gamma}$ for a positive constant $\gamma < 1$. We provide extensions (Theorem D.2) and simulations (Section 4) which show that qualitatively similar results hold when this condition is relaxed. Section 4 provides several other robustness checks for the theorem using simulations based on the theoretical model as well as Monte Carlo experiments using NYC school choice data.

Intuition and the main ideas behind the proof are discussed in Section 3.2. Before providing intuition for the main results, we discuss an alternative interpretation of the results from the perspective of individual students.

### 3.1 Alternative interpretations of the results

A variation of Theorem 3.1, namely Theorem E.1 in the appendix, allows us to interpret the findings from the perspective of individual students. For every $\tau \in \{\mathsf{MTB}, \mathsf{STB}\}$, define $\pi_{\tau} = (\varphi, \psi_{\tau})$ to be a market profile such that $\varphi$ and $\psi_{\tau}$ are drawn independently and uniformly at random from $\Phi^{n,m}$ and $\Psi_{\tau}^{n,m}$, respectively. Part (i) of Theorem E.1 shows that, for every constant $\epsilon > 0$, there exists $m_{\epsilon} > 0$ such that for all $m > m_{\epsilon}$ and $n > m$

$$\mathbb{E}_{\pi_{\mathsf{MTB}}}[R_{\pi_{\mathsf{MTB}}}] \preceq^{\epsilon} \mathbb{E}_{\pi_{\mathsf{STB}}}[R_{\pi_{\mathsf{STB}}}]. \tag{1}$$

To see how (1) relates to an individual student, let us first compute the probability that a student $s$ is assigned to the school that she ranks $i$-th, under either of the tie-breaking rules. Define the vector $R_{\tau} = \mathbb{E}_{\pi_{\tau}}[R_{\pi_{\tau}}]$. Since the students' preference orders are iid, then for every $i \leq m$, $\frac{R_{\tau}(i)}{\min\{n,m\}}$ is the probability that student $s$ has rank $i$ in $\mu_{\pi_{\tau}}$, i.e., the probability that $s$ has rank $i$ in the generated assignment conditional on being assigned. (Note that this probability remains the same even if we condition on the preference order of $s$, since the students' preference orders are iid.) Denote this probability by $p_{\tau}(i)$. Then, (1) implies that student $s$ "prefers" STB to MTB in an over-demanded market, in the sense that for every rank $r$

$$\sum_{i=1}^{r} p_{\mathsf{STB}}(i) \geq \sum_{i=1}^{r} p_{\mathsf{MTB}}(i)$$

holds when $m > m_{\epsilon}$, unless possibly when $\sum_{i=1}^{r} p_{\mathsf{STB}}(i) > 1 - \frac{(\log m)^{1+\epsilon}}{m}$.[16]

The notion of social inequity as well can be interpreted alternatively as a notion that concerns an individual student. Consider a student $s$ and fix her preference order. Let the random variable $r_{\tau}$ denote the rank of $s$ in DA when the tie-breakers are drawn according to rule $\tau$, and the other students' preference orders are drawn independently and uniformly at random from the set of all strict linear orders over the schools. Then, the expected social inequity in the assignment generated

---

[16]We note that the inequalities $\sum_{i=1}^{r} p_{\mathsf{STB}}(i) > 1 - \frac{(\log m)^{1+\epsilon}}{m}$ and $(\log m)^{1+\epsilon} > \sum_{i=r+1}^{m} R_{\mathsf{STB}}(i)$ are equivalent since $m(1 - \sum_{i=1}^{r} p_{\mathsf{STB}}(i)) = \sum_{i=r+1}^{m} R_{\mathsf{STB}}(i)$.

by DA is equal to the conditional variance of $r_\tau$ given that $s$ is assigned. This is a direct consequence of the fact that the students' preference orders are iid and is formally proved in Lemma D.4.

## 3.2 Intuition and proof ideas

### 3.2.1 Rank-wise dominance

To gain some intuition it is useful to first consider a simple example. Consider an over-demanded market with students $s_1, s_2, s_3$ and schools $c_1, c_2$ with unit capacities. Students $s_1$ and $s_2$ prefer $c_1$ to $c_2$, whereas $s_3$ prefers $c_2$ to $c_1$. All students belong to the same priority class in every school.

Under both multiple and single tie-breaking, in the first round of DA, $s_1$ and $s_2$ propose to $c_1$, whereas $s_3$ proposes to $c_2$. Therefore, $c_1$ has to reject one of $s_1, s_2$. Suppose $s_2$ is rejected. Then, in the second round of DA, $s_2$ would propose to $c_2$ (Figure 1a). The key observation is that it is more likely that $s_2$ gets rejected in the second round under STB than under MTB. The reason is that, since $s_2$ got rejected in the first round, under STB she is more likely to have a worse "lottery number" (i.e., her position in the common tie-breaking order) than $s_3$. Under MTB, however, $s_2$ is given a new lottery number when she is considered at $c_2$. Therefore, it is more likely that $s_2$ is accepted at $c_2$ in the second round under MTB than under STB. For the sake of intuition, suppose that indeed $s_2$ is accepted at $c_2$ in the second round under MTB but not under STB (Figures 1b and 1c). So all assigned students under STB have a rank of 1, whereas this is not the case under MTB. This implies that the rank distribution under STB rank-wise dominates the one under MTB.



(a) $s_2$ proposes to $c_2$ in the 2nd round.

(b) $s_2$ is accepted under MTB.
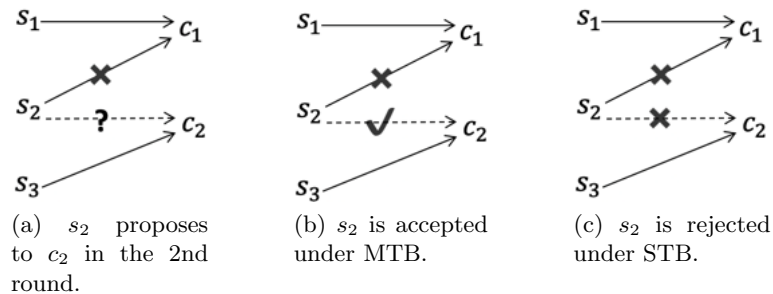
(c) $s_2$ is rejected under STB.

Figure 1: The solid and dotted arrows respectively correspond to students applying in rounds 1 and 2.

More generally, consider our model with a shortage of one seat. Under STB the student with the lowest lottery number will be unassigned by the end of DA. Under MTB, however, there is no student with the lowest lottery number. In the course of DA, students compete by displacing each other, so as not to end up as the sole unassigned student. This excessive competition worsens students' ranks overall, leading to the rank-wise dominance relation.

When there is an excess of seats, all students are eventually assigned. Hence, under MTB, the chain of students displacing each other in the course of DA ends much "faster" than when there is a shortage of seats, and the excessive competition between students gets alleviated. Thus, the rank-wise dominance relation does not occur. To illustrate this, consider the example above

10

without student $s_3$. Note that, then, $s_2$ is accepted in the second round without displacing any other student, and therefore all students would be assigned to their first choice under both tie-breaking rules. For further intuition, recall the result of Ashlagi et al. (2017) which implies that, under MTB, the total number of proposals made in the student-proposing DA is on average much smaller in an under-demanded market than in an over-demanded market. If there is an excess of one seat, the average number of proposals is proportional to $n \log n$, but with a shortage of one seat, this number is proportional to $\frac{n^2}{\log n}$.

Finally, we note that the formal proof approach is different from this intuitive argument, as we briefly discuss next.

**The proof approach.**  In over-demanded markets, the main idea is to show that there exists some rank $r > 1$ for which the following conditions hold *with high probability*:[17] (i) all but $(\log m)^{1+\epsilon}$ students have rank better than $r$ under STB, (ii) close to half of the students have rank 1 under STB, and (iii) at most about $0.4m$ students have rank better than $r$ under MTB.

To prove that conditions (i) and (ii) hold, we use the fact that the stochastic process governing DA-STB is equivalent to that of the Random Serial Dictatorship mechanism, under the assumption that there is a single priority class in each school (Appendix A.1).

The stochastic process governing DA-MTB is more complex to analyze. The idea for proving condition (iii) is to analyze the *school-proposing* DA rather than DA-MTB. The school-proposing DA is similar to DA-MTB, with the difference that students and schools switch roles: the schools propose and the students reject or tentatively accept proposals. In over-demanded markets, school-proposing DA is more tractable than DA-MTB, since it involves a significantly smaller total number of proposals.

The reason that, under MTB, we can analyze the school-proposing DA instead of DA-MTB follows from a result by Ashlagi et al. (2017). They consider imbalanced marriage markets, in which preference lists are generated independently and uniformly at random. They find that all but a vanishingly small fraction of men and women (students and schools, in our setup) have the same partner in all stable assignments. Employing this result, we prove condition (iii) for school-proposing DA: we show that each student receives at most about $\log m$ proposals *with high probability*, from which we infer that at most about $0.4m$ students have rank better than $r$.

### 3.2.2  Social inequity

To establish the results about social inequity we also analyze school-proposing DA rather than DA-MTB. When there are fewer students than schools, we prove that each student receives at least $\frac{n}{2 \log n}$ proposals, *with high probability*. The intuition is that the shortage of students creates a harsh competition between the schools, as schools compete over not remaining unassigned. This results in schools making many proposals in aggregate (Ashlagi et al., 2017; Pittel, 2019). We prove that,

---

[17]I.e., the probability that all of these conditions hold approaches 1 as $m$ approaches infinity. The asymptotic notions are formally defined and used in the appendix.

because of this effect, each student receives at least $\frac{n}{2\log n}$ proposals *with high probability*,[18] which implies that the variance of her rank is small (i.e., of the order of $\log^2 n$ or lower). The proof then concludes by noting the equivalence between the notions of variance and expected social inequity (as shown by Lemma D.4).

When there are excess students, the competition among schools is alleviated. Schools make significantly fewer proposals (in aggregate) compared to the previous case. We prove that, when there is an excess of students, each student receives at most about $\log m$ proposals *with high probability*, which sharply increases the variance of her rank. (That is, the variance would be of the order of $\frac{n^2}{\log^2 n}$ or higher.)

# 4    Computational experiments

This section presents findings from computational experiments that complement our theoretical results. A first set of simulations is based on the theoretical model and a second set of simulations uses NYC school choice data.

## 4.1    Simulations of the model

The first simulations illustrate the effect of market imbalance on the students' rank distributions under STB and MTB. For each market size, we draw multiple market profiles as follows. First, a preference list for each student is drawn independently and uniformly at random. Then, the schools' tie-breakers are generated as follows: Under MTB, for each school a strict linear order over students is drawn independently and uniformly at random. Under STB, a single strict linear order over students is drawn uniformly at random, and is used by all schools as a tie-breaker. (Because all students belong to the same priority class in every school, a school's tie-breaker coincides with its strict priority order after ties are resolved.) This completes the construction of a sample.

We construct 1000 samples independently and, for each sample, compute the student-optimal stable assignment. We then compute the *average cumulative rank distribution* which, for each rank $r$, gives the average number of students assigned to a rank at least as good as $r$ in the student-optimal stable assignment. The average is taken over all samples. For brevity, we refer to the average cumulative rank distributions as *cumulative rank distributions*.

Figure 2 presents the cumulative rank distribution under each tie-breaking rule in four markets, each with 1000 students and an excess or shortage of 1 or 100 seats, i.e., four markets with $1000 \pm 1$ and $1000 \pm 100$ seats. Each school has a capacity of one seat. Observe that when there is a shortage of seats (left panel), the cumulative rank distribution under STB rank-wise dominates the one under MTB. When there is a surplus of seats (right panel), there is no rank-wise dominance.

---

[18]We remark that, although this observation is reminiscent of the result in Ashlagi et al. (2017), the proof is through a different approach, since we need to provide a guarantee that holds for every student.

(a) Shortage of 1 seat  (b) Surplus of 1 seat

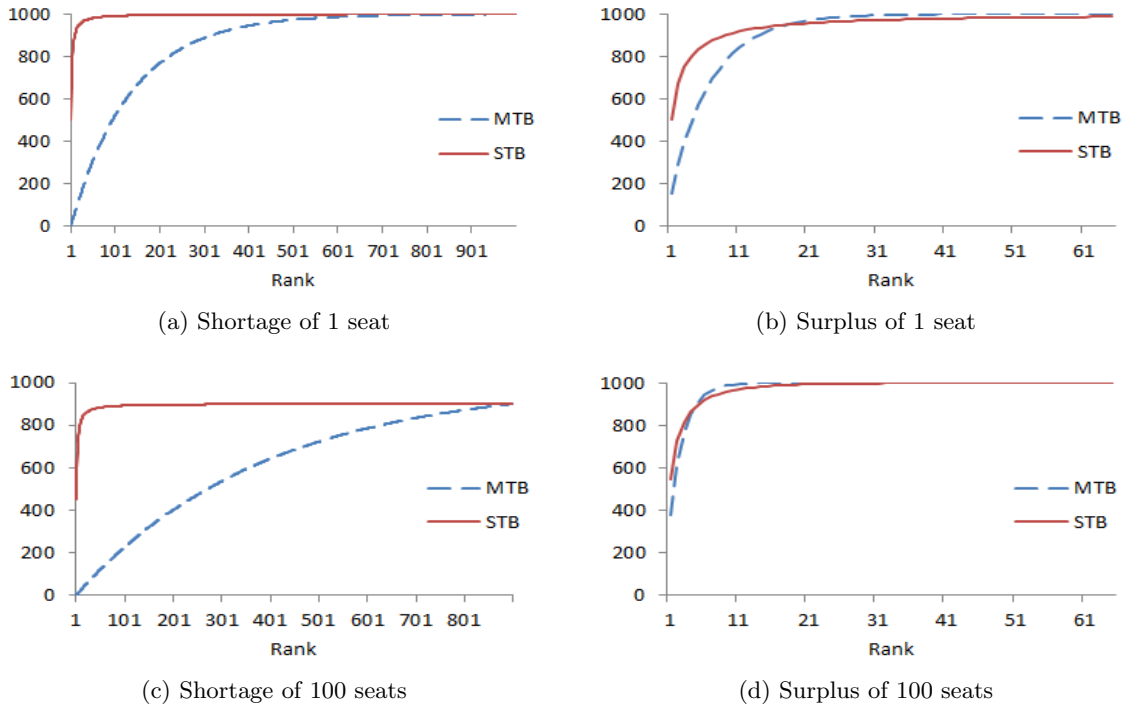(c) Shortage of 100 seats  (d) Surplus of 100 seats

Figure 2: The cumulative rank distributions under MTB and STB. The horizontal axis represents the rank, and the vertical axis the number of students. Each graph is generated by taking the average over rank distributions in 1000 random school choice markets. There are 1000 students in each market. Panels 2a and 2c plot the distributions in markets with a shortage of 1 seat and 100 seats, respectively. Panels 2b and 2d plot the distributions in markets with a surplus of 1 seat and 100 seats, respectively. The dashed and solid lines indicates the rank distributions under MTB and STB, respectively.

The next set of simulations illustrate the effect of market imbalance on the social inequity (which equals the variance of a student's rank conditional on the student being assigned). As in the previous simulations for each market size, we draw 1000 samples independently and compute for each sample the student-optimal stable assignment. We then compute the *average social inequity*, by taking the average of the social inequities of the student-optimal assignments, with the average taken over all samples.

Table 1 reports the average rank of the assigned students and the average social inequity for markets with varying imbalances and a single seat in each school. Observe that the average social inequity is larger under MTB (than under STB) when there is a shortage of seats and that it increases significantly as the shortage grows from 1 seat to 10 seats. Furthermore, notice that the variance of the rank is smaller under MTB when there is a surplus of seats.

| $n$ | $n - m$ | -10 | -1 | 1 | 10 |
|---|---|---|---|---|---|
| 100 | $\mathcal{A}r(\mu_{\text{STB}})/\mathcal{A}r(\mu_{\text{MTB}})$ | 2.52/2.54 | 3.78/4.1 | 4.14/20.3 | 4.23/32.3 |
| | $\mathcal{S}i(\mu_{\text{STB}})/\mathcal{S}i(\mu_{\text{MTB}})$ | 9.47/3.87 | 49.8/12.6 | 76.6/340.1 | 77.3/632.6 |
| 1000 | $\mathcal{A}r(\mu_{\text{STB}})/\mathcal{A}r(\mu_{\text{MTB}})$ | 4.53/4.59 | 6/6.46 | 6.46/137 | 6.48/209 |
| | $\mathcal{S}i(\mu_{\text{STB}})/\mathcal{S}i(\mu_{\text{MTB}})$ | 144.4/16.51 | 628.9/36.07 | 919/18606 | 949/37346 |

Table 1: Average rank ($\mathcal{A}r$) and social inequity under STB and MTB in the student-optimal stable assignment.

Next, we perform a robustness check for the findings of Theorem 3.1 about social inequity in under-demanded markets. Figure 3 compares the social inequities under MTB and STB in markets with a linear excess of seats. We fix the ratio of the number of seats to the number of students and report the average social inequity under both tie-breaking rules while varying the number of students from 2000 to 20,000. Figures 3a and Figure 3b present the results for the cases in which the ratio of seats to students is 1.1 and 1.5, respectively. Observe that the average social inequity is higher under STB than under MTB (as predicted by Theorem 3.1), and that the quantities remain essentially unchanged as the number of students increases.
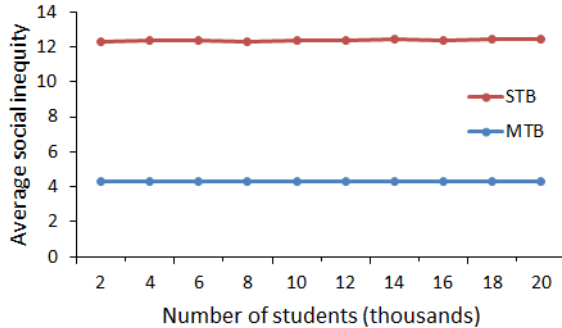


(a) The ratio of the number of seats to the number of students is 1.1.
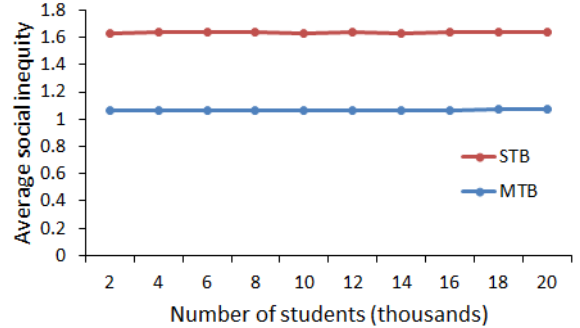
(b) The ratio of the number of seats to the number of students is 1.5.

Figure 3: Social inequities under MTB and STB in markets with a linear surplus of seats.

## 4.2 NYC school choice

Every year in New York City, approximately 90,000 students are assigned to roughly 700 public high school programs through a centralized matching mechanism. Until 2010 the matching process included three rounds of assignments; we focus on the main (second) round, in which about 80,000 students were assigned to schools using student-proposing DA.[19]

---

[19]The first round assigns students only to specialized exam schools. In the main round there were 79,524 seats and 79,403 students.

Each student that participated in this round submitted a rank-ordered list that included at most 12 schools, which we call the student's preference list. Programs at this round assign *coarse priorities* to students, and ties are broken exogenously using the STB rule. That is, every student was assigned a single lottery number, and whenever a school had to reject a subset of students from a set of students with equal priority, only then were the lottery numbers used to break the ties.

In each experiment we run DA-MTB and DA-STB, which take as input students' reported preferences and the *actual priorities* that schools assign to students.[20] In particular, if student $s$ belongs to a lower-priority class than student $s'$ in school $c$, then $c$ prioritizes $s$ over $s'$ under any tie-breaker. So, as done in practice, schools use lottery numbers generated by STB and MTB only to break the ties between students who belong to the same priority class.

There are three main deviations from the theoretical model in these experiments: (i) schools have different priorities over students, (ii) students' preferences are not generated randomly, but taken as given in the data, and (iii) schools have more than 1 seat.

### 4.2.1    A measure of school popularity

In the NYC data, schools are not naturally tiered as popular versus non-popular (as in the two-tiered market discussed in Section 2.1). We adopt a simple notion that defines whether a school is *popular* in the data, as follows. The *popularity* of a school $c$ is the ratio of the number of students for whom school $c$ is their top choice to the capacity of school $c$. Formally, given the actual preference lists of students in the NYC dataset, let $p_1(c)$ denote the number of students who list $c$ as their top choice and let $q_c$ be the capacity at school $c$. The popularity of a school $c$ is therefore $\alpha_c = \frac{p_1(c)}{q_c}$.[21] Note that every school $c$ with $\alpha_c \geq 1$ is fully assigned under student-proposing DA. It is also worth noting that if the students' preferences are drawn from a multinomial-logit discrete choice model (a common assumption when estimating students' preferences for schools; see, e.g., Pathak and Shi (2017)), this measure is an unbiased estimator for the "weight" of a school in that model normalized by its capacity. The formal relation with multinomial-logit discrete choice models is discussed in Appendix F.

A *popularity threshold* $\alpha$ determines a set $P_\alpha^\uparrow = \{c : c \in C, \alpha_c \geq \alpha\}$, which contains all schools with a popularity of at least $\alpha$. Such schools are called *popular* with respect to popularity threshold $\alpha$, and simply *popular* when $\alpha$ is clearly known from the context. The schools in the set $P_\alpha^\downarrow = C \backslash P_\alpha^\uparrow$ are called non-popular schools. Figure 4 reports the distribution of schools' popularity in the NYC data.

---

[20]Since preference lists are bounded in NYC, the mechanism is not strategyproof. For simplicity, however, we assume that students' observed preferences are sincere.

[21]We have repeated our experiments with a different notion of popularity, the *applicant-seat ratio*, formally defined as $\alpha_c' = \frac{p(c)}{q_c}$, where $p(c)$ is the total number of students who have school $c$ on their list. The NYC department of education suggests this notion to students as one way of figuring out how competitive a school is. The observations that we make in our experiments do not change qualitatively under this alternative notion of popularity.
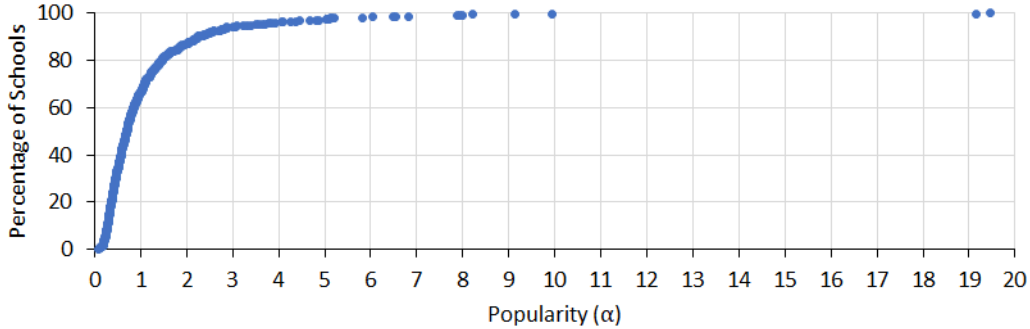
Figure 4: We first index schools in increasing order of popularity; let $c_1, \ldots, c_m$ be this ordering, where $m = 670$ is the total number of schools (i.e., programs) in our data from NYC. Then, for every school $c_i$, we place a dot in the above graph at position $(\alpha_{c_i}, \frac{i}{m} \times 100)$.

### 4.2.2 Rank-wise dominance

The *cumulative rank distribution in a set of schools* $D \subseteq C$ under assignment $\mu$ is a function $R_\mu^D : \{1, \ldots, m\} \to [0, 1]$ that, for each rank $r$, gives the total number of students that are assigned by $\mu$ to a school in $D$ that has a rank $r$ or better on their list. In the remainder of the section we assume that the assignment $\mu$ is the outcome of the student-proposing DA in a market that will be clearly known from the context.

We calculate the cumulative rank distributions in popular and non-popular schools for two different popularity thresholds. After fixing $\alpha$, for each $r \in \{1, \ldots, 12\}$ and each $D \in \{P_\alpha^\uparrow, P_\alpha^\downarrow\}$ we report the average value of $R_\mu^D(r)$, with the average taken over 50 samples that are drawn independently for each tie-breaking rule. We emphasize that: (i) both popular and non-popular schools are included in each sample, but we report the average cumulative rank distribution for the sets of popular and non-popular schools separately; and (ii) in every sample, students' preference lists and schools' priorities are taken as given in the data, and the (randomly drawn) tie-breaking rule is used only to resolve ties between students with equal priorities.

