Electronic Companion to: What matters in school choice tie-breaking? How competition guides design

This electronic companion provides the proofs for the theoretical results (Sections 1-5) and stylized simulations for robustness checks (Section 6).

1 Roadmap for proofs

We break Theorem 3.2 into several smaller (sub-)theorems and prove each one separately. This is done in sections 1.1, 1.2, and 1.3, which state and discuss the theorems about stochastic dominance, Pareto improving pairs, and variance, respectively. The theorems stated in these sections are then proved separately in the later sections. Section 2 contains the proofs for stochastic dominance. Section 3 contains some preliminary results that will be used in Sections 4 and 5, which contain the proofs about Pareto improving pairs and variance, respectively.

1.1 Stochastic dominance

We focus on the over-demanded case in the next theorem, and on the under-demanded market in the theorem that comes after that.

Theorem 1.1. *Consider a sequence of random school choice problems with n students and m schools where $n = m + 1$. Then, with high probability, $\mathcal{R}_{\mathsf{STB}}$ almost stochastically dominates $\mathcal{R}_{\mathsf{MTB}}$.*[1]

Theorem 1.2. *Consider a sequence of random school choice problems with n students and m schools with $n = m - 1$. Then, with very high probability (wvhp), $\mathcal{R}_{\mathsf{STB}}$ does not almost stochastically*

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

dominate $\mathcal{R}_{\mathsf{MTB}}$.[2] *Furthermore, wvhp $\mathcal{R}_{\mathsf{STB}}$ does not stochastically dominate $\mathcal{R}_{\mathsf{MTB}}[k]$ for any $k = o(n/ln^2 n)$, where $\mathcal{R}_{\mathsf{MTB}}[k]$ is the rank distribution resulting from the removal of the bottom k students from $\mathcal{R}_{\mathsf{MTB}}$.*[3]

The proofs for both theorems are given in Section 2. The same proofs imply that the theorem holds when the imbalance is larger than one. (i.e. the theorems hold for $n > m$ and $m < n$ respectively in the over-demanded and under-demanded cases)

1.2 Pareto improving pairs

In the next theorem, we show that deferred acceptance paired with MTB generates many Pareto improving pairs when there is a shortage of seats, and very few Pareto improving pairs when there is a surplus of seats. The proof is given in Section 4.

Theorem 1.3. *Consider a sequence of random school choice problems with n students and m schools, let $\mu = \mu_{\mathsf{MTB}}$, and let s be an arbitrary student.*

1. *If $n > m$,*

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

2. *If $n < m$,*

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

1.3 Variance

The next lemma says that, under either tie-breaking rule, if students' preference are i.i.d, then the expected social inequity is equal to the expected variance of the rank of an arbitrarily fixed student. The proof appears in Appendix 5.1.

---

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

[3] For any two functions $f, g : \mathbb{Z}_+ \to \mathbb{R}_+$ we adopt the notation $g = o(f)$ when $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$, $g = O(f)$ when $f \neq o(g)$, $g = \Theta(f)$ when $f = O(g)$ and $g = O(f)$, and finally, $g = \Omega(f)$ when $f = O(g)$.

Lemma 1.4. *For any student $s \in S$*

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

where expectation on the left-hand side is taken over all students' preferences except s, and expectation on the right-hand side is taken over all students' and schools' preferences with schools' preferences generated by either the STB or the MTB rule.

The lemma further shows that the expected variance of rank of a student s is equal to the expected variance of rank of a student s', for any $s, s' \in S$. Therefore, we sometimes refer to this notion as the *variance of student rank*, without specifying s.

The next theorem shows that the imbalance in the market determines whether MTB or STB results in a larger variance.

Theorem 1.5. *Consider a sequence of random school choice problems with n students and m schools.*

1. *If $n = m$ or $n = m - 1$, then $\lim_{n\to\infty} \frac{\mathit{Var}[\mu_{\mathsf{STB}}]}{\mathit{Var}[\mu_{\mathsf{MTB}}]} = \infty$.*
2. *If $n = m + 1$, then $\lim_{n\to\infty} \frac{\mathit{Var}[\mu_{\mathsf{STB}}]}{\mathit{Var}[\mu_{\mathsf{MTB}}]} = 0$.*[4]

Theorem 1.5 follows directly from the next result, which quantifies the social inequities in our model.

Lemma 1.6. *Consider a sequence of random school choice problems with n students and m schools.*

1. *If $n = m + 1$, the expected social inequity under MTB is $\Omega(\frac{n^2}{\log^2 n})$ and under STB is $\Theta(n)$.*
2. *If $n = m$, the expected social inequity under MTB is $O(\log^4 n)$, and under STB is $\Theta(n)$.*
3. *If $n = m - 1$, the expected social inequity under MTB is $O(\log^2 n)$ and under STB is $\Theta(n)$.*

The proof for Lemma 1.6 is given in Appendix 5.

We briefly discuss how our results on variance are affected by varying the size of the imbalance, length of preference lists, and correlation in preferences. Before this, we note that our empirical findings using NYC data support the theoretical findings (see Section 4).

[4] Expectations are taken over students' preferences and the tie-breaking lotteries.

In an over-demanded market with m schools, for any $n > m + 1$, the variance under STB remains the same as when $n = m+1$; however, the variance under MTB remains at least as high as in the case $n = m + 1$ due to the harsher competition (this is implied by the proof of the first part of Lemma 1.6). Thus, part 2 of Theorem 1.5 always holds as long as $n > m$. In an under-demanded market with n students, one should expect the variance under STB to decrease as the surplus of seats grows larger, since an increasing number of students will be assigned to their top choices. Nevertheless, we show in the next theorem that the variance under STB remains strictly larger than the variance under MTB, even when the surplus of seats is of the same order as the number of students. (The proof is given in Appendix 5.3.)

Theorem 1.7. *Suppose $m = n + \lambda n$ for any positive $\lambda \leq 0.01$. Then, $\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{S}i(\mu_{\mathsf{MTB}})]} > 1$, where the expectations are taken over preferences and the tie-breaking rules.*

We conjecture that this theorem holds for any positive fixed λ. To avoid unnecessary technicalities, we only prove it for $\lambda \leq 0.01$. We quantify the ratio between social inequities for different values of λ in our computational experiments in the Online Appendix. (For instance, we see that this ratio is around 3 for $\lambda = 0.1$.)

In another set of experiments (see the Online Appendix) we show that the gap between the variances persists even when the preference lists are short. To test how correlation in preferences affects our results, we conduct experiments (see the Online Appendix), in which students' preferences are drawn independently from a discrete choice model (one may think of these preferences as drawn proportionally with respect to publicly known schools' qualities). We see that in an under-demanded market, the variance under STB is larger than the variance under MTB, unless students' preferences are extremely correlated, in which case the rank distributions will become similar.

2 Proofs for Section 1.1

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

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

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

For any rank distribution $\mathcal{R}$, let $\mathcal{R}^+$ denote the corresponding cumulative rank distribution, i.e. $\mathcal{R}^+(k) = \sum_{i=1}^{k} \mathcal{R}(i)$ is the number of students who are assigned to one of their top k choices under $\mathcal{R}$.

2.1 Proof of Theorem 1.1

We need the following lemmas before proving this theorem.

2.1.1 Computing $\mathcal{R}_{\mathsf{MTB}}$

Lemma 2.1. *When $n = m + 1$, wvhp there at most $\frac{3n \log n}{t}$ students who receive more than t proposals in the school-proposing DA.*

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

Definition 2.2. *Let $\bar{t} = 3\theta \log m$, where $\theta > 1$ is a large constant that we set later.*

Proposition 2.3. *At most n/θ students receive more than $\bar{t}$ offers in school-proposing DA wvhp. This is a direct consequence of 2.1.*

Lemma 2.4. *Suppose a student s receive t proposals in the school-proposing DA such that $1 \leq t \leq \bar{t}$. Then, for any constant $\alpha > 2$*

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

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

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

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

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

Lemma 2.5. *For any constant $\alpha > 4$, $\mathcal{R}^{+}_{\mathsf{MTB}}\left(\lfloor\frac{m}{\alpha \bar{t}}\rfloor\right) \leq 0.4n + o(n)$, wvhp.*

Proof. To compute $\mathcal{R}_{\mathsf{MTB}}$, first we run the school-proposing DA and prove the lemma statement for the school-optimal matching. Then, using the fact that almost every student has the same match in the student-optimal matching Ashlagi et al. (2017), we establish the lemma statement (which holds for the student-optimal matching).

For any student s, let x_s be a binary random variable that is 1 iff $\eta^{\#}(s) > \frac{m}{\alpha \bar{t}}$. Also, let S' denote the subset of students who received at least one but no more than $\bar{t}$ offers. For any $s \in S'$, Lemma 2.4 implies

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

since $\alpha > 4$. This means $\mathbb{E}\left[\sum_{s\in S'} x_s\right] \geq e^{-1/2}\cdot|S'| \geq 0.606m$. Now, applying a standard Chernoff concentration bound implies that wvhp $\sum_{s\in S'} x_s \geq 0.6n$. This fact, together with the fact that $|S\backslash S'| = o(n)$ (which holds by 2.3, there are at most $0.4n+o(n)$ students s for whom $\eta^{\#}(s) \leq \frac{m}{\alpha \bar{t}}$.

