College Admission in Three Chinese Provinces: Boston Mechanism vs. Deferred Acceptance Mechanism*

Yun Pu†

This version: January 31, 2018

*I thank Ian Sheldon, Javier Donna, Yongyang Cai, Lixin Ye, and Jason Blevins for the advice and guidance. I thank Hengshui High School for providing the data of the number of students achieving each score in Hebei. This work is partially conducted on the Ohio Supercomputer Center. All errors are mine.

†Department of Agricultural, Environmental, and Development Economics, The Ohio State University, email: pu.61@osu.edu

Abstract

In college admissions in China over the last fifteen years, the Boston mechanism (BM, aka the sequential mechanism or Shunxu Zhiyuan) has been replaced by the deferred acceptance mechanism (DA, aka the parallel mechanism or Pingxing Zhiyuan). In this paper, we compare the empirical performance of these two mechanisms in the Chinese context. We construct a BM model and apply it in the provinces of Guangxi, Hebei, and Sichuan. This model only employs public data, instead of the micro data normally used in the literature, because the micro data are highly confidential in college admissions in China. Then, we conduct counterfactuals to empirically compare BM and DA in these three provinces for given years. We find that not only is BM superior to DA in terms of total welfare but also that most students receive lower utility after the switch from BM to DA.

JEL classification: C51, C78, D47, D61.

Keywords: College admission; Boston mechanism; deferred acceptance mechanism; welfare; China.

1. Introduction

On June 6-8 every year, around ten million high school seniors take the College Entrance Exam (Gaokao) in China. For most of these young adults, it is the most important exam they will ever take, and its results will have a profound influence on the rest of their lives. In each Chinese province, there is a "student placement office" (Zhu, 2014), which ranks the students in the province based on their exam results and asks them to submit their college preferences. Each college has an admission quota for each province. Then, the office uses a variant of either the Boston mechanism[1] (BM, aka the sequential mechanism or Shunxu Zhiyuan) or the deferred acceptance mechanism (DA, aka the parallel mechanism or Pingxing Zhiyuan) to assign students to colleges according to their exam-grade-based ranking, their preferences, and the admission quotas.

Under BM, students first-choice colleges are important. If a student is rejected by her first-choice college, she is very likely to be assigned to a much worse college or even to lose the opportunity of admission. Thus, students must think carefully before submitting their preferences, and many complain about the risks of the process. In contrast, under DA, if a student is rejected by one of her choices, her next choice will be considered. Since her admission is guaranteed only if her exam score is above the cutoff threshold for one of the colleges in her preference list, students feel much safer under this mechanism. In addition, BM also has inferior theoretical properties according to the literature (Abdulkadiroglu and Sönmez, 2003): DA is Pareto optimum for college admission in China, while BM is not. Furthermore, BM is not strategy-proof. Under BM students strategically select their preferences, so that

[1] See details in Appendix A in page 30.

we are unable to learn the true college preferences of the students directly from the data. All these drawbacks of BM led this mechanism to be largely abandoned in China: while before 2001 BM was implemented in all the provinces, in 2001 DA was first introduced, in Hunan province, and became more and more popular thereafter. By 2012, BM was applied only in three out of thirty-one provinces(Chen and Kesten, 2017).

However, all these drawbacks of BM can be called into question. First, both mechanisms will assign the same number of students to each college, while the admission quotas of the colleges also remain the same under the two mechanisms. If the preferences of students are homogeneous past a certain degree[2], not all students will be able to benefit equally from the switch from BM to DA-at least some of the students will sacrifice other benefits to enjoy the safety of DA. Second, though DA is Pareto optimum, it may not yield highest social welfare when the preferences are cardinal: that is, where students can not just prefer one college to another, but *much* prefer one college to another. The overall performance of DA under cardinal preference is still unclear, but the literature (Abdulkadiroğlu et al., 2011, 2015) has demonstrated that DA is not the best strategy when the priorities of the students are coarse. Third, although under BM we cannot learn the true preferences of the students directly from the data, we can still learn them after constructing a model of this mechanism.

Therefore, can we estimate colleges quality (or more accurately, their attractiveness) by using manipulated preference data under BM? Can we compare the empirical performance of BM to that of DA in China? Can we estimate the welfare loss (or gain) for an individual student when the mechanism switches from BM to DA?

[2]In China, most students would like to attend "good" colleges because better colleges not only provide better education, but also charge lower tuition fees. In addition, the quality of the different colleges is common knowledge; thus, the preferences of the students should be homogeneous enough.

To attempt this, first, we need to construct a model of BM for structural estimation. As far as we know, all the extant models in the literature (Agarwal and Somaini, Forthcoming; Calsamiglia et al., 2017; He, 2017; Hwang, 2016) require preference data of individual students (that is, micro data). However, in college admissions in China such data are highly confidential. For example, Li et al. (2010) used such micro data but anonymized the names of the provinces due to the sensitivity of the data. In addition, the use of micro data also restricts the reusability of the model in the future. In our model, we do not use micro data but instead consult the admission quotas and cutoff thresholds of the colleges in order to estimate the attractiveness of each college under BM. Nevertheless, computing the model is a heavy task: we need to estimate over 100 highly nonlinear parameters without the assumptions of contraction mapping being satisfied. To do so, we invent a new estimation method, FQR, which uses the DA model to assist with solving the BM model, based on the theorem proved in this paper stating that the two models are equivalent when the admission quotas are large.

Then, we apply the BM model to Guangxi, Hebei, and Sichuan provinces [3] and conduct counterfactuals calculating social and individual welfare under DA. In these three provinces for given years, the students apply the colleges after receiving their scores and are assigned to the colleges based on BM. We find that total welfare under DA is 1.73% - 6.63% lower than that under BM. Given that DA is Pareto optimum and thus yields the best result under ordinal preference, the welfare loss from BM to DA must be related to the property of cardinal preference. We raise an example to illustrate this relationship. Suppose that there are two colleges M and N and five students x, y, z, w, and r. M will admit one student while

[3] We plot the location of the three provinces in Figure 1 on Page 47.

N will admit three students. x is the top student, y the second, z the third, w the fourth, and r the bottom. It is common knowledge that M is a better school than N, but the students do not know others private preferences. Thus, y, z, w and r know that x is more likely to choose M but do not know her actual decision. Further suppose that x, w and r prefer N, while y and z prefer M. Under DA, students will be assigned to the best available colleges: x, z, and w will be admitted to N and y to M. All other mechanisms are dominated by DA if the preferences are ordinal; in contrast, if the preferences are cardinal, we can further assume that y prefers M *a little*, while z prefers M *a lot*. Under this assumption, DA is no longer optimum, because assigning z to M and y to N will yield higher social welfare. Under BM, y will choose N because the probability of her admission to M is much lower and she is nearly indifferent between the two colleges. Meanwhile, z will choose M because she prefers M greatly. Even if the admission probability is low, she wants to try. In the end, x, y, and w will be assigned to N while z will be assigned to M under BM, yielding higher social welfare than under DA. Intuitively, this holds up: since students are allowed to express their cardinal preferences under BM but not under DA, BM potentially yields better social welfare than does DA.