Figure 5 reports these statistics for popular schools with popularity thresholds $\alpha \in \{1, 1.5\}$. Observe that, for both popularity thresholds, the rank distribution under STB rank-wise dominates the one under MTB.

Figure 6 reports similar statistics for non-popular schools with popularity thresholds $\alpha \in \{1, 1.5\}$. Observe that for each $\alpha$, neither rank distribution rank-wise dominates the other. Although the plots for each $\alpha$ seem close to each other, the differences can be large, since many students are assigned to non-popular schools.[22] For instance, at $\alpha = 1.5$, STB assigns, on average, about 1800 more students to their top choice than MTB does.

We remark that in every school $c$ (regardless of its popularity), the number of students assigned to $c$ that rank it as their first choice is larger under STB than under MTB.[23] So, for any popularity

---

[22] For example, for $\alpha = 1.5$, 15,000 and 60,000 students are assigned to popular and non-popular schools, respectively.

[23] To see why this holds intuitively, observe that a student who is tentatively assigned to her first choice is less
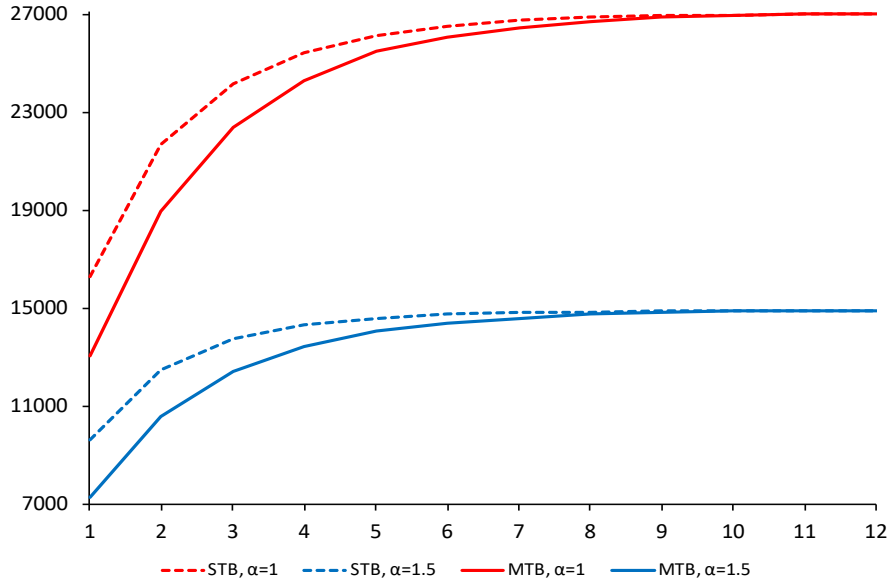
Figure 5: The cumulative rank distributions under MTB and STB for popular schools with different popularity thresholds $\alpha$ (schools with popularity above $\alpha$ are popular). The dashed and solid lines indicate the rank distributions under STB and MTB, respectively. The horizontal axis represents the rank, and the vertical axis the number of students.

threshold, the rank distributions in non-popular schools cross each other at a rank of at least 2.

### 4.2.3   Social inequity

Table 2 reports the average social inequity in popular schools. After fixing the popularity threshold, we draw 50 samples independently for each tie-breaking rule. For each sample, we compute the outcome of the student-proposing DA, and compute the *social inequity in popular schools* in that sample as follows: We take the average, over all students assigned to popular schools, of the squares of the differences between the rank of a student and the average rank of students in popular schools. (This quantity is just the sample variance of the rank of the students assigned to popular schools.) We report the average of the social inequities computed in the 50 samples. Similarly, Table 3 reports the average social inequity in non-popular schools. Consistent with Theorem 3.1, we observe that MTB results in higher social inequity than STB in popular schools, but lower social inequity in non-popular schools.

## 5   Discussion of practical implications

In over-demanded markets the trade-off between a single lottery and multiple independent lotteries disappears: STB outperforms MTB in both notions we consider, rank-wise dominance and social inequity. Hence, using multiple lotteries in "popular" schools leads to students being assigned to

---

likely to be rejected under STB than under MTB, as we discussed in Section 3.2.1.
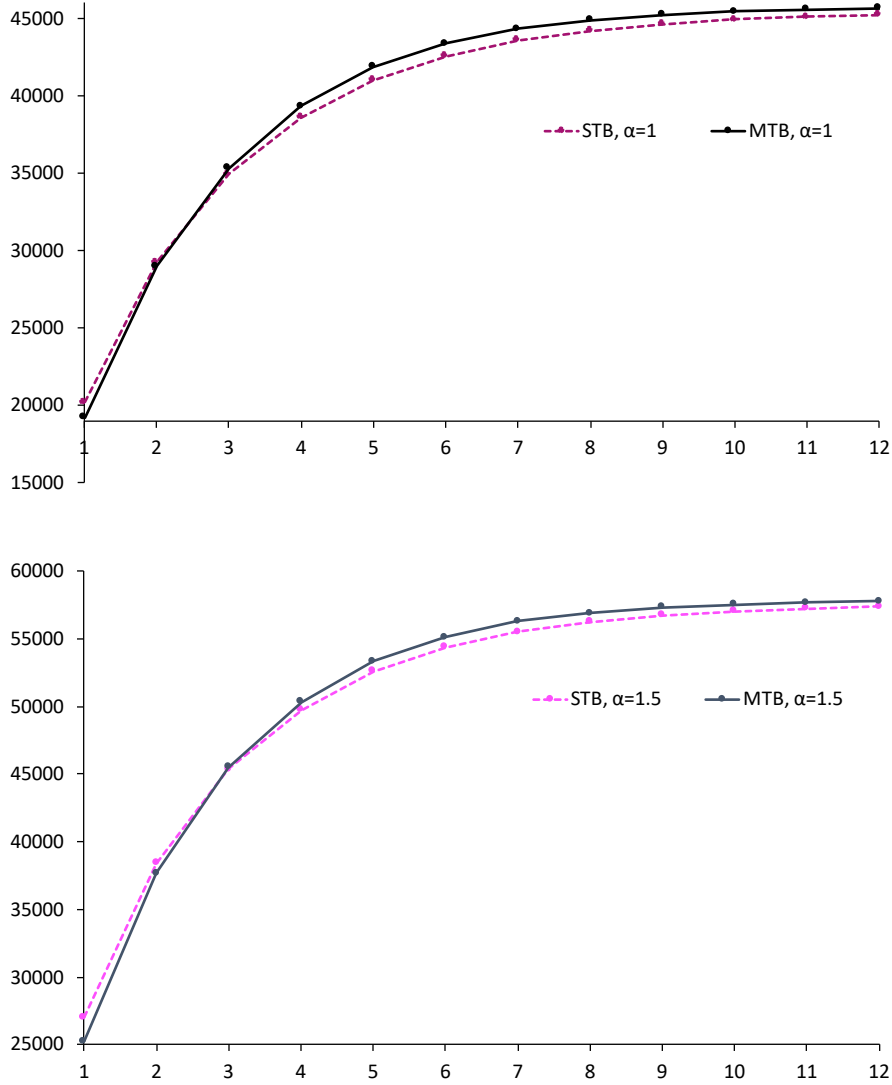
Figure 6: The average cumulative rank distributions under MTB and STB for non-popular schools. The dashed and solid lines indicate the rank distributions under STB and MTB, respectively. The horizontal axis represents the rank and the vertical axis the number of students. At $\alpha = 1$ and $\alpha = 1.5$ about $57,300$ and $57,700$ students are assigned to non-popular schools, respectively.

| $\alpha$ | | STB | MTB |
|---|---|---|---|
| 1 | Social inequity | 2.10 | 2.99 |
| | Average rank | 1.83 | 2.21 |
| 1.5 | Social inequity | 1.47 | 2.87 |
| | Average rank | 1.65 | 2.18 |

Table 2: Social inequity and the average rank in popular schools.

relatively worse ranks in those schools than using a single lottery.

One concern that arises when implementing an inefficient assignment is that students may

| $\alpha$ | | STB | MTB |
|---|---|---|---|
| 1 | Social inequity | 4.22 | 3.69 |
| | Average rank | 2.52 | 2.50 |
| 1.5 | Social inequity | 3.90 | 3.58 |
| | Average rank | 2.41 | 2.44 |

Table 3: Social inequity and the average rank in non-popular schools.

realize that they are part of a Pareto-improving cycle. This indeed happened in Amsterdam; two families filed a lawsuit to exchange their children's assignments (De Haan et al. (2015)). Although the judge ruled against the case, the following year Amsterdam adopted STB. We ran numerical experiments using NYC data to examine how common is it to be part of a pair of students who wish to swap schools. These are referred to as *Pareto-improving pairs*. In these experiments (Appendix G.2) we focus on cycles of length two because these are arguably easier for families to detect after the assignment. We find that there are no Pareto-improving pairs under STB, while under MTB there are "many" Pareto-improving pairs in popular schools and "few" in non-popular schools. This is consistent with the findings of Theorem 3.1 regarding the inefficiency of MTB in popular schools.

Finally, for policymakers who favor MTB, e.g., due to equity reasons,[24] a possible alternative would be to use independent lotteries in non-popular schools and a common lottery in popular schools. Such a hybrid rule can lead to (i) lower social inequity in non-popular schools than under STB, and (ii) fewer Pareto-improving pairs in popular schools than when using MTB alone. For instance, a hybrid rule may have been an attractive alternative to MTB in Amsterdam, which has four "very popular" schools (De Haan et al., 2015); such a rule may have also prevented the lawsuit. These insights are confirmed by our computational experiments using the NYC data (Appendices G.1 and G.2).

## 6 Conclusion

This paper revisits the impact of tie-breaking rules on students' assignments in school choice. Splitting the market into popular and non-popular schools proves useful in explaining the source of the differences and trade-offs between STB and MTB with respect to our notions of efficiency and equity. The trade-offs vanish within the set of popular schools but persist in the set of non-popular schools. These insights reduce the ambiguity documented by previous studies (Abdulkadiroğlu et al., 2009; De Haan et al., 2015).

The assumptions in the model do not typically hold in practice. The model assumes that schools have unit capacities and a single priority class. Nevertheless, the computational experiments with NYC data indicate the robustness of the predictions of the model.

---

[24]As in Amsterdam and Chile.

The popularity measure adopted in the empirical experiments is heuristic. It remains an interesting direction to develop well-grounded empirical measures for popularity.

This study adds another rationale for selecting a single lottery for breaking ties (Pathak, 2017). Policymakers who favor MTB may find attractive a hybrid tie-breaking rule (a common lottery in popular schools and independent lotteries in non-popular schools), in that it is likely to reduce the inefficiencies of MTB in popular schools while maintaining some of its benefits, such as fewer unassigned students.

# 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.

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

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

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. ACM Conference on Economics and Computation (EC), 2015.

Itai Ashlagi, Yash Kanoria, and Jacob D. Leshno. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy*, 125(1):69–98, 2017.

Itai Ashlagi, Afshin Nikzad, and Assaf Romm. Assigning more students to their top choices: A comparison of tie-breaking rules. *Games and Economic Behavior*, 115:167–187, 2019.

Pierre Bremaud. *Discrete Probability Models and Methods*. Springer, 2017.

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. Payoff equivalence of efficient mechanisms in large matching markets. *Theoretical Economics*, 13(1):239–271, 2018.

Yeon-Koo Che and Olivier Tercieux. Efficiency and stability in large matching markets. *Journal of Political Economy*, 127(5):2301–2342, 2019.

Monique De Haan, Pieter A. Gautier, Hessel Oosterbeek, and Bas Van der Klaauw. The performance of school assignment mechanisms in practice. Working paper, 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.

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

Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 53–62. Society for Industrial and Applied Mathematics, 2005.

Kennislink. Why exchange is not allowed, even if both parties benefit from it. 2015. URL https://www.nemokennislink.nl/publicaties/schoolstrijd-in-amsterdam/.

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

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

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

SangMok Lee and Leeat Yariv. On the efficiency of stable matchings in large markets. Working paper, 2014.

David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

Parag Pathak and Peng Shi. How well do structural demand models work? counterfactual predictions in school choice. Technical report, National Bureau of Economic Research, Inc, 2017.

Parag A Pathak. What really matters in designing school choice mechanisms. *Advances in Economics and Econometrics*, 1:176–214, 2017.

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

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

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

Boris Pittel. On likely solutions of the stable matching problem with unequal numbers of men and women. *Mathematics of Operations Research*, 44(1):122–146, 2019.

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

Mark B Villarino. Sharp bounds for the harmonic numbers. *arXiv preprint math/0510585*, 2005.

Neil A Weiss. *A course in probability*. Addison-Wesley, 2006.

# A    Preliminaries: notations and definitions

We will analyze *simple school choice markets*, as defined in Definition 2.1. For brevity, we often use the term *market* instead throughout our analysis in the appendix.

We recall the following definitions: $\Phi^{n,m}$ and $\Psi^{n,m}$ respectively denote the set of preference profiles and the set of tie-breaking profiles of all simple school choice markets with $n$ students and $m$ schools. Define $\Psi^{n,m}_{\mathsf{MTB}} = \Psi^{n,m}$ and

$$\Psi^{n,m}_{\mathsf{STB}} = \{\psi : \psi \in \Psi^{n,m} \text{ and } \psi(c) = \psi(c') \text{ for all } c, c' \in C\}.$$

A market profile is a tuple $\pi = (\phi, \psi)$ where $\phi$ is a preference profile and $\psi$ is a tie-breaking profile. For a student $s$, we let $\pi(s)$ denote $\phi(s)$, and for a school $c$, we let $\pi(c)$ denote $\psi(c)$. The set of all market profiles in a simple school choice market with $n$ students and $m$ schools is denoted by $\Pi^{n,m}$. Define $\Pi^{n,m}_{\mathsf{MTB}} = \Pi^{n,m}$. Also, define

$$\Pi^{n,m}_{\mathsf{STB}} = \{\pi : \pi \in \Pi^{n,m} \text{ and } \pi(c) = \pi(c') \text{ for all } c, c' \in C\}.$$

In a market with market profile $\pi$, the student-optimal and the school-optimal stable assignments are denoted by $\mu_\pi$ and $\eta_\pi$, respectively. Recall that these assignments are the outcomes of the student- and school-proposing DA, respectively (Gale and Shapley, 1962).

Given an assignment $\mu$, let $\mu(s)$ denote the school assigned to student $s$, and $\mu(c)$ denote the student assigned to school $c$ in $\mu$. If $s$ is unassigned under $\mu$, we define $\mu(s) = \varnothing$. Similarly, we define $\mu(c) = \varnothing$ if no student is assigned to $c$. For any $X \subseteq S \cup C$, let $\mu(X) = \cup_{x \in X}\mu(x)$.

For every student $s$, let $\mu^{\#}(s)$ denote the rank of school $\mu(s)$ on the preference order of $s$. That is, $\mu^{\#}(s) = 1$ if $s$ is assigned to her first choice, $\mu^{\#}(s) = 2$ if $s$ is assigned to her second choice, and so on. When $s$ is unassigned in $\mu$, we define her rank in $\mu$ to be $\infty$, i.e., $\mu^{\#}(s) = \infty$. By convention, we say that a rank $r$ is *better* (*worse*) than a rank $r'$ if $r$ is smaller (larger) than $r'$.

Let $\mathcal{A}r(\mu)$ denote the average rank of students that are assigned in $\mu$, i.e., $\mathcal{A}r(\mu) = \frac{1}{|\mu(C)|} \cdot \sum_{s \in \mu(C)} \mu^{\#}(s)$.

A rank distribution is a function from positive integers to nonnegative integers. For any rank distribution $R$, denote by $R^+$ its corresponding *cumulative rank distribution,* which is defined by

$R^+(k) = \sum_{i=1}^{k} R(i)$ for every integer $k > 0$. The rank distribution of an assignment $\mu$ is a vector $R_\mu \in \mathbb{Z}_+^m$ where $R_\mu(i)$ denotes the number of students who are assigned to their $i$-th choice in the assignment $\mu$. The cumulative rank distribution of $\mu$ then is denoted by $R_\mu^+$.

For a finite set $X$, let $U(X)$ denote the uniform distribution over $X$.

## A.1 Equivalence of DA-STB and Random Serial Dictatorship (RSD)

We will see here that, in a simple school choice market, the assignment under DA-STB can be attained using a very simple process. For exposition, interpret the common tie-breaker used by DA-STB as priorities (with lower numbers indicating higher priorities).

Assume all seats are initially empty and all students are unassigned. When running DA-STB, assume that the student with the highest priority among unassigned students is the next student to propose.[25] Observe that the student with the highest priority (who proposes first) will never be rejected and therefore will be assigned to her top choice. Consider the student with the second-highest priority. Either her first choice has an empty seat and she is assigned to it, or her first choice is tentatively assigned to the highest priority student, in which case she will be assigned to her second choice. This process continues until all students are assigned or there are no more empty seats. Observe that this process is equivalent to the algorithm in which students choose schools one at a time only from schools with empty seats according to a (random) priority order. This is the definition of Random Serial Dictatorship (Abdulkadiroğlu and Sönmez, 1998).

## A.2 The Principle of Deferred Decisions

The idea behind The Principle of Deferred Decisions is that the entire set of random choices required to run an algorithm are not made in advance, but rather during the run of the algorithm. We use this principle when analyzing the school-proposing DA algorithm.

Recall that students' preference orders are drawn i.i.d from the set of all strict linear orders over the schools. One way to run school-proposing DA then would be to determine the students' preferences in the course of running DA, rather than drawing their preferences before DA is run. That is, we can assume that, upon receiving each proposal, the student assigns a distinct *rank* to the proposing school, which is an integer between 1 and $m$; this rank is chosen uniformly at random from the set of ranks that the student has not assigned so far. The assignment generated in this way has the same distribution as when the student's preferences are generated before running DA.

## A.3 Asymptotic notions

We say a statement $\mathcal{S}(i)$ holds for *sufficiently large i* if there exists $i_0$ such that $\mathcal{S}(i)$ holds for all $i > i_0$.

Let $E(i)$ be an event parameterized by a positive integer $i$. We say that $E(i)$ occurs with *high probability as i grows large* if $\lim_{i \to \infty} \mathbb{P}[E(i)] = 1$. When $i$ is clearly known from the context, we

---

[25]The order in which students propose has no impact on the final assignment (Gale and Shapley, 1962).

simply say that $E(i)$ occurs *with high probability* or, briefly, $E(i)$ occurs *whp.*

Furthermore, we say that $E(i)$ occurs *with very high probability as i grows large* if there exists $\alpha > 1$ such that $\lim_{i \to \infty} \frac{1 - \mathbb{P}[E(i)]}{\exp(-(\log i)^\alpha)} = 0$. When $i$ is clearly known from the context, we simply say that $E(i)$ occurs *with very high probability*; or briefly, $E(i)$ occurs *wvhp.*

For any two functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$ we adopt the notation $f = o(g)$ when for every positive constant $\epsilon$ there exists a constant $i_\epsilon$ such that $f(i) \le \epsilon g(i)$ holds for all $i > i_\epsilon$.

We define $f = O(g)$ if there exist positive constants $i_0, \Delta$ such that $f(i) \le g(i)\Delta$ holds for all $i > i_0$. We write $g = \Theta(f)$ when $f = O(g)$ and $g = O(f)$, and $g = \Omega(f)$ when $f = O(g)$.

## A.4  Inequalities

**Fact A.1.** *When $A \le B$, the function $f(x) = \frac{A-x}{B-x}$ is decreasing at all $x < B$.*

*Proof.* Observe that $f'(x) = \frac{A-B}{(B-x)^2} \le 0$. $\qquad\square$

**Fact A.2.** *For every $m > 0$, the function $f(x) = \frac{\log(1 + \frac{m}{x})}{\log(m+1)} \cdot \frac{m+x}{m}$ is decreasing at all $x \ge 1$.*

*Proof.* Observe that $f'(x) = \frac{x \log\left(\frac{m+x}{x}\right) - m}{mx \log(m+1)}$. Since $\log(1 + m/x) \le m/x$ holds, then $x \log(1 + m/x) \le m$. Therefore, $f'(x) \le 0$. $\qquad\square$

**Fact A.3.** *For any positive integers $m, x$ we have $(m+x) \log(1 + \frac{m}{x}) \ge m$.*

*Proof.* Define $g(x) = (m+x) \log(1 + \frac{m}{x})$. By Fact A.2, $g(x)$ is decreasing in $x$. Therefore, to prove the claim, it suffices to show that $\lim_{x \to \infty} \frac{g(x)}{m} = 1$. To this end, observe that

$$\lim_{x \to \infty} \frac{g(x)}{m} = \lim_{x \to \infty} \frac{\log(1 + \frac{m}{x})}{\frac{m}{m+x}} = \lim_{x \to \infty} \frac{m+x}{x} = 1,$$

where the second equality follows from L'Hopital's rule. This concludes the proof. $\qquad\square$

**Fact A.4.** *For any positive integers $n, m$ with $n > m$,*

$$\frac{m}{n} \cdot \frac{1}{\log \frac{n}{n-m}} \ge \frac{1}{\log \frac{n}{n-m}} - \frac{n-m}{m}.$$

*Proof.* We apply Fact A.3 with $x = n - m$, which implies that

$$(m+x) \log(1 + \frac{m}{x}) \ge m$$
$$\Leftrightarrow \frac{x}{m} \log(1 + \frac{m}{x}) \ge \frac{x}{m+x}$$
$$\Leftrightarrow 1 - \frac{x}{m+x} + \frac{x}{m} \log(1 + \frac{m}{x}) \ge 1$$

Plugging $x = n - m$ into the above inequality implies that

$$\frac{m}{n} + \frac{n-m}{m} \log \frac{n}{n-m} \ge 1.$$

This implies that

$$\frac{m}{n} \geq 1 - \frac{n-m}{m} \log \frac{n}{n-m}.$$

Dividing both sides of the above inequality by $\log \frac{n}{n-m}$ concludes the proof.

$\square$

**Fact A.5.** *Suppose that $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are real numbers belonging to the unit interval such that $y_i \leq x_i$ holds for all $i$. Then*

$$\prod_{i=1}^{n}(x_i - y_i) \geq \prod_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i.$$

*Proof.* The proof is by induction on $n$. The induction basis for $n = 1$ is trivial. For the induction step, we show that the claim holds, assuming that

$$\prod_{i=1}^{n-1}(x_i - y_i) \geq \prod_{i=1}^{n-1} x_i - \sum_{i=1}^{n} y_i. \tag{2}$$

By the above inequality, we can write

$$\prod_{i=1}^{n}(x_i - y_i) \geq \left( \prod_{i=1}^{n-1} x_i - \sum_{i=1}^{n-1} y_i \right)(x_n - y_n) \tag{3}$$

$$\geq \prod_{i=1}^{n} x_i - y_n \prod_{i=1}^{n-1} x_i - x_n \sum_{i=1}^{n-1} y_i \tag{4}$$

$$\geq \prod_{i=1}^{n} x_i - y_n - \sum_{i=1}^{n-1} y_i = \prod_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i, \tag{5}$$

which completes the induction step.

$\square$

# B   Preliminaries: concentration bounds

## B.1   Chernoff bounds

Chernoff bounds are well-known concentration inequalities that bound the deviations of a weighted sum of Bernouli random variables from its mean. Here we present their multiplicative form (see, e.g., Bremaud (2017)). Let $X_1, \ldots, X_n$ be a sequence of $n$ independent random binary variables such that $X_i = 1$ with probability $p_i$ and $X_i = 0$ with probability $1 - p_i$. Let $\alpha_1, \ldots, \alpha_n$ be arbitrary

real numbers in the unit interval. Also, let $\mu = \sum_{i=1}^{n} \alpha_i \mathbb{E}[X_i]$. Then for any $\epsilon$ with $0 \le \epsilon \le 1$:

$$\Pr\left[\sum_{i=1}^{n} \alpha_i X_i > (1+\epsilon)\mu\right] \le e^{-\epsilon^2 \mu/3} \tag{6}$$

$$\Pr\left[\sum_{i=1}^{n} \alpha_i X_i < (1-\epsilon)\mu\right] \le e^{-\epsilon^2 \mu/2}. \tag{7}$$

## B.2  Concentration bounds for stable assignments

The following three theorems are immediate corollaries from Pittel (2019) translated to our setting.