It is straight-forward to imply a similar result for the student-optimal matching, μ. Note that the number of students who have different matches in μ and η is at most $n/\sqrt{\log n}$, wvhp Ashlagi et al. (2017). Consequently, there are at most $0.4n + o(n) + n/\sqrt{\log n}$ students s for whom $\mu^{\#}(s) \leq \frac{m}{\alpha \bar{t}}$. □

2.1.2 Computing $\mathcal{R}_{\mathsf{STB}}$

Lemma 2.6. *Suppose student $s \in S$ has priority number $n-x$. Then, the probability that s is not assigned to one of her top i choices is at most $(1-\frac{x}{n})^i$*

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

Lemma 2.7. *A student s who has priority number $n-x$ is assigned to one her top $\frac{2n\log(n)}{x}$ choices with probability at least $1-1/n^2$.*

Proof. Set $i=\frac{2n}{x}\log(n)$ and apply Lemma 2.6. Noting that $(1-\frac{x}{n})^i \le e^{-\frac{xi}{n}}$ proves the claim. $\square$

Lemma 2.8. *For any positive constant $\alpha > 1$, $\mathcal{R}^+_{\mathsf{STB}}\left(\lfloor \frac{m}{\alpha\bar{t}} \rfloor\right) \ge n - O(\log n \cdot \log\log n)$*

Proof. Define $x=\frac{\alpha\bar{t}2n\log n}{m}$. Let S' be the subset of students who have priority numbers better than $n-x$. First, we apply Lemma 2.7 on each student in S'. Lemma 2.7 implies that a student with priority number $n-x$ or better, gets assigned to one of her top $\frac{m}{\alpha\bar{t}}$ choices with probability at least $1-n^{-2}$. Taking a union bound over all students with priority number no worse than $n-x$, implies that at least $n-x$ students are assigned to one of their top $\frac{m}{\alpha\bar{t}}$ choices, with probability at least $1-1/n$. This means $\mathcal{R}^+_{\mathsf{STB}}\left(\frac{m}{\alpha\bar{t}}\right) \ge n-x = n-O(\log^2 n)$ holds with probability at least $1-1/n$. To prove the sharper bound in the lemma statement, we need to take the students in $S\backslash S'$ into account.

Let $S'' \subset S\backslash S'$ denote the subset of students who have priority number between $n-\beta\bar{t}\cdot\log\log n$ and $n-x$, where $\beta = 2\alpha^2\bar{t}/\log n$. Lemma 2.6 implies that for any $s\in S''$,

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

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

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

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

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

where in the last inequality we have used the fact that $x=\frac{\alpha\bar{t}2n\log n}{m}=O(\log^2 n)$.[5]

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

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

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

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

Lemma 2.9. *Let $\epsilon > 0$ be an arbitrary constant. Then, wvhp, at least $(1-\epsilon)n/2$ students are assigned to their top choice in STB.*

Proof. Consider the following implementation of STB. 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 $X = \sum_{i=1}^n X_i$. Observe that $\mathbb{P}[X_i = 1] = (i-1)/n$. Therefore, $\mathbb{E}[X] = \sum_{i=1}^n \frac{i-1}{n} = \frac{n-1}{2}$. A standard application of Chernoff bound then implies that for any $\epsilon > 0$, we have

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

This proves the claim. □

Now, we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. Lemma 2.5 says that $\mathcal{R}^+_{\mathsf{MTB}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \leq 0.4n + o(n)$ wvhp. This and Lemma 2.9 together imply that $\mathcal{R}^+_{\mathsf{MTB}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) < \mathcal{R}_{\mathsf{STB}}(1)$ wvhp. On the other hand, Lemma 2.8 says that $\mathcal{R}^+_{\mathsf{STB}}\left(\lfloor \frac{m}{\alpha \bar{t}} \rfloor\right) \geq n - (\log n)^{1+\epsilon}$ with high probability. The two latter facts, by definition, imply that $\mathcal{R}_{\mathsf{STB}}$ almost stochastically dominates $\mathcal{R}_{\mathsf{MTB}}$. □

2.2 Proof of Theorem 1.2

Lemma 2.10. *When $n = m - 1$, at least $\frac{n(1-\epsilon)}{16 \log^2 n}$ students are not assigned to one of their top $3 \log^2 n$ choices in STB, wvhp, for any $\epsilon > 0$.*

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

Applying Lemma 2.11 implies that any student with priority number below $n-t$ is assigned to one of her top x choices with probability at most 3/4; in other words, it implies $\mathbb{P}[X_s=1]\geq 1/4$. Now, let S_t denote the set of students with lowest t priority numbers in STB. A standard application of Chernoff bound implies that $\sum_{s\in S_t} X_s \geq |S_t|(1-\epsilon)/4$, wvhp, for any $\epsilon>0$. $\square$

Lemma 2.11. *A student with priority number $n-t$ in STB is assigned to one of her top x choices with probability at most $\frac{tx}{n-t+1}$.*

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

3 Preliminary findings: concentration lemmas

Lemma 3.1. *Suppose $n=m+1$, and fix a student s. Then, under MTB, in the school-proposing DA, the number of offers received by s is whp at most $(1+\epsilon)\log n$ for any constant $\epsilon>0$.*

Proof. The proof idea is defining another stochastic process that we denote by $\mathcal{B}$. Process $\mathcal{B}$ is defined by a sequence of binary random variables $X_1,\ldots,X_k$, where $k=(1-\delta)n\log n$ for some arbitrary small constant $\delta>0$. Each random variable in this sequence takes the value 1 with probability $\frac{1}{n-3\log^2 n}$, and 0 otherwise. For convenience, we also refer to these random variables by *coins*, and the process that determines the value of a random variable by *coin-flip*.

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

The problem with this argument is that the proposal-making processes of schools are not independent of each other, and we have to account for the dependencies. We have to define a new random process, $\mathcal{B}$, which is a simple coin-flipping process: it flips a number of coins independently, all with success probabilities $\frac{1}{n-3\log^2 n}$. Then, we define a new random process $(\mathrm{DA}, \mathcal{B})$, which is a coupling of the random processes $\mathrm{DA}, \mathcal{B}$. The coupled process would have two components, one for each of the original random processes. Each component behaves (statistically) identical to its corresponding original process, but there is no restriction on the joint behavior of the components. It is straight-forward to define a simple coupling in which in almost all sample paths (i.e. wvhp), the number of successful coin flips is an upper bound on the number of proposals made to s. Whenever a school wants to make a proposal during the DA, process $\mathcal{B}$ flips the next coin. Then:

1. If c has made a proposal to s before, ignore the coin flip, and let c pick a school uniformly at random from the set of students whom it has not proposed to yet.

2. If c has made a proposal to s before, then let c make a proposal to the rest of the students that she has not proposed to yet, uniformly at random.

   (a) Suppose c has made $d \leq \log^2 n$ proposals so far. (Otherwise, ignore this sample path)

   (b) With probability $\frac{n-3\log^2 n}{n-d}$, let c make a proposal to s, otherwise, let c make a proposal to the rest of the students that she has not proposed to yet, uniformly at random.

It is straight-forward to verify that this defines a valid coupling of $\mathrm{DA}, \mathcal{B}$. Now, note that the total number of successful coin flips in $\mathcal{B}$ is an upper bound on the total number of proposals made to s in the coupled DA process, in almost all sample paths (i.e. wvhp). Therefore, we can apply the argument that we mentioned in the beginning of the proof to conclude the lemma. $\square$

Lemma 3.2. *Suppose $m = n + \lambda n$. Then, for any positive constant ϵ, the number of proposals received by a fixed student in the school-proposing DA is wvhp at least $(1-\epsilon)\kappa$, where $\kappa = \frac{n}{2(1+K)} + \frac{\lambda n}{2}$ and $K = (1+\lambda)\log(1+1/\lambda)$.*

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

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

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

Claim 3.3. *In $(DA, \mathcal{B})$, wvhp we have $d_s \geq X$.*

Claim 3.4. *For any constant $\delta' > 0$, $X \geq (1-\delta')(1-\delta)\kappa$ holds wvhp.*

The proofs of these claims are stated after the proof of the lemma. First, we verify that if we are given a valid coupling and the above claims, then proof of the lemma is almost complete: In Claim 3.4, we show that for any constant $\delta' > 0$,the inequality

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

holds wvhp. Therefore by Claim 3.3, $d_s \geq (1-\epsilon)\kappa$ holds wvhp for any constant $\epsilon > 0$.

To complete the proof, it remains to define our coupling. As mentioned before, this involves defining a new process, $(\mathrm{DA}, \mathcal{B})$, which is in fact a coupling of the processes $\mathrm{DA}, \mathcal{B}$. First, we define the coupling formally, and after that we prove Claim 3.3.