In all three provinces we look at, we also find that the cutoff thresholds under BM are looser than under DA. In actual fact, many provinces switching from BM to DA saw cutoff thresholds become stricter after switching. The phenomenon was reported on in the news media[4]. In the last example, the cutoff for M is 3 (at z) under BM and 2 (at y) under DA. This reveals the cost of students enjoying safety under DA: It becomes more difficult to receive admission if one's rank is not competitive.

[4]Please see one example here (http://news.sohu.com/20140713/n402175085.shtml).

Intuitively, "good" students (like y) will benefit from increased safety under DA while "bad" students (like z) suffer from cutoff thresholds being stricter after switching from BM to DA. However, how good should a "good" student be? From our results, only 0.64% - 10.65% of students above the key cutoff threshold[5] benefit from mechanism switching in Round 1 admission; most students suffer from switching.

Although the results suggest that BM is superior to DA on the whole, we also discover a potential drawback of BM: the cutoff thresholds of some colleges may become stricter as their attractiveness decreases. The attractiveness of a college is common knowledge, and so if one college becomes less attractive, students will realize that they may face less competition if they choose it. As a result, more students may apply to this college and the cutoff threshold become stricter as a result. That is, although the cutoff threshold is an indicator of the quality of a college, a college may nevertheless strengthen its cutoff threshold by decreasing its attractiveness in such a situation. This will result in worse actual aggregate quality of the higher education. We raise a theoretical instance to prove that this situation may actually occur. Fortunately, empirically in the three provinces for given years, only Tsinghua University suffered this for Guangxi science major students in 2008. Tsinghua University still needs to improve itself to attract better students from all other provinces.

Below, we review the literature in Section 2 and construct the theoretical models in Sections 3. We illustrate the estimation method and simulate it in Section 4. We describe the data and present the results in Section 5. Finally, we conclude in Section 6.

[5]The key cutoff threshold is the threshold for the students to be considered for admission in Round 1.

2. Literature

Abdulkadiroglu and Sönmez (2003) analyzed the school choice problem in terms of mechanism design. They defined the "justified envy" that occurs when a student prefers another school to her assigned school while the preferred school admits someone with lower priority than her priority. They argue that any mechanism without justified envy is Pareto dominated by DA and any mechanism, including DA, is Pareto dominated by the top trading cycle (TTC) mechanism. In college admission in China, a given student has the same priority in any college since colleges rank students only through their total scores. That is, DA and TTC are equivalent, and therefore DA is Pareto efficient and justified-envy free. Ergin and Sönmez (2006) demonstrate that DA is more efficient than BM. They raise an example with two regions (region M and region N) and three schools (school L, school M, and school N). All students prefer school M and N to school L, and students prefer school M (or N) if they live in region N (or M respectively). However, students who live in region M (or N) have higher priority in school M (or N respectively). In DA, students report their preferences truthfully[6]. Hence, students living in region M (or N) will be admitted by school N (or M respectively). In BM, students are afraid to report their true preferences, and thus do not have the highest priority in their favorite schools. They also do not want to be admitted by school L. Therefore, to ensure a seat in school M or N, a student in region M (or N) would choose school M (or N) as her first choice. However, not all students will be admitted by their favorite schools, and BM is not as efficient as DA. This example relies heavily on the

[6]Dubins and Freedman (1981), showing that students are unable to improve their utility by lying under DA.

assumption that the schools can rank students differently. If a student has the same priority in all schools, the two mechanisms are equivalent to a mechanism in which students with higher priority choose schools before than the others do[7]. Thus, Ergin and Sönmez (2006) is unable to show in fact that DA performs better than BM in college admission in China. Moreover, these two papers (Abdulkadiroglu and Sönmez, 2003; Ergin and Sönmez, 2006) consider a scenario of complete information and ordinal preference, in which all information is public and in which students can prefer one school to another but cannot *much* prefer one school to another. These assumptions are not realistic.

Abdulkadiroğlu et al. (2011) considered incomplete information and cardinal preference, and assumed that schools have no priority and that, while students have the same ordinal preference, their cardinal preferences may be different. They showed BM performing slightly better than DA. Abdulkadiroğlu et al. (2015) generalized this idea, illustrating that DA is not Pareto efficient when the priorities are coarse under complete information and cardinal preference. That is, these papers assume that the priority is not strict, an assumption that is essential for their results. Unfortunately, priority in college admission in China is strict. Therefore, these theoretical papers leave us with a mystery for the Chinese case.

Our paper is related to four previous papers using structural models to compare the two mechanisms empirically, all of which found that BM is more efficient than DA. He (2017) studied the Chinese high school admission, using a model based on the assumption that students behave at Nash equilibrium. He found that both naïve and sophisticated students suffer if he switches BM to DA at equilibrium. However, multiple Nash equilibria may exist, harming the reliability of the results. The other three papers do not use the game structure.

[7]This is called the serial dictatorship mechanism.

Calsamiglia et al. (2017) analyzed public school admission in Barcelona, Spain, using a counterfactual analysis showing that average welfare decreases by 1020 euros from BM to DA. Agarwal and Somaini (Forthcoming) scrutinized public elementary schools admission in Cambridge, MA. They found that the Cambridge controlled choice mechanism (a variant of BM) performed better than DM. Hwang (2016) proved that both naïve and sophisticated students follow a simple rule, a rule he uses to partially identify the model. In his empirical application, he found that ex-ante welfare is higher in BM.

However, all these papers rely on individual preference data, which are highly confidential and unavailable in our case. Thus, departing from the literature, we created a BM model that does not need micro data, only the admission quotas and cutoff thresholds of the colleges, which are public and can be found in college application guides. The looser requirements of the data can expand the future usability of the model.

3. Model

3.1. Boston Mechanism

Suppose that there are L colleges in China. College l has quota A_l for a given province[8]. This means that it can admit up to A_l students. Students are ranked; student i is the ith-highest-ranked student. We observe that the lowest-ranked student admitted by college l is student N_l. Thus, college l has a cutoff threshold (at student) N_l, with no lower-ranked students admitted.

[8] As described in the Appendix A (page 30), we consider the admission process for each province independently, so the model is for one province only. We consider only Round 1 admission.

Each college l has a fixed effect ξ_l. This effect is the attractiveness of the college, and is public information. Each student i has a private preference in relation to each college ε_{il}. The students cannot observe others private preferences; thus, for them, others private preference is random but its distribution remains common knowledge. If student i is able to be admitted by the college l, she receives $\xi_l + \varepsilon_{il}$ in utility. However, the student may also be rejected. College admission is quite competitive in China, and all spaces at most colleges are filled in the first step of BM. If a student is rejected in the first step, she may be either rejected overall or admitted by a much lower-ranked college in this round. We assume that she gets zero utility if she is rejected in the first step of BM. If student i has a $\mathbb{P}_l^a(i, \boldsymbol{A})$ chance to be admitted by college l, the expected utility for her of listing college l as her first choice is $(\xi_l + \varepsilon_{il})\mathbb{P}_l^a(i, \boldsymbol{A})$. If the admission probability becomes too small ($\mathbb{P}_l^a(i, \boldsymbol{A}) < \alpha$), the student will not consider this college. In addition, $l = 0$ is the students outside option. Student i getting ε_{i0}[9] from this option. Therefore, each student i solves a maximization problem[10],

$$\max_l \begin{cases} (\xi_l + \varepsilon_{il})\mathbb{P}_l^a(i, \boldsymbol{A}) - \inf \mathbb{1}\Big(\mathbb{P}_l^a(i, \boldsymbol{A}) < \alpha\Big), l \geq 1 \\ \varepsilon_{il}, l = 0 \end{cases} \tag{1}$$

where α is a small positive number. We emphasize that students are unable to observe N_l when they submit their preferences. We also solve $\mathbb{P}_l^a(i, \boldsymbol{A})$ recursively.