**Theorem B.1** (Theorem 1.4 in Pittel (2019))**.** *Consider a simple school choice market with $n$ students and $m = n + \lambda(n)$ schools where $\lambda : \mathbb{N} \to \mathbb{N}$ is an arbitrary function. Let the market profile be drawn from $\Pi_{\mathrm{MTB}}^{n,m}$ uniformly at random. Also, let $M$ denote the set of all assignments that are stable given the drawn market profile. For any $\mu \in M$, let $s(\mu)$ denote the sum of ranks of the assigned students in $\mu$. Then, for any constant $\epsilon > 0$, the event*

$$\max_{\mu \in M} \left| \frac{s(\mu)}{m \log \frac{m}{m-n}} - 1 \right| < \epsilon$$

*holds with very high probability as $n$ grows large.*

**Theorem B.2** (Theorem 1.5 in Pittel (2019))**.** *Consider a simple school choice market with $m$ schools and $n = m + \lambda(m)$ students where $\lambda : \mathbb{N} \to \mathbb{N}$ is such that $\lambda(m) \le m^{3/2-d}$ holds for all $m$ where $d \in (0, \frac{1}{2})$ is a constant. Let the market profile be drawn from $\Pi_{\mathrm{MTB}}^{n,m}$ uniformly at random. Also, let $M$ denote the set of all assignments that are stable given the drawn market profile. For any $\mu \in M$, let $s(\mu)$ denote the sum of ranks of the assigned students in $\mu$. Then, for any constant $\epsilon > 0$, the event*

$$\max_{\mu \in M} \left| \frac{s(\mu)}{m^2 f(\log \frac{n}{n-m})} - 1 \right| < \epsilon$$

*holds with very high probability as $m$ grows large, where $f(x) = \frac{1}{x} - \frac{1}{e^x - 1}$.*

**Theorem B.3** (Theorem 1.6 in Pittel (2019))**.** *Consider a simple school choice market with $m$ schools and $n = m + \lambda(m)$ students where $\lambda : \mathbb{N} \to \mathbb{N}$ is such that $\lambda(m) \ge m^\alpha$ holds for every $m$ where $\alpha > 0$ is a constant. Let the market profile be drawn from $\Pi_{\mathrm{MTB}}^{n,m}$ uniformly at random. Also, let*

$$M_\pi = |\{s \in S : \mu_\pi(s) \ne \eta_\pi(s)\}|,$$

*where $\mu_\pi$ and $\eta_\pi$ denote the student- and school-optimal stable assignments, respectively. Then, for every constant $\epsilon > 0$ there exists $m_\epsilon > 0$ such that for all $m > m_\epsilon$,*

$$\mathbb{E}_\pi[M_\pi] \le m^{3/2-\alpha+\epsilon}.$$

**Lemma B.4.** *Consider a simple school choice market in which every school has an arbitrary tie-breaker and the preference order of every student is drawn independently and uniformly at random from the set of all strict linear orders over the schools. For a student $s$, let $d_s$ denote the number of proposals that $s$ receives when the school-proposing DA is run in this market. Also, let $r_s$ denote the rank of $s$ in the assignment constructed by the school-proposing DA. Then,*

$$\mathbb{P}\left[r_s > x | d_s = t\right] \leq e^{-\frac{tx}{m}}.$$

*Furthermore, for $m \geq 3$, any positive integer $t < m$, and any $\alpha \geq 3$,*

$$\mathbb{P}\left[r_s > \left\lfloor \frac{m}{\alpha t} \right\rfloor \Big| d_s = t\right] \geq \exp\left(-\frac{2m}{\alpha(m-t)}\right).$$

*Proof.* By the Principle of Deferred Decisions, we let students rank proposals upon receiving them. We assume that, upon receiving each proposal, the student assigns a distinct *rank* to the proposing school, which is an integer between 1 and $m$; this rank is chosen uniformly at random from the set of ranks that the student has not assigned so far.

The proof for the first inequality is as follows. The probability that $s$ ranks the first school that proposes to her lower (worse) than $x$ is $1 - \frac{x}{m}$. Conditional on that, the probability that the second school that proposes to $s$ is ranked lower than $x$ is at most $1 - \frac{x}{m}$. Conditional on the first and second schools being ranked lower than $x$ by $s$, the third school that proposes to $s$ is ranked lower than $x$ with probability at most $1 - \frac{x}{m}$, and so on. Therefore, conditional on $s$ receiving $t$ proposals, the probability that all of the proposing schools are ranked lower than $x$ is at most $(1 - \frac{x}{m})^t$, which is at most $e^{-\frac{tx}{m}}$. This proves the first inequality.

We next prove the second inequality in the lemma statement. The probability that the first school is ranked lower than $\lfloor \frac{m}{\alpha t} \rfloor$ is at least $1 - \frac{m/\alpha t}{m}$. More generally, the probability that $s$ ranks the $i$-th school who proposes to her lower than $\frac{m}{\alpha t}$ is at least $1 - \frac{m/\alpha t}{m-i}$. Thus, we have

$$\mathbb{P}_\varphi\left[r_s > \left\lfloor \frac{m}{\alpha t} \right\rfloor \Big| d_s = t\right] \geq \prod_{i=1}^{t} 1 - \frac{1}{\alpha t(1 - i/m)}$$

$$\geq \exp\left(-\sum_{i=1}^{t} \frac{2}{\alpha t(1 - i/m)}\right) \geq \exp\left(-\frac{2m}{\alpha(m-t)}\right)$$

where in the second inequality we used the fact that $1 - x \geq e^{-2x}$ for any $x \leq 1/2$. Note that this inequality is applicable, because $\alpha t(1 - i/m) \geq 2$ holds for all $m \geq 3$, $i \leq t$ and $\alpha \geq 3$. $\qquad\square$

**Lemma B.5.** *Consider a simple school choice market with $m$ schools and $n = m + \lambda(m)$ students where $\lambda : \mathbb{N} \to \mathbb{N}$ is an arbitrary function. Let the market profile be drawn from $\Pi_{\mathrm{MTB}}^{n,m}$ uniformly at random. Fix a student $s$ and an arbitrary constant $\epsilon > 0$. Then, in the school-proposing DA, the following event holds with high probability as $m$ grows large: the number of proposals received by $s$ is at most $(1 + \epsilon) \log m$.*

*Proof.* The stochastic process governing the school-proposing DA is complicated. The proof idea is simplifying that process by coupling it with a simpler process, which we denote by $\mathcal{B}$.[26] Process $\mathcal{B}$ is defined by a sequence of binary random variables $X_1, \ldots, X_k$, where $k = (1 + \delta)m \log m$ for an arbitrary constant $\delta \in (0, \epsilon)$.[27] The idea behind the choice of $k$ is that, whp, this is an upper bound on the total number of proposals in the school-proposing DA (due to Theorem B.1 and Fact A.2). Each random variable in this sequence takes the value 1 with probability $\frac{1}{n - 3\log^2 m}$, and 0 otherwise. For convenience, we also refer to these random variables as *coins*, and the process that determines the value of a random variable as a *coin flip*.

The coupling of the school-proposing DA and the coin-flipping process will be defined such that the number of successful coin flips would be an upper bound on the number of proposals that $s$ receives in almost all of the sample paths of the coupled process. This will conclude the proof, since computing an upper bound on the number of successful coin flips can be done by a Chernoff concentration bound.

The coupled process, namely $(\mathsf{DA}, \mathcal{B})$, has two components. The first component, $\mathsf{DA}$, corresponds to the stochastic process governing school-proposing DA. The second component, $\mathcal{B}$, corresponds to the coin-flipping process. In the coupled process, the school-proposing DA runs as usual, and whenever a school $c$ is about to propose, the following rules determine to whom the proposal is made.

R0. If there are no more coins left to flip in $\mathcal{B}$, then the $\mathsf{DA}$ process runs independently. Otherwise, if the the school-proposing DA stops (i.e., finds an assignment), then stop $\mathcal{B}$ as well.

R1. If $c$ has already proposed to $s$, let $c$ propose to a student drawn uniformly at random from the set of students to whom it has not yet proposed.

R2. If $c$ has not yet proposed to $s$, then $\mathcal{B}$ flips the next coin.

    (a) Let $d_c$ denote the number of proposals that $c$ has made so far. If $d_c > 3\log^2 m$, then let the school-proposing DA and $\mathcal{B}$ be run independently from this point forward. Otherwise,

        i. if the coin flip was successful: with probability $\frac{n - 3\log^2 m}{n - d_c}$ let $c$ propose to $s$, and otherwise (with probability $1 - \frac{n - 3\log^2 m}{n - d_c}$) let $c$ propose to a student drawn uniformly at random from the set of students that $c$ has not already proposed to minus $\{s\}$.

        ii. If the coin flip was not successful: $c$ proposes to a student that is drawn uniformly at random from the set of students that $c$ has not already proposed to minus $\{s\}$.

Observe that, every time a school $c$ is about to propose, if $c$ has not proposed to $s$ before and if $d_c \leq 3\log^2 m$, then $s$ will receive a proposal with probability $\frac{1}{n - 3\log^2 m} \cdot \frac{n - 3\log^2 m}{n - d_c} = \frac{1}{n - d_c}$, which coincides with the probability that she receives an offer from $c$ in the school-proposing DA. On

---

[26] We recall the definition of *coupling* in probability theory; see, e.g., Levin and Peres (2017).

[27] We defined $k = (1 + \delta)m \log m$, assuming that $k$ is an integer. To not to clutter the notation with floors and ceilings, we drop these operators. This does not change the proofs but simplifies notation significantly.

the other hand, if $d_c > 3\log^2 m$, the school-proposing DA and the process $\mathcal{B}$ continue to run independently. Hence, the distribution of the number of proposals received by $s$ in this coupling is identical to the distribution of the number of proposals received by $s$ in the school-proposing DA.

If rule R0 is executed with no more coins left to flip in $\mathcal{B}$, or if the condition $d_c > 3\log^2 m$ is met at some point in a sample path of the coupled process, we call that sample path an *ignored* sample path. Observe that in any sample path that is *not* ignored, student $s$ receives at most one offer per successful coin flip. Therefore, the number of successful coin flips in that sample path gives an upper bound on the number of proposals that $s$ receives in the same sample path.

Recall that $k$ is, whp, an upper bound on the total number of proposals in the school-proposing DA (due to Theorem B.1 and Fact A.2). In addition, Pittel (1992) shows that, whp, no school makes more than $3\log^2 m$ proposals in the school-proposing DA. That means the probability that a sample path is ignored approaches 0 as $m$ approaches infinity. To complete the proof, we need to prove the following claim.

**Claim B.6.** *For any constant $\epsilon > 0$, it holds whp that the number of successful coin flips in the process $\mathcal{B}$ is at most $(1 + \epsilon)\log m$.*

*Proof.* Choose $\delta$ to be a constant such that $\delta \in (0, \epsilon)$. The expected number of successful coin flips in $\mathcal{B}$ is at most $\frac{(1+\delta)m\log m}{n-3\log^2 m}$. A standard application of Chernoff concentration bounds (as stated in Appendix B.1) then shows that, for any constant $\delta' > 0$, the number of successful coin flips in $\mathcal{B}$ is smaller than $(1 + \delta')\frac{(1+\delta)m\log m}{n-3\log^2 m}$, whp. Choosing $\delta'$ such that $(1 + \delta)(1 + \delta') < 1 + \epsilon$ proves the claim. $\qquad\square$

Claim B.6 shows that the number of successful coin flips is at most $(1 + \epsilon)\log m$, whp. Recall that the number of successful coin flips bounds from above the number of proposals that $s$ receives, in sample paths that are not ignored. We showed the probability that a sample path is ignored approaches 0 as $m$ approaches infinity. Therefore, for any constant $\epsilon > 0$, the number of proposals that $s$ receives is at most $(1 + \epsilon)\log m$, whp.

$\qquad\square$

**Lemma B.7.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$, and consider a simple school choice market $\mathcal{M}$ with $n$ students and $m = n + \lambda(n)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\mathrm{MTB}}^{n,m}$. Then, for any constant $\epsilon > 0$, the following event holds with very high probability as $n$ grows large: the number of proposals received by a fixed student in the school-proposing DA is at least $(1 - \epsilon)nf(\log \frac{m}{m-n})/2$, where $f(x) = \frac{1}{x} - \frac{1}{e^x-1}$.*

*Proof.* Let $P$ denote the total number of proposals made in the school-proposing DA. First, we define a parameter $L$, which, wvhp, will be a lower bound on $P$. By Theorem B.2, for any constant $\epsilon' > 0$, wvhp it holds that $P \geq (1 - \epsilon')L^*$, where $L^* = n^2 f(\log \frac{m}{m-n})$. Throughout this proof, we define $\epsilon' = \epsilon/4$ and $L = (1 - \epsilon/2)L^*$. We emphasize that the idea behind the choice of $L$ is that, whp, $L$ is a lower bound on the total number of proposals in the school-proposing DA (due to Theorem B.2).

We prove the lemma for a fixed student $s \in S$. The stochastic process governing the school-proposing DA is a complicated process. The proof idea is simplifying that process by coupling it with a simpler process, which we denote by $\mathcal{B}$. Process $\mathcal{B}$ is defined by a sequence of binary random variables $X_1, \ldots, X_k$. Each random variable in this sequence takes the value 1 independently with probability $1/n$, and 0 otherwise. For convenience, we also refer to these random variables by *coins*. The process $\mathcal{B}$ is formally defined as follows.

**Definition B.8** (Definition of Process $\mathcal{B}$)**.**

1. *Let $k = L$.*

2. *Let $i = 1$.*

3. *While $i \leq k$ do*

    (a) *Let $X_i = 1$ independently with probability $1/n$, and let $X_i = 0$ otherwise. (I.e., flip a coin.)*

    (b) *If $X_i = 1$ then $k \leftarrow k - n$. (I.e., remove $n$ coins if the coin flip is successful.)*

    (c) *$i \leftarrow i + 1$.*

Note that the process stops when the condition $i \leq k$ is not met; then, we say *there are no coins left to flip*.

Next, we will define a coupling of the process $\mathcal{B}$ and the school-proposing DA. We denote this coupled process by $(\mathsf{DA}, \mathcal{B})$, where $\mathsf{DA}$ stands for the school-proposing DA. This coupling is defined such that, in almost all sample paths of the coupling (i.e., wvhp), the number of successful coin flips in $\mathcal{B}$ is a lower bound on the number of proposals that $s$ has received in the school-proposing DA, which we denote by $d_s$.

**Definition of the Coupling** $(\mathsf{DA}, \mathcal{B})$

Recall that we fixed a student $s$, and aim to provide a lower bound on the number of proposals made to $s$ in the school-proposing DA algorithm. We do this by defining a coupled process, $(\mathsf{DA}, \mathcal{B})$. The first component, $\mathsf{DA}$, corresponds to the stochastic process governing school-proposing DA, and the second component, $\mathcal{B}$, corresponds to the coin-flipping process. In the coupled process, the results of the coin flips in $\mathcal{B}$ would be used to decide whether each proposal in $\mathsf{DA}$ is made to $s$ or not.

In the coupled process $(\mathsf{DA}, \mathcal{B})$, a coin flip corresponds to a new proposal from a school to a student. If there are no coins left to flip in $\mathcal{B}$, then $(\mathsf{DA}, \mathcal{B})$ continues to run $\mathsf{DA}$ but stops running $\mathcal{B}$. If $\mathsf{DA}$ stops (by finding a stable assignment), the coupled process stops as well.

To complete the definition of the coupling, we need to define how $\mathsf{DA}$ and $\mathcal{B}$ interact. To this end, we first need a few definitions. While running the school-proposing DA, let $S_c$ denote the set of students to whom $c$ has proposed so far. In the coupling $(\mathsf{DA}, \mathcal{B})$, each school could have three possible states: *active*, *idle*, and *inactive*. At the beginning of the process, all schools are active.

As the school-proposing DA process evolves, a school might change its state from active to idle, from active to inactive, or from idle to inactive.

We are now ready to state the formal definition of the coupled process $(\mathsf{DA}, \mathcal{B})$. In the coupled process, the school-proposing DA runs as usual, and whenever a school $c$ is about to propose, the following rules determine to whom the proposal is made.

R0. If there are no more coins left to flip (i.e., the process $\mathcal{B}$ is stopped because the condition in line 3 of Definition B.8 is not met), then continue the process $\mathsf{DA}$ independently. Otherwise, if the school-proposing DA stops (i.e., finds an assignment), then stop $\mathcal{B}$ as well.

R1. If $c$ is active, then flip the next coin (this action corresponds to line 3-a of Definition B.8). If the coin flip is successful, then let $c$ propose to $s$ and then make $c$ inactive; moreover, dismiss $n$ of the unflipped coins (this action corresponds to line 3-b of Definition B.8). If the coin flip is not successful, with probability $1 - \frac{1 - 1/|S \backslash S_c|}{1 - 1/n}$, $c$ proposes to $s$ and becomes idle, and with probability $\frac{1 - 1/|S \backslash S_c|}{1 - 1/n}$, $c$ proposes to one of the students in $S \backslash (S_c \cup \{s\})$ uniformly at random (without changing its state).

R2. If $c$ is idle, then flip the next coin (this action corresponds to line 3-a of Definition B.8). If the coin flip is successful, make $c$ inactive and dismiss $n$ of the unflipped coins (this action corresponds to line 3-b of Definition B.8); otherwise, do not change the state of $c$. Let $c$ propose to one of the students in $S \backslash S_c$ uniformly at random.

R3. If $c$ is inactive, then do not flip any coins. Let $c$ propose to one of the students in $S \backslash S_c$ uniformly at random.

First, we verify that the probability that $s$ receives a proposal from $c$ in the coupled process is identical to the probability that she receives a proposal from $c$ when the school-proposing DA is run independently. To see this, observe that in the coupled process, every time that a school $c$ wants to make a proposal, if it has not proposed to $s$ before, then it will propose to $s$ with probability

$$\frac{1}{n} + (1 - \frac{1}{n}) \left( 1 - \frac{1 - 1/|S \backslash S_c|}{1 - 1/n} \right) = \frac{1}{|S \backslash S_c|},$$

which coincides with the probability that $s$ receives a proposal from $c$ if the school-proposing DA is run independently.

The intuition behind the three states (active, idle, inactive) is as follows. All schools are active in the beginning of the process. Every time an active school is about to make a proposal, it flips a coin. If the coin flip is successful, the school proposes to $s$ and becomes inactive. If the coin flip is not successful, then the school, namely $c$, proposes to $s$ with a positive probability of $1 - \frac{1 - 1/|S \backslash S_c|}{1 - 1/n}$; if the proposal is made to $s$, the school becomes idle, and otherwise the school remains active. Idle schools are thus schools who have proposed to $s$ without flipping a coin successfully. These schools continue to flip a coin every time they are about to make a proposal, and become inactive once

one of these coin flips is successful. Finally, we note that whenever a school becomes inactive, $n$ of the unflipped coins are dismissed (which, as we will see, account for the proposals that the inactive school may make from then on).

We will show that the coupling satisfies the following properties: (i) the sum of the flipped and dismissed coins equals $L$, whp (Claim B.9) (ii) every successful coin flip corresponds to a proposal to $s$ (Claim B.10), and (iii) the number of successful coin flips is "large" (Claim B.11). The rest of the proof formalizes this argument.

**Claim B.9.** *The sum of the number of flipped and dismissed coins in the process* $(\mathsf{DA}, \mathcal{B})$ *equals $L$, wvhp.*

*Proof.* Consider the time when the process $(\mathsf{DA}, \mathcal{B})$ stops. By rule R0, there are two possibilities: either there was a time in the coupled process when no coins were left to flip in $\mathcal{B}$, after which the process $\mathsf{DA}$ was run independently, or not (in which case the process $\mathsf{DA}$ stopped when there were still coins left to flip in $\mathcal{B}$). To prove the claim, we will show that the first possibility occurs wvhp.

To this end, consider the scenario where the second possibility occurs. By the definition of the coupling, a coin is flipped whenever an active or idle school makes a proposal. On the other hand, $n$ coins are dismissed whenever a school becomes inactive (see rules R1 and R2), and every inactive school makes at most $n$ proposals. Thus, when the second possibility occurs, the sum of the number of flipped and dismissed coins is larger than the total number of proposals made in $\mathsf{DA}$. Since there exist only $L$ coins, then the total number of proposals made in $\mathsf{DA}$ is smaller than $L$. However, recall that $L$ is, wvhp, a lower bound on the total number of proposals in $\mathsf{DA}$ (Theorem B.2). Therefore, wvhp, the second possibility does not occur. This implies that the first possibility occurs wvhp. □

**Claim B.10.** *Let the random variable $X^*$ denote the number of successful coin flips in $(\mathsf{DA}, \mathcal{B})$. Then, $d_s \geq X^*$.*

*Proof.* For the proof, we show that there is an injection from the set of successful coin flips in $(\mathsf{DA}, \mathcal{B})$ to the set of proposals made to $s$ in $(\mathsf{DA}, \mathcal{B})$. First, observe that in the definition of the coupling, only rules R1 and R2 allow for coin flips. In R1, immediately after a successful coin flip, a proposal is made to $s$. We associate the successful coin flip with the proceeding proposal in the injection that we construct. It remains to define the injection for the successful coin flips that are allowed by R2. For any such coin flip $f$, there exists a school $c$ satisfying the following condition: In the course of the coupled process, at a time before $f$ happens, R1 has allowed the school $c$ to make a proposal to $s$ even though the preceding coin flip was unsuccessful. In this case, associate the successful coin flip in R2 with the proposal that $c$ made to $s$ in R1. This completes the construction of the injection and completes the proof.

□

**Claim B.11.** *For any constant $\delta > 0$, $X^* \geq (1 - \delta)L/(2n)$ holds wvhp.*

*Proof.* It is useful to partition the $L$ coins that are used by the process $\mathcal{B}$ into two subsets, namely $A, B$, with subset $A$ containing $L(1+\delta)/2$ of the coins, and subset $B$ containing the $L(1-\delta)/2$ remaining coins. (For example, assume $A = \{x_1, \ldots, x_{L(1+\delta)/2}\}$ and $B = \{x_{1+L(1+\delta)/2}, x_L\}$.) Without loss of generality, suppose that whenever process $\mathcal{B}$ is about to flip a coin (line 3-a of Definition B.8), it flips a coin from $A$ if $A$ is non-empty; otherwise, it flips a coin from $B$. Any coin that is flipped is removed from the subset that contains it (i.e., the subset $A$ or $B$). On the other hand, if the process is about to remove $n$ coins (line 3-b of Definition B.8), it removes the coins from $B$ as long as $B$ is non-empty.

By Claim B.9, both $A, B$ will be empty by the end of the coupled process $(\mathsf{DA}, \mathcal{B})$, wvhp. We will show that, wvhp, $B$ empties before $A$ does. This would imply that at least $|B|/n = L(1-\delta)/(2n)$ of the coins in $A$ are flipped successfully, which would conclude the proof.

It remains to show that set $B$ empties before set $A$ does. This is a direct consequence of Chernoff concentration bounds. (Recall their definition from Appendix B.1.) These bounds imply that if $|A| = L(1+\delta)/2$ coins are flipped independently, each with a success probability $1/n$, then wvhp at least $L/(2n)$ of the coin flips are successful. Since for each successful coin flip from $A$ there are $n$ removed from $B$, and since $nL/(2n) > |B|$, then wvhp $B$ empties before $A$. $\qquad\square$

Claims B.10 and B.11 together imply that $d_s \geq (1 - \delta)L/(2n)$ holds wvhp. Recalling that $L = (1 - \epsilon/2)L^*$ and setting $\delta = \epsilon/2$ concludes the proof. $\qquad\square$

**Lemma B.12.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$, and consider a simple school choice market with $n$ students and $m = n + \lambda(n)$ schools. Suppose that the market profile is drawn from $\Pi_{\mathsf{MTB}}^{n,m}$ uniformly at random. Then, for any positive constant $\epsilon$ and any student $s \in S$, the following event holds with very high probability as $n$ grows large: the number of proposals received by $s$ in the school-proposing DA is at least $(1 - \epsilon)\frac{n}{2\log n}$.*

*Proof.* The proof is a consequence of Lemma B.7, which states that for any constant $\epsilon > 0$, the number of proposals received by a fixed student in the school-proposing DA is wvhp at least $(1 - \epsilon)nf(\log \frac{m}{m-n})/2$, where $f(x) = \frac{1}{x} - \frac{1}{e^x - 1}$.