Definition of the Coupling

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

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

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

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

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

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

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

Proof of Claim 3.3. For any school c who has made a proposal to s, there is at most one successful coin-flip corresponding to c. This holds since

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

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

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

Proof of Claim 3.4. First, we show that wvhp $(\mathrm{DA}, \mathcal{B})$ terminates with no coins left. To see this, note that in $(\mathrm{DA}, \mathcal{B})$, the number of proposals that are made is at most equal to the number of flipped or dismissed coins. On the other hand, by the results of Ashlagi et al. (2017), the number of proposals made by the school-proposing DA is at least $k = (\frac{n^2}{1+K} + \lambda n^2)(1-\delta)$, wvhp (To see why, note that the number of proposals made by empty schools and the number of proposals made by non-empty schools respectively are at least $\lambda n^2(1-\delta)$ and $(\frac{n^2}{1+K})(1-\delta)$, wvhp). Since $\mathcal{B}$ starts with k coins, then, wvhp, $(\mathrm{DA}, \mathcal{B})$ ends when there are no coins left.

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

Lemma 3.5. *Suppose $n = m + 1$. Then, for any positive constant ϵ, the number of proposals received by a fixed school in the student-proposing DA is wvhp at least $(1-\epsilon)\kappa$, where $\kappa = \frac{n}{2\log n}$.*

Proof. The proof is similar to the proof of Lemma 3.2. The only adjustments are swapping the roles of schools and students and using the new definition of κ stated in this lemma. $\square$

Lemma 3.6. *Suppose $n = m + 1$. Fix an arbitrary small constant $\epsilon > 0$. Then, in the school-proposing DA, the number of proposals received by a fixed student in the school-proposing algorithm is whp at least $(1-\epsilon) \cdot \kappa$, where $\kappa = \frac{\log n}{2}$.*

Proof. The proof is similar to our proof for Lemma 3.2, with the exception that we should use the new definition of κ that we state in this lemma. $\square$

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

Definition 3.7. *Let* $\underline{r}, \overline{r}$ *respectively denote* $n/(\log n)^{1+\epsilon}, n/(\log n)^{1-\epsilon}$.

Lemma 3.8. *Suppose* $n = m + 1$ *and fix a student* $s \in S$*. Then, for any constant* $\epsilon > 0$ *we have*

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

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

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

Ashlagi et al. (2017) show that $\mathbb{P}\left[\mu(s) \neq \eta(s)\right] \leq \frac{\sqrt{\log n}}{n}$. Therefore it is sufficient to show that (1) holds.

We use Lemma 3.6 to prove (1). Let d denote the number of proposals received by s. Lemma 3.6 implies that

$$\mathbb{P}\left[d < \underline{\alpha} \log n\right] = o(1), \tag{2}$$

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

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

$$\begin{aligned}\mathbb{P}\left[\eta^{\#}(s) \geq \overline{r}\right] &= \prod_{i=1}^{d} 1 - \frac{\overline{r}}{m - i + 1} \\ &\leq \left(1 - \frac{\overline{r}}{m}\right)^{d} \leq e^{-\frac{d\overline{r}}{m}} \leq \mathsf{exp}(-\underline{\alpha}(\log n)^{\epsilon}) = o(1).\end{aligned} \tag{3}$$

In the other hand, we have

$$
\begin{aligned}
\mathbb{P}\left[\eta^{\#}(s) \leq \underline{r}\right] &= 1 - \mathbb{P}\left[\eta^{\#}(s) > \underline{r}\right] \\
&\leq 1 - \prod_{i=1}^{d} 1 - \frac{\underline{r}}{m-i+1} \\
&\leq 1 - \left(1 - \frac{\underline{r}}{m-d}\right)^{d} \leq 1 - \left(1 - \frac{d\underline{r}}{m-d}\right) \leq O\left(\frac{\overline{\alpha}}{(\log n)^{\epsilon}}\right) = o(1). \qquad (4)
\end{aligned}
$$

Taking a union bound over the bounds (2), (3), and (4) completes the proof.

□

4 Proofs for Section 1.2

Consider a matching γ. Let $\gamma(x)$ be the agent to which x is matched, and for any subset of agents $A \subseteq S \cup C$, let $\gamma(A)$ be the set of agents matched to agents in A. Therefore, $\gamma(C)$ is the set of students who are assigned under γ.

Theorem 4.1. *Suppose $n = m + 1$ and fix a student $s \in S$. Then, under MTB, we have*

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

for any constant $\epsilon > 0$.

Next, we will define a random variable $\Pi(s)$, which we will use in the proof of Theorem 4.1. Recall that $\underline{r}, \overline{r}$ respectively denote $n/(\log n)^{1+\epsilon}, n/(\log n)^{1-\epsilon}$. For a fixed student s, we will define the random variable $\Pi(s)$, which represent a preference profile that is constructed by fixing the the interval $[\underline{r}, \overline{r}]$ of the preference list of s, while letting the rest of the preference profile be constructed randomly. This notion is formally defined below.

Definition 4.2. *For a fixed student s, we define a random variable $\Pi(s)$, which is a subset of preference profiles. We define $\Pi(s)$ by constructing it, this would implicitly define the corresponding support and probability mass function (PMF); we denote the PMF by $\mathcal{P}(s)$. We define $\Pi(s)$ by first defining a* partial preference profile $\hat{\pi}$*, as follows:*

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

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

$\Pi(s)$ *contains the set of all preference profiles π who are consistent with $\hat{\pi}$ (i.e. agree with $\hat{\pi}$ on the positions where $\hat{\pi}$ is defined). Given a realization $\Pi(s)$, let $\mathcal{U}(\Pi(s))$ denote the uniform distribution over the elements of $\Pi(s)$.*

Lemma 4.3. *Suppose $\Pi(s) \sim \mathcal{P}(s)$. Also, suppose π, π' are preference profiles that are drawn independently uniformly at random from $\Pi(s)$. Then, whp $\mu_\pi = \mu_{\pi'}$. (i.e., almost all student-optimal matchings in $\Pi(s)$ are identical, whp)*

Proof. By definition, π, π' are selected so that they are identical everywhere except on a fixed student, namely s. So, π, π' coincide on the interval $[\underline{r}, \overline{r}]$ of the preference list of s, but they are constructed independently (and uniformly at random) everywhere else in the preference list of s. (In other words, the schools listed in the interval $[\underline{r}, \overline{r}]$ of $\pi'(s)$ are the same as $\pi(s)$, but in all other schools in $\pi'(s)$ are shuffled randomly)

Using lemma 3.8 and a simple union bound we obtain that

$$\begin{aligned}
&\mathbb{P}\left[\mu_\pi^{\#}(s) \notin [\underline{r}, \overline{r}] \bigvee \mu_{\pi'}^{\#}(s) \notin [\underline{r}, \overline{r}]\right] \\
\leq &\mathbb{P}\left[\mu_\pi^{\#}(s) \notin [\underline{r}, \overline{r}]\right] + \mathbb{P}\left[\mu_{\pi'}^{\#}(s) \notin [\underline{r}, \overline{r}]\right] = o(1). \qquad (5)
\end{aligned}$$

The preference list of each student $s' \neq s$ is the same in π, π'; also, whp, $\mu_\pi(s), \mu_{\pi'}(s)$ are both in the interval $[\underline{r}, \overline{r}]$ of the preference list of s. If this holds, then since the preference lists $\pi(s), \pi'(s)$ are identical in this interval, we get $\mu_\pi = \mu_{\pi'}$ (It is straight-forward to verity this). Therefore, $\mu_\pi = \mu_{\pi'}$, whp.

□

Proof of Theorem 4.1. For a preference profile π, define $B_\pi(s)$ to be the subset of students s' for which $\mu_\pi(s) \succ_{s'} \mu_\pi(s')$. Define $A_\pi(s)$ to be the subset of students s' for which $s' \in B_\pi(s)$, and moreover, $\mu_\pi(s') \succ_s \mu_\pi(s)$. The proof is done in two steps. In Step 1, we show that $|B_\pi(s)|$ is "large", whp. In Step 2, we show that $|A_\pi(s)|$ is "large", whp; this would prove the lemma.

Step 1. Consider an arbitrary school $c \in C$. We will show that wvhp, there are "many" students who rank c above their match in the student-optimal matching. Then, taking a union bound over all schools $c \in C$ would show that wvhp, many students rank $\mu_\pi(s)$ above their current match, implying that $|B_\pi(s)|$ is large. Instead of showing that many students rank c above their match in the student-optimal matching, we can equivalently show that c receives many proposals in the student-proposing DA. This is what we proved in Lemma 3.5.

We now formalize this idea. By Lemma 3.5, for any constant $\epsilon > 0$, each school receives at least $\frac{n(1-\epsilon)}{2\log n}$ proposals wvhp, which also implies that all schools receive at least $\frac{n(1-\epsilon)}{2\log n}$ proposals wvhp. Thus, $\mu_\pi(s)$ receives at least $\frac{n(1-\epsilon)}{2\log n}$ proposals wvhp, which means for any constant $\epsilon > 0$, wvhp we have $|B_\pi(s)| > \frac{n(1-\epsilon)}{2\log n}$. This completes Step 1.

Observe that in Step 1 we showed that

$$\mathbb{P}_{\pi\sim\mathcal{P}}\left[|B_\pi(s)| > \frac{n(1-\epsilon)}{2\log n}\right] \geq 1 - o(1), \tag{6}$$

where $\mathcal{P}$ denotes the uniform distribution over all preference profiles. Next, we write an alternative version of (6), which will be used later in Step 2.