Lemma 1. $\mathbb{P}_l^a(i, A_l; A_{-l})^{11} = \mathbb{P}_l^a(i-1, A_l; A_{-l})(1-\mathbb{P}_l^c(i-1, \boldsymbol{A}))+\mathbb{P}_l^a(i-1, A_l-1; A_{-l})\mathbb{P}_l^c(i-$

[9] $\varepsilon_{i0} = 0$ is the most realistic outcome here, because a student should receive the same utility for rejection and for the outside option. To make the model more generic, we also allow ε_{i0} to be a random variable in this section.

[10] We define $\inf \times 0 = 0$.

[11] $\mathbb{P}_l^a(i, A_l; A_{-l}) = \mathbb{P}_l^a(i, \boldsymbol{A})$

$1, \boldsymbol{A})$ *for* $i \geq 2$ *and* $A_l \geq 1$ *where* $\mathbb{P}_l^c(i, \boldsymbol{A})$ *is the chance of student* i *choosing college* l*. In addition,* $\mathbb{P}_l^a(1, A_l; A_{-l}) = 1$ *for* $A_l \geq 1$ *and* $\mathbb{P}_l^a(i, 0) = 0$ *for* $i \geq 1$*.*

The proof is in Appendix B.1 of page 32. Intuitively, for instance, suppose that college l admits one student ($A_l = 1$). Then, we can simplify the equation as

$$\mathbb{P}_l^a(i, 1; A_{-l}) = \mathbb{P}_l^a(i-1, 1; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) = \prod_{j=1}^{i-1} \Big(1 - \mathbb{P}_l^c(j, \boldsymbol{A})\Big) \tag{2}$$

In other words, the probability of student i being admitted equals the probability of the students ranked higher than i not choosing college l if this college admits only one student.

In Lemma 1, we treat $\mathbb{P}_l^c(j, \boldsymbol{A})$ $\forall j < i$ as given. The admission probability $\mathbb{P}_l^a(i, \boldsymbol{A})$ is recursively determined by the first $i-1$ students choice probabilities. So is student is choice. This does not mean that student is choice depends on the first $i-1$ students choices; since ε_{il} is private and independent, the choices are also independent.

We now want to estimate the colleges attractiveness ξ_l. Then we will know the preferences of the students. The next theorem establishes mapping from the quota $\boldsymbol{A}$ and the cutoff threshold $\boldsymbol{N}$ to $\boldsymbol{\xi}$

Theorem 1. *For all* l, $A_l/N_l - 1/N_l \sum_{i=1}^{N_l} \mathbb{P}_l^c(i, \boldsymbol{A}) \overset{a.s.}{\to} 0$

The proof is provided in Appendix B.2 on page 33. Intuitively A_l/N_l is from the actual choices, while $1/N_l \sum_{i=1}^{N_l} \mathbb{P}_l^c(i, \boldsymbol{A})$ is from the expected choices. In the large sample, the two terms are equivalent due to the Law of Large Numbers. This theorem links $\boldsymbol{A}$, $\boldsymbol{N}$ and $\boldsymbol{\xi}$ in the large sample. Thousands or even tens of thousands of students are admitted in Round 1. The assumption of the large sample is valid. $\boldsymbol{\Psi} : \boldsymbol{\xi}, \boldsymbol{A} \to \boldsymbol{N}$ denotes the mapping from $\boldsymbol{\xi}$ and $\boldsymbol{A}$ to $\boldsymbol{N}$, while $\boldsymbol{\Psi^{-1}} : \boldsymbol{N}, \boldsymbol{A} \to \boldsymbol{\xi}$ denotes the mapping from $\boldsymbol{N}$ and $\boldsymbol{A}$ to $\boldsymbol{\xi}$. Notably,

$\boldsymbol{N}$ are assumed to be integers in the theorem. If $\boldsymbol{N}$ are not only integers, we extend the theorem to be

$$\frac{A_l}{N_l} - \frac{1}{N_l}\Big(\sum_{i=1}^{\lfloor N_l \rfloor} \mathbb{P}_l^c(i, \boldsymbol{A}) + (N_l - \lfloor N_l \rfloor)\mathbb{P}_l^c(\lfloor N_l \rfloor + 1, \boldsymbol{A})\Big) = 0, \forall l \tag{3}$$

where $\lfloor N_l \rfloor$ is the largest integer not larger than N_l. This ensures that both mappings will be continuous. We learn more features of the mappings in the following theorem.

Theorem 2 (Uniqueness)**.** *Given that $\boldsymbol{A}$ and $\boldsymbol{\xi}$, $\boldsymbol{N}$ are uniquely generated by $\boldsymbol{\Psi}$ if α is small enough.*

The proof is in Appendix B.3, on page 34.

Theorem 3 (Existence)**.** *Rank N_l from the smallest to the largest as $N_{(1)}$, $N_{(2)}$, ..., $N_{(L)}$. When α is small enough and given $\boldsymbol{A}$ and $\boldsymbol{N}$, (1) $\boldsymbol{\Psi}^{-1} = \varnothing$ if $\exists l$ $N_{(l)} \leq \sum_{j=1}^{l} A_{(j)}$, where $A_{(j)}$ is the admission quota of the college with the cutoff threshold $N_{(j)}$; (2) $\boldsymbol{\Psi}^{-1} \neq \varnothing$ if $N_l > \sum_{j=1}^{L} A_j$ for all l.*

The proof is in Appendix B.4 on page 35.

Theorem 4 (Non-Uniqueness)**.** *$\boldsymbol{\Psi}^{-1}$ may be multi-valued.*

Corollary 1. *For college l, its cutoff threshold N_l may increase (i.e. be looser) when it becomes more attractive.*

The proof is in Appendix B.5 on page 38.

Corollary 1 indicates that a college may admit worse students when it becomes more attractive, as students are more afraid to apply to an attractive college since they know their admission probabilities are low. This results in the quota being more difficult to fill. To the best of our knowledge, the literature has not addressed this property of BM. In fact, this property may be a major drawback, as the cutoff threshold is very important to a college as an indicator of its quality. Change in the line may affect funding and future students applications. In addition, worse students may also find worse jobs after their graduation, also affecting the reputation of the college. Thus, colleges want to have stricter cutoff thresholds (i.e. N_l is smaller). Therefore, if a college is in the situation of Corollary 1, it may decrease its attractiveness (or quality) to reach the goal. In fact, if most colleges are in this situation, they may compete to lower their quality. Thus, BM may result in low college quality. We will check this in our empirical analysis.

Theorem 2 indicates that cutoff thresholds will not change if attractiveness and quotas stay the same. We will estimate attractiveness given certain quotas and cutoff thresholds; then, we can generate cutoff thresholds based on the quotas and the estimated attractiveness. The distance between the generated cutoff thresholds and the real cutoff thresholds is our goodness-of-fit.