**Claim B.13.** *The function $f(x) = \frac{1}{x} - \frac{1}{e^x - 1}$ is decreasing in $x$ for $x \geq 0$.*

*Proof.* Observe that $f'(x) = \frac{e^x}{(e^x-1)^2} - \frac{1}{x^2}$. Algebraic manipulation of $f'(x) \leq 0$ reveals that this inequality holds if and only if

$$x^2 + 2 \leq e^x + e^{-x}. \tag{8}$$

To see that the above inequality holds for every $x \geq 0$, define $h(x) = e^x + e^{-x} - (x^2 + 2)$. We then observe that $h''(x) = e^{-x} + e^x - 2 \geq 0$, $h'(0) = h''(0) = h'''(0) = 0$, and $h^4(0) = 2$, where $h^4(\cdot)$ denotes the fourth derivative of $h$. Thus, $h(x)$ is increasing at every $x \geq 0$. $\qquad\square$

Now, recall that $m = n + \lambda(n)$. Hence,

$$f(\log \frac{m}{m-n}) = f(\log(1 + \frac{n}{\lambda(n)})) \geq f(\log(1+n)) = \frac{1}{\log(n+1)} - \frac{1}{n}, \tag{9}$$

where the inequality follows from the fact that $f(x)$ is decreasing at all $x \geq 0$. By Lemma B.7 and by (9), to prove the claim of the lemma, it suffices to prove its claim for the case that $\lambda(n) = 1$ for all $n \in \mathbb{N}$. In this case, observe that by Lemma B.7, the number of proposals received by a fixed student in the school-proposing DA is wvhp at least $(1-\delta)\frac{n}{2}(\frac{1}{\log(n+1)} - \frac{1}{n})$, for any constant $\delta > 0$. Choosing $\delta = \epsilon/2$ and noting that $(1-\delta)\frac{n}{2}(\frac{1}{\log(n+1)} - \frac{1}{n}) > (1-\epsilon)\frac{n}{2\log n}$ holds for sufficiently large $n$ concludes the proof. □

**Lemma B.14.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$, and consider a simple school choice market $\mathcal{M}$ with $m$ schools and $n = m + \lambda(m)$ students. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\mathsf{MTB}}^{n,m}$. Then, for every constant $\kappa > 0$, there exists $m_\kappa$ such that for all $m > m_\kappa$, the number of proposals received by a fixed student in the school-proposing DA is less than $0.5(1-\kappa)\log \frac{n}{n-m}$ with probability at most $\frac{1}{m} + e^{-\frac{\kappa^2(1-\kappa/2)\log(\frac{n}{n-m})}{16(1+\kappa/2)}}$.*

*Proof.* Let $P$ denote the total number of proposals made in the school-proposing DA. By Theorem B.1, for any constant $\kappa > 0$, wvhp it holds that $P \geq (1 - \kappa/2)L^*$, where $L^* = n\log(\frac{n}{n-m})$. By the definition of the notion of wvhp, the latter fact also also implies the existence of a constant $m_\kappa > 0$ such that $P \geq (1-\kappa/2)L^*$ holds with probability at least $1 - \frac{1}{m}$. Define $L = (1-\kappa/2)L^*$.

We prove the lemma for a fixed student $s \in S$. The stochastic process governing the school-proposing DA, namely, the process $\mathsf{DA}$, is a complicated one. The proof idea is simplifying $\mathsf{DA}$ by coupling it with a simpler process, denoted by $\mathcal{B}$. The process $\mathcal{B}$ and the coupled process $(\mathsf{DA}, \mathcal{B})$ are defined identical to the proof of Lemma B.7 (with the difference that, here, the number of coins $L$ is given as above).

As in the proof of Lemma B.7, we let $d_s$ denote the number of proposals that student $s$ receives in $(\mathsf{DA}, \mathcal{B})$. Also, let $X^*$ denote the number of successful coin flips in $(\mathsf{DA}, \mathcal{B})$.

**Claim B.15.** $d_s \geq X^*$.

*Proof.* The proof is identical to the proof of Claim B.10. □

**Claim B.16.** *For any constant $\delta > 0$, $X^* < (1-\delta)\frac{L}{2n}$ holds with probability at most $\frac{1}{m} + e^{-\frac{\delta^2 L}{4(1+\delta)n}}$.*

*Proof.* Let $x_1, \ldots, x_L$ be iid Bernoulli random variables with mean $1/n$. We interpret these variables as coins: when $(\mathsf{DA}, \mathcal{B})$ is about to flip a coin, the variable with the smallest index whose value has not been *used* yet will be used to determine the result of the coin flip. On the other hand, when $(\mathsf{DA}, \mathcal{B})$ is about to dismiss a coin, the variable with the largest index that has not been dismissed yet will be dismissed. (The same structure is used in the proof of Claim B.11.)

By Theorem B.1, there exists $m_\kappa > 0$ such that for every $m > m_\kappa$, $P < L$ holds with probability at most $\frac{1}{m}$. That is, the probability that, by the end of $(\mathsf{DA}, \mathcal{B})$, there exists a variable $x_i$ that is not used or dismissed is at most $\frac{1}{m}$

Partition the variables to two subsets, $A = \{x_1, \ldots, x_{L(1+\delta)/2}\}$ and $B = \{x_{1+L(1+\delta)/2}, x_L\}$. We note that, by Chernoff concentration bounds (Appendix B.1), $n\sum_{i \in A} x_i \geq \frac{L}{2}$ with probability at least $1 - e^{-\frac{\delta^2 L}{4(1+\delta)n}}$.

Next, we note that if the event $n\sum_{i \in A} x_i \geq \frac{L}{2}$ holds, so does the event $n\sum_{i \in A} x_i > |B|$. If the latter event and the event $P \geq L$ hold, then the number of successful coin flips $X^*$ is at least $\frac{|B|}{n} = (1-\delta)\frac{L}{2n}$ (because for every successful coin flip in $A$, $n$ coins are dismissed from $B$). Therefore, a union bound implies that the probability that $X^* < (1 - \delta)\frac{L}{2n}$ holds is bounded by the sum of the probabilities of the events $P < L$ and $n\sum_{i \in A} x_i < \frac{L}{2}$, which is at most $\frac{1}{m} + e^{-\frac{\delta^2 L}{4(1+\delta)n}}$. $\qquad\square$

Claim B.15 shows that $d_s \geq X^*$, and Claim B.16 shows that $X^* \geq (1 - \delta)L/(2n)$ holds with probability at least $1 - \frac{1}{m} - e^{-\frac{\delta^2 L}{4(1+\delta)n}}$. Therefore, $d_s \geq (1 - \delta)L/(2n)$ holds with probability at least $1 - \frac{1}{m} - e^{-\frac{\delta^2 L}{4(1+\delta)n}}$. That is, $d_s < (1 - \delta)L/(2n)$ holds with probability $at\ most\ \frac{1}{m} + e^{-\frac{\delta^2 L}{4(1+\delta)n}}$. Recalling that $L = (1 - \kappa/2)L^*$ and $L^* = n\log(\frac{n}{n-m})$, and setting $\delta = \kappa/2$, implies that $d_s < \frac{1}{2}(1 - \kappa/2)^2 \log(\frac{n}{n-m})$ holds with probability at most

$$\frac{1}{m} + e^{-\frac{\kappa^2(1-\kappa/2)\log(\frac{n}{n-m})}{16(1+\kappa/2)}}.$$

Observing that $d_s < \frac{1}{2}(1 - \kappa)\log(\frac{n}{n-m})$ implies $d_s < \frac{1}{2}(1 - \kappa/2)^2 \log(\frac{n}{n-m})$ concludes the proof. $\quad\square$

**Lemma B.17.** *Consider a market $\mathcal{M}$ with $m$ schools and $n$ students. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\mathsf{STB}}^{n,m}$. Let $\epsilon > 0$ be an arbitrary constant. Then,*

*i. if $n \geq m$, with high probability as $m$ grows large,*

$$R_{\pi_{\mathsf{STB}}}(1) \geq \frac{(1 - \epsilon)m}{2};$$

*ii. if $n < m$, with high probability as $n$ grows large, $R_{\pi_{\mathsf{STB}}}(1) \geq \frac{(1-\epsilon)n}{2}$.*

*Proof.* We first prove part (i). The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) We fix the common tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random from the set of all linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Let $X_i$ be a binary random variable which is 1 iff the student with priority number $i$ is assigned to her first choice, and let $X = \sum_{i=1}^m X_i$. Observe that $\mathbb{P}_\varphi[X_i = 0] = (i - 1)/m$. Therefore, $\mathbb{E}_\varphi[X] = \sum_{i=1}^m \frac{m-i+1}{m} = \frac{m+1}{2}$. Also, observe that the random variables $X_1, \ldots, X_m$ are independent. A

standard application of Chernoff bounds (as stated in Section B.1) then implies that for any $\epsilon > 0$, we have

$$\mathbb{P}_\varphi \left[ X < (1 - \epsilon) \cdot \mathbb{E}_\varphi \left[ X \right] \right] \le \exp\left( -\frac{\epsilon^2 \mathbb{E}_\varphi \left[ X \right]}{2} \right).$$

This proves part (i).

The proof for part (ii) is identical to the above proof, but $m$ is replaced with $n$.

$\square$

# C  Proof of Theorem 3.1: rank-wise dominance

For expositional purposes we state and prove in this section only the parts of Theorem 3.1 that concern rank-wise dominance.

**Theorem C.1.** *Let* $\lambda : \mathbb{N} \to \mathbb{N}$.

(i) *Consider a simple school choice market with $m$ schools and $n = m + \lambda(m)$ students. Let the preference profile $\varphi$ be drawn independently and uniformly at random from $\Phi^{n,m}$, and let the tie-breaking profiles $\psi_{\mathsf{MTB}}, \psi_{\mathsf{STB}}$ be drawn independently and uniformly at random from $\Psi^{n,m}_{\mathsf{MTB}}, \Psi^m_{\mathsf{STB}}$, respectively. Also, define the market profiles $\pi_{\mathsf{MTB}} = (\varphi, \psi_{\mathsf{MTB}})$ and $\pi_{\mathsf{STB}} = (\varphi, \psi_{\mathsf{STB}})$. Then, for every constant $\epsilon > 0$*

$$\lim_{m \to \infty} \mathbb{P} \left[ R_{\pi_{\mathsf{MTB}}} \preceq^\epsilon R_{\pi_{\mathsf{STB}}} \right] = 1.$$

(ii) *Consider a simple school choice market with $n$ students and $m = n + \lambda(n)$ schools. Let the preference profile $\varphi$ be drawn independently and uniformly at random from $\Phi^{n,m}$, and the tie-breaking profiles $\psi_{\mathsf{MTB}}, \psi_{\mathsf{STB}}$ be drawn independently and uniformly at random from $\Psi^{n,m}_{\mathsf{MTB}}, \Psi^m_{\mathsf{STB}}$, respectively. Also, define the market profiles $\pi_{\mathsf{MTB}} = (\varphi, \psi_{\mathsf{MTB}})$ and $\pi_{\mathsf{STB}} = (\varphi, \psi_{\mathsf{STB}})$. Then, for every constant $\epsilon > 0$,*

$$\lim_{n \to \infty} \mathbb{P} \left[ R_{\pi_{\mathsf{MTB}}} \preceq^\epsilon R_{\pi_{\mathsf{STB}}} \right] = 0, \ \lim_{n \to \infty} \mathbb{P} \left[ R_{\pi_{\mathsf{STB}}} \preceq^\epsilon R_{\pi_{\mathsf{MTB}}} \right] = 0$$

*hold when there exists a constant $\gamma < 1$ such that $\lambda(n) \le n^\gamma$ for all $n$.*

## C.1  Proof of Theorem C.1, part (i)

The following definitions and lemmas will be used in the proof. Denote by $\mathcal{M}_{\mathsf{MTB}}$ the simple school choice market defined in the theorem statement which has $m$ schools, $n = m + \lambda(m)$ students, and market profile $\pi_{\mathsf{MTB}}$. Similarly, let $\mathcal{M}_{\mathsf{STB}}$ denote the simple school choice market defined in the theorem statement which has $m$ schools, $n = m + \lambda(m)$ students, and market profile $\pi_{\mathsf{STB}}$. We recall the definitions of $\pi_{\mathsf{MTB}}, \pi_{\mathsf{STB}}$ from the theorem statement.

### C.1.1 Lemmas related to $\mathcal{M}_{\mathsf{MTB}}$

**Lemma C.2.** *In the market $\mathcal{M}_{\mathsf{MTB}}$, for any $t > 0$, the following holds with high probability as $m$ grows large: at most $\frac{2m \log m}{t}$ students receive at least $t$ proposals in the school-proposing DA.*

*Proof.* By Theorem 1 in Ashlagi et al. (2017), the total number of proposals in the school-proposing DA is at most $2m \log m$, whp. This implies that, whp, there are at most $\frac{2m \log m}{t}$ students who receive at least $t$ proposals in the school-proposing DA. $\square$

**Lemma C.3.** *Let $\alpha > 4$ be a constant and define $\bar{t} = 200 \log m$. Then, with high probability as $m$ grows large,*

$$R^+_{\pi_{\mathsf{MTB}}} \left( \lfloor \frac{m}{\alpha \bar{t}} \rfloor \right) \leq 0.4m + m/\sqrt{\log m}.$$

*Proof.* To prove the claim, first we run the school-proposing DA and prove the lemma statement for the school-optimal assignment. Then, using the fact that almost every student has the same match in the student-optimal assignment as shown in Ashlagi et al. (2017), we establish the lemma statement (which holds for the student-optimal assignment).

For any student $s$, let the random variable $d_s$ denote the number of proposals that $s$ receives in the school-optimal assignment. Also, define the random vector $\mathbf{d} = (d_1, \ldots, d_n)$. For any student $s$, let $r_s$ denote the rank of $s$ in the school-optimal assignment. Also, let $x_s$ be a binary random variable that is equal to 1 iff $r_s > \frac{m}{\alpha \bar{t}}$.

**Claim C.4.** $\lim_{m \to \infty} \mathbb{P}_{\pi_{\mathsf{MTB}}} \left[ \sum_{s \in S} x_s \geq 0.6m \right] = 1$.

*Proof.* Let $S'$ denote the subset of students who received at least one but no more than $\bar{t}$ offers. Lemma B.4 implies that

$$\mathbb{P}_{\pi_{\mathsf{MTB}}} \left[ r_s > \left\lfloor \frac{m}{\alpha \bar{t}} \right\rfloor \middle| s \in S' \right] \geq e^{-1/2}$$

holds for sufficiently large $m$, since $\alpha > 4$. This means that $\mathbb{E}_{\pi_{\mathsf{MTB}}} \left[ \sum_{s \in S'} x_s \right] \geq e^{-1/2} |S'|$ holds for sufficiently large $m$. Let $M = e^{-1/2} |S'|$. Observe that, conditional on $\mathbf{d}$, the random variables $\{r_s\}_{s \in S}$ are independent. Therefore, conditional on $\mathbf{d}$, the random variables $\{x_s\}_{s \in S}$ are also independent. Hence, Chernoff concentration bounds imply that, for any positive constant $\delta < 1$, for any $\mathbf{d}$, and for sufficiently large $m$,

$$\mathbb{P}_{\pi_{\mathsf{MTB}}} \left[ \sum_{s \in S'} x_s \leq (1 - \delta) e^{-1/2} |S'| \middle| \mathbf{d} \right] \leq e^{-\frac{\delta^2 M}{2}}. \tag{10}$$

On the other hand, Lemma C.2 implies that $|S'| \geq 0.99m$ holds whp. Therefore, (10) implies that, for any constant $\delta < 1$,

$$\mathbb{P}_{\pi_{\mathsf{MTB}}} \left[ \sum_{s \in S'} x_s \leq (1 - \delta) e^{-1/2} |S'| \right]$$

approaches 0 as $m$ approaches infinity. The facts that $S' \subseteq S$ and $0.99 e^{-1/2} > 0.6$ conclude the proof. $\square$

Recall that Claim C.4 was proved for the school-proposing DA. To complete the proof of the lemma, we will show that a slightly weaker version of Claim C.4 holds for the student-proposing DA as well. Let $\mu, \eta$ respectively denote the assignments generated by the student- and school-proposing DA in the market $\mathcal{M}_{\mathsf{MTB}}$. The number of students who have different matches under $\mu, \eta$ is at most $m/\sqrt{\log m}$, whp, as shown by Ashlagi et al. (2017). Therefore, Claim C.4 implies that, whp, there are at most $0.4m + m/\sqrt{\log m}$ students $s$ for whom $\mu^{\#}(s) \leq \frac{m}{\alpha \bar{t}}$. $\qquad \square$

### C.1.2  Lemmas related to $\mathcal{M}_{\mathsf{STB}}$

**Lemma C.5.** *In the market $\mathcal{M}_{\mathsf{STB}}$, suppose that student $s \in S$ is the $(m-x)$-th highest priority student in the common tie-breaker, with $x \geq 0$. (That is, in the common tie-breaker that is used by all schools, the student has strictly higher priority than $x + n - m$ other students.) Then, the probability that in the student-optimal assignment $s$ is not assigned to one of her top $i$ choices is at most $(1 - \frac{x}{m})^i$ for any positive integer $i > 0$.*

*Proof.* This can be seen by the equivalence between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) The probability that $s$ is not assigned to her top choice is $1 - \frac{x+1}{m}$, which is at most $1 - \frac{x}{m}$. The probability that $s$ is not assigned to her second top choice is $(1 - \frac{x+1}{m})(1 - \frac{x}{m-1})$, which is at most $(1 - \frac{x}{m})^2$. Similarly, the probability that $s$ is not assigned to her $i$-th top choice is at most $(1 - \frac{x}{m})^i$, for any positive integer $i > 0$. $\qquad \square$

**Lemma C.6.** *In the market $\mathcal{M}_{\mathsf{STB}}$, suppose that the student $s \in S$ is the $(m-x)$-th highest priority student in the common tie-breaker, with $x \geq 0$. Then, in the student-optimal assignment, $s$ is assigned to one of her top $\frac{2m \log(m)}{x}$ choices with probability at least $1 - 1/m^2$.*

*Proof.* Set $i = \frac{2m}{x} \log(m)$ and apply Lemma C.5. Noting that $(1 - \frac{x}{m})^i \leq e^{-\frac{xi}{m}}$ proves the claim. $\qquad \square$

For the next lemma, we recall that, for a market profile $\pi$, we denote the rank distribution $R_{\mu_\pi}$ by $R_\pi$, for notational simplicity.

**Lemma C.7.** *Let $\bar{t} = 3 \log m$. For any constant $\alpha > 1$, there exists a constant $\beta_\alpha > 0$ (independent of $m$) such that the following event holds with high probability as $m$ grows large:*

$$R^{+}_{\pi_{\mathsf{STB}}} \left( \lfloor \frac{m}{\alpha \bar{t}} \rfloor \right) \geq m - \beta_\alpha \log m \cdot \log \log m.$$

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) We fix the common tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random from the set of all linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Define $x = 2\alpha \bar{t} \log m$. Let $S'$ be the subset of students in $\mathcal{M}_{\mathsf{STB}}$ who receive priority numbers better (smaller) than $m - x$. First, we apply Lemma C.6 to each student in $S'$. Lemma C.6 implies that a student with priority number $m - x$ or better gets assigned to one of her top $\frac{m}{\alpha \bar{t}}$ choices

38

with probability at least $1 - m^{-2}$. Taking a union bound over all students with priority number no worse than $m - x$ implies that at least $m - x$ students are assigned to one of their top $\frac{m}{\alpha t}$ choices, with probability at least $1 - 1/m$. This means $R^+_{\pi_{\mathsf{STB}}}\left(\frac{m}{\alpha \bar{t}}\right) \geq m - x$ holds with probability at least $1 - 1/m$. To prove the sharper bound in the lemma statement, we need to take the students in $S \backslash S'$ into account.

Let $\mu$ denote the student-optimal assignment in $\mathcal{M}_{\mathsf{STB}}$. Also, let $S'' \subset S \backslash S'$ denote the subset of students who have a priority number between $m - x$ and $m - \beta \bar{t} \cdot \log \log m$, where $\beta = 2\alpha^2 \bar{t} / \log m$. Lemma C.5 implies that for any $s \in S''$,

$$
\mathbb{P}_\varphi \left[ \mu^{\#}(s) > \frac{m}{\alpha \bar{t}} \right] \leq \left( 1 - \frac{\beta \bar{t} \cdot \log \log m}{m} \right)^{\frac{m}{\alpha \bar{t}}} \leq \exp\left( -\frac{\beta}{\alpha} \cdot \log \log m \right).
$$

Having $\beta = 2\alpha^2 \bar{t} / \log m$ implies

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

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

$$
\mathbb{P}_\varphi \left[ \max_{s \in S''} \mu^{\#}(s) > \frac{m}{\alpha \bar{t}} \right] \leq |S''| \cdot (\log m)^{-\frac{2\alpha \bar{t}}{\log m}}
$$

$$
\leq x (\log m)^{-6} \leq \frac{6\alpha}{(\log n)^4}.
$$

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

$$
\mathbb{P}_\varphi \left[ \max_{s \in S' \cup S''} \mu^{\#}(s) > \frac{m}{\alpha \bar{t}} \right] \leq 1/m + \frac{6\alpha}{(\log n)^4}.
$$

Consequently, $R^+_{\pi_{\mathsf{STB}}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \geq |S' \cup S''|$ holds, whp. To conclude the proof, we note that $|S' \cup S''| = \frac{2\alpha^2 \bar{t}^2 \log \log m}{\log m} = 18\alpha^2 \log m \cdot \log \log m$. Setting $\beta_\alpha = 18\alpha^2$ completes the proof. $\qquad \square$

**Proof of Theorem C.1, part (i).** Fix the constant $\epsilon > 0$. By Lemma C.3,

$$
R^+_{\pi_{\mathsf{MTB}}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \leq 0.4m + m/\sqrt{\log m} \tag{11}
$$

holds with high probability as $m$ grows large. On the other hand, by Lemma C.7, whp it holds that

$$
R^+_{\pi_{\mathsf{STB}}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \geq m - (\log m)^{1+\epsilon}. \tag{12}
$$

Also, by Lemma B.17, whp it holds that

$$
R_{\pi_{\mathsf{STB}}}(1) \geq (1 - \epsilon)m/2. \tag{13}
$$

A union bound then implies that, whp, (11), (12), and (13) hold simultaneously. The latter fact, together with the fact that there exists $m_\epsilon > 0$ such that $(1 - \epsilon)m/2 > 0.4m + m/\sqrt{\log m}$ holds for all $m > m_\epsilon$, implies that $R_{\pi_{\text{MTB}}} \preceq^\epsilon R_{\pi_{\text{STB}}}$ holds whp. $\qquad\square$

## C.2 Proof of Theorem C.1, part (ii)

Let $\mathcal{M}_{\text{MTB}}$ denote the simple school choice market defined in the theorem statement which has $n$ students, $m = n + \lambda(n)$ schools, and market profile $\pi_{\text{MTB}}$. Similarly, we use $\mathcal{M}_{\text{STB}}$ to denote the same market but with market profile $\pi_{\text{STB}}$, as defined in the theorem statement.