Recall Definition 4.2, by which $\Pi(s)$ is a random variable containing the set of all the possible placements of schools $[m]\backslash[\underline{r},\overline{r}]$ in positions $[m]\backslash[\underline{r},\overline{r}]$. Note that, without loss of generality, we can assume that schools listed on positions $\underline{r},\ldots,\overline{r}$ of $\pi(s)$ are schools $\underline{r},\ldots,\overline{r}$, respectively. Thus, we can rewrite (6) as

$$\mathbb{P}_{\Pi(s)\sim\mathcal{P}(s),\pi\sim\mathcal{U}(\Pi(s))}\left[|B_\pi(s)| > \frac{n(1-\epsilon)}{2\log n}\right] \geq 1 - o(1). \tag{7}$$

Step 2 Lemma 4.3 shows that, when $\Pi(s) \sim \mathcal{P}(s)$, almost all student-optimal matchings in $\Pi(s)$ (i.e. a fraction $1 - o(1)$ of them) are the same whp. Let μ denote this matching. Suppose that, for $\pi, \pi' \in \Pi(s)$, we have $\mu_\pi = \mu_{\pi'} = \mu$. Then, see that by the definition of $\Pi(s)$, we have $B_\pi(s) = B_{\pi'}(s)$. Thus, we let $B(s)$ denote $B_\pi(s)$ for any $\pi \in \Pi(s)$ for which $\mu_\pi = \mu$. Now, (7) implies that $|B(s)|$ is large, whp. This means, if $\pi \sim \mathcal{U}(\Pi(s))$, then, both of the events $\mu_\pi = \mu$ and $|B_\pi(s)| \geq \frac{n(1-\epsilon)}{2\log n}$ hold whp. We use this fact to prove that $|A_\pi(s)|$ is large, whp. This would conclude Step 2.

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

by Lemma 4.3. Therefore, we assume that the event $\mu_\pi = \mu$ holds in the rest of this analysis.

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

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

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

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

Thus, $|A_\pi(s)|$ is at least $\frac{(1-\delta)\cdot|B(s)|}{(\log n)^{1+\epsilon}}$ whp. In Step 1, (7) shows that $|B(s)|$ is large whp. Consequently,

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

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

Theorem 4.4. *Fix a student s. Under MTB, if $n < m$*

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

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

$$\mathbb{P}\left[|L(s) \cap L(s')| \geq 2\right] \leq O\left(\frac{\log^4 n}{n^2}\right). \tag{8}$$

That is, the probability that (s, s') is a Pareto improving pair is very small. Assuming (8) holds the proof is completed by taking a union bound over all $s' \neq s$ since the union bound implies that

$$\mathbb{P}[\ddot{\mu}(s) \geq 1] \leq n \cdot O\left(\frac{\log^4 n}{n^2}\right) = o(1).$$

It remains to show that (8) holds. First fix $L(s)$ and then start constructing $L(s')$ randomly (we are using the principle of deferred decisions). It is straight-forward to verify that

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

□

5 Proofs for Section 1.3

5.1 Equivalence of social inequity and variance

Proof of Lemma 1.4. Let $q = \min\{m, n\}$ be the number of assigned students, which is the same in all stable matchings. Then,

$$\begin{aligned}
\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &= \mathbb{E}_\pi\left[\frac{1}{|\mu_\pi(C)|} \cdot \sum_{t \in \mu_\pi(C)} \left(\mathcal{A}r(\mu_\pi) - \mu_\pi^\#(t)\right)^2\right] \\
&= \mathbb{E}_\pi\left[\frac{1}{q} \cdot \sum_{t \in \mu_\pi(C)} \mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t)\right] \\
&= \sum_{t \in S} \mathbb{P}_\pi\left[t \in \mu_\pi(C)\right] \cdot \mathbb{E}_\pi\left[\frac{1}{q} \cdot \mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t) \middle| t \in \mu_\pi(C)\right] \\
&= \sum_{t \in S} \frac{q}{n} \cdot \mathbb{E}_\pi\left[\frac{1}{q} \cdot \mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t) \middle| t \in \mu_\pi(C)\right] \\
&= \frac{1}{n} \cdot \sum_{t \in S} \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(t)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(t) \middle| t \in \mu_\pi(C)\right] && (9) \\
&= \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 + \mu_\pi^\#(s)^2 - 2\mathcal{A}r(\mu_\pi)\mu_\pi^\#(s) \middle| s \in \mu_\pi(C)\right] && (10) \\
&= \mathbb{E}_{\pi(s)}\mathbb{E}_{\{\pi(s'): s' \in S, s' \neq s\}}\left[\mathsf{Var}\left[r_s\right]\right] && (11) \\
&= \mathbb{E}_{\{\pi(s'): s' \in S, s' \neq s\}}\left[\mathsf{Var}\left[r_s\right]\right]. && (12)
\end{aligned}$$

In the above inequalities, (10) holds because the term inside the expectation in (9) is equal for all students by symmetry. (12) holds since, by symmetry, the inner expectation in (11) is equal for all preference profiles of s. □

5.2 Proof of Lemma 1.6

5.2.1 preliminaries

Proposition 5.1. *Suppose $d \leq n$, and define the random variable $X = \min\{X_1, \ldots, X_d\}$, where $X_1, \ldots, X_d$ respectively represent the first d elements of a permutation over $[n]$ that is chosen uniformly at random. Then, $\mathbb{E}\left[X^2\right] = \frac{d(n+1)(n-d)}{(d+1)^2(d+2)} + \frac{(n+1)^2}{(d+1)^2}$.*

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

Lemma 5.2. *Suppose $n \leq m$. Then, a student s with priority number $n - t$ is assigned to one of her top $\frac{n\log(n)}{t}$ choices with probability at least $1 - 1/n$.*

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

Lemma 5.3. *Suppose $|n - m| = 1$. Then, under STB, for any student s,*

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

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

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

$$\begin{aligned}\mathbb{E}_\pi\left[\mu^{\#}_\pi(s)^2 \middle| \mu_\pi(s) \neq \emptyset\right] &= \mathbb{P}\left[p_s \leq n-t\right] \cdot \mathbb{E}\left[\mu^{\#}_\pi(s)^2 \middle| p_s \leq n-t\right] \\ &\quad + \mathbb{P}\left[n-t < p_s\right] \cdot \mathbb{E}\left[\mu^{\#}_\pi(s)^2 \middle| n-t < p_s\right]. \end{aligned} \tag{13}$$

We provide an upper bound for each of the terms in the right-hand side of (13).

By Lemma 5.2, we have:

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

which implies that

$$\mathbb{P}\left[p_s \le n-t\right] \cdot \mathbb{E}\left[\mu_\pi^{\#}(s)^2 \middle| p_s \le n-t\right] \le 2n. \tag{14}$$

Also, we have that

$$\begin{aligned}\mathbb{P}\left[n-t < p_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| n-t < p_s\right] &\le \frac{t}{n} \cdot \sum_{i=1}^{t} \frac{1}{t} \mathbb{E}\left[r_s^2 \middle| p_s = n-i+1\right] \\ &\le \frac{1}{n} \cdot \sum_{i=1}^{t} 2(n/i)^2. \qquad (15) \\ &\le n \cdot \frac{\pi^2}{3}. \qquad (16)\end{aligned}$$

where (15) holds since for a geometric random variable X with mean p we have $\mathbb{E}\left[X\right] = \frac{2-p}{p^2}$.

Finally, putting (14) and (16) together implies

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

□

5.2.2 Proof of Lemma 1.6 - Part 1

The proof for Part 1 of Lemma 1.6 is directly implied by Lemmas 5.4 and 5.5.

Lemma 5.4. *When $n = m+1$, expected social inequity in MTB is $\Omega(\frac{n^2}{\log^2 n})$.*

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

Step 1. Since the social inequity and the expected variance in the rank of a fixed student are equal by Lemma 1.4, there is no harm in analyzing the latter notion (we switch to the former notion in Step 2). We are interested in providing a lower bound on $\mathbb{E}\left[(r_s - r)^2\right]$, where r_s is a random variable denoting the rank for student s and $r = \mathcal{A}r(\eta)$ (note that r is also equal to the average rank of s, conditioned on being assigned). Since $\mathbb{E}\left[(r_s - r)^2\right] = \mathbb{E}\left[r_s^2\right] - r^2$, we can instead provide a lower bound on the RHS of the equality.

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

$$\mathbb{E}\left[r_s^2\right] \geq \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}\left[E_s\right]) \cdot 0. \tag{17}$$

To give a lower bound on the RHS of (17), we provide a lower bound on $\mathbb{E}\left[(r_s - r)^2 \middle| E_s\right]$. If student s receives d_s proposals in school-proposing DA, then it chooses the best out of these d_s proposals, which means its rank is the first order statistic among the proposals that she had received. In Proposition 5.1, we calculate $\mathbb{E}\left[r_s^2 \middle| d_s\right]$ (which is the expected rank squared for s conditioned on receiving d_s proposals).

Using Proposition 5.1 and (17) together we can write

$$\begin{aligned}\mathbb{E}\left[r_s^2 | E_s\right] &\geq \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}\left[E_s\right]) \cdot 0 \\ &\geq (1 - o(1)) \cdot \frac{3n^2}{2\log^2 n} + o(1) \cdot 0, \end{aligned} \tag{18}$$

where (18) follows from Lemma 3.1, which shows event E_s happens whp.