However, we do not know much about $\boldsymbol{\Psi^{-1}}$. Theorem 3 demonstrates that $\boldsymbol{\Psi^{-1}}$ does not exist with the small $\boldsymbol{N}$ but does exist with the large $\boldsymbol{N}$; we do not know whether $\boldsymbol{\Psi^{-1}}$ exists if $\boldsymbol{N}$ is in-between. Even worse, Theorem 4 finds that $\boldsymbol{\Psi^{-1}}$ may be multi-valued. If we work on the model directly and estimate the attractiveness from $\boldsymbol{\Psi^{-1}}$, the computation is quite heavy. We do not know whether the colleges attractiveness is identified in most cases. The objective function based on $\boldsymbol{\Psi^{-1}}$ may be non-convex, since $\boldsymbol{\Psi^{-1}}$ may be multi-valued. In addition, we have more than 100 variables (i.e. $\boldsymbol{\xi}$) to estimate, while the function is highly non-linear. We will also show that the assumptions of contraction mapping fail in

Section 5.2. In the simplest case, we take ten values of each variable and look for the best fit of the objective function. We need to compute more than 10^{100} times. We propose a method to solve this problem in the following sections.

3.2. Deferred Acceptance Mechanism

DA is equivalent to the serial dictatorship mechanism in college admission in China, because a student has the same priority in any college (Abdulkadiroglu and Sönmez, 2003). Therefore, in DA, students are assigned to the best available colleges. For instance, if a student has preference $l_1 \succ l_2 \succ l_3 \succ l_4 \cdots$, and colleges l_1 and l_2 have admitted enough students prior to that student, while l_3 has not, in DA the student is assigned to l_3. Any mechanism assigning students to the best available colleges is equivalent to DA.

Let us construct an imaginary mechanism. This mechanism is the same as BM, except that the students are able to observe the cutoff thresholds *before* submitting their preferences, which is impossible in the real world. In this case, students will know which college will reject them based on their ranking. Thus, the students will list the best available colleges as their first choices, and they will be admitted by these colleges. This indicates that this imaginary mechanism is equivalent to DA, and so we can use this imaginary mechanism to study DA.

Theorem 1 requires independence among choices across students (i.e. $\forall l$, $\mathbb{1}_l(1)$, $\mathbb{1}_l(2)$, ..., $\mathbb{1}_l(N_l)$ are independent). We emphasize that this assumption still holds given freely available information on the cutoff thresholds. To see why, we denote whether a student i chooses college l as her first choice by $\mathbb{1}_l(i|\boldsymbol{N})$; $\boldsymbol{N}$ is a random vector determined by the actual choices of the students, meaning that the choice of a student may be affected by the choices of other students through $\boldsymbol{N}$. However, $\mathbb{P}_l^c(i, \boldsymbol{A}|\boldsymbol{N})\mathbb{P}(\boldsymbol{N}) = \mathbb{P}_l^c(i, \boldsymbol{A}, \boldsymbol{N}) = \mathbb{P}_l^c((i, \boldsymbol{A}|\boldsymbol{N}), \boldsymbol{N})$. $\mathbb{1}_l(i|\boldsymbol{N})$ and $\boldsymbol{N}$ are independent, so $\mathbb{1}_l(i|\boldsymbol{N})$ is independent with $\mathbb{1}_l(-i|\boldsymbol{N})$. Intuitively, we may provide fake cutoff thresholds $\boldsymbol{N}^f$ to the students; in such a situation, their choices will be unrelated due to the fake information, while the fake cutoff thresholds $\boldsymbol{N}^f$ may coincide with the real ones $\boldsymbol{N}$.

In this imaginary mechanism, we rank cutoff thresholds N_l from smallest to largest, as $N_{(1)}$, $N_{(2)}$, ..., $N_{(L)}$. The top $N_{(1)}$ students can be admitted by any college, and understand this (i.e. $\forall l$ $\mathbb{P}_l^a(i, \boldsymbol{A}) = 1$). Thus, they compare the utility from each of L colleges as well as from the outside option and choose the best one. From student $N_{(1)} + 1$ to student $N_{(2)}$, they will be rejected by college (1) but accepted by other colleges (i.e. $\forall l \neq (1)$ $\mathbb{P}_l^a(i, \boldsymbol{A}) = 1$ and $\mathbb{P}_{(1)}^a(i, \boldsymbol{A}) = 0$). Then, they compare the utility from each of the $L - 1$ colleges as well as from the outside option. ... Next, student $N_{(L-1)} + 1$ to student $N_{(L)}$ will be rejected by any college except college (L) (i.e. $\forall l \neq (L)$ $\mathbb{P}_l^a(i, \boldsymbol{A}) = 0$ and $\mathbb{P}_{(L)}^a(i, \boldsymbol{A}) = 1$); thus, they compare the utility from college (L) and from the outside option. They remaining students will be rejected by any college (i.e. $\forall l$ $\mathbb{P}_l^a(i, \boldsymbol{A}) = 0$), and fully understand this. They choose the outside option.

The general expression of this model is tedious. We only consider $\varepsilon_{il} + \gamma$ being i.i.d Extreme Value Type 1 distributed, where $\gamma = 0.5772156649...$ is the Euler constant. This ensures that the mean of ε_{il} is zero. From Theorem 1, we get the relationship between average behavior and mean behavior.

$$\frac{A_{(1)}}{N_{(1)}} - \frac{\exp\left(\xi_{(1)}\right)}{1+\sum\limits_{l=1}^{L}\exp\left(\xi_{(l)}\right)} \xrightarrow{a.s.} 0$$

$$\frac{A_{(2)}}{N_{(2)}} - \frac{N_{(1)}}{N_{(2)}}\frac{\exp\left(\xi_{(2)}\right)}{1+\sum\limits_{l=1}^{L}\exp\left(\xi_{(l)}\right)} - \frac{N_{(2)}-N_{(1)}}{N_{(2)}}\frac{\exp\left(\xi_{(2)}\right)}{1+\sum\limits_{l=2}^{L}\exp\left(\xi_{(l)}\right)} \xrightarrow{a.s.} 0$$

$$\dots$$

$$\frac{A_{(L)}}{N_{(L)}} - \frac{N_{(1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\sum\limits_{l=1}^{L}\exp\left(\xi_{(l)}\right)} - \frac{N_{(2)}-N_{(1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\sum\limits_{l=2}^{L}\exp\left(\xi_{(l)}\right)} - \dots - \frac{N_{(L)}-N_{(L-1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\exp\left(\xi_{(L)}\right)} \xrightarrow{a.s.} 0 \tag{4}$$

where A_l/N_l is the average behavior, while other terms are the mean behavior. Similar to the full model, the simplified model relates $\boldsymbol{\xi}$, $\boldsymbol{A}$ and $\boldsymbol{N}$ by Equation 4. $\boldsymbol{\Phi} : \boldsymbol{\xi}, \boldsymbol{A} \to \boldsymbol{N}$ denotes the mapping from $\boldsymbol{\xi}$ and $\boldsymbol{A}$ to $\boldsymbol{N}$, and $\boldsymbol{\Phi^{-1}} : \boldsymbol{N}, \boldsymbol{A} \to \boldsymbol{\xi}$ denotes the mapping from $\boldsymbol{N}$ and $\boldsymbol{A}$ to $\boldsymbol{\xi}$, in the simplified model.