**Lemma C.8.** *In the market $\mathcal{M}_{\text{STB}}$, suppose that the student $s \in S$ is the $(n - t + 1)$-th highest priority student in the common tie-breaker of the schools, with $t \geq 0$. Then, $s$ is assigned to one of her top $x$ choices with probability at most $\frac{x\Delta}{m-x+1}$, where $\Delta = t + m - n$.*

*Proof.* By the equivalence between the DA-STB and the RSD mechanism (which we recall from Section A.1), the probability that $s$ is not assigned to her top choice is $1 - \frac{\Delta}{m}$. The probability that $s$ is not assigned to her top two choices is $(1 - \frac{\Delta}{m})(1 - \frac{\Delta}{m-1})$. Similarly, the probability that $s$ is not assigned to her top $i$ choices is $\Pi_{j=1}^{i}(1 - \frac{\Delta}{m-j+1})$. Observing that $\Pi_{j=1}^{x}(1 - \frac{\Delta}{m-j+1}) \geq 1 - \frac{\Delta x}{m-x+1}$ completes the proof. $\qquad\square$

**Lemma C.9.** *For any $\epsilon > 0$, in the student-optimal assignment in the market $\mathcal{M}_{\text{STB}}$, the following event holds with high probability as $n$ grows large: at least $\frac{n(1-\epsilon)}{25 \log^2 n}$ students are not assigned to one of their top $3 \log^2 n$ choices.*

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) We fix the common tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random from the set of all linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Let $s$ be a student with priority number $n - t + 1$, and let $X_s$ be a binary random variable which is 1 iff student $s$ is not assigned to one of her top $x$ choices. Let $x = 3 \log^2 n$ and $t = \frac{n}{5 \log^2 n}$. Applying Lemma C.8 implies that any student with a priority number below $n - t$ is assigned to one of her top $x$ choices with probability at most

$$\frac{x(t + m - n)}{m - x + 1} = \frac{3n/5 + 3(m - n) \log^2 n}{m - 3 \log^2 n + 1},$$

which is at most $4/5$ for sufficiently large $n$. Hence, $\mathbb{P}_\varphi [X_s = 1] \geq 1/5$ holds for sufficiently large $n$. Now, let $S_t$ denote the set of students with the $t$ lowest priority numbers. Since the students' preference lists are drawn independently, Chernoff bounds apply, and imply that for any constant $\epsilon > 0$, whp it holds that $\sum_{s \in S_t} X_s \geq (1 - \epsilon)|S_t|/5$. Observing that the right-hand side of the latter inequality is equal to $\frac{n(1-\epsilon)}{25 \log^2 n}$ concludes the proof. $\qquad\square$

*Proof of Theorem C.1, part (ii).* Fix $\epsilon > 0$. First, we show that

$$\lim_{n \to \infty} \mathbb{P}\left[R_{\pi_{\mathsf{MTB}}} \preceq^\epsilon R_{\pi_{\mathsf{STB}}}\right] = 0.$$

To this end, we recall a result from Pittel (1992) that shows, with high probability as $n$ grows large, no student has a rank worse than $3\log^2 n$ in the student-optimal assignment in the market $\mathcal{M}_{\mathsf{MTB}}$. On the other hand, Lemma C.9 shows that at least $\frac{n(1-\epsilon)}{25\log^2 n}$ students are not assigned to one of their top $3\log^2 n$ choices, whp. Therefore, whp it holds that $R_{\pi_{\mathsf{STB}}}$ does not $\epsilon$-rank-wise dominate $R_{\pi_{\mathsf{MTB}}}$. This implies that the above equation holds.

Next, we show that $\lim_{n \to \infty} \mathbb{P}\left[R_{\pi_{\mathsf{STB}}} \preceq^\epsilon R_{\pi_{\mathsf{MTB}}}\right] = 0$. Lemma B.17 implies that, whp,

$$R_{\pi_{\mathsf{STB}}}(1) \geq (1-\epsilon)n/2. \tag{14}$$

On the other hand, Proposition 3.1 from Ashlagi et al. (2019) shows that

$$\lim_{n \to \infty} \mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R^+_{\pi_{\mathsf{MTB}}}(1)\right]/n = 0.$$

By the Markov inequality, this also implies that $\lim_{n \to \infty} \mathbb{P}_{\pi_{\mathsf{MTB}}}\left[R^+_{\pi_{\mathsf{MTB}}}(1) > n/3\right] = 0$. This, together with (14), implies that $\lim_{n \to \infty} \mathbb{P}\left[R_{\pi_{\mathsf{STB}}} \preceq^\epsilon R_{\pi_{\mathsf{MTB}}}\right] = 0$.

$\square$

# D  Proof of Theorem 3.1: social inequity

For expositional simplicity, here we state and prove the parts of Theorem 3.1 that are about social inequity.

**Theorem D.1.** *Let* $\lambda : \mathbb{N} \to \mathbb{N}$.

(i) *Suppose that there exists a constant* $\gamma < \frac{3}{2}$ *such that* $\lambda(n) \leq n^\gamma$ *for all* $n$. *Consider a simple school choice market with* $m$ *schools and* $n = m + \lambda(m)$ *students. Let the market profiles* $\pi_{\mathsf{MTB}}$ *and* $\pi_{\mathsf{STB}}$ *be drawn independently and uniformly at random from* $\Pi^{n,m}_{\mathsf{MTB}}$ *and* $\Pi^{n,m}_{\mathsf{STB}}$, *respectively. Then,*

$$\lim_{m \to \infty} \frac{\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[\mathcal{S}i(\mu_{\pi_{\mathsf{MTB}}})\right]}{\mathbb{E}_{\pi_{\mathsf{STB}}}\left[\mathcal{S}i(\mu_{\pi_{\mathsf{STB}}})\right]} = \infty.$$

(ii) *Suppose that there exists a constant* $\gamma < 1$ *such that* $\lambda(n) \leq n^\gamma$ *for all* $n$. *Consider a simple school choice market with* $n$ *students and* $m = n + \lambda(n)$ *schools. Let the market profiles* $\pi_{\mathsf{MTB}}$ *and* $\pi_{\mathsf{STB}}$ *be drawn independently and uniformly at random from* $\Pi^{n,m}_{\mathsf{MTB}}$ *and* $\Pi^{n,m}_{\mathsf{STB}}$, *respectively. Then,*

$$\lim_{n \to \infty} \frac{\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[\mathcal{S}i(\mu_{\pi_{\mathsf{MTB}}})\right]}{\mathbb{E}_{\pi_{\mathsf{STB}}}\left[\mathcal{S}i(\mu_{\pi_{\mathsf{STB}}})\right]} = 0.$$

In under-demanded markets, we also consider a larger excess number of seats, as follows.

**Theorem D.2.** *Consider a simple school choice market $\mathcal{M}$ with $n$ students and $m = (1 + \lambda)n$ schools. Let the market profiles $\pi_{\mathsf{MTB}}$ and $\pi_{\mathsf{STB}}$ be drawn independently and uniformly at random from $\Pi_{\mathsf{MTB}}^{n,m}$ and $\Pi_{\mathsf{STB}}^{n,m}$, respectively. Then,*

$$\lim_{n \to \infty} \frac{\mathbb{E}_{\pi_{\mathsf{STB}}}\left[Si(\mu_\pi)\right]}{\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[Si(\mu_\pi)\right]} > 1$$

*holds when $\lambda$ is a positive constant no larger than $0.01$.*

## D.1 Preliminaries

**Fact D.3.** *Let $k$ and $d \leq k$ be positive integers. Define the random variable $X = \min\{X_1, \ldots, X_d\}$, where $X_1, \ldots, X_d$ respectively represent the first $d$ elements in a permutation of $1, \ldots, k$ which is drawn uniformly at random. Then,*

$$\mathsf{Var}[X] = \frac{d(k+1)(k-d)}{(d+1)^2(d+2)},$$

$$\mathbb{E}\left[X^2\right] = \frac{d(k+1)(k-d)}{(d+1)^2(d+2)} + \frac{(k+1)^2}{(d+1)^2} \geq \frac{2k(k-d)}{(d+2)^2}.$$

*Furthermore, the functions $f(d) = \frac{d(k+1)(k-d)}{(d+1)^2(d+2)}$ and $g(d) = \frac{d(k+1)(k-d)}{(d+1)^2(d+2)} + \frac{(k+1)^2}{(d+1)^2}$ are decreasing at all $d \in (0, k)$.*

*Proof.* The proof for variance is given, e.g., in Arnold et al. (1992), page 55. There, it is also shown that $\mathbb{E}[X] = \frac{k+1}{d+1}$. Plugging this equality and the expression for variance into $\mathbb{E}\left[X^2\right] = \mathsf{Var}[X] + \mathbb{E}[X]^2$ proves the claim for $\mathbb{E}\left[X^2\right]$.

Observe that

$$\frac{d(k+1)(k-d)}{(d+1)^2(d+2)} + \frac{(k+1)^2}{(d+1)^2} = \frac{(k+1)(2k+2-d)}{(d+1)(d+2)} \geq \frac{2k(k-d)}{(d+2)^2}.$$

To prove the last part of the lemma, we compute

$$f'(d) = \frac{d^3 - d^2(2k+1) - 2d(k+2) + 2k}{(d+1)^3(d+2)^2},$$

$$g'(d) = \frac{(k+1)\left(d^2 - 4dk - 4d - 6k - 8\right)}{(d+1)^2(d+2)^2}.$$

Since $2d(k+2) > 2k$ and $d^2(2k+1) > d^3$, then $f'(d) < 0$. Since $d^2 - 4dk - 4d - 6k - 8 < 0$ holds for all $d \in (0, k]$, then $g'(d) < 0$. □

**Lemma D.4.** *Let $\pi = (\varphi, \psi)$ be a market profile such that the preference profile $\varphi$ is drawn independently and uniformly at random from $\Phi^{n,m}$ and the tie-breaking profile $\psi$ is drawn independently*

*and uniformly at random from* $\Psi_{\mathsf{MTB}}^{n,m}$. *Then,*

$$\mathsf{Var}_\pi[\mu_\pi^\#(s)\big|\mu_\pi(s) \neq \varnothing] = \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right],$$
$$\mathsf{Var}_\pi[\eta_\pi^\#(s)\big|\eta_\pi(s) \neq \varnothing] = \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right].$$

*The same holds if* $\psi$ *is drawn independently and uniformly at random from* $\Psi_{\mathsf{STB}}^{n,m}$, *instead of* $\Psi_{\mathsf{MTB}}^{n,m}$.

*Proof.* We prove the first equality by observing that

$$
\begin{aligned}
\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &= \mathbb{E}_\pi\left[\frac{1}{|\mu_\pi(C)|} \cdot \sum_{t \in \mu_\pi(C)} (\mathcal{A}r(\mu_\pi) - \mu_\pi^\#(t))^2\right] \\
&= \mathbb{E}_\pi\left[\frac{1}{|\mu_\pi(C)|} \cdot \sum_{t \in \mu_\pi(C)} \mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t)\right] \\
&= \frac{1}{n} \cdot \sum_{t \in S} \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t)\Big|t \in \mu_\pi(C)\right] \quad (15) \\
&= \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(s)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(s)\Big|s \in \mu_\pi(C)\right] \quad (16) \\
&= \mathsf{Var}_\pi[\mu_\pi^\#(s)\big|s \in \mu_\pi(C)] \\
&= \mathsf{Var}_\pi[\mu_\pi^\#(s)\big|\mu_\pi(s) \neq \varnothing]
\end{aligned}
$$

where (16) holds because the term inside the summation in (15) is equal for all students, by symmetry. This proves the first equality in the lemma statement.

The second equality in the lemma statement holds by an identical argument as above but with $\mu$ replaced with $\eta$. $\qquad\square$

## D.2 Proof of Theorem D.1, part (i)

**Lemma D.5.** *Let* $\lambda : \mathbb{N} \to \mathbb{N}$ *be such that* $\lambda(i) \leq i^{3/2-\delta}$ *for a positive constant* $\delta$ *and every* $i$. *Consider a random market,* $\mathcal{M}$, *with* $m$ *schools and* $n = m+\lambda(m)$ *students. Suppose that the market profile for* $\mathcal{M}$, *namely* $\pi$, *is drawn uniformly at random from* $\Pi_{\mathsf{MTB}}^{n,m}$. *Then,* $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] = \Omega(\frac{m^2}{\log^2 m})$.

*Proof.* The proof has two steps. In Step 1, we show that $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] = \Omega(\frac{m^2}{\log^2 m})$. In Step 2, we show that $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]$ is close to $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right]$. In Step 3 we conclude the proof.

**Step 1.** Recall that $\eta_\pi$ is the outcome of the school-proposing DA. Let $r_\pi(s)$ denote the rank of a student $s$ in the assignment $\eta_\pi$. Fix a student $s \in S$. Since $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] = \mathsf{Var}_\pi[r_\pi(s)|r_\pi(s) \neq \infty]$ holds by Lemma D.4, we will provide a lower bound on the right-hand side. Let $d_s$ denote the number of proposals that student $s$ receives in the school-proposing DA. By the law of total variance

(Weiss, 2006),

$$\text{Var}_\pi[r_\pi(s) \mid r_\pi(s) \neq \infty]$$
$$=\mathbb{E}_\pi\left[\text{Var}_\pi[r_\pi(s) \mid (d_s = d) \wedge (r_\pi(s) \neq \infty)]\right] + \text{Var}_\pi[\mathbb{E}_\pi\left[r_\pi(s) \mid (d_s = d) \wedge (r_\pi(s) \neq \infty)\right]].$$

This implies that

$$\text{Var}_\pi[r_\pi(s) \mid r_\pi(s) \neq \infty] \geq \mathbb{E}_\pi\left[\text{Var}_\pi[r_\pi(s) \mid (d_s = d) \wedge (r_\pi(s) \neq \infty)]\right].$$

Since the event $r_\pi(s) \neq \infty$ is the same as the event $d_s > 0$, then we can rewrite the above inequality as

$$\text{Var}_\pi[r_\pi(s) \mid r_\pi(s) \neq \infty] \geq \mathbb{E}_\pi\left[\text{Var}_\pi[r_\pi(s) \mid (d_s = d) \wedge (d > 0)]\right],$$

which is equivalent to

$$\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] \geq \mathbb{E}_\pi\left[\text{Var}_\pi[r_\pi(s) \mid (d_s = d) \wedge (d > 0)]\right]. \tag{17}$$

In the remainder of the proof, we fix $\epsilon$ to be an arbitrary constant in the interval $(0, 0.1)$.

**Claim D.6.** *Let $\bar{d} = 2(1+\epsilon)\frac{n}{m}\log\frac{n}{n-m}$. There exists $m_\epsilon > 0$ such that for all $m > m_\epsilon$,*

$$\mathbb{P}_\pi\left[d_s > \bar{d} \mid d_s > 0\right] < \frac{1}{2}.$$

*Proof.* Let $P_\pi$ denote the total number of proposals made under the school-proposing DA. Hence,

$$\mathbb{E}_\pi\left[d_s \mid d_s > 0\right] = \mathbb{E}_\pi\left[P_\pi/m\right].$$

Let $K$ denote the right-hand side of the above equation. The Markov inequality then implies that

$$\mathbb{P}_\pi\left[d_s > 2K \mid d_s > 0\right] < \frac{1}{2}. \tag{18}$$

By Theorem B.1, the event

$$(1 - \epsilon/2)n\log\frac{n}{n-m} \leq P_\pi \leq (1 + \epsilon/2)n\log\frac{n}{n-m}$$

holds with very high probability as $m$ grows large.[28] Therefore, there exists a constant $m_\epsilon$ such that for all $m > m_\epsilon$,

$$\mathbb{E}_\pi\left[P_\pi\right] \leq (1 + \epsilon)n\log\frac{n}{n-m},$$

---

[28]While we are concerned with the number of proposals in the school-proposing DA here, Theorem B.1—which provides bounds for the number of proposals in the student-proposing DA—is applicable by symmetry.

which means that

$$K \leq (1 + \epsilon) \frac{n}{m} \log \frac{n}{n - m}.$$

The above inequality, together with (18), concludes the proof.

□

By Claim D.6 and Fact D.3, (17) implies that for every constant $\epsilon \in (0, 0.1)$,

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] \geq \frac{1}{2} \cdot \frac{\bar{d}(m + 1)(m - \bar{d})}{(\bar{d} + 1)^2 (\bar{d} + 2)} \tag{19}$$

holds for all $m > m_\epsilon$, where recall that $m_\epsilon$ is the constant given in Claim D.6 and $\bar{d} = 2(1 + \epsilon) \frac{n}{m} \log \frac{n}{n - m}$. This completes Step 1.[29]

**Step 2.** In this step we show that a lower bound similar to (19) holds for the student-optimal assignment. (Recall that (19) holds for the school-optimal assignment.) To this end, we will show that the expected social inequities in the school-optimal and the student-optimal assignments are "close", as follows.

Let $q_\pi(s)$ denote the rank of student $s$ in the student-optimal assignment. By Lemma D.4, and by the fact that the set of assigned students are the same in the student- and school-optimal stable assignments, we have

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] = \mathrm{Var}_\pi[q_\pi(s) \mid q_\pi(s) \neq \infty] - \mathrm{Var}_\pi[r_\pi(s) \mid r_\pi(s) \neq \infty]$$

$$= \frac{1}{m} \mathbb{E}_\pi \left[ \sum_{t \in \mu_\pi(C)} q_\pi(t)^2 - r_\pi(t)^2 \right] - \left( \mathbb{E}_\pi \left[ q_\pi(s) \mid q_\pi(s) \neq \infty \right]^2 - \mathbb{E}_\pi \left[ r_\pi(s) \mid r_\pi(s) \neq \infty \right]^2 \right)$$

$$= \frac{1}{m} \mathbb{E}_\pi \left[ \sum_{t \in \mu_\pi(C)} q_\pi(t)^2 - r_\pi(t)^2 \right] - \left( \mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 - \mathbb{E}_\pi \left[ \mathcal{A}r(\eta_\pi) \right]^2 \right). \tag{20}$$

Let $\Upsilon_m = \frac{1}{m} \mathbb{E}_\pi \left[ \sum_{t \in \mu_\pi(C)} q_\pi(t)^2 - r_\pi(t)^2 \right]$ and $\Xi_m = - \left( \mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 - \mathbb{E}_\pi \left[ \mathcal{A}r(\eta_\pi) \right]^2 \right)$. To complete Step 2, we provide lower bounds for $\Upsilon_m$ and $\Xi_m$.

**A lower bound for $\Xi_m$.** We note that, for every $\pi$ and every student $s$, $q_\pi(s) \leq r_\pi(s)$, because the rank of a student $s$ in the student-optimal matching is not worse than her rank in the school-optimal matching. Thus, $\mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right] \leq \mathbb{E}_\pi \left[ \mathcal{A}r(\eta_\pi) \right]$. This implies that

$$\Xi_m \geq 0. \tag{21}$$

---

[29]Later in the proof, Claim D.9 formally shows that the right-hand side of (19) is $\Omega(\frac{m^2}{\log^2 m})$.

**A lower bound for $\Upsilon_m$.** We provide a lower bound that holds for all $m > \tilde{m}$, where $\tilde{m}$ is a sufficiently large constant that will be set at the end of this step. First, we write

$$-\Upsilon_m = \frac{1}{m}\mathbb{E}_\pi\left[\sum_{t\in\mu_\pi(C)} r_\pi(t)^2 - q_\pi(t)^2\right]. \tag{22}$$

We will provide an upper bound for (22) by considering two cases in the following two claims.

**Claim D.7.** *There exists $m_1 > 0$ such that for all $m > m_1$, if $n \geq m + m^{\frac{2}{3}}$, then*

$$-\Upsilon_m \leq m^{\frac{11}{6}+\frac{1}{100}}.$$

*Proof.* For a market profile $\pi$, define

$$M_\pi = |\{t \in S : \mu_\pi(t) \neq \eta_\pi(t)\}|.$$

Let $\delta = 1/100$. Then, by Theorem B.3, there exists $m_\delta > 0$ such that

$$\mathbb{E}_\pi\left[M_\pi\right] \leq m^{3/2-2/3+\delta} = m^{5/6+\delta}$$

holds for all $m > m_\delta$. Then, by the above inequality and by (22) we can write

$$-\Upsilon_m \leq \frac{1}{m}\mathbb{E}_\pi\left[M_\pi m^2\right] \leq m^{11/6+\delta}.$$

Setting $m_1 = m_\delta$ concludes the proof. $\quad\square$

**Claim D.8.** *There exists $m_2 > 0$ such that for all $m > m_2$, if $n < m + m^{2/3}$, then $-\Upsilon_m \leq \frac{m^2}{500(\log m)^2}$.*

*Proof.* First, we need a few definitions. Define

$$\alpha = 1/321, \beta = 12\alpha,$$
$$r = \frac{m\beta}{(\log m)^{3/4}},$$
$$\theta = 2ne^{-\alpha(\log m)^{1/4}}(\log m)^3.$$

Let $x_\pi$ denote the number of students with rank worse than $r$ in $\eta_\pi$. (We recall that, by definition, an unassigned student has rank $\infty$, and that such a rank is *worse than* any finite rank.) Let $E$ be the event $x_\pi \leq \theta$. (We recall that the random variable $\pi$ is drawn uniformly at random from $\Pi^{n,m}_{\text{MTB}}$).

For market profile $\pi$, define

$$M_\pi = |\{s \in S : \mu_\pi(s) \neq \eta_\pi(s)\}|.$$

Let $F$ be the event $M_\pi \leq \frac{m}{\sqrt{\log m}}$.

Using the above definitions we can write

$$-\Upsilon_m =$$

$$\mathbb{P}_\pi\left[\bar{E}\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m|\bar{E}\right] \tag{23}$$

$$+\mathbb{P}_\pi\left[E \cap \bar{F}\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m|E \cap \bar{F}\right] \tag{24}$$

$$+\mathbb{P}_\pi\left[E \cap F\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m|E \cap F\right]. \tag{25}$$

To prove the claim, we will provide upper bounds for (23), (24), and (25). For every student $t$, let $d_\pi(t)$ denote the number of proposals that $t$ receives in the school-proposing DA when the market profile is $\pi$.

**An upper bound for** (23). To provide an upper bound for (23), we apply Lemma B.14 with $\kappa = 1/2$. This implies that there exists $m_\kappa$ such that, for all $m > m_\kappa$,

$$\mathbb{P}_\pi\left[d_\pi(s) < \frac{1}{4}\log\frac{n}{n-m}\right] \leq \frac{1}{m} + e^{-\frac{1}{107}\log(\frac{n}{n-m})}.$$

The above inequality, together with Lemma B.4, implies that there exists a constant $\tilde{m}_k$ such that for every $m > \tilde{m}_k$,

$$\mathbb{P}_\pi\left[r_\pi(s) > r\right] \leq \frac{1}{m} + e^{-\frac{1}{107}\log(\frac{n}{n-m})} + e^{-\frac{\frac{r}{4}\log\frac{n}{n-m}}{m}}$$

$$\leq \frac{1}{m} + e^{-\frac{\log m}{321}} + e^{-\frac{r\log m}{12m}}$$

$$\leq e^{-\log m} + e^{-\frac{\log m}{321}} + e^{-\alpha(\log m)^{1/4}} \leq 2e^{-\alpha(\log m)^{1/4}}$$

holds, where the second inequality follows from the fact that $n < m + m^{\frac{2}{3}}$, the third inequality follows from the definition of $r$, and the last inequality follows from the definition of $\alpha$ and the fact that $e^{-\log m} + e^{-\frac{\log m}{321}} \leq e^{-\alpha(\log m)^{1/4}}$ holds for sufficiently large $m$. Recall that $x_\pi$ denotes the number of assigned students with rank worse than $r$ in $\eta_\pi$. By the above inequality, for every $m > \tilde{m}_\kappa$

$$\mathbb{E}_\pi\left[x_\pi\right] \leq 2ne^{-\alpha(\log m)^{1/4}}.$$

Hence, by the Markov inequality, for every $m > \tilde{m}_\kappa$,

$$\mathbb{P}_\pi\left[x_\pi > 2ne^{-\alpha(\log m)^{1/4}}(\log m)^3\right] < (\log m)^{-3}.$$

Recall that $\theta = 2ne^{-\alpha(\log m)^{1/4}}(\log m)^3$, by definition. Also, recall that $E$ is the event $x_\pi \leq \theta$, by definition. Hence, we can rewrite the above inequality as

$$\mathbb{P}_\pi\left[\bar{E}\right] < (\log m)^{-3}.$$

The above bound, together with the fact that $-\Upsilon_m \leq m^2$ (which follows directly from (22)), implies that

$$\mathbb{P}_\pi\left[\bar{E}\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m|\bar{E}\right] \leq \frac{m^2}{(\log m)^3} \tag{26}$$

holds for every $m > \tilde{m}_\kappa$. This is the promised upper bound for (23). Define $m_{2,1} = \tilde{m}_\kappa$.

**An upper bound for** (24). We first note that $\mathbb{P}_\pi\left[E \cap \bar{F}\right] \leq \mathbb{P}_\pi\left[\bar{F}\right]$. Then, we recall Theorem 1 from Ashlagi et al. (2017) that shows there exists a constant $m_{2,2}$ such that, for all $m > m_{2,2}$,

$$\mathbb{P}_\pi\left[\bar{F}\right] \leq e^{-(\log m)^{0.4}}.$$

On the other hand, $-\Upsilon_m \leq m^2$ (which follows directly from (22)). The two latter inequalities imply that, for all $m > m_{2,2}$,

$$\mathbb{P}_\pi\left[E \cap \bar{F}\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m|E \cap \bar{F}\right] \leq e^{-(\log m)^{0.4}}m^2, \tag{27}$$

which is the promised upper bound for (24).