It is known that, $r \in [\frac{(1-\delta)n}{\log n}, \frac{(1+\delta)n}{\log n}]$ for any constant $\delta > 0$ and large enough n (see Ashlagi et al. (2017)). Therefore, together with (18),

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

This finishes Step 1.

Step 2. In this step, instead of working with the notion of expected variance in the rank of a fixed student, we switch to its equivalent notion, expected social inequity. Step 1 and Lemma 1.4

together imply that $\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]$ is $\Omega(\frac{n^2}{\log^2 n})$. In this step, we show that moving from the school-optimal matching to the student-optimal matching does not change the social inequity much in expectation, and as the result, we would prove that $\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right]$ is also $\Omega(\frac{n^2}{\log^2 n})$. This is done as follows.

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

$$\begin{aligned}&=m\cdot\mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2-\mathcal{A}r(\eta_\pi)^2\right]+\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)}\mu_\pi^{\#}(s)^2-\eta_\pi^{\#}(s)^2\right]\\&-2\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)}\mu_\pi^{\#}(s)\mathcal{A}r(\mu_\pi)-\sum_{s\in\eta_\pi(C)}\eta_\pi^{\#}(s)\mathcal{A}r(\eta_\pi)\right].\end{aligned}\tag{19}$$

We can rewrite the above inequality by simplifying (19) as

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

which together with the previous equation implies that

$$m\cdot\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)-\mathcal{S}i(\eta_\pi)\right]= \tag{20}$$
$$-m\cdot\mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2-\mathcal{A}r(\eta_\pi)^2\right] \tag{21}$$
$$+\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)}\mu_\pi^{\#}(s)^2-\eta_\pi^{\#}(s)^2\right]. \tag{22}$$

To prove the lemma, we provide lower bounds for (21) and (22). When we move from the school-optimal matching to the student-optimal matching, each student gets assigned to a school at least as good as before. Let $\Delta_\pi(s)=\eta_\pi^{\#}(s)-\mu_\pi^{\#}(s)$, and $\Delta_\pi=\sum_{s\in\mu_\pi(C)}\Delta_\pi(s)$.

A Lower Bound for (21). First, we note that $m\left(\mathcal{A}r(\eta_\pi) - \mathcal{A}r(\mu_\pi)\right) = \Delta_\pi$ is "small" wvhp. This is a direct consequence of Theorem 5 in Ashlagi et al. (2017); they show that there exist constants $n_0, \delta > 0$ such that for $n > n_0$, we have

$$\mathbb{P}_{\pi\sim\Pi}\left[\Delta_\pi \geq \delta n \log n\right] < \mathsf{exp}\left\{-(\log n)^{0.4}\right\}. \tag{23}$$

According to this bound, we have that

$$m \cdot \left(\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right) = m \cdot \left((\mathcal{A}r(\eta_\pi) - \Delta_\pi/m)^2 - \mathcal{A}r(\eta_\pi)^2\right) = \Delta_\pi^2/m - 2\Delta_\pi \mathcal{A}r(\eta_\pi).$$

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

$$\begin{aligned} m \cdot \mathbb{E}_\pi\left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2\right] &= \mathbb{E}_\pi\left[\Delta_\pi^2/m - 2\Delta_\pi \mathcal{A}r(\eta_\pi)\right] \\ &\leq (\bar{\delta} n \log n)^2/m, \end{aligned} \tag{24}$$

where the last inequality follows from (23), for any constant $\bar{\delta} > \delta$ and sufficiently large n. This implies a lower bound of $-(\bar{\delta} n \log n)^2/m$ for (21).

A Lower Bound for (22). First, we rewrite (22) as follows.

$$\begin{aligned} \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \mu_\pi^{\#}(s)^2 - \eta_\pi^{\#}(s)^2\right] &= \mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} (\eta_\pi^{\#}(s) - \Delta_\pi(s))^2 - \eta_\pi^{\#}(s)^2\right] \\ &\geq -2\mathbb{E}_\pi\left[\sum_{s\in\mu_\pi(C)} \eta_\pi^{\#}(s)\Delta_\pi(s)\right]. \end{aligned} \tag{25}$$

We proceed by providing a lower bound on (25). First, we use the Cauchy-Schwarz inequality to write

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

Taking expectation from both sides of the above inequality implies

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

Using (23), we can rewrite the above upper bound:

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

which holds for any constant $\overline{\delta} > \delta$. According to (25), this upper bound can be directly translated into a lower bound $-2m^{3/2} \cdot n(\overline{\delta} \log n)^{1/2}$ for (22).

Using the lower bounds that we provided for (21) and (22), we can rewrite equation (20) as follows:

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

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

$$\mathbb{E}_\pi \left[\mathcal{Si}(\mu_\pi) \right] = \mathbb{E}_\pi \left[\mathcal{Si}(\eta_\pi) \right] + \mathbb{E}_\pi \left[\mathcal{Si}(\mu_\pi) - \mathcal{Si}(\eta_\pi) \right] \geq \Omega(n^2 / \log^2 n).$$

This completes the proof. $\square$

Lemma 5.5. *Suppose $|n - m| = 1$. Then, under STB, the expected social inequity is $\Theta(n)$.*

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

$$\mathbb{E} \left[\mathcal{Si}(\mu_{\mathsf{STB}}) \right] = \mathbb{E} \left[\mathsf{Var} \left[r_{\underline{s}} \right] \right] \geq \frac{1}{n} \cdot \left(\mathcal{Ar}(\mu_{\mathsf{STB}}^{\#}(s)) - n \right)^2.$$

It is proved by Knuth (1995) that $\mathcal{Ar}(\mu_{\mathsf{STB}}^{\#}(s)) = \Theta(\log n)$. Plugging this into the above inequality implies that $\mathbb{E}\left[\mathcal{Si}(\mu_{\mathsf{STB}})\right] \geq \Omega(n)$. On the other hand, by Lemma 5.3 we have that

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

which completes the proof. □

5.2.3 Proof of Lemma 1.6 - Part 2

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

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

This implies that the expected social inequity under MTB is $O(\log^4 n)$. On the other hand, Lemma 5.5 implies that the expected social inequity under STB is $\Theta(n)$.

5.2.4 Proof of Lemma 1.6 - Part 3

First note that Part 2 implies a weaker version of Part 3. That is, If $n = m - 1$, the expected social inequity under MTB is still $O(\log^4 n)$, by the same analysis for $n = m$. On the other hand, by Lemma 5.5 the expected social inequity under STB is $\Theta(n)$. This gap is large enough that Theorem 1.5 still holds, even with this weaker version of Part 3.

We prove here that the gap is even larger, by showing how the bound on the expected social inequity under MTB can be improved to $O(\log^2 n)$. The proof follows the same steps as the proof of Lemma 5.7, where we provide an upper bound on $\mathbb{E}\left[\mathcal{Si}(\mu_{\mathsf{MTB}})\right]$ when the imbalance is linear. During the proof, we will also use Lemma 3.5, which was proved in Section 4.

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

Step 1. First, we rewrite the following equality from the proof of Lemma 5.4.

$$
\begin{aligned}
m \cdot \mathbb{E}_\pi \left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] = & \qquad (26) \\
& - m \cdot \mathbb{E}_\pi \left[\mathcal{A}r(\mu_\pi)^2 - \mathcal{A}r(\eta_\pi)^2 \right] \qquad (27) \\
& + \mathbb{E}_\pi \left[\sum_{s \in \mu_\pi(C)} \mu_\pi^{\#}(s)^2 - \eta_\pi^{\#}(s)^2 \right]. \qquad (28)
\end{aligned}
$$

To complete Step 1, we need to provide upper bounds for (27) and (28).

An upper bound for (27) We will use the following relation between average ranks, provided by Theorem 3 of Ashlagi et al. (2017): wvhp we have

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

Consequently, $m \cdot o(1) \cdot \mathbb{E}_\pi \left[\mathcal{A}r(\mu_\pi) \right]$ is a valid upper bound for (27).

An upper bound for (28) 0 is a valid upper bound since, by the definition of μ, η, we always have $\mu_\pi^{\#}(s) \leq \eta_\pi^{\#}(s)$.

Plugging the provided upper bounds into (26) implies

$$\mathbb{E}_\pi \left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] \leq o(1) \cdot \mathbb{E}_\pi \left[\mathcal{A}r(\mu_\pi) \right].$$

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

$$\mathbb{E}_\pi \left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi) \right] \leq o(1), \qquad (29)$$

which concludes Step 1.

Step 2. Suppose we are running the school-proposing DA. First, see that

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

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

$$\mathbb{E}_\pi\left[\mathcal{Si}(\eta_{\mathsf{MTB}})\right] \leq \mathbb{E}\left[r_s^2\right]. \tag{30}$$

Next, we provide an upper bound on $\mathbb{E}\left[r_s^2\right]$. Fix an arbitrary small constant $\epsilon > 0$. Let E_s denote the event in which student s receives at least $\kappa = \frac{(1-\epsilon)n}{2\log n}$ proposals. Lemma 3.5 shows that E_s happens wvhp. Consequently,

$$\mathbb{E}\left[r_s^2 | E_s\right] \lesssim \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] \leq O(\log^2 n), \tag{31}$$

where we used Proposition 5.1 to bound $\mathbb{E}\left[r_s^2 \middle| E_s\right]$.