Theorem 5. *Based on* $\boldsymbol{\Phi^{-1}}$

$$\xi_{(l)} = \log \frac{\left(1+\sum\limits_{k=l+1}^{L}\exp\left(\xi_{(k)}\right)\right)A_l}{N_{(l)} - \sum\limits_{k=1}^{l} A_{(k)}} \tag{5}$$

for all $l < L$ *and*

$$\xi_{(L)} = \log \frac{A_{(L)}}{N_{(L)} - \sum\limits_{k=1}^{L} A_{(k)}} \tag{6}$$

Theorem 6. *Rank* $A_l/\exp\left(\xi_l\right)$ *from smallest to largest as* $A_{(1)}/\exp\left(\xi_{(1)}\right)$, $A_{(2)}/\exp\left(\xi_{(2)}\right)$,

..., $A_{(L)}/\exp\left(\xi_{(L)}\right)$. *Based on* $\mathbf{\Phi}$

$$N_{(l)}=\sum_{k=1}^{l}A_{(k)}+\frac{1+\sum_{k=l+1}^{L}\exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l)}\right)}A_{(l)} \tag{7}$$

for all $l<L$ *and*

$$N_{(L)}=\sum_{k=1}^{L}A_{(k)}+\frac{1}{\exp\left(\xi_{(L)}\right)}A_{(L)} \tag{8}$$

The proof is provided in Appendix B.6 on page 41. The two theorems generate one-to-one mapping between $\boldsymbol{\xi}$ and $\boldsymbol{N}$ given $\boldsymbol{A}$. The calculation is simple. However, what is the relationship between BM and DA?

Theorem 7. $\Psi_l/A_l=\Phi_l/A_l$ *and* $\Psi_l^{-1}=\Phi_l^{-1}$ *assuming (1)* $N_l\to\infty$ *and* $A_l/N_l>0$ *for* $\forall l$*; (2)* $N_l/N_{l'}$ *is finite for any* l *and* l'*; (3)* α *is small enough; and (4)* ε_{il} *of* Ψ_l *has the same distribution with that of* Φ_l*.*

The proof is provided in Appendix B.7 on page 43. This theorem indicates that the two models are equivalent when $\boldsymbol{A}$ is large. Intuitively, there are two types of knowledge relevant here: the actual cutoff thresholds and the expected cutoff thresholds. We know the actual cutoff thresholds, while the students do not; however, the students can calculate the expected cutoff thresholds. The two types of knowledge are similar in the large sample. Most students know which colleges they can be admitted to, and only a small proportion of the students are guessing. Thus, BM collapses to DA. We emphasize that while Theorem 1 requires $\boldsymbol{N}$ to be large, the present theorem requires $\boldsymbol{A}$ to be large. For example, let us consider a college that plans to admit one student. After the admission, the admitted student is ranked 3000. In this case, $A_l=1$, which is small, while $N_l=3000$, which is large. This is common in college admission in China because the number of students and that of colleges are both large. Therefore, $\boldsymbol{A}$s being large is harder to hold. Nonetheless, the two models are similar even if $\boldsymbol{A}$ is not large, based on the theorem. We will employ this similarity to estimate BM with the assistance of DA.

4. Estimation and Simulation

4.1. Estimation

We cannot simply use $\boldsymbol{\Phi^{-1}}$ to approximate $\boldsymbol{\Psi^{-1}}$, for three reasons. First, $\boldsymbol{A}$ being large is unrealistic in most situations, and the two models are not equivalent in a finite sample. In addition, $\varepsilon_{i0} = 0$ is most reasonable in BM, while $\varepsilon_{i0} + \gamma$ is Extreme Value Type 1 distributed in the closed-form solution of DA. Thus, the specifications of the two models are not exactly the same. Furthermore, one purpose of the paper is to compare BM with DA, and so it is unreasonable to assume that the two models are equivalent in the beginning. Instead, we propose a finite quota remedy (FQR) procedure to estimate $\boldsymbol{\Psi^{-1}}$.

Step 1: Get $\hat{\boldsymbol{\xi}} = \boldsymbol{\Phi^{-1}}(\boldsymbol{N}, \boldsymbol{A})$. Here, $\boldsymbol{A}$ is the real admission quotas and $\boldsymbol{N}$ is initialized as the real cutoff thresholds. $\hat{\boldsymbol{\xi}}$ is the estimated attractiveness of the DA model. The calculation of $\hat{\boldsymbol{\xi}}$ is simple due to the closed-form expression of $\boldsymbol{\Phi^{-1}}$.

Step 2: Get $\hat{\boldsymbol{N}} = \boldsymbol{\Psi}(\hat{\boldsymbol{\xi}}, \boldsymbol{A})$. We generate a new set of cutoff thresholds $\hat{\boldsymbol{N}}$ based on the estimated attractiveness from the last step and the BM model. If the two models are not equivalent, $\hat{\boldsymbol{N}} \neq \boldsymbol{N}$ and thus $\boldsymbol{\Psi^{-1}}(\boldsymbol{N}, \boldsymbol{A}) \neq \hat{\boldsymbol{\xi}}$

Step 3: Set $\boldsymbol{N}_\zeta = \boldsymbol{N} + \zeta(\boldsymbol{N} - \hat{\boldsymbol{N}})$

Step 4: Get $\boldsymbol{\xi}_\zeta = \boldsymbol{\Phi^{-1}}(\boldsymbol{N}_\zeta, \boldsymbol{A})$

Step 5: Calculate the distance between $\boldsymbol{\Psi}(\boldsymbol{\xi}_\zeta, \boldsymbol{A})$ and $\boldsymbol{N_0}$, and choose the best ζ as ζ^*. Here, $\boldsymbol{N_0}$ is the real cutoff thresholds. Steps 3-5 yield the best $\boldsymbol{\xi}_\zeta$ by modifying $\boldsymbol{N}$ with the direction $\boldsymbol{N} - \hat{\boldsymbol{N}}$, in the light of the line search. $\boldsymbol{\xi}_\zeta$ is weakly better than $\hat{\boldsymbol{\xi}}$, because $\boldsymbol{\xi}_0 = \hat{\boldsymbol{\xi}}$.

Step 6: Set new $\boldsymbol{N} = \boldsymbol{N}_{\zeta^*}$ and go to Step 1.

In Step 5, we do not need perfect optimization; in practice, we randomly select 11 ζs within $[-1, 1]$ and choose the best one as ζ^*. The procedure stops after 100 iterations; we record $\boldsymbol{\xi}_{\zeta^*}$ in each iteration, and the best one is denoted by $\boldsymbol{\xi}^*$.

4.2. Simulation

In this section, we will study the performance of FQR and compare it with $\boldsymbol{\Phi^{-1}}$, which uses the DA model directly to approximate BM. In Figure 2(a) on page 48, we suppose that there are two colleges: $\xi_1 = 3$ for college 1 while $\xi_2 = 5$ for college 2. Each college admits the same number of students (i.e. $A_1 = A_2$.). We generate their cutoff thresholds using BM for the different quotas (i.e. $A_1 = A_2 = 1, 2, \ldots, 50$.); then, we use the cutoff thresholds $\boldsymbol{N}$ and the quotas $\boldsymbol{A}$ to estimate the attractiveness $\boldsymbol{\xi}$ by either $\boldsymbol{\Phi^{-1}}$ or FQR.