**An upper bound for** (25). Recall that (22) is

$$-\Upsilon_m = \frac{1}{m}\mathbb{E}_\pi\left[\sum_{t \in \mu_\pi(C)} r_\pi(t)^2 - q_\pi(t)^2\right].$$

Then, let $S' \subseteq S$ denote the set of students who have different ranks under the school- and student-optimal assignments. Since event $F$ holds, $|S'| \leq \frac{m}{\sqrt{\log m}}$. Let $S'' \subseteq S'$ contain every student $s \in S'$ for whom $r_\pi(s') > r$. Since event $E$ holds, we must have $|S''| \leq \theta$. Therefore, by (22),

$$\mathbb{E}_\pi\left[-\Upsilon_m|E \cap F\right] \leq \frac{1}{m}\left(m^2|S''| + r^2|S'\backslash S''|\right) \leq \frac{1}{m}\left(m^2\theta + \frac{r^2m}{\sqrt{\log m}}\right)$$
$$\leq \frac{4m^2(\log m)^3}{e^{\alpha(\log m)^{1/4}}} + \frac{m^2\beta^2}{(\log m)^2}$$

48

holds, where recall that $\alpha = 1/321$ and $\beta = 12\alpha$. Since $\frac{4m^2(\log m)^3}{e^{\alpha(\log m)^{1/4}}} = o(\frac{m^2\beta^2}{(\log m)^2})$, then there exists a constant $m_{2,3}$ such that, for all $m > m_{2,3}$,

$$\mathbb{E}_\pi\left[-\Upsilon_m | E \cap F\right] \le \frac{m^2}{500(\log m)^2}.$$

By the above inequality,

$$\mathbb{P}_\pi\left[E \cap F\right] \cdot \mathbb{E}_\pi\left[-\Upsilon_m | E \cap F\right] \le \frac{m^2}{500(\log m)^2} \tag{28}$$

holds for all $m > m_{2,3}$. This is the promised upper bound for (25).

We are now ready complete the proof of the claim by providing an upper bound on $-\Upsilon_m$. Recall that we wrote $-\Upsilon_m$ as the sum of three summands, namely, (23), (24), and (25). The three upper bounds that we provided above for these three summands, namely, (26), (27), and (28) together imply that for every $m > \max\{m_{2,1}, m_{2,2}, m_{2,3}\}$,

$$-\Upsilon_m \le \frac{m^2}{500(\log m)^2}.$$

Setting $m_2 = \max\{m_{2,1}, m_{2,2}, m_{2,3}\}$ concludes the proof of the claim.

$\square$

By Claim D.7 and Claim D.8, for every $m > \max\{m_1, m_2\}$ we have

$$-\Upsilon_m \le \max\left\{m^{\frac{11}{6} + \frac{1}{100}}, \frac{m^2}{500(\log m)^2}\right\}.$$

Observe that $m^{\frac{11}{6} + \frac{1}{100}} = O(\frac{m^2}{500(\log m)^2})$. Therefore, there exists $\tilde{m}$ such that for all $m > \tilde{m}$,

$$\Upsilon_m \ge -\frac{m^2}{500(\log m)^2}. \tag{29}$$

This is the promised lower bound on $\Upsilon_m$.

We are now ready to complete Step 2. Recall that by (20) we have

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] = \Upsilon_m + \Xi_m.$$

By (21) and (29),

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] \ge -\frac{m^2}{500(\log m)^2} \tag{30}$$

holds for $m > \tilde{m}$, where we recall the definition of $\tilde{m}$ from (29). This completes Step 2.

**Step 3.**  In this step, we will use (19) and (30) to conclude the proof of the lemma.

We recall (19) by which, for any constant $\epsilon \in (0, 0.1)$, there exists $m_\epsilon > 0$ such that

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] = \text{Var}_\pi[r_\pi(s) \mid r_\pi(s) \neq \infty] \geq \underbrace{\frac{1}{2} \cdot \frac{\bar{d}(m+1)(m-\bar{d})}{(\bar{d}+1)^2(\bar{d}+2)}}_{X_m} \tag{31}$$

holds for all $m > m_\epsilon$, where recall that $\bar{d} = 2(1+\epsilon)\frac{n}{m}\log\frac{n}{n-m}$. We denote the right-hand side of the above inequality by $X_m$.

**Claim D.9.** *There exists $m^*$ such that for all $m > m^*$,*

$$X_m/4 \geq \frac{m^2}{500(\log m)^2} + 1.$$

*Proof.* We first provide an upper bound for $\bar{d}$ by applying Fact A.2. This implies that

$$\bar{d} \leq 2(1+\epsilon)\frac{m+1}{m}\log\frac{(m+1)}{(m+1)-m} \leq 2(1+\epsilon)(1+\frac{1}{m})\log(m+1).$$

Let $\mathsf{d} = 2(1+\epsilon)(1+\frac{1}{m})\log(m+1)$. The above bound, (31), and Fact D.3 together imply that

$$\begin{aligned}
X_m &\geq \frac{1}{2} \cdot \frac{\mathsf{d}(m+1)(m-\mathsf{d})}{(\mathsf{d}+1)^2(\mathsf{d}+2)} \\
&\geq \frac{1}{2} \cdot \frac{\mathsf{d}}{\mathsf{d}+2} \cdot \frac{(m+1)(m-\mathsf{d})}{(\mathsf{d}+1)^2} \\
&\geq \frac{1}{2} \cdot (1 - \frac{2}{\mathsf{d}+2})(1 - \frac{1}{\mathsf{d}+1})^2(1 - \frac{1}{m}) \cdot \frac{m(m-\mathsf{d})}{\mathsf{d}^2}.
\end{aligned} \tag{32}$$

Now, we observe that there exists a constant $m^* > 10$ such that, for all $m > m^*$ we have

$$\mathsf{d} \leq \frac{m}{10}, \tag{33}$$

$$1 - \frac{1}{\mathsf{d}+1}, 1 - \frac{2}{\mathsf{d}+2} \geq \frac{9}{10}, \tag{34}$$

$$\frac{4}{500}\frac{m^2}{(\log m)^2} \geq 1, \tag{35}$$

$$(1 + \frac{1}{m})\log(m+1) \leq 1.1\log m. \tag{36}$$

(32), (33), and (34) together imply that

$$\begin{aligned}
X_m &\geq \frac{1}{2} \cdot (1 - \frac{2}{\mathsf{d}+2})(1 - \frac{1}{\mathsf{d}+1})^2(1 - \frac{1}{m}) \cdot \frac{m(m-\mathsf{d})}{\mathsf{d}^2} \\
&\geq \frac{0.9^4}{2} \cdot \frac{m(m-m/10)}{\mathsf{d}^2} \\
&\geq \frac{0.9^5}{2} \cdot \frac{m^2}{\mathsf{d}^2}
\end{aligned}$$

50

The above bound implies that

$$X_m/4 \geq \frac{0.9^5}{8} \cdot \frac{m^2}{\mathsf{d}^2} \geq \frac{0.9^5}{8(2.2(1+\epsilon))^2} \cdot \frac{m^2}{(\log m)^2} > \frac{5}{500} \cdot \frac{m^2}{(\log m)^2} \geq \frac{1}{500} \cdot \frac{m^2}{(\log m)^2} + 1,$$

where the second inequality follows from (35) and (36), and the third inequality follows from the fact that $\epsilon < 0.1$. This concludes the proof. $\qquad\square$

We are now ready to prove the lemma. For $m > \max\{\tilde{m}, m^*, m_\epsilon\}$ it holds that

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

$$\geq X_m + \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] \tag{37}$$

$$\geq X_m - \frac{m^2}{500 \log m} \tag{38}$$

$$\geq \frac{m^2}{500(\log m)^2} \tag{39}$$

where (37) follows from (31); (38) follows from (30); and (39) follows from Claim D.9. Hence, $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] = \Omega(m^2/(\log m)^2)$. $\qquad\square$

**Lemma D.10.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$. Consider a market $\mathcal{M}$ with $m$ schools and $n = m + \lambda(m)$ students. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\mathrm{STB}}^{n,m}$. Then,*

$$\mathbb{E}_\pi\left[\mu_\pi^\#(s)^2 \big| \mu_\pi(s) \neq \varnothing\right] \leq m(2 + \pi^2/3).$$

*Proof.* Let $t = \sqrt{m}\log m$. The proof uses the equivalence relation between the DA-STB and the RSD mechanism. Recall this equivalence relation from Section A.1. For any student $s$, let $p_s$ denote the *priority number* of $s$ in the mechanism. We consider two cases: either $p_s \leq m - t$ or not. Observe that

$$\mathbb{E}_\pi\left[\mu_\pi^\#(s)^2 \big| \mu_\pi(s) \neq \varnothing\right] = \mathbb{P}_\pi\left[p_s \leq m - t \big| \mu_\pi(s) \neq \varnothing\right] \cdot \mathbb{E}_\pi\left[\mu_\pi^\#(s)^2 \big| p_s \leq m - t\right]$$

$$+ \mathbb{P}_\pi\left[m - t < p_s \leq m \big| \mu_\pi(s) \neq \varnothing\right] \cdot \mathbb{E}_\pi\left[\mu_\pi^\#(s)^2 \big| m - t < p_s \leq m\right]. \tag{40}$$

We will provide an upper bound for each of the terms on the right-hand side of (40).

**Claim D.11.** *A student $s$ with priority number $m - t$ is assigned to one of her top $\frac{m \log m}{t}$ choices with probability at least $1 - 1/m$.*

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

By Claim D.11, we have:

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

which implies that

$$\mathbb{P}_\pi \left[ p_s \leq m - t \big| \mu_\pi(s) \neq \varnothing \right] \cdot \mathbb{E} \left[ \mu_\pi^\#(s)^2 \big| p_s \leq m - t \right] \leq 2m. \tag{41}$$

Also, we have that

$$\mathbb{P}_\pi \left[ m - t < p_s \leq m \big| \mu_\pi(s) \neq \varnothing \right] \cdot \mathbb{E}_\pi \left[ \mu_\pi^\#(s)^2 \big| m - t < p_s \leq m \right] \leq \frac{t}{m} \cdot \sum_{i=1}^t \frac{1}{t} \mathbb{E} \left[ r_s^2 \big| p_s = m - i + 1 \right]$$

$$\leq \frac{1}{m} \cdot \sum_{i=1}^t \frac{2 - i/m}{(i/m)^2} \tag{42}$$

$$\leq \frac{1}{m} \cdot \sum_{i=1}^t 2(m/i)^2$$

$$\leq m \cdot \frac{\pi^2}{3}, \tag{43}$$

where (42) holds because (i) conditional on $p_s = m - i + 1$, the distribution of $r_s$ is stochastically dominated by the geometric distribution with mean $m/i$; this holds because for any integer $k \in [0, m - i]$, we have

$$\mathbb{P}_\pi \left[ r_s > k \big| p_s = m - i + 1 \right] = \frac{m - i}{m} \cdot \frac{m - i - 1}{m - 1} \cdot \ldots \cdot \frac{m - i - k + 1}{m - k + 1} \leq \left( \frac{m - i}{m} \right)^k,$$

and (ii) for a geometric random variable $X$ with mean $1/p$, we have $\mathbb{E} \left[ X^2 \right] = \frac{2-p}{p^2}$.

Finally, putting (40), (41), and (43) together concludes the proof:

$$\mathbb{E}_\pi \left[ \mu_\pi^\#(s)^2 \big| \mu_\pi(s) \neq \varnothing \right] \leq m(2 + \frac{\pi^2}{3}).$$

$\square$

**Lemma D.12.** *Let* $\lambda : \mathbb{N} \to \mathbb{N}$. *Consider a market* $\mathcal{M}$ *with* $m$ *schools and* $n = m + \lambda(m)$ *students. Suppose that the market profile for* $\mathcal{M}$, *namely* $\pi$, *is drawn uniformly at random from* $\Pi_{\mathrm{STB}}^{n,m}$. *Then,* $\mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) \right] = \Theta(m)$.

*Proof.* For notational brevity, let the random variable $r_s$ denote the rank of a student $s$ in the assignment $\mu_\pi$. Since the expected social inequity and the conditional variance of the rank of a fixed student given that she is assigned are equal (by Lemma D.4), we can analyze the latter notion instead of the former. Hence, we will show that $\mathbb{E}_\pi \left[ (r_s - r)^2 \big| r_s \neq \infty \right] = \Theta(m)$, where $r = \mathbb{E}_\pi \left[ r_s \big| r_s \neq \infty \right] = \mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]$.

We observe that, with probability at least $1/2$, the student with the priority number $m$ has a rank worse than $m/2$ in $\mu_\pi$. So, for any student $s \in S$ we can write:

$$\mathbb{E}_\pi \left[ (r_s - r)^2 | r_s \neq \infty \right] \geq \frac{1}{m} \cdot \frac{1}{2} \cdot (r - m/2)^2 .$$

As shown by Knuth (1995), $r = \Theta(\log m)$. Thus, by the above inequality, $\mathbb{E}_\pi \left[ (r_s - r)^2 | r_s \neq \infty \right] = \Omega(m)$.[30] On the other hand,

$$\mathbb{E}_\pi \left[ (r_s - r)^2 | r_s \neq \infty \right] = \mathbb{E}_\pi \left[ r_s^2 | r_s \neq \infty \right] - r^2$$
$$\leq \mathbb{E}_\pi \left[ r_s^2 | r_s \neq \infty \right] \leq m(2 + \frac{\pi^2}{3}),$$

where the last inequality holds by Lemma D.10. Therefore, $\mathbb{E}_\pi \left[ (r_s - r)^2 | r_s \neq \infty \right] = \Theta(m)$. $\square$

*Proof of Theorem D.1, part (i).* Observe that Lemma D.5 and Lemma D.12 together imply that

$$\frac{\mathbb{E}_{\pi \sim U(\Pi_{\text{MTB}}^{n,m})} \left[ Si(\mu_\pi) \right]}{\mathbb{E}_{\pi \sim U(\Pi_{\text{STB}}^{n,m})} \left[ Si(\mu_\pi) \right]} = \Omega \left( \frac{m^2/\log^2 m}{m} \right),$$

which proves the claim. $\square$

## D.3  Proof of Theorem D.1, part (ii)

**Lemma D.13.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$. Consider a market $\mathcal{M}$ with $n$ students and $m = n + \lambda(n)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\text{MTB}}^{n,m}$. Then, $\mathbb{E}_\pi \left[ Si(\mu_\pi) \right] = O(\log^2 n)$.*

*Proof.* The proof has two steps. Loosely speaking, in Step 1, we show that $\mathbb{E}_\pi \left[ Si(\mu_\pi) \right]$ cannot be "much larger" than $\mathbb{E}_\pi \left[ Si(\eta_\pi) \right]$. In Step 2, we show that $\mathbb{E}_\pi \left[ Si(\eta_\pi) \right] = O(\log^2 n)$.

**Step 1.** First, write the following equality. For every student $s \in S$,

$$n \cdot \mathbb{E}_\pi \left[ Si(\mu_\pi) - Si(\eta_\pi) \right] = n \operatorname{Var}_\pi [\mu_\pi^\#(s) \mid \mu_\pi(s) \neq \varnothing] - n \operatorname{Var}_\pi [\eta_\pi^\#(s) \mid \eta_\pi(s) \neq \varnothing] \quad (44)$$
$$=$$
$$- n \cdot \mathbb{E}_\pi \left[ Ar(\mu_\pi)^2 - Ar(\eta_\pi)^2 \right] \quad (45)$$
$$+ \mathbb{E}_\pi \left[ \sum_{t \in S} \mu_\pi^\#(t)^2 - \eta_\pi^\#(t)^2 \right], \quad (46)$$

where the first equality holds by Lemma D.4. We will complete Step 1 by providing upper bounds for (45) and (46).

---

[30]To see why this holds, we recall that $f = \Omega(g)$ if there exist constants $m_0, c > 0$ such that for every $m > m_0$, $f(m) \geq cg(m)$.

**An upper bound for** (45). By Theorem B.1, the event $\mathcal{A}r(\eta_\pi) \leq \mathcal{A}r(\mu_\pi)(1+\zeta)$ holds with very high probability as $n$ grows large. Hence, for any constant $\zeta > 0$, there exists a constant $n_\zeta$ such that for $n > n_\zeta$,

$$\mathbb{P}_\pi \left[ \mathcal{A}r(\eta_\pi) \leq \mathcal{A}r(\mu_\pi)(1+\zeta) \right] \geq 1 - \frac{1}{m^2}.$$

This implies the following upper bound on (45), which holds for $n > n_\zeta$ :

$$-n \cdot \mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2 \right] \leq n(2\zeta + \zeta^2)\mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 + nm^2 \cdot \frac{1}{m^2}$$
$$= n(2\zeta + \zeta^2)\mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 + n.$$

**An upper bound for** (46). By the definition of $\mu, \eta$, we always have $\mu_\pi^\#(s) \leq \eta_\pi^\#(s)$. Therefore, 0 is a valid upper bound for (46), i.e.,

$$\mathbb{E}_\pi \left[ \sum_{t \in S} \mu_\pi^\#(t)^2 - \eta_\pi^\#(t)^2 \right] \leq 0.$$

The upper bounds that we provided above for (45) and (46) imply that, for every constant $\zeta > 0$ and every $n > n_\zeta$,

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] \leq (2\zeta + \zeta^2)\mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 + 1. \tag{47}$$

**Step 2.** We will show that $\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] = O(\log^2 n)$. First, see that for any $s \in S$,

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] = \mathbb{E}_\pi \left[ (\mathcal{A}r(\eta_\pi) - \eta_\pi^\#(s))^2 \big| \eta_\pi(s) \neq \varnothing \right]$$
$$= \mathbb{E}_\pi \left[ \eta_\pi^\#(s)^2 \big| \eta_\pi(s) \neq \varnothing \right] - \mathbb{E}_\pi \left[ \mathcal{A}r(\eta_\pi) \right]^2$$
$$\leq \mathbb{E}_\pi \left[ \eta_\pi^\#(s)^2 \big| \eta_\pi(s) \neq \varnothing \right].$$

For notational simplicity, let $r_s$ denote $\eta_\pi^\#(s)$. Also, observe that the event $r_s \neq \infty$ holds in all realizations of $\pi$, because $n < m$. Hence, we can rewrite the above bound as

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] \leq \mathbb{E}_\pi \left[ r_s^2 \right]. \tag{48}$$

Next, we provide an upper bound on $\mathbb{E}_\pi \left[ r_s^2 \right]$. Fix an arbitrary small constant $\epsilon > 0$. Let $E_s$ denote the event in which student $s$ receives at least $\bar{d} = \frac{(1-\epsilon)n}{2 \log n}$ proposals in the school-proposing DA.

**Claim D.14.** $\mathbb{E}_\pi \left[ r_s^2 \right] = O(\log^2 n)$.

*Proof.* We start by writing

$$\mathbb{E}_\pi \left[ r_s^2 \right]$$
$$= \mathbb{P}_\pi \left[ E_s \right] \mathbb{E}_\pi \left[ r_s^2 \middle| E_s \right] + (1 - \mathbb{P}_\pi \left[ E_s \right]) \mathbb{E}_\pi \left[ r_s^2 \middle| \overline{E_s} \right]$$
$$\leq \mathbb{E}_\pi \left[ r_s^2 \middle| E_s \right] + (1 - \mathbb{P}_\pi \left[ E_s \right]) m^2. \tag{49}$$

We can provide an upper bound on the right-hand side of (49) by observing the following two facts. First, we have

$$\mathbb{E} \left[ r_s^2 \middle| E_s \right] = O(m^2/\bar{d}^2) = O(\log^2 n), \tag{50}$$

which holds by Fact D.3. Second, by Lemma B.12, $E_s$ happens with very high probability as $n$ grows large. Therefore, $1 - \mathbb{P}_\pi \left[ E_s \right] \leq n^{-2}$ holds for sufficiently large $n$. This implies that

$$(1 - \mathbb{P}_\pi \left[ E_s \right]) m^2 \leq \frac{m^2}{n^2} \leq 4 \tag{51}$$

holds for sufficiently large $n$, where the last inequality follows from the fact that $m \leq 2n$.

Finally, observe that (49), (50), and (51) together imply that $\mathbb{E}_\pi \left[ r_s^2 \right] = O(\log^2 n)$.

$\square$

We are now ready to finish the proof. Observe that (48) and Claim D.14 together imply that

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right] = O(\log^2 n). \tag{52}$$

Therefore, (47) and (52) imply that, for every constant $\zeta$ and all $n > n_\zeta$,

$$\mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) \right] \leq \mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] + \mathbb{E}_\pi \left[ \mathcal{S}i(\eta_\pi) \right]$$
$$\leq (2\zeta + \zeta^2) \mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right]^2 + 1 + O(\log^2 n).$$

Theorem B.1, together with Fact A.2, implies that, for any constant $\delta > 0$, $\mathcal{A}r(\mu_\pi) \leq (1 + \delta) \log n$ holds wvhp. This fact, together with the above inequality, implies that $\mathbb{E}_\pi \left[ \mathcal{S}i(\mu_\pi) \right] = O(\log^2(n))$.

$\square$

So far, we have shown that the expected social inequity under MTB is "small". In the following lemmas, we show that the expected social inequity under STB is "large", and then we conclude the proof.

**Lemma D.15.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$. Consider a market $\mathcal{M}$ with $n$ students and $m = n + \lambda(n)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{STB}^{n,m}$. Then, $\mathbb{E}_\pi \left[ \mathcal{A}r(\mu_\pi) \right] \leq \frac{m}{n} (\frac{1}{2m} + \log(\frac{m}{m-n}))$.*

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism.

(Recall this equivalence relation from Section A.1.) For any student $s$, let $p_s$ denote the priority number of $s$. Let $r_s$ denote the rank of a student $s$ in $\mu_\pi$.

Observe that, for any integer $k \geq 0$ and any integer $i \in [0, k]$, we have

$$\mathbb{P}_\pi [r_s > i | p_s = k + 1] = \frac{k}{m} \cdot \frac{k-1}{m-1} \cdot \ldots \cdot \frac{k-i+1}{m-i+1} \leq (k/m)^i.$$

Let $X$ denote a geometric random variable with parameter (i.e., success probability) $1 - \frac{k}{m}$. The above inequality implies that

$$\mathbb{E}_\pi [r_s | p_s = k + 1] \leq \mathbb{E} [X] = \frac{m}{m-k}.$$

Hence,

$$\mathbb{E}_\pi [\mathcal{A}r(\mu_\pi)] \leq \frac{1}{n} \cdot \sum_{k=0}^{n-1} \frac{m}{m-k} = \frac{m}{n}(H_m - H_{m-n}), \tag{53}$$

where $H_i$ denotes the $i$-th harmonic number. Finally, we recall the following bound on harmonic numbers (see, e.g., Villarino (2005)),

$$\gamma + \log i \leq H_i \leq \gamma + \frac{1}{2i} + \log i,$$

where $\gamma$ is the Euler–Mascheroni constant. This bound, together with (53), concludes the lemma:

$$\mathbb{E}_\pi [\mathcal{A}r(\mu_\pi)] \leq \frac{m}{n}(H_m - H_{m-n}),$$
$$\leq \frac{m}{n} \cdot \left( \gamma + \frac{1}{2m} + \log m - \gamma - \log(m-n) \right) = \frac{m}{n}(\frac{1}{2m} + \log \frac{m}{m-n}).$$

$\square$

**Lemma D.16.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$ be such that $\lambda(n) \leq n^\gamma$ for a positive constant $\gamma < 1$. Consider a market $\mathcal{M}$ with $n$ students and $m = n + \lambda(n)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn from $\Pi_{\text{STB}}^{n,m}$ uniformly at random. Then, there exist positive constants $c_\gamma, n_\gamma$ (independent of $n$) such that $\mathbb{E}_\pi [\mathcal{S}i(\mu_\pi)] \geq n^{c_\gamma}$ holds for all $n > n_\gamma$.*

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) For any student $s$, let $p_s$ denote the priority number of $s$. Let $r_s$ denote the rank of a student $s$ in $\mu_\pi$.