Now we are ready to finish the proof of Part 3. See that (30) and (31) together imply that

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

Therefore, together with Step 1, we have that

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

5.3 Proof of Theorem 1.7

We first prove a weaker version of Theorem 1.7 (Theorem 5.6) and at the end of this section, we explain how our proof for Theorem 5.6 can be adapted to work for Theorem 1.7.

Theorem 5.6. *Suppose $m = n + \lambda n$ for any positive $\lambda \leq 0.008$. Then,* $\lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{Si}(\mu_{\mathsf{STB}})]}{\mathbb{E}[\mathcal{Si}(\mu_{\mathsf{MTB}})]} > 1$, *where the expectations are taken over preferences and the tie-breaking rules.*

To prove this theorem, we need the following lemmas, the proofs for which appears after the proof of the theorem.

Lemma 5.7. *Suppose $m = n + \lambda n$. Then, under MTB we have*

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

where $K = (1+\lambda)\log(1+1/\lambda)$ and $T = \frac{2(1+\lambda)}{\lambda+1/(1+K)}$.

Lemma 5.8. *Suppose $m = n + \lambda n$. Then, under STB we have*

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

Proof of Theorem 5.6. The proof is directly implied by Lemmas 5.7 and 5.8 below.

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

where $K = (1+\lambda)\log(1+1/\lambda)$ and $T = \frac{2(1+\lambda)}{\lambda+1/(1+K)}$. For $\lambda \leq 0.008$, RHS of the above inequality is strictly greater than one. $\square$

Next, we prove the two lemmas that we used in the proof of this theorem. To simplify algebraic calculations, we use the notions $\approx, \gtrsim$ which respectively mean equality and inequality up to vanishingly small terms.

Proof of Lemma 5.7. We use Lemma 1.4, by which the expected social inequity and the expected variance of the rank of a fixed student are equal. So, to prove the lemma, we fix a student s and show that

$$\lim_{n\to\infty} \mathbb{E}_{\{\pi(s'):s'\in S, s'\neq s\}} \left[\mathsf{Var}\left[r_s\right]\right] \leq T(2T-1) - K^2. \tag{32}$$

We prove (32) in 2 Steps. In Step 1, we show that that the variance of the rank of student s in the student-proposing DA is approximately equal to the variance of its rank in the school-proposing

DA. Then, in Steps 2, we provide an upper bound $T(2T-1) - K^2$ on the variance of rank in the school-proposing DA. Steps 1,2 then together will imply that (32) holds.

To prove the lemma, it remains to prove each of the steps separately.

Step 1. This step is identical to Step 1 in the proof of Part 3 of Lemma 1.6, which was presented in Section 5.2.4).

Step 2. This Step is similar to Step 1 in the proof of Lemma 5.4.

Since the expected social inequity and the expected variance of the rank of a fixed student are equal by Lemma 1.4, in this step we use the latter notion. We will switch to the former notion in Step 2. We are interested in providing an upper bound on $\mathbb{E}\left[(r_s - r)^2\right]$, where r_s is a random variable denoting the rank for student s and $r = \mathcal{A}r(\eta)$ (note that r is also equal to the average rank of s, conditioned on being assigned). Since $\mathbb{E}\left[(r_s - r)^2\right] = \mathbb{E}\left[r_s^2\right] - r^2$, we can instead provide an upper bound on the RHS of the equality.

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

$$\mathbb{E}\left[r_s^2\right] \leq \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 \middle| E_s\right] + (1 - \mathbb{P}\left[E_s\right]) \cdot (n + \lambda n)^2. \tag{33}$$

We proceed by providing an upper bound on the RHS of (33). Lemma 3.2 implies E_s happens wvhp, and so, we can ignore the second term in the RHS of (33) since it is a lower order term. We provide an upper bound on the first term in the RHS of (33), i.e. on $\mathbb{E}\left[r_s^2 \middle| E_s\right]$. If student s receives d_s proposals in school-proposing DA, then it chooses the best out of these d_s proposals, which means its rank is the first order statistic among the proposals that she had received. In Proposition 5.1, we calculate $\mathbb{E}\left[r_s^2 \middle| d_s\right]$ (which is the expected rank squared for s conditioned on receiving d_s proposals).

Using Proposition 5.1 and (33) together we can write

$$\begin{aligned}\mathbb{E}\left[r_s^2|E_s\right] &\lesssim \mathbb{P}\left[E_s\right]\cdot\mathbb{E}\left[r_s^2\middle|E_s\right]\\ &\lesssim (\frac{n(1+\lambda)}{\kappa})(\frac{2n(1+\lambda)}{\kappa}-1) &(34)\\ &= \left(\frac{1+\lambda}{\frac{1}{2(1+K)}+\frac{\lambda}{2}}\right)\left(\frac{2(1+\lambda)}{\frac{1}{2(1+K)}+\frac{\lambda}{2}}-1\right) = T(2T-1). &(35)\end{aligned}$$

Now, (35) implies that

$$\lim_{n\to\infty}\mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] = \lim_{n\to\infty}\mathbb{E}_\pi\left[r_s^2 - r^2\right] = T(2T-1) - K^2. \tag{36}$$

This completes Step 2.

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

$$\begin{aligned}\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] &\le \mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi) - \mathcal{S}i(\eta_\pi)\right] + \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right]\\ &\le o(1) + \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] &(37)\\ &\approx T(2T-1) - K^2, &(38)\end{aligned}$$

where (37) follows from Step 1, and (38) follows from (36). □

Next, we show how the proof works for Lemma 5.8.

Proof of Lemma 5.8. Suppose students indexed with respect to their priority number in STB, i.e. the student with the highest priority number is indexed 1, and the student with the lowest priority number is indexed with n. Fix a student s. Using Lemma 1.4, we can write

$$\mathbb{E}_\pi\left[\mathcal{S}i(\mu_\pi)\right] = \mathsf{Var}\left[r_s\right] = \mathbb{E}\left[r_s^2\right] - \mathbb{E}\left[r_s\right]^2, \tag{39}$$

where r_s denotes the rank assigned to student s.

To provide a lower bound for (39), we lower bound $\mathbb{E}\left[r_s^2\right]$ and upper bound $\mathbb{E}\left[r_s\right]^2$.

Upper bound for $\mathbb{E}\left[r_s\right]^2$. First, we state the following claim.

Claim 5.9. *Suppose $m = (1+\lambda)n$. Then, $\mathbb{E}\left[r_s\right] \approx (1+\lambda)\log(1+\frac{1}{\lambda})$.*

Proof. This follows from Ashlagi et al. (2017). □

By Claim 5.9, we have that

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

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

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

Next, we state the following claim; its proof comes after the proof of this lemma.

Claim 5.10. *Suppose $m = (1+\lambda)n$. Then, $\mathbb{E}\left[r_{k+1}^2\right] \geq \frac{2-p}{p} - O(\frac{\log^5 m}{m})$, where $p = \frac{m-k}{m}$.*

Now, we use Claim 5.10 to calculate an upper bound on the RHS of the above inequality:

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

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

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

By combining the above bounds, we can provide the promised lower bound on (39).

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

□

Proof of Claim 5.10. A straight-forward calculation gives

$$\mathbb{E}\left[r_{k+1}^2\right] = \sum_{j=0}^{k}(j+1)^2 \cdot (1 - \frac{k-j}{m-j}) \cdot \prod_{l=0}^{j-1} \frac{k-l}{m-l}. \tag{40}$$

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

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

Now, using the above inequality, we provide the following upper bound on (40):

$$\mathbb{E}\left[r_{k+1}^2\right] \geq \left(\sum_{j=0}^{\bar{t}}(j+1)^2 \cdot (1 - \frac{k}{m})(\frac{k}{m})^j\right) - \frac{\lambda \bar{t}^5}{2m}. \tag{41}$$

We are almost done. In the RHS of (41), we bound the first term from below by

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

which holds because of the following well-known fact: $\mathbb{E}\left[Z^2\right] = \frac{2-q}{q}$ where Z is a geometric random variable with success probability q. Using the above bound, we can rewrite (41) as

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

which completes the proof. □

5.4 Proof Sketch for Theorem 1.7

Finally, we describe how proof of Theorem 5.6 can be adapted to prove Theorem 1.7. The main difference is in Lemma 3.2. By proving a stronger version of Lemma 3.2, the same proof would work for $\lambda > 0$. Some of the less important details are omitted from this proof.

We define the stronger version of Lemma 3.2 simply by using, the variable

$$\kappa' = \left(\frac{n}{1+K} + \lambda n\right) \cdot \left(\frac{1}{2 - \left(\frac{1}{(1+K)(1+\lambda)} + \frac{\lambda}{1+\lambda}\right) \cdot \frac{1}{2}}\right)$$

instead of a variable κ. Replacing κ with κ' in the lemma statement would give the stronger version of the lemma. To show why the stronger version holds, we need to consider again the coupling $(\mathrm{DA}, \mathcal{B})$ which we defined in the proof of Lemma 3.2. There, for each successful coin-flip (a proposal made to s), we removed n coins. However, instead of doing that, here we remove $n-y$ coins, where y is the number of proposals made by the proposer so far. Everything else in the coupling remains the same, e.g. the number of coins that we flip will remain $2n\kappa(1-\delta)$). We will follow the same proof that we gave for Lemma 3.2, with some adjustments. We sketch the proof below.