We find FQR outperforming the DA model. FQR works well even when the quotas are small (e.g. $A_1 = A_2 = 5$.), while $\boldsymbol{\Phi^{-1}}$ does not perform well when $\boldsymbol{A}$ are small. The model works better when colleges admit more students, which coincides with Theorem 7. If we suppose that there are three, four, or five colleges instead of two colleges, they all get similar results to Figure 2(a), as reported in Figure 2(b)-2(d).

5. Empirical Analysis

5.1. Data

We collect data for Guangxi, Hebei, and Sichuan from different sources. In Guangxi, the Guangxi Provincial Academy of Recruitment and Examination (Gvangjsih Cauhswngh Gaujsi Yen in Zhuang language) composed guides for the college entrance examination ("Gaokao Zhinan") in 2007, 2008, and 2009. These guides include the quota for each college, the lowest score for the students admitted to each college, and the number of students achieving each score in 2006, 2007, and 2008. We calculate the cutoff threshold for each college from the lowest score for admitted students to each college and the number of students on each score. In addition, admission is divided into 11 rounds. The first four rounds are Round 0, a round for arts and physical education, Round 1, and Round 1 for college-preparatory education. Only a small proportion of the students are eligible to apply to colleges in the second and the fourth round, while the choices of major and college are limited in Round 0. Thus, most highly ranked students will apply in in Round 1. In this paper, we combine Round 1 and Round 1 for college prep into one round and study this round only. We also assume that all highly ranked students will apply to college in this round.

In Hebei, the Hebei Education Examinations Authority compiled "Statistics of Admission Score Distribution in Hebei of China's Colleges and Universities from 2005 to 2007" ("Quanguo Putong Gaoxiao zai Hebei Zhaosheng Luqu Fenshu Fenbu Tongji (2005-2007)"). These statistics include the quota for each college and the lowest score of the students admitted to each college in 2005, 2006, and 2007. Unfortunately, we miss page 166 of this statistics for science major students, and so we do not have the quota or the lowest score of the students admitted to China University of Mining in 2007 or China University of Mining (Beijing) in 2005. We assume therefore that these two colleges did not admit students for given years. In addition, the lowest score of students admitted to Xi'an International Studies University was 570 for science major students in 2007, lower than the key cutoff threshold (587). This may be an error in the data; regardless, we address it by again assuming that this college did not admit science major students in 2007. Since in fact Xi'an International Studies University only admitted five science majors from Hebei that year, this presumption does not significantly affect the results. We also collect the number of students achieving each score in these three years from Hengshui High School. In admissions, the first three rounds are Round 0, Round 1A, and Round 1B; most highly ranked students apply in Rounds 1A and 1B. Because Round 1B is conducted after the completion of Round 1A, the existence of Round 1B does not affect the study of Round 1A. Therefore, in this paper, we only analyze Round 1A.

In Sichuan, the Sichuan Recruitment and Examination Information Co. (Sichuan Zhaosheng Kaoshi Xinxi Zixun Youxiangongsi), a state-owned enterprise supervised by the Sichuan Educational Examination Authority, composed guides for the college entrance examination ("Gaokao Zhinan") in 2007 and 2008. These guides include the quota for each college, the lowest score of students admitted to each college, and the number of students achieving each score, presented in five-score increments, in 2006 and 2007. In the data, the cutoff threshold ($N_{(l)}$) of a given college may be smaller than the sum of the quotas of the colleges with cutoff threshold stricter than that college (i.e. $\sum_{j=1}^{l} A_{(j)}$), leading to non-existent results (i.e. $\mathbf{\Psi^{-1}} = \varnothing$ and $\mathbf{\Phi^{-1}} = \varnothing$) based on Theorem 3 and 5. This may be caused by errors in the data and/or our simplifications. To solve this issue, we presume that the top one or two colleges in terms of cutoff threshold do not admit students. This approach does not significantly affect the results, since these colleges do not admit many students, but it ensures that $N_{(l)} > \sum_{j=1}^{l} A_{(j)}$ for $\forall l$. Specifically, we presume that the Chinese University of Hong Kong and Peking University did not admit arts majors from Sichuan in 2006, while in reality they admitted 1 and 31, respectively; that Tsinghua University did not admit science majors in 2006, while in fact it admitted 78 students; that Tsinghua University and Peking University did not admit arts majors in 2007, while in fact they admitted 12 and 29, respectively; and that Tsinghua University did not admit science majors in 2007, while in fact it again admitted 78.

In addition, the lowest score of Sichuan science majors admitted to Tianjin University was given as 518 in 2006, much lower than the key cutoff (560). This is an error; based on other information provided in the guide, this score should be between 618 and 619, so we corrected it to 618 from 518. In the admission process, the first two rounds are Round 0 and Round 1; we only consider Round 1 in this paper.

The student placement office in each province conducts admissions for science and arts majors separately; thus, we analyze the two groups separately in the model. Furthermore, in the years under consideration, students in all three provinces received their exam scores and the distribution of the scores before applying to college.

5.2. *Results*

The results are similar for all three provinces; thus, we only report the results for science majors in Guangxi for 2008 here (the other results are reported in Table 1-4 on Page 52-54). For the BM model, Figure 3 on page 49 presents the top ten colleges in terms of attractiveness ξ_l. The attractiveness of Tsinghua University and Peking University, the top two colleges on the Chinese mainland, is much higher than that of all other colleges. Most students will definitely choose one of these two if they have a reasonable chance to be admitted. In addition, the attractiveness of all colleges other than the top seven is negative. We emphasize that attractiveness is the average preference of the students; if a given student chooses a college, she must receive non-negative utility from it, because she receives zero from the outside option. A student may be only interested in some small number of colleges (such as high-ranking ones) and thus may receive positive utility only from these colleges. This leads average preference (or attractiveness) to be negative for most colleges.

Then, we simulate and compare BM and DA based on estimated attractiveness. The cumulative welfare change from BM to DA is reported in Figure 4, on page 49, and the individual change is reported in Figure 5, on page 50. In these two graphs, the x-axis is the rank of a student while the y-axis is the average utility change for all students ranked weakly better than the student (Figure 4) or the utility change of this student (Figure 5). The two graphs both start from 0% on the x-axis. The two mechanisms are equivalent for the top students, who are able to choose any college without worrying about rejection. In addition, the two graphs also show that well-ranked students benefit from the switch while badly ranked students suffer from it. However, Figure 7 on page 51 indicates further that only 129 (0.9%) students in fact benefit from the switch, while 14241 students are above the key cutoff threshold. All of them are among the top 197 students, as seen in Figure 5. Social welfare would increase after the switch only if fewer than 335 students exist, as suggested in Figure 4.

In the two graphs, "A" is the sum of the quotas of all colleges, while "L" is the key cutoff threshold. In reality, students will be considered for admission in this round only when they have scores higher than that of the key cutoff threshold. Therefore, the welfare of all eligible students decreases 1.96% after the switch from BM to DA. If we assume that all students can apply to these colleges, their utility decreases 2.49% after the switch. We emphasize here that the former estimate (1.96%) underestimates the real welfare loss, because we only consider the first step of BM in our model, whereas rejected in the first step may be also admitted in the second step. Thus, students may receive higher utility in real BM than that in our BM model. In the latter estimate (2.49%), colleges admit enough students under our BM model, which then collapses to real BM.