Fix a student $s$. Using Lemma D.4, we can write

$$\mathbb{E}_\pi [\mathcal{S}i(\mu_\pi)] = \mathbb{E}_\pi \left[ (r_s - r)^2 | r_s \neq \infty \right] = \mathbb{E}_\pi \left[ r_s^2 | r_s \neq \infty \right] - r^2, \tag{54}$$

where $r$ denotes $\mathbb{E}_\pi [r_s | r_s \neq \infty]$ and $r_s$ denotes the rank of $s$ in $\mu_\pi$. To provide a lower bound for (54), we provide a lower bound for $\mathbb{E}_\pi \left[ r_s^2 | r_s \neq \infty \right]$ and use Lemma D.15, which provides an upper

bound for $r$.

**Lower bound for $\mathbb{E}_\pi\left[r_s^2|r_s \neq \infty\right]$.** First, see that

$$\mathbb{E}_\pi\left[r_s^2|r_s \neq \infty\right] = \frac{1}{n} \cdot \sum_{i=0}^{n-1} \mathbb{E}\left[r_s^2|p_s = i+1\right].$$

**Claim D.17.** *Let $\Delta_k = m - n + k$. Then, $\mathbb{E}_\pi\left[r_s^2|p_s = n - k + 1\right] \geq \left(\frac{m}{5\Delta_k}\right)^2$.*

*Proof.* By Lemma C.8, a student $s$ who receives a priority number $n - k + 1$ is assigned to one of her top $x$ choices with probability at most $\frac{x\Delta_k}{m-x+1}$. For any $\epsilon \in (0,1)$, choosing $x = \frac{\epsilon m}{\Delta_k}$ implies that a student $s$ who receives a priority number $n - k + 1$ is assigned to one of her top $\frac{\epsilon m}{\Delta_k}$ choices with probability at most $\frac{\epsilon}{1-\epsilon}$. Letting $\epsilon = 1/3$ implies that

$$\mathbb{P}_\pi\left[r_s \geq \frac{\epsilon m}{\Delta_k}\middle|p_s = n - k + 1\right] \geq \frac{1}{2}.$$

Therefore,

$$\mathbb{E}_\pi\left[r_s^2\middle|p_s = n - k + 1\right] \geq \left(\frac{\epsilon m}{\sqrt{2}\Delta_k}\right)^2 \geq \left(\frac{m}{5\Delta_k}\right)^2.$$

$\square$

Observe that by Claim D.17, for any positive integer $\overline{k}$, we have

$$\mathbb{E}_\pi\left[r_s^2|r_s \neq \infty\right] \geq \frac{1}{n} \cdot \sum_{k=0}^{\overline{k}}\left(\frac{m}{5\Delta_k}\right)^2$$

$$\geq \frac{\overline{k}}{n} \cdot \left(\frac{m}{5\Delta_{\overline{k}}}\right)^2$$

$$\geq \frac{m}{25} \cdot \frac{\overline{k}}{(\Delta_{\overline{k}})^2} \geq \frac{\overline{k}n}{25(n^\gamma + \overline{k})^2},$$

where the last inequality uses the fact that $m - n \leq n^\gamma$ for a constant $\gamma \in (0,1)$. Let $b = \frac{\overline{k}n}{25(n^\gamma + \overline{k})^2}$. Choose a positive constant $\zeta < 1 - \gamma$, and set $\overline{k} = n^{2\gamma - 1 + \zeta}$. Then, observe that

$$\mathbb{E}_\pi\left[r_s^2|r_s \neq \infty\right] \geq b \geq \frac{n^{2\gamma+\zeta}}{25(2n^\gamma)^2} = \frac{n^\zeta}{100}.$$

Recall that, by Lemma D.15, $\mathbb{E}_\pi\left[r_s|r_s \neq \infty\right] \leq \frac{m}{n}\left(\frac{1}{2m} + \log\left(\frac{m}{m-n}\right)\right)$. This and the above inequality together imply that

$$\mathbb{E}_\pi\left[Si(\mu_\pi)\right] = \mathbb{E}_\pi\left[r_s^2|r_s \neq \infty\right] - \left(\mathbb{E}_\pi\left[r_s|r_s \neq \infty\right]\right)^2 = \Omega(n^\zeta),$$

which concludes the proof.

$\square$

*Proof of Theorem D.1, part (ii).* By Lemma D.13, we have that $\mathbb{E}_{\pi \sim U(\Pi_{\text{MTB}}^{n,m})}[Si(\mu_\pi)] = O(\log^2 n)$, and by Lemma D.16 we have that $\mathbb{E}_{\pi \sim U(\Pi_{\text{STB}}^{n,m})}[Si(\mu_\pi)] = \Omega(n^{c_\gamma})$ for some constant $c_\gamma > 0$. Therefore,

$$\lim_{n \to \infty} \frac{\mathbb{E}_{\pi \sim U(\Pi_{\text{STB}}^{n,m})}[Si(\mu_\pi)]}{\mathbb{E}_{\pi \sim U(\Pi_{\text{MTB}}^{n,m})}[Si(\mu_\pi)]} = \infty.$$

$\square$

## D.4 Proof of Theorem D.2

**Lemma D.18.** *Let $\lambda > 0$ be a constant. Consider a market $\mathcal{M}$ with $n$ students and $m = n(1 + \lambda)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\text{MTB}}^{n,m}$. Then,*

$$\mathbb{E}_\pi[Si(\eta_\pi)] \le \frac{8(1 + \lambda)^2}{f(\log \frac{1+\lambda}{\lambda})^2} - \left((1 + \lambda) \log \frac{1 + \lambda}{\lambda}\right)^2 + o(1),$$

*where $f(x) = \frac{e^x - 1 - x}{x(e^x - 1)}$ and $o(1)$ suppresses a term that approaches $0$ as $n$ approaches infinity.*

*Proof.* For every student $s \in S$, let $r_s = \eta_\pi^{\#}(s)$. Fix a student $s \in S$. Define $r = \mathbb{E}_\pi[r_s | r_s \ne \infty]$. First, we use Lemma D.4 to write

$$\mathbb{E}_\pi[Si(\eta_\pi)] = \mathbb{E}_\pi\left[(r_s - r)^2 | r_s \ne \infty\right] = \mathbb{E}_\pi\left[r_s^2 | r_s \ne \infty\right] - r^2.$$

The proof works by providing an upper bound on $\mathbb{E}_\pi\left[r_s^2 | r_s \ne \infty\right]$ and a lower bound on $\mathbb{E}_\pi[r_s | r_s \ne \infty]$.

**Upper bound on $\mathbb{E}_\pi\left[r_s^2 | r_s \ne \infty\right]$.** Let the random variable $d_s$ denote the number of proposals that student $s$ receives when the school-proposing DA is run in the market $\mathcal{M}$. By Fact D.3, for any positive integer $d \le m$, we have

$$\mathbb{E}_\pi\left[r_s^2 | d_s = d\right] = \frac{d(m+1)(m-d)}{(d+1)^2(d+2)} + \frac{(m+1)^2}{(d+1)^2} \le 2\left(\frac{m}{d}\right)^2, \tag{55}$$

because each summand in the middle term is at most $(\frac{m}{d})^2$. (In particular, the first summand is bounded by $(\frac{m}{d})^2$ because $(m+1)(m-d) \le m^2$.) In addition, by Lemma B.7, for any constant $\epsilon > 0$, $d_s \ge (1-\epsilon)nf(\log \frac{m}{m-n})/2$ holds with high probability as $n$ grows large, where $f(x) = \frac{e^x - 1 - x}{x(e^x - 1)}$. This fact together with (55) imply that, for any constant $\epsilon \in (0, 1)$, there exists $n_\epsilon > 0$ such that for all $n > n_\epsilon$,

$$\mathbb{E}_\pi\left[r_s^2 | r_s \ne \infty\right] \le \frac{8(1 + \lambda)^2}{(1 - \epsilon)^2 f(\log \frac{1+\lambda}{\lambda})^2}. \tag{56}$$

This is the promised upper bound.

58

**Lower bound on** $r = \mathbb{E}_\pi[r_s | r_s \neq \infty]$. Theorem B.1 shows that, for any constant $\delta > 0$, wvhp, $r_s \geq \frac{m(1-\delta)}{n} \log \frac{m}{m-n}$. Therefore, for any constant $\delta \in (0,1)$, there exists a constant $n_\delta$ such that

$$r = \mathbb{E}_\pi[r_s | r_s \neq \infty] \geq (1-\delta)(1+\lambda) \log \frac{1+\lambda}{\lambda} \tag{57}$$

holds for all $n > n_\delta$.

Finally, observe that (56) and (57) together imply that for any constant $\gamma > 0$, there exists $n_\gamma > 0$ such that

$$\mathbb{E}_\pi[Si(\eta_\pi)] = \mathbb{E}_\pi[r_s^2 | r_s \neq \infty] - r^2$$

$$\leq \frac{8(1+\lambda)^2}{(1-\gamma)^2 f (\log \frac{1+\lambda}{\lambda})^2} - \left( (1-\gamma)(1+\lambda) \log \frac{1+\lambda}{\lambda} \right)^2.$$

This just means that

$$\mathbb{E}_\pi[Si(\eta_\pi)] \leq \frac{8(1+\lambda)^2}{f (\log \frac{1+\lambda}{\lambda})^2} - \left( (1+\lambda) \log \frac{1+\lambda}{\lambda} \right)^2 + o(1),$$

where $o(1)$ suppresses a term that approaches $0$ as $n$ approaches infinity.

$\square$

**Lemma D.19.** *Let $\lambda > 0$ be a constant. Consider a market $\mathcal{M}$ with $n$ students and $m = n(1+\lambda)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn uniformly at random from $\Pi_{\text{MTB}}^{n,m}$. Then,*

$$\lim_{n \to \infty} \mathbb{E}_\pi[Si(\mu_\pi) - Si(\eta_\pi)] = 0.$$

*Proof.* First, write the following equality. For every student $s \in S$,

$$\mathbb{E}_\pi[Si(\mu_\pi) - Si(\eta_\pi)] = \text{Var}_\pi[\mu_\pi^\#(s) \mid \mu_\pi(s) \neq \varnothing] - \text{Var}_\pi[\eta_\pi(s) \mid \eta_\pi(s) \neq \varnothing] \tag{58}$$

$$=$$

$$- \mathbb{E}_\pi\left[ Ar(\mu_\pi)^2 - Ar(\eta_\pi)^2 \right] \tag{59}$$

$$+ \frac{1}{n} \mathbb{E}_\pi\left[ \sum_{t \in S} \mu_\pi^\#(t)^2 - \eta_\pi^\#(t)^2 \right], \tag{60}$$

where the first equality holds by Lemma D.4. We next provide upper bounds for (59) and (60).

**An upper bound for** (59). Recall that $m = (1+\lambda)m$. By Theorem B.1, there exists a constant $c_\lambda > 0$ such that for every constant $\zeta > 0$, there exists a constant $n_\zeta$ such that

$$\mathbb{P}_\pi\left[ Ar(\eta_\pi), Ar(\mu_\pi) \in [(1-\zeta)c_\lambda, (1+\zeta)c_\lambda] \right] \geq 1 - \frac{1}{n^3}$$

holds for $n > n_\zeta$. This implies the following upper bound on (59), which holds for every constant $\zeta > 0$ and all $n > n_\zeta$:

$$-\mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right] \leq 4\zeta c_\lambda^2 + n^2 \cdot \frac{1}{n^3}.$$

**An upper bound for** (60). By the definition of $\mu, \eta$, we always have $\mu_\pi^\#(s) \leq \eta_\pi^\#(s)$. Therefore, 0 is a valid upper bound for (60), i.e.,

$$\mathbb{E}_\pi\left[\sum_{t \in S} \mu_\pi^\#(t)^2 - \eta_\pi^\#(t)^2\right] \leq 0.$$

The upper bounds that we provided above for (45) and (46) imply that, for every constant $\zeta > 0$ and all $n > n_\zeta$,

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] \leq 4\zeta c_\lambda^2 + \frac{1}{n},$$

which means that $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] = o(1)$. $\qquad\square$

**Lemma D.20.** *Let $\lambda > 0$ be a constant. Consider a market $\mathcal{M}$ with $n$ students and $m = n(1 + \lambda)$ schools. Suppose that the market profile for $\mathcal{M}$, namely $\pi$, is drawn from $\Pi_{\mathsf{STB}}^{n,m}$ uniformly at random. Then,*

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

*where $o(1)$ suppresses the terms that approach 0 as $n$ approaches infinity.*

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) For a student $s$, let $p_s$ denote the *priority number* of $s$. Also, let $r_s$ denote the rank of $s$ in $\mu_\pi$.

Fix a student $s$, and let $r$ denote $\mathbb{E}_\pi\left[r_s | r_s \neq \infty\right]$. Using Lemma D.4, we can write

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] = \mathbb{E}_\pi\left[(r_s - r)^2 | r_s \neq \infty\right] = \mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right] - r^2. \tag{61}$$

To provide a lower bound for (61), we provide a lower bound for $\mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right]$ and use Lemma D.15, which provides an upper bound for $r$.

**Lower bound for** $\mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right]$. First, see that

$$\mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right] = \frac{1}{n} \cdot \sum_{i=0}^{n-1} \mathbb{E}_\pi\left[r_s^2 | p_s = i + 1\right].$$

60

**Claim D.21.** *For all integers $k \in [2\log_{1+\lambda} n, n]$, it holds that*

$$\mathbb{E}_\pi\left[r_s^2 | p_s = k+1\right] \geq \frac{z+1-(2+2\log_{1+\lambda}n)^2/n^2}{(1-z)^2} - \frac{(\bar{t}+1)^5}{2m},$$

*where $z = \frac{k}{m}$ and $\bar{t} = 2\log_{1+\lambda} n$.*

*Proof.* First, we observe that

$$\mathbb{E}_\pi\left[r_s^2 | p_s = k+1\right] = \sum_{j=0}^{k}(j+1)^2 \cdot (1 - \frac{k-j}{m-j}) \cdot \prod_{l=0}^{j-1}\frac{k-l}{m-l}, \tag{62}$$

which holds because $(1-\frac{k-j}{m-j})\cdot\prod_{l=0}^{j-1}\frac{k-l}{m-l}$ is the probability that $r_s = j+1$ conditional on $p_s = k+1$.

For any nonnegative integer $t \leq k$, we provide a lower bound for the summand on the right-hand side of (62) corresponding to $j = t$. This summand contains the term $\prod_{l=0}^{t-1}\frac{k-l}{m-l}$, which we bound from below as follows:

$$\prod_{l=0}^{t-1}\frac{k-l}{m-l} \geq \prod_{l=0}^{t-1}\frac{k}{m} - \sum_{l=0}^{t-1}\left(\frac{k}{m} - \frac{k-l}{m-l}\right)$$

$$\geq \prod_{l=0}^{t-1}\frac{k}{m} - \sum_{l=0}^{t-1}\frac{l}{m} \geq (k/m)^t - \frac{t^2}{2m},$$

where the first inequality holds by Fact A.5. Now, using the above inequality and (62), we can write

$$\mathbb{E}_\pi\left[r_s^2 | p_s = k+1\right] \geq \sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1 - \frac{k}{m}) \cdot \left((k/m)^j - \frac{j^2}{2m}\right)$$

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

$$\geq \sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1 - \frac{k}{m}) \cdot (k/m)^j - \frac{(\bar{t}+1)^5}{2m}. \tag{63}$$

To provide a lower bound for (63), let $z = k/m$. We will show that

$$\sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1 - \frac{k}{m}) \cdot (k/m)^j \geq \frac{z+1-(2+2\log_{1+\lambda}n)^2/n^2}{(1-z)^2}, \tag{64}$$

61

which will conclude the proof. To this end, we observe that

$$\sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1-z) \cdot z^j$$

$$= \frac{-(\bar{t}+1)^2 z^{\bar{t}+3} + \left(2\bar{t}^2 + 6\bar{t} + 3\right) z^{\bar{t}+2} - (\bar{t}+2)^2 z^{\bar{t}+1} + z + 1}{(1-z)^2}$$

$$\geq \frac{-(\bar{t}+2)^2 z^{\bar{t}+1} + z + 1}{(1-z)^2} \geq \frac{z + 1 - (2 + 2\log_{1+\lambda} n)^2/n^2}{(1-z)^2}$$

holds.[31] Hence, (64) holds. This completes the proof. $\qquad\square$

We now provide the lower bound on $\mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right]$ using Claim D.21, as follows. Let $\underline{k} = 2\log_{1+\lambda} n$. (Hence, $\underline{k} = \bar{t}$.)

$$\mathbb{E}_\pi\left[r_s^2 | r_s \neq \infty\right] \geq \frac{1}{n} \cdot \sum_{k=\underline{k}}^{n-1} \mathbb{E}\left[r_s^2 | p_s = k + 1\right]$$

$$\geq \frac{1}{n} \cdot \sum_{k=\underline{k}}^{n-1} \left(\frac{k/m + 1 - (2 + 2\log_{1+\lambda} n)^2/n^2}{(1 - k/m)^2} - \frac{(\bar{t}+1)^5}{2m}\right)$$

$$\geq \frac{1}{n} \cdot \left(\sum_{k=\underline{k}}^{n-1} \frac{2 - (1 - k/m)}{(1 - k/m)^2}\right) - \frac{(2 + 2\log_{1+\lambda} n)^2/n^2}{(1 - 1/(1+\lambda))^2} - \frac{(\bar{t}+1)^5}{2m}$$

$$= \frac{m^2}{n} \cdot \sum_{k=\underline{k}}^{n-1} \frac{2}{(m-k)^2} - \frac{m}{n} \cdot \sum_{k=\underline{k}}^{n-1} \frac{1}{m-k} - o(1)$$

$$\geq \frac{m^2}{n} \cdot \sum_{k=\underline{k}}^{n-1} \frac{2}{(m-k)^2} - \frac{m}{n} \cdot \left(\frac{1}{2m} + \log \frac{m}{m-n}\right) - o(1), \tag{65}$$

where $o(1)$ suppresses the term $\frac{(2+2\log_{1+\lambda} n)^2/n^2}{(1-1/(1+\lambda))^2} + \frac{(\bar{t}+1)^5}{2m}$, which approaches $0$ as $n$ approaches infinity. The last inequality uses the fact that $\gamma + \log i \leq H_i \leq \frac{1}{2i} + \gamma + \log i$, where $H_i$ is the $i$-th harmonic number and $\gamma$ is the Euler–Mascheroni constant (see, e.g., Villarino (2005)).

Next, we use the inequality $\frac{1}{x^2} \geq \frac{1}{x} - \frac{1}{x+1}$ to bound the summation in (65) from below:

$$\frac{m^2}{n} \cdot \sum_{k=\underline{k}}^{n-1} \frac{2}{(m-k)^2} \geq \frac{2m^2}{n} \cdot \left(\frac{1}{m-n+1} - \frac{1}{m-\underline{k}+1}\right) = \frac{2(1+\lambda)}{\lambda} - o(1),$$

where $o(1)$ suppresses a term that approaches $0$ as $n$ approaches infinity. Plugging the above

---

[31] We note that the we computed the equality using Mathematica 12. A direct derivation is possible by writing the left-hand side in terms of $\sum_j z^j$, $\sum_j jz^j$, and $\sum_j j^2 z^j$.

inequality into (65) implies that

$$\mathbb{E}_\pi \left[ r_s^2 | r_s \neq \infty \right] \geq \frac{2(1+\lambda)}{\lambda} - (1+\lambda)\log(1+1/\lambda) - o(1), \tag{66}$$

where $o(1)$ suppresses a term that approaches $0$ as $n$ approaches infinity.

**Claim D.22.** $\mathbb{E}\left[r_s | r_s \neq \infty\right] \leq (1+\lambda)\log(1+\frac{1}{\lambda}) + \frac{1}{2n}$.

*Proof.* The proof follows immediately from Lemma D.15, which shows that

$$\mathbb{E}_{\pi \sim U(\Pi_{\mathsf{STB}}^{n,m})} \left[\mathcal{A}r(\mu_\pi)\right] \leq \frac{m}{n}\left(\frac{1}{2m} + \log(\frac{m}{m-n})\right).$$

$\square$

Claim D.22 and (66) together conclude the proof of the lemma:

$$\mathbb{E}_\pi \left[\mathcal{S}i(\mu_\pi)\right] = \mathbb{E}\left[r_s^2 | r_s \neq \infty\right] - \mathbb{E}\left[r_s | r_s \neq \infty\right]^2$$

$$\geq \frac{2(1+\lambda)}{\lambda} - (1+\lambda)\log(1+1/\lambda) - (1+\lambda)^2 \log(1+\frac{1}{\lambda})^2 - o(1).$$

$\square$

*Proof of Theorem D.2.* The proof is a direct consequence of Lemmas D.18, D.19, and D.20. The first lemma provides an upper bound for expected social inequality under MTB in school-proposing DA. The second lemma shows that, under MTB, the expected social inequalities in school- and student-proposing DA are essentially the same, in the limit. The third lemma provides a lower bound for expected social inequality under STB in student-proposing DA. By these lemmas we have:

$$\lim_{n \to \infty} \frac{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{STB}})\right]}{\mathbb{E}\left[\mathcal{S}i(\mu_{\mathsf{MTB}})\right]} \geq \frac{\frac{2(1+\lambda)}{\lambda} - (1+\lambda)\log(1+1/\lambda) - (1+\lambda)^2 \log(1+\frac{1}{\lambda})^2}{\frac{8(1+\lambda)^2}{f(\log\frac{1+\lambda}{\lambda})^2} - \left((1+\lambda)\log\frac{1+\lambda}{\lambda}\right)^2}, \tag{67}$$

where $f(x) = \frac{e^x - 1 - x}{x(e^x - 1)}$. It is straightforward to verify that the right-hand side of the above inequality is always greater than 1 when $\lambda \leq 0.01$. We omit the algebra and instead demonstrate this by plotting the right-hand side in Figure 7. $\square$

# E  Alternative version of the main theorem

**Theorem E.1.** *Let $\lambda : \mathbb{N} \to \mathbb{N}$.*

(i) *Consider a simple school choice market with $m$ schools and $n = m + \lambda(m)$ students. Let the market profiles $\pi_{\mathsf{MTB}}, \pi_{\mathsf{STB}}$ be drawn independently and uniformly at random from $\Pi_{\mathsf{MTB}}^{n,m}, \Pi_{\mathsf{STB}}^{n,m}$,*

(a) $\lambda \in (0, 0.01)$
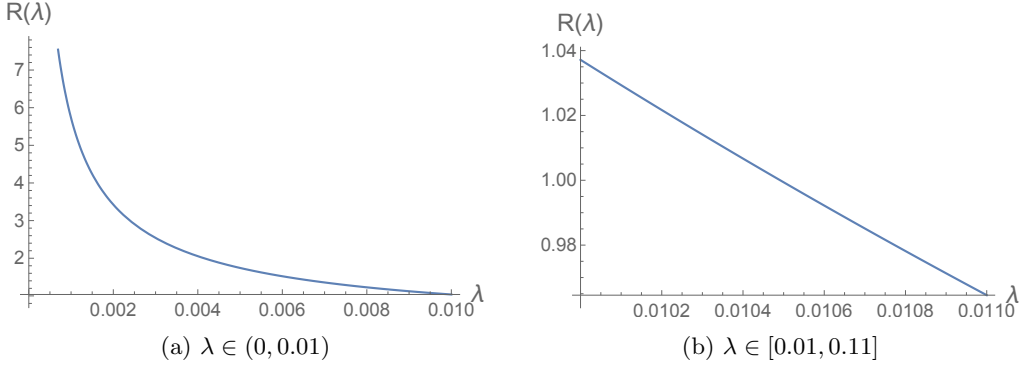
(b) $\lambda \in [0.01, 0.11]$

Figure 7: We plot the right-hand side of (67), namely $R(\lambda)$, for $\lambda \in (0, 0.1)$ and $\lambda \in [0.01, 0.011]$.

respectively. Then, for any $\epsilon > 0$, there exists $m_\epsilon > 0$ such that, for all $m > m_\epsilon$,

$$\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R_{\pi_{\mathsf{MTB}}}\right] \preceq^\epsilon \mathbb{E}_{\pi_{\mathsf{STB}}}\left[R_{\pi_{\mathsf{STB}}}\right].$$

(ii)   Consider a simple school choice market with $n$ students and $m = n + \lambda(n)$ schools. Let the market profiles $\pi_{\mathsf{MTB}}, \pi_{\mathsf{STB}}$ be drawn independently and uniformly at random from $\Pi^{n,m}_{\mathsf{MTB}}, \Pi^{n,m}_{\mathsf{STB}}$, respectively. Then, for any $\epsilon > 0$, there exists $n_\epsilon > 0$ such that, for all $n > n_\epsilon$,

$$\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R_{\pi_{\mathsf{MTB}}}\right] \not\succeq^\epsilon \mathbb{E}_{\pi_{\mathsf{STB}}}\left[R_{\pi_{\mathsf{STB}}}\right] \quad and \quad \mathbb{E}_{\pi_{\mathsf{STB}}}\left[R_{\pi_{\mathsf{STB}}}\right] \not\succeq^\epsilon \mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R_{\pi_{\mathsf{MTB}}}\right]$$

hold if there exists a constant $\gamma < 1$ such that $\lambda(n) = O(n^\gamma)$.