Let X be a random variable that denotes the total number of successful coin flips in the coupling. Our goal is showing that $X \geq \kappa'(1-\delta)$ holds wvhp.

Claim 5.11. *Wvhp, $X \geq \kappa'(1-\delta)$.*

First, we verify that the lemma is proved by the above claim, and after that we prove the claim itself. To prove the lemma, we follow the proof of Lemma 5.7 by rewriting (34) and (35) as follows. Let E_s denote the event at which s receives at least $\kappa'(1-\delta)$ proposals. Then,

$$\begin{aligned} \mathbb{E}\left[r_s^2 | E_s\right] &\lesssim \mathbb{P}\left[E_s\right] \cdot \mathbb{E}\left[r_s^2 | E_s\right] \\ &\lesssim \frac{n(1+\lambda)}{\kappa'} \cdot \left(\frac{2n(1+\lambda)}{\kappa'} - 1\right) \end{aligned} \tag{42}$$

Now, (42) implies that

$$\lim_{n\to\infty} \mathbb{E}_\pi\left[\mathcal{S}i(\eta_\pi)\right] = \lim_{n\to\infty} \mathbb{E}_\pi\left[r_s^2 - r^2\right] \leq \frac{n(1+\lambda)}{\kappa'} \cdot \left(\frac{2n(1+\lambda)}{\kappa'} - 1\right) - K^2. \tag{43}$$

Note that (43) is an improved upper bound. On the other hand, as we showed in Step 1 of the proof of Lemma 5.7,

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

Consequently,

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

where $K = (1+\lambda)\log(1+1/\lambda)$. The RHS of the above inequality is strictly greater than one for any positive constant $\lambda \leq 0.01$. (Note that the RHS is only a function of λ) This would prove the lemma. It remains to prove Claim 5.11.

First, we will argue that the claim holds in expectation, i.e. $\mathbb{E}[X] \geq \kappa'(1-\delta)$. Recall that in the (new) coupling, after each successful coin-flip, i.e. a proposal made to s by a school c, only z_c coins are removed where $z_c = n - y_c$ and y_c is the number of proposals that c has made so far. Let d_c be the total number of proposals made by school c. Also, let F_c denote the event in which school c makes a proposal to s. Conditioning on school c making exactly d_c proposals, we get

$$\mathbb{E}\left[y_c \middle| d_c, F_c\right] = \frac{d_c+1}{2},$$

which holds for any arbitrary school $c \in C$. This holds simply because we can relabel the students (using a consistent permutation of the labels), without changing the student-optimal matching (up to relabeling). This equality, together with

$$\mathbb{E}\left[d_c \middle| F_c\right] \approx \frac{1}{1+\lambda} \cdot \frac{n}{1+K} + \frac{\lambda}{1+\lambda} \cdot n$$

(which follows from Ashlagi et al. (2017)) imply

$$\mathbb{E}\left[y_c \middle| F_c\right] \approx \frac{n}{2(1+\lambda)} \cdot \left(\frac{1}{1+K} + \lambda\right). \tag{44}$$

Now, since all of the $2n\kappa(1-\delta)$ coins will be flipped wvhp, the following holds wvhp as well:

$$\mathbb{E}[X] \cdot n + \mathbb{E}[X] \cdot \left(n - \mathbb{E}\left[y_c \middle| F_c\right]\right) \approx 2n\kappa, \tag{45}$$

$$\implies \mathbb{E}[X] \left(2n - \mathbb{E}\left[y_c \middle| F_c\right]\right) \approx 2n\kappa,$$

$$\implies \mathbb{E}[X] \approx \frac{2n\kappa}{2n - \mathbb{E}\left[y_c \middle| F_c\right]}$$

$$= \frac{2\kappa}{2 - \mathbb{E}\left[y_c \middle| F_c\right]/n} = \frac{\frac{n}{(1+K)} + \lambda n}{2 - \frac{1}{2(1+\lambda)} \cdot \left(\frac{1}{1+K} + \lambda\right)} \tag{46}$$

where (45) holds since, on average, for any n unsuccessful coin flips, we have 1 successful one, which results in removal of $\mathbb{E}\left[z_c \middle| F_c\right]$ coins in expectation, and also, (46) holds by (44). So, the weaker version of Claim 5.11 that we mentioned holds, i.e. when *wvhp* is replaced with expectation. Following the same approach, we can prove Claim 5.11. We explain the high-level idea here. Note that if the random variables $\{y_c\}$ were known to be independent, we could simply apply the Chernoff bound, which would imply that the sum $\sum_c y_c$ taken over all c that propose to s is concentrated around its mean, $X \cdot \mathbb{E}[y_1|F_1]$. This would let us write a stronger version of (45) (which holds wvhp, and not in expectation), which then proves Claim 5.11. Although $\{y_c\}$ are not independent, they are "almost" independent, roughly speaking, because preferences of schools are constructed independently. A careful treatment of these dependencies let us write the same concentration bounds. We omit the details.

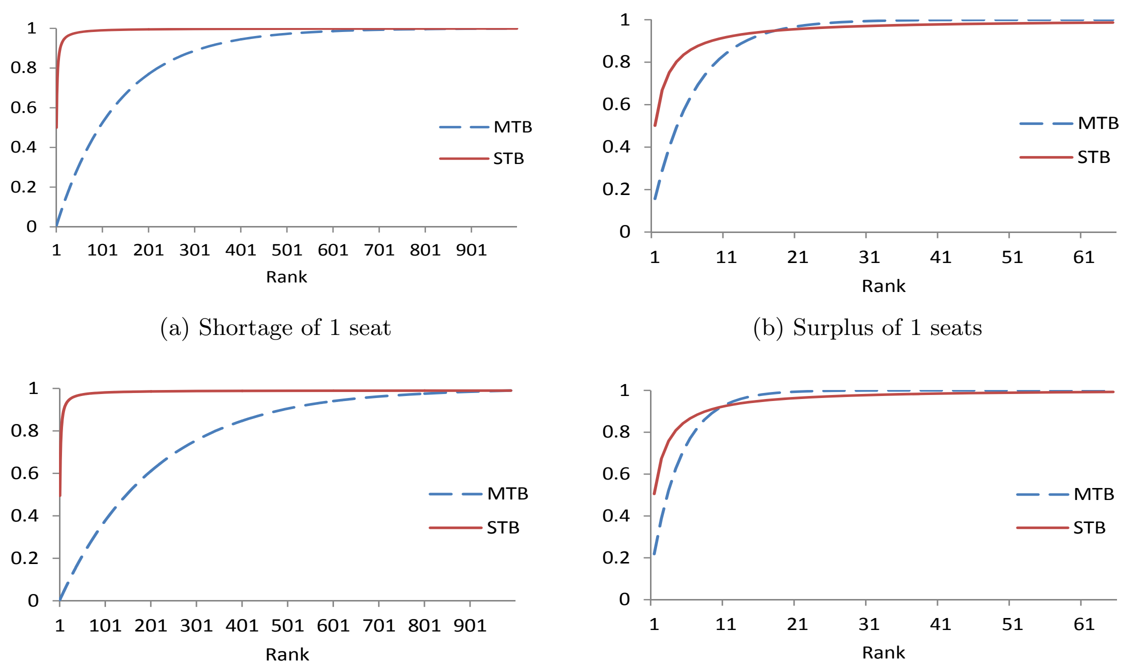
6 Computational experiments (not for publication)

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

6.1 Numerical results for our model

The first computational experiments illustrates the effect of the imbalance in the market on the students' rank distributions under STB and MTB and the relationship between the two. For each instance that we consider,[6] we sample realizations by drawing complete preference lists uniformly at random and independently for each student. In addition, under MTB, for each market realization we draw a complete order over students for each school, independently and uniformly at random. Under STB, for each market realization we draw a single ranking over students uniformly at random. Then, we compute the student optimal stable matching. The plots and the tables that we present here are generated by taking average over several (between 100 to 1000) samples for each instance.

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



(a) Shortage of 1 seat

(b) Surplus of 1 seats

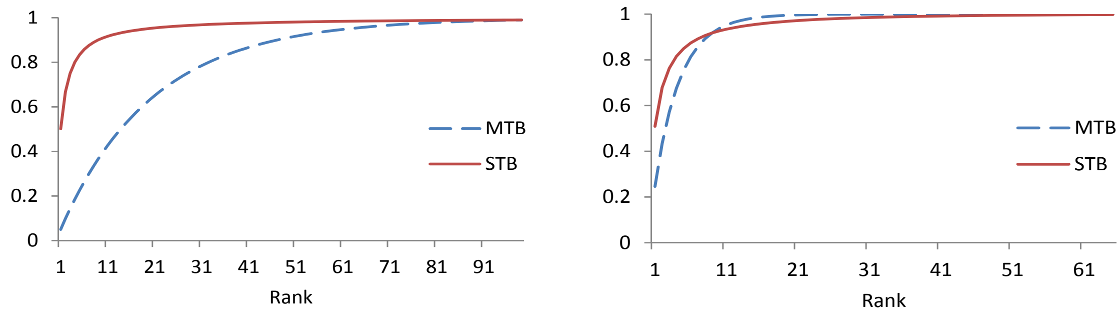
(c) Shortage of 10 seats

(d) Surplus of 10 seats

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

[6] An instance contains the information regarding market characteristics (size, capacities, list length), and the choice of tie-breaking rule.