In Figure 6, on page 50, the cutoff thresholds of all colleges become stricter after the switch. This is a reason for welfare loss: a student who can be admitted under BM may be rejected under DA. As seen in Figure 4 and 5, students who receive scores a little bit higher than the key cutoff threshold suffer most from the switch. In Guangxi (as well as in Sichuan), the quality of colleges in this round is much higher than that in the latter rounds; students around the key cutoff threshold can be admitted under BM but not under DA because of the stricter cutoff, so they receive much lower utility under DA.

In Figure 5, the results are noisy for the bottom students. The two mechanisms are equivalent for these students because the bottom students receive rejection and zero utility in both mechanisms. The relative error of our simulation becomes larger when utility is close to zero, which contributes to the noisiness of the results.

We now check whether the problem indicated in Corollary 1 exists in college admission in China. We find that the cutoff threshold of each college becomes looser when the colleges attractiveness decreases, except for science majors in Guangxi applying to Tsinghua University in 2008. In other words, Tsinghua University can attract better students after its attractiveness (or quality) decreases for this given province and year. The reason is that Tsinghua is one of the top colleges in China; most students realize that the probability of being admitted by Tsinghua is low, and apply for the alternatives instead. If Tsinghua lowers its quality, students realize that the competition to enter will be milder and are more likely to apply there. Although this situation is possible, it is rare. For most colleges, students are less likely to apply for them when their attractiveness decreases, because the students would receive lower utility from the college. Thus, due to the rareness of this situation, it does not reduce the superiority of BM in college admission in China. Tsinghua University would not plausibly choose to decrease its quality only to lure students in from only one province.

Finally, we explain why we are unable to directly use contraction mapping in the estimation. In Appendix I of Berry et al. (1995), the proof of the contraction mapping requires the same sign for all $\partial \log(\Psi_l)/\partial \xi_{l'}$ where $l \neq l'$. However, our results indicate that this assumption fails in all cases.

6. Conclusion

In this paper, we simulated and compared the empirical performance of BM and DA in college admission in China. We constructed a model of BM and employed it to estimate the attractiveness of Chinese colleges in three Chinese provinces. Then, we conducted counterfactuals to empirically compare BM and DA in these three provinces for given years. We find that not only is BM superior to DA in terms of total welfare, but also that most students suffer from the switch from BM to DA.

This paper makes the following contributions. First and most importantly, this paper shows that BM is a better approach than DA to college admission in China from a social welfare perspective. Historically, BM was implemented in all the provinces of the Chinese mainland, whereas currently DA is employed by most provinces. The results indicate that this switch from BM to DA has been costly: the total welfare of students has decreased 1.73% - 6.63% due to the switch. On the Chinese mainland, there are around 150 colleges which admit students in Round 1 (aka key colleges). In the case of a 1.73% - 6.63% welfare loss, these colleges need to improve their quality by 1.73% - 6.63% to compensate, equivalent to constructing 2.595-9.945 more key colleges, assuming that the cost of each unit of the quality is the same. Further, if we assume that one key college is worth 1 billion dollars, the switch costs 2.595-9.945 billion dollars.

Second, departing from the literature, our model does not need micro data on the submitted preferences of individual students; instead, it requires only the admission quota and cutoff threshold for each college. Micro data are difficult to obtain and may be restricted for privacy reasons; our use of public data makes our results easier to replicate and makes the model potentially more widely usable.

Third, theoretically, we have proven that the performances of DA and BM are equivalent when admission quotas $\boldsymbol{A}$ are large enough. As far as we know, this equivalence has not been pointed out in the literature. We also learn from this equivalence that DA and BM are similar in most situations; for example, in our case, the difference in total welfare of these two mechanisms is only 1.73% - 6.63%, although the difference is significant when applying to the ten million students. In addition, currently each province in China conducts the admission separately. A college may admit a few students in a province. If all the provinces conduct the admission together, the quota of a college will be the sum of its quotas in these provinces. Thus $\boldsymbol{A}$ will become larger and the theorem indicates that the performances of BM and DA will be closer. Therefore, conducting BM nationwide may weaken its superiority.

Fourth, we point out a potential drawback of BM: under BM, a college may need to lower its quality to attract better students, and if most colleges are in this situation, BM may result in lower quality of higher education. Future studies may want to investigate this drawback when analyzing the performance of BM.

Fifth, we derived a closed-form expression of the DA model. Given the quotas and cutoff thresholds, we can then derive attractiveness, and given the attractiveness and the quotas we can predict cutoffs. Currently, most Chinese provinces use DA for college admission, as noted; we can predict cutoff thresholds based on the previous years cutoffs and the present and previous years admission quotas. As far as we know, this is the first study allowing such a prediction in the Chinese market based on a solid model.

To sum up, in this paper, we find that BM performs better than DA in college admission in China. However, is BM the best possible mechanism? Future studies may want to propose new mechanisms better than BM or conversely to prove that BM yields the highest welfare in college admission in China.

References

Abdulkadiroğlu, A., Che, Y.-K., Yasuda, Y., 2011. Resolving conflicting preferences in school choice: The boston mechanism reconsidered. The American Economic Review 101, 399–410.

Abdulkadiroğlu, A., Che, Y.-K., Yasuda, Y., 2015. Expanding choice in school choice. American Economic Journal: Microeconomics 7, 1–42.

Abdulkadiroglu, A., Sönmez, T., 2003. School choice: A mechanism design approach. The American Economic Review 93, 729–747.

Agarwal, N., Somaini, P., Forthcoming. Demand analysis using strategic reports: An application to a school choice mechanism. Econometrica .

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica: Journal of the Econometric Society pp. 841–890.

Calsamiglia, C., Fu, C., Güell, M., 2017. Structural estimation of a model of school choices: the boston mechanism vs. its alternatives. Resubmitted at the Journal of Political Economy .

Chen, Y., Kesten, O., 2016. Chinese college admissions and school choice reforms: An experimental study. Tech. rep., Working Paper, University of Michigan.

Chen, Y., Kesten, O., 2017. Chinese college admissions and school choice reforms: A theoretical analysis. Journal of Political Economy 125, 000–000.

Dubins, L. E., Freedman, D. A., 1981. Machiavelli and the gale-shapley algorithm. The American Mathematical Monthly 88, 485–494.

Ergin, H., Sönmez, T., 2006. Games of school choice under the boston mechanism. Journal of public Economics 90, 215–237.

Fonda, A., Gidoni, P., 2016. Generalizing the poincaré–miranda theorem: the avoiding cones condition. Annali di Matematica Pura ed Applicata (1923-) 195, 1347–1371.

Gale, D., Shapley, L. S., 1962. College admissions and the stability of marriage. The American Mathematical Monthly 69, 9–15.

Haeringer, G., Klijn, F., 2009. Constrained school choice. Journal of Economic theory 144, 1921–1947.

He, Y., 2017. Gaming the boston school choice mechanism in beijing .