## E.1   Proof of Theorem E.1, part (i)

We use $\mathcal{M}_{\mathsf{MTB}}$ to denote the simple school choice market defined in the theorem statement, which has $m$ schools, $n = m + \lambda(m)$ students, and market profile $\pi_{\mathsf{MTB}}$. Similarly, we use $\mathcal{M}_{\mathsf{STB}}$ to denote the simple school choice market defined in the theorem statement, which has $m$ schools, $n = m + \lambda(m)$ students, and market profile $\pi_{\mathsf{STB}}$.

The proof follows the same steps as the proof for part(i) of Theorem C.1, with the difference that we provide the following counterparts for Lemmas C.7 and B.17.

**Lemma E.2.** Let $\bar{t} = 3 \log m$. For any constant $\alpha > 1$, there exists a constant $\beta_\alpha > 0$ (independent of $m$) such that the following holds:

$$\mathbb{E}_{\pi_{\mathsf{STB}}}\left[R^+_{\pi_{\mathsf{STB}}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right)\right] \geq m - \beta_\alpha \log m \cdot \log \log m.$$

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) We fix the common tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random

64

from the set of all strict linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Define $x = 2\alpha\bar{t}\log m$. Let $S'$ be the subset of students in $\mathcal{M}_{\mathsf{STB}}$ who receive priority numbers better (smaller) than $m - x$. First, we apply Lemma C.6 to each student in $S'$. Lemma C.6 implies that a student with priority number $m - x$ or better gets assigned to one of her top $\frac{m}{\alpha\bar{t}}$ choices with probability at least $1 - m^{-2}$. For every student $s$, let the binary random variable $z_s$ be equal to 1 iff $s$ is assigned to one of her top $\frac{m}{\alpha\bar{t}}$ choices. Linearity of expectation then implies that

$$\mathbb{E}_\varphi \left[ \sum_{s \in S'} z_s \right] \geq (1 - m^{-2}) \cdot |S'|. \tag{68}$$

To prove the bound in the lemma statement, we need to take the students in $S \backslash S'$ into account. Also, let $S'' \subset S \backslash S'$ denote the subset of students who have a priority number between $m - x$ and $m - \beta\bar{t} \cdot \log\log m$, where $\beta = 2\alpha^2\bar{t}/\log m$. Lemma C.5 implies that, for any $s \in S''$,

$$\mathbb{P}_\varphi [z_s = 0] \leq \left( 1 - \frac{\beta\bar{t} \cdot \log\log m}{m} \right)^{\frac{m}{\alpha\bar{t}}} \leq \exp\left( -\frac{\beta}{\alpha} \cdot \log\log m \right).$$

Having $\beta = 2\alpha^2\bar{t}/\log m$ implies that, for any $s \in S''$,

$$\mathbb{P}_\varphi [z_s = 0] \leq (\log m)^{-\frac{2\alpha\bar{t}}{\log m}}.$$

The above bound and linearity of expectation imply that

$$\mathbb{E}_\varphi \left[ \sum_{s \in S''} z_s \right] \geq |S''| \cdot (1 - (\log m)^{-\frac{2\alpha\bar{t}}{\log m}}). \tag{69}$$

Finally, (68) and (69) imply that

$$\mathbb{E}_\varphi \left[ \sum_{s \in S} z_s \right] \geq |S'| \cdot (1 - m^{-2}) + |S''| \cdot (1 - (\log m)^{-6})$$

$$= (m - x)(1 - m^{-2}) + (x - \beta\bar{t} \cdot \log\log m)(1 - (\log m)^{-6})$$

$$\geq m - x - \frac{1}{m} + x - \beta\bar{t} \cdot \log\log m - x(\log m)^{-6}$$

$$= m - 18\alpha^2 \log m \cdot \log\log m - 6\alpha(\log m)^{-4} - \frac{1}{m}$$

$$= m - 18\alpha^2 \log m \cdot \log\log m - o(1).$$

Observing that $\mathbb{E}_\varphi \left[ \sum_{s \in S} z_s \right] = \mathbb{E}_{\pi_{\mathsf{STB}}} \left[ R_{\pi_{\mathsf{STB}}}^+ \left( \lfloor \frac{m}{\alpha\bar{t}} \rfloor \right) \right]$ completes the proof. □

**Lemma E.3.** $\mathbb{E}_{\pi_{\mathsf{STB}}} \left[ R_{\pi_{\mathsf{STB}}}^+ (1) \right] = \frac{m+1}{2}$.

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism.

(Recall this equivalence relation from Section A.1.) We fix the common tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random from the set of all strict linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Suppose that, first, the student with the highest priority number chooses her favorite school. Then, the student with the next highest priority number chooses, and so on. We call the student with the $i$-th highest priority number student $i$. Let $X_i$ be a binary random variable which is 1 iff student $i$ is assigned to her first choice, and let $X = \sum_{i=1}^{m} X_i$. Observe that $\mathbb{P}_\varphi [X_i = 0] = (i-1)/m$ holds for $i \leq m$. Therefore, $\mathbb{E}_\varphi [X] = \sum_{i=1}^{m} \frac{m-i+1}{m} = \frac{m+1}{2}$. This proves the claim. $\qquad \square$

Now, we are ready to prove part(i) of Theorem E.1.

*Proof of Theorem E.1, part (i).* By Lemma C.3, $R^+_{\pi_{\mathsf{MTB}}} \left( \lfloor \frac{m}{\alpha \bar{t}} \rfloor \right) \leq 0.4m + m/\sqrt{\log m}$ holds with high probability as $m$ grows large, which implies that

$$\mathbb{E}_{\pi_{\mathsf{MTB}}} \left[ R^+_{\pi_{\mathsf{MTB}}} \left( \lfloor \frac{m}{\alpha \bar{t}} \rfloor \right) \right] \leq 0.4m + o(m).$$

On the other hand, by Lemma E.2, for every constant $\epsilon > 0$ there exists $\hat{m}_\epsilon$ such that, for all $m > \hat{m}_\epsilon$,

$$\mathbb{E}_{\pi_{\mathsf{STB}}} \left[ R^+_{\pi_{\mathsf{STB}}} \left( \lfloor \frac{m}{\alpha \bar{t}} \rfloor \right) \right] \geq m - (\log m)^{1+\epsilon}.$$

Moreover, by Lemma E.3,

$$\mathbb{E}_{\pi_{\mathsf{STB}}} \left[ R^+_{\pi_{\mathsf{STB}}}(1) \right] = \frac{m+1}{2}.$$

The above three bounds together imply that, for any $\epsilon > 0$, there exists $m_\epsilon$ such that

$$\mathbb{E}_{\pi_{\mathsf{MTB}}} \left[ R_{\pi_{\mathsf{MTB}}} \right] \preceq^\epsilon \mathbb{E}_{\pi_{\mathsf{STB}}} \left[ R_{\pi_{\mathsf{STB}}} \right]$$

holds for $m > m_\epsilon$. $\qquad \square$

## E.2  Proof of Theorem E.1, part (ii)

We use $\mathcal{M}_{\mathsf{MTB}}$ to denote the simple school choice market defined in the theorem statement, which has $n$ students, $m = n + \lambda(n)$ schools, and market profile $\pi_{\mathsf{MTB}}$. Similarly, we use $\mathcal{M}_{\mathsf{STB}}$ to denote the same market but with market profile $\pi_{\mathsf{STB}}$, as defined in the theorem statement.

The proof follows the same steps as the proof for part(ii) of Theorem C.1, with the difference that we provide the following counterpart for Lemma C.9.

**Lemma E.4.** *There exists $n_0 > 0$ such that for all $n > n_0$, in the student-optimal assignment in market $\mathcal{M}_{\mathsf{STB}}$, the expected number of students that are not assigned to one of their top $3\log^2 n$ choices is at least $\frac{n}{25\log^2 n}$.*

*Proof.* The proof uses the equivalence relation between the DA-STB and the RSD mechanism. (Recall this equivalence relation from Section A.1.) We fix the tie-breaker of the schools and let the preference order of every student be drawn independently and uniformly at random from the set of all strict linear orders over the schools. Let the random variable $\varphi$ denote the preference profile of the students. (Hence, $\varphi$ is drawn uniformly at random from $\Phi^{n,m}$.)

Define $x = 3\log^2 n$ and $t = \frac{n}{5\log^2 n}$. Let $s$ be a student with priority number $n - t + 1$, and let $X_s$ be a binary random variable which is 1 iff student $s$ is not assigned to one of her top $x$ choices. Applying Lemma C.8 implies that a student with a priority number $n - t + 1$ or worse (i.e., a larger number) is assigned to one of her top $x$ choices with probability at most

$$\frac{x(t + m - n)}{m - x + 1} = \frac{(3\log^2 n)(m - n + \frac{n}{5\log^2 n})}{m + 1 - 3\log^2 n} = \frac{3n/5 + 3(m - n)\log^2 n}{m + 1 - 3\log^2 n},$$

which is at most $4/5$ for sufficiently large $n$. That is, there exists $n_0 > 0$ such that for all $n > n_0$, the above expression is at most $4/5$. Hence, $\mathbb{P}[X_s = 1] \geq 1/5$ holds for sufficiently large $n$. Now, let $S_t$ denote the set of students with priority numbers $n - t + 1, \dots, n$. Linearity of expectation then implies that $\mathbb{E}_\varphi\left[\sum_{s \in S_t} X_s\right] \geq |S_t|/5$ holds for $n > n_0$. Observing that the right-hand side of the latter inequality is equal to $\frac{n}{25\log^2 n}$ concludes the proof. $\qquad\square$

*Proof of Theorem E.1, part (ii).* Fix $\epsilon > 0$. First, we show that $\mathbb{E}_{\pi_{\mathsf{MTB}}}[R_{\pi_{\mathsf{MTB}}}] \not\succeq^\epsilon \mathbb{E}_{\pi_{\mathsf{STB}}}[R_{\pi_{\mathsf{STB}}}]$. To this end, we recall a result in Pittel (1992) that shows there exists a constant $\gamma > 0$ such that, with probability at least $1 - n^{-\gamma}$, no student has a rank worse than $3\log^2 n$ in the student-optimal assignment in the market $\mathcal{M}_{\mathsf{MTB}}$. This implies that

$$\mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R^+_{\pi_{\mathsf{MTB}}}(3\log^2 n)\right] \geq n(1 - n^{-\gamma}) = n - \frac{n}{n^\gamma}. \tag{70}$$

On the other hand, Lemma E.4 shows that

$$\mathbb{E}_{\pi_{\mathsf{STB}}}\left[R^+_{\pi_{\mathsf{MTB}}}(3\log^2 n)\right] \leq n - \frac{n}{25\log^2 n}. \tag{71}$$

Finally, observe that (70) and (71) together imply that there exists $n_\epsilon \geq n_0$ such that for all $n > n_\epsilon$, $\mathbb{E}_{\pi_{\mathsf{MTB}}}[R_{\pi_{\mathsf{MTB}}}] \not\succeq^\epsilon \mathbb{E}_{\pi_{\mathsf{STB}}}[R_{\pi_{\mathsf{STB}}}]$.

It remains to show that there exists $n'_\epsilon > 0$ such that $\mathbb{E}_{\pi_{\mathsf{STB}}}[R_{\pi_{\mathsf{STB}}}] \not\succeq^\epsilon \mathbb{E}_{\pi_{\mathsf{MTB}}}[R_{\pi_{\mathsf{MTB}}}]$ holds for all $n > n'_\epsilon$. To this end, we recall the proof of Lemma B.17, which shows that $\mathbb{E}_{\pi_{\mathsf{STB}}}\left[R^+_{\pi_{\mathsf{STB}}}(1)\right] = \frac{n-1}{2}$. On the other hand, Proposition 3.1 of Ashlagi et al. (2019) shows that

$$\lim_{n \to \infty} \mathbb{E}_{\pi_{\mathsf{MTB}}}\left[R^+_{\pi_{\mathsf{MTB}}}(1)\right]/n = 0.$$

The latter two equations conclude the proof. □

# F  School choice markets with a discrete choice model

Consider a school choice market with $n$ students and $m$ schools with identical capacities. Students' preferences are based on a symmetric multinomial-logit discrete choice model, defined as follows. Each school $c$ has a quality factor $w_c > 0$, and each student's preference list is generated as follows. A student's first choice is drawn proportionally to the schools' quality factors. That is, her top choice is school $c$ with probability $\frac{w_c}{\sum_{c'} w_{c'}}$. Her second choice is drawn in a similar way after her top choice is removed. That is, conditional on her top choice being school $c_1$, her second top choice is school $c$ with probability $\frac{w_c}{\sum_{c' \neq c_1} w_{c'}}$. Her third choice is drawn in a similar way after her two top choices are removed, and so on.

Given one realization of the students' preferences drawn from the above discrete choice model, let $p_1(c)$ denote the number of students who list school $c$ as their top choice. Then, $p_1(c)$ is an unbiased estimator for $w_c$. We note that every school $c$ with $\alpha_c = \frac{p_1(c)}{q_c} \geq 1$ is assigned as many students as its capacity, $q_c$, in the student-proposing DA algorithm.

# G  Further computational experiments

## G.1  A hybrid tie-breaking rule

We present computational experiments regarding a hybrid tie-breaking rule that uses a common lottery in popular schools and independent lotteries in non-popular schools.

### G.1.1  A hybrid tie-breaking rule in a two-tier model

Consider a simple school choice market as described in Section 2.1. That is, some schools are considered as *top* (popular) schools and others are considered as *bottom* (non-popular) schools. Every student prefers every top school to every bottom school, and preferences within a tier are drawn independently and uniformly at random. Three tie-breaking rules are considered: STB, MTB, and Hybrid Tie-Breaking rule (HTB), in which all top schools use a common priority order as their tie-breaker which is drawn independently and uniformly at random, and every bottom school uses a tie-breaker that is drawn independently and uniformly at random.

The markets we consider have 1000 students and 1000 schools, 100 of which are top schools. For each of the three tie-breaking rules, we draw 50 independent samples, compute the assignment generated by the student-proposing DA in each sample, and then compute the cumulative rank distribution in that assignment. Finally, we compute and report the average cumulative rank distribution, with the average taken over the 50 samples (as in Section 4). Figure 8 reports the average cumulative rank distribution under the three tie-breaking rules. Observe that the distributions under STB and HTB coincide in top schools (i.e., schools ranked 100 or better) and perform better

68

than the one under MTB. On the other hand, the distributions under MTB and HTB coincide in bottom schools (i.e., schools ranked 101 or worse). Note that there is no rank-wise dominance in this range.

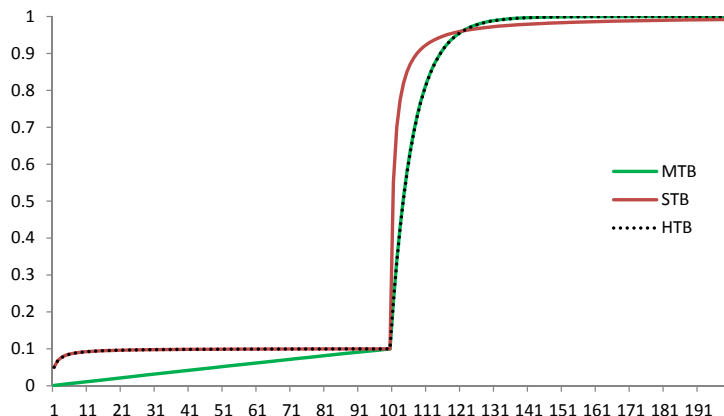So HTB inherits the properties of STB in top schools and those of MTB in non-popular schools.



Figure 8: Cumulative rank distributions under MTB, STB, and HTB. The horizontal axis represents rank and the vertical axis fraction of students.

### G.1.2 Experimenting with a hybrid rule using NYC data

In NYC the market is not perfectly tiered as in the previous experiment. Therefore, to test the hybrid tie-breaking rule (HTB), we heuristically select a set of popular schools.[32] For each popularity threshold $\alpha$, we define the set of popular schools as the set of schools with a popularity threshold higher than $\alpha$.

Table 4 reports the (average) number of students assigned to each rank under each tie-breaking rule for different popularity thresholds. First, we fix a popularity threshold $\alpha \in \{1, 1.5, 2\}$. Then, for each tie-breaking rule we draw 50 independent samples, run the student-proposing DA in each sample, and compute the number of students assigned to each rank. Then, for each rank $r$ we compute the average number of students assigned to a school ranked $r$ or better on his or her list with the average taken over the 50 samples. (Since the students are allowed to rank up to 12 schools, $r \in \{1, \ldots, 12\}$.)

Observe that none of the tie-breaking rules rank-wise dominates the other two. When we measure the number of assigned students, MTB performs better than HTB, and HTB performs better than STB. However, when we measure how many students are assigned to their first choice, STB performs better than HTB, and HTB better than MTB. Naturally, the lower the popularity threshold $\alpha$, the "closer" the rank distribution under HTB is to the rank distribution under STB.

To understand the effect of each tie-breaking rule within each tier, we plot the rank distribution in popular and non-popular schools for a fixed popularity threshold $\alpha = 2$. Figures 9a and 9b

---

[32]A careful study on classifying schools based on their popularity is an essential prerequisite of using the hybrid rule in practice.

| $\alpha$ | Tie-breaking | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | STB | 36670 | 50976 | 59117 | 64059 | 67141 | 69066 | 70314 | 71057 | 71570 | 71892 | 72093 | 72234 |
| | HTB | 35234 | 49927 | 58588 | 63889 | 67175 | 69197 | 70491 | 71245 | 71758 | 72070 | 72263 | 72392 |
| | MTB | 32379 | 48076 | 57778 | 63715 | 67347 | 69510 | 70844 | 71605 | 72097 | 72384 | 72552 | 72667 |
| 1.5 | STB | 36675 | 50987 | 59133 | 64069 | 67157 | 69086 | 70333 | 71080 | 71595 | 71913 | 72109 | 72254 |
| | HTB | 34196 | 49272 | 58302 | 63825 | 67232 | 69307 | 70600 | 71351 | 71851 | 72157 | 72340 | 72462 |
| | MTB | 32389 | 48094 | 57784 | 63728 | 67343 | 69501 | 70833 | 71592 | 72084 | 72370 | 72538 | 72654 |
| 2 | STB | 36672 | 50983 | 59124 | 64072 | 67154 | 69077 | 70325 | 71069 | 71582 | 71903 | 72100 | 72243 |
| | HTB | 33283 | 48716 | 58065 | 63787 | 67291 | 69400 | 70715 | 71475 | 71968 | 72263 | 72437 | 72556 |
| | MTB | 32381 | 48069 | 57765 | 63715 | 67341 | 69503 | 70837 | 71601 | 72094 | 72381 | 72548 | 72660 |

Table 4: The average cumulative rank distribution under each tie-breaking rule. The numbers $1, \ldots, 12$ in the table header correspond to ranks. The digits after the decimal points are truncated.

report the average cumulative rank distributions. Observe that the rank distribution in popular schools under HTB rank-wise dominates the one under MTB, whereas these rank distributions almost coincide in non-popular schools.
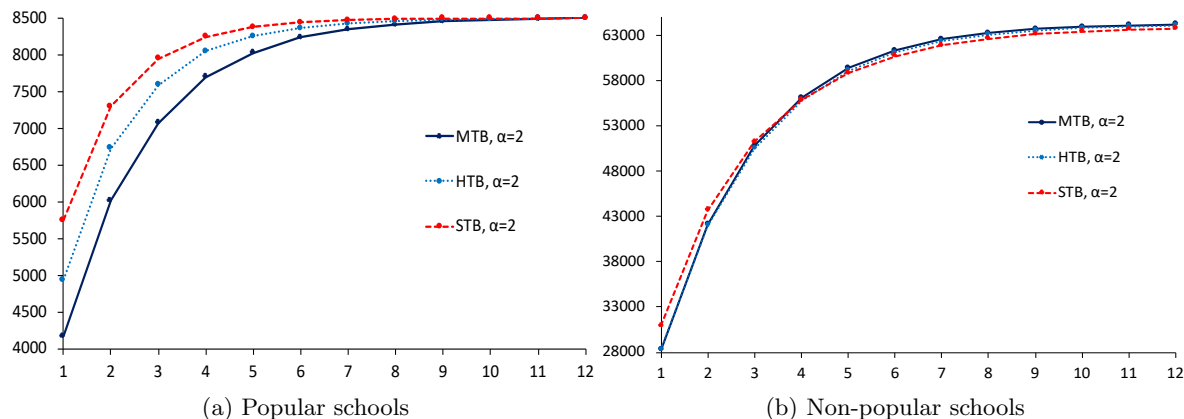


(a) Popular schools  (b) Non-popular schools

Figure 9: Students' rank distributions in popular schools (left) and non-popular schools (right) under STB, MTB, and HTB with popularity threshold $\alpha = 2$. The horizontal axis represents the rank and the vertical axis the number of students.

## G.2    Pareto-improving pairs

We ran experiments to measure the average number of students involved in Pareto-improving (PI) pairs. These are pairs of students who both would be better off if they swap their seats. Each statistic is calculated by taking an average over 50 samples for each tie-breaking rule. We find no PI pairs under STB in any of these samples. We report the results about MTB and HTB below.

A student is a *PI student* if she is involved in at least one PI pair. Figure 10 reports the fraction of students involved in a PI pair for various $\alpha$'s under MTB and HTB. The following three statistics are reported for each $\alpha$: (i) *overall*: the fraction of PI students, (ii) *in popular schools*: the fraction of PI students in popular schools, i.e., the number of PI students in popular schools divided by the

total number of students assigned to popular schools, and (iii) *in non-popular schools*: the fraction of PI students in non-popular schools.

Under MTB (Figure 10a), there exists a clear gap between the fraction of PI students in popular and non-popular schools. At $\alpha = 1$, these fractions are 20% and 1%, respectively. At $\alpha = 2$, this gap naturally shrinks, because fewer schools are categorized as popular as $\alpha$ goes up. Under HTB (Figure 10b), there are no PI students in popular schools. However, there are such students in non-popular schools, as expected, since MTB is used in non-popular schools. The fraction of PI students increases with the popularity threshold, because more schools are categorized as non-popular (and run independent lotteries) at higher values of $\alpha$.
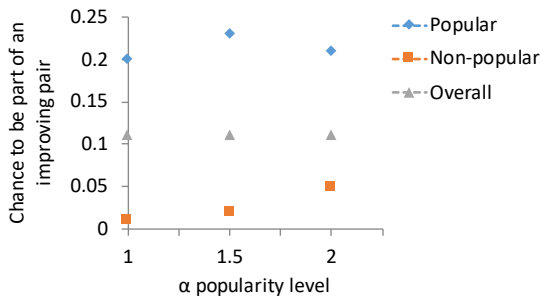


(a) Fraction of the PI students under MTB.
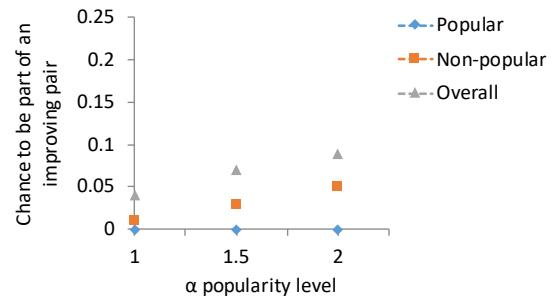


(b) Fraction of the PI students under HTB.

Figure 10: The horizontal axes represent popularity thresholds and the vertical axes the fraction of PI students.