Figure 2 illustrates similar findings for a market with only 100 students, unit capacities, and a shortage or surplus of a single seat.



(a) Shortage of seats

(b) Surplus of seats

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

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

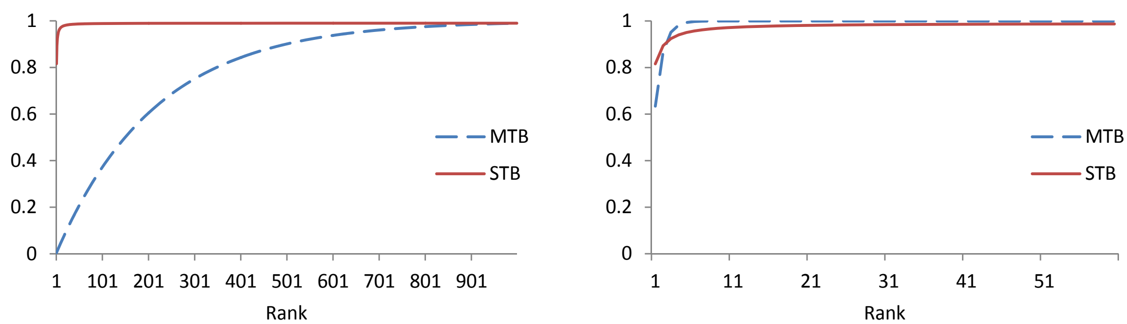
m \ $n-m$		-10	-1	1	10
100	$\mathcal{A}r(\mu_{\mathsf{STB}})/\mathcal{A}r(\mu_{\mathsf{MTB}})$	2.52/2.54	3.78/4.1	4.14/29.5	4.23/19.79
	$\mathcal{S}i(\mu_{\mathsf{STB}})/\mathcal{S}i(\mu_{\mathsf{MTB}})$	9.47/3.87	49.8/12.6	69.6/516.9	78.2/322.9
1000	$\mathcal{A}r(\mu_{\mathsf{STB}})/\mathcal{A}r(\mu_{\mathsf{MTB}})$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

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

6.2 Robustness to large imbalances and capacities

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

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



(a) Shortage of seats

(b) Surplus of seats

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

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

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

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

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

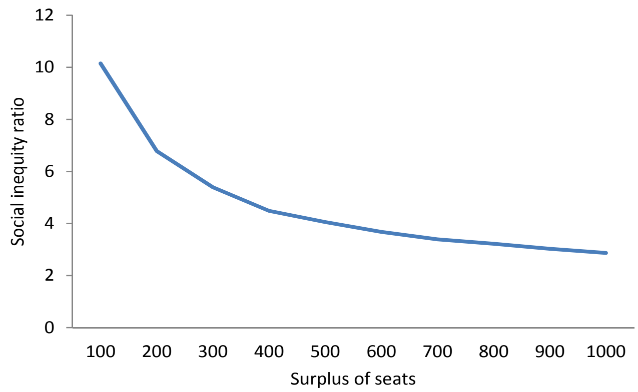


Figure 4: The ratio between $\mathcal{S}i(\mu_{\text{STB}})$ to $\mathcal{S}i(\mu_{\text{MTB}})$ in a random market with 10000 students, unit capacities, and a surplus of seats. (the x-axis denotes the surplus of schools)

6.3 Short preference lists

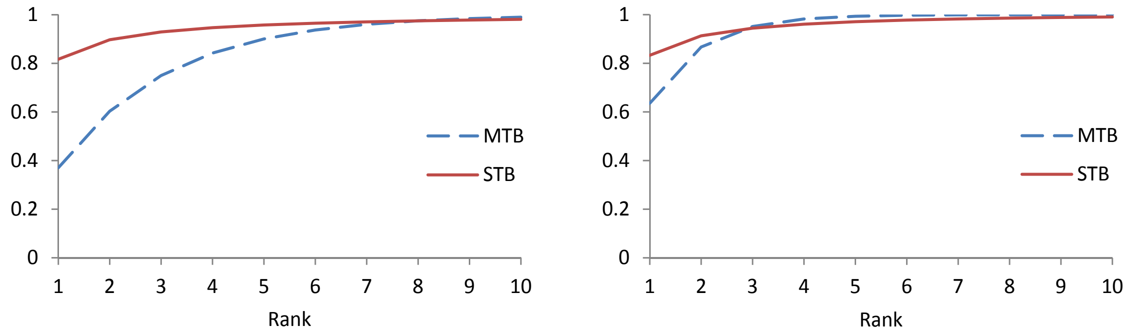
In this section, we present simulations to illustrate the effect of shortening the students' preference lists on our results.

Figure 5 presents the rank distribution in random market with 1000 schools, each with capacity of 10. In addition there are either 10100 or 9900 students, each of which ranks independently uniformly at random 10 schools. (Note that we consider the same instance with complete preference lists in Appendix 6.2, Table 3). When there is a shortage of seats and the preference lists are complete, our simulations reveal that the rank distribution under STB stochastically dominates the rank distribution under MTB; when the preference lists are short, stochastic dominance "almost" holds.

Shortening the lists reduces competition among students (see Ashlagi et al. (2015)), which impacts the market balance, i.e. whether students are "effectively" on the long side or the short side of the market. Therefore, whether there is a surplus or shortage in the market, as the preference lists become shorter, the crossing point of the rank distributions moves to the left (if the crossing happens at all).[7] In overdemanded markets, shortlists and large capacities act as two forces pushing in opposite directions: the former reduces competition and the latter increases it: When the capacities are large in an overdemanded market, MTB creates a much harsher competition relative to when the capacities are small, i.e. rejection chains become longer. On the other hand, under STB, a rejection reveals much more information about the rejected student's priority number, and thus,

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

that student is less likely to initiate rejection chains. Consequently, as the capacities increase, the crossing point moves to the right (if crossing happens at all).



(a) Shortage of seats

(b) Surplus of seats

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

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

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

6.4 Comparison to a hybrid tie-breaking rule

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

Example: unit capacity Figure 6 shows the rank distribution under the three tie-breaking rules in a market with 1000 students and 1000 schools, each with unit capacity. We consider 100 schools to be the top schools. Notice that up to rank 100, the STB and HTB plots coincide and are above the MTB plot. Conditioning on being above the 100 rank, the MTB and HTB coincide and note that there is no stochastic dominance in this range.

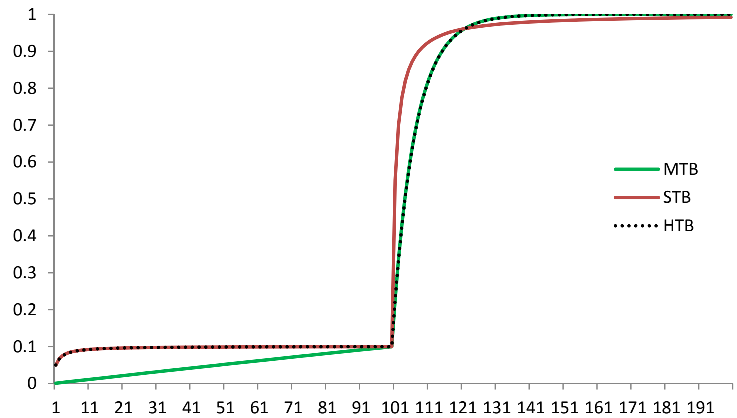


Figure 6: Students' rank distribution under STB, MTB and HTB. The market consists of $n = m = 1000$ and 100 top schools.

We list down the expected average rank and social inequity under the three tie-breaking rules below.

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

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

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

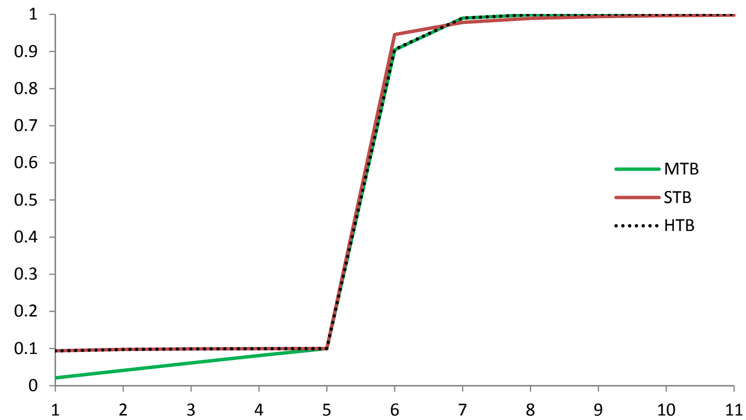


Figure 7: Students' rank distribution under STB, MTB and HTB. The market consists of 1000 students, 26 schools each with 50 seats and 5 top schools.

We list down the expected average rank and social inequity under the three tie-breaking rules below.

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

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

References

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

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

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

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