Hwang, I. M., 2016. A robust redesign of High School Match. Ph.D. thesis, THE UNIVERSITY OF CHICAGO.

Kallenberg, O., 1997. Foundations of modern probability .

Li, F., Gan, L., Yang, X., 2010. Research on application timing and admission mechanism in national college entrance exam. Educational Research 10.

Ohio Supercomputer Center, 1987. Ohio supercomputer center. http://osc.edu/ark:/19495/f5s1ph73.

Pathak, P. A., Sönmez, T., 2013. School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. The American Economic Review 103, 80–106.

Wu, B., Zhong, X., 2014. Matching mechanisms and matching quality: Evidence from a top university in china. Games and Economic Behavior 84, 196–215.

Zhu, M., 2014. College admissions in china: A mechanism design perspective. China Economic Review 30, 618–631.

Appendix A. Boston Mechanism, Deferred Acceptance Mechanism, Serial Dictatorship Mechanism And Their Variants in China's College Admission

A.1. Boston Mechanism

Step 1 College l has quota A_l^1. The student placement office sends each college a list containing all students who choose the college as their first choice[12]. If the list contains more than A_l^1 students, college l admits the top A_l^1 students and rejects the remaining students. The quota for the next step (A_l^2) is zero. Otherwise, college l admits all the students on the list and the remaining quota is A_l^2.

Step k College l has quota A_l^k. The student placement office sends each college a list containing all students who are rejected in Step $k-1$ and choose the college as their kth choice. If the list contains more than A_l^k students, college l admits the top A_l^k students and rejects the remaining students. The quota for the next step (A_l^{k+1}) is zero. Otherwise, college l admits all the students on the list and the remaining quota is A_l^{k+1}

The mechanism stops when all lists are blank.

A.2. Deferred Acceptance Mechanism

Step 1 College l has quota A_l. The student placement office sends each college a list containing all students who choose the college as their first choice. If the list contains more than A_l students, college l *tentatively* admits the top A_l students and rejects the remaining students. Otherwise, college l *tentatively* admits all the students on the list.

Step k The student placement office sends each college a list containing all students who are rejected in Step $k-1$ and choose the college as their kth choice. College l compares the students on the list and the ones that have been *tentatively* admitted. If there are more than A_l students, college l *tentatively* admits the top A_l students and rejects the remaining students. Otherwise, college l *tentatively* admits all the students.

The mechanism stops when all lists are blank. All *tentatively* admitted students are confirmedly admitted.

A.3. Serial Dictatorship Mechanism

Step 1 College l has quota A_l. The student placement office sends the information of the top student to her first choice - college l_1^1. Since $A_{l_1^1} > 0$, the student will be admitted by the college and the remaining quota of this college for the next step is $A_{l_1^1}^2 = A_{l_1^1} - 1$.

[12]In practice, the office does not send the full list to the colleges. Instead, it sends a list containing slightly more than A_l^1 students to college l if more than A_l^1 students choose college l as their first choice.

Step k The student placement office sends the information of the kth ranked student to her first choice - college l_k^1. If this college has admitted enough student (that is $A_{l_k^1}^k = 0$), the student will be rejected and the office sends her information to her second choice. If she is rejected again, the office sends her information to her next choice. If she is admitted by her h choice - college l_k^h, the remaining quota of this college for the next step is $A_{l_k^h}^{k+1} = A_{l_k^h}^k - 1$. If she is rejected by all the choices on her preference list, she is rejected in this round.

The mechanism stops when the office sends the information of all the students to the colleges. Deferred acceptance mechanism (DA) is equivalent to serial dictatorship mechanism (SD) because the students are assigned to the best available college based on their preference lists in both mechanisms. We will use DA and SD interchangeably in this paper, while SD is actually implemented in college admission in China.

A.4. Variants in China's College Admission

In China, college admission has several rounds. In each round, a mechanism is applied. Most good colleges are in and only in Round 1, and most highly ranked students apply colleges in Round 1. Thus, we only consider Round 1 for simplicity. In addition, the ranking is strict. If one student has higher total score than the other, she is ranked higher than the other. If two students have the same total score, their scores of each part are compared to break the tie.

The students submit their preferences at different times in different provinces. In some provinces, they shall submit them before taking the exam. In some provinces, they shall submit them after the exam but before the ranking is published. In other provinces, they shall submit them after the ranking is published. In this paper, we only consider the provinces in the last case. Moreover, the students shall not submit the full list of their preference. They can submit their first two to eighty choices depending on the province.

The mechanism also depends on the province. It can be either the Boston mechanism (BM), DA or a mixture of these two. For example, some provinces apply BM in the first step and apply the DA in the following steps. The literature (Haeringer and Klijn, 2009; Wu and Zhong, 2014) indicates that the first choice is the most important choice in BM. We consider a mechanism as BM if this mechanism applies BM in its first step.

Moreover, the students submit their major preferences for each college that they choose. A college uses a mechanism to assign students to the majors. Students can reject the assignment but the rejection may result them getting to much worse colleges. Wu and Zhong (2014) indicate that almost all the students accept the assignment. Therefore, we can separate the admission into two stages. In the first stage, the students are assigned to the colleges, which are considered as composite goods. In the second stage, the students are assigned to the majors. This argument was proposed by Wu and Zhong (2014). In this paper, we only consider the first stage.

Appendix B. Proofs

B.1. Proof of lemma 1 on page 9

Lemma 1. $\mathbb{P}_l^a(i, A_l; A_{-l})$[13] $= \mathbb{P}_l^a(i-1, A_l; A_{-l})(1-\mathbb{P}_l^c(i-1, \boldsymbol{A}))+\mathbb{P}_l^a(i-1, A_l-1; A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A})$ *for* $i \geq 2$ *and* $A_l \geq 1$ *where* $\mathbb{P}_l^c(i, \boldsymbol{A})$ *is the chance of student* i *choosing college* l*. In addition,* $\mathbb{P}_l^a(1, A_l; A_{-l}) = 1$ *for* $A_l \geq 1$ *and* $\mathbb{P}_l^a(i, 0) = 0$ *for* $i \geq 1$*.*

Proof. The student i does not need to consider the choices of the students ranked lower than her. She considers the first $i-1$ students' choices. $\mathbb{P}_l^o(i, k, A_{-l})$ denotes the probability of k slots of the school l having been taken by the first $i-1$ students. We decompose $\mathbb{P}_l^a(i, A_l; A_{-l})$ as

$$\mathbb{P}_l^a(i, A_l; A_{-l}) = \sum_{k=0}^{A_l-1} \mathbb{P}_l^o(i, k, A_{-l}), A_l \geq 1 \tag{9}$$

If the college l admits the student i, the quota must be not filled up by the first $i-1$ students. This case can be broken down into that the first $i-1$ students take zero slot, that these students take one slot, ..., that these students take $A_l - 1$ slots. This derives Equation 9.

In addition, we can express $\mathbb{P}_l^o(i, k, A_{-l})$ as

$$\mathbb{P}_l^o(i, k, A_{-l}) = \begin{cases} \mathbb{P}_l^o(i-1, k, A_{-l})(1-\mathbb{P}_l^c(i-1, \boldsymbol{A})); k = 0, i \geq 2 \\ \mathbb{P}_l^o(i-1, k, A_{-l})(1-\mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^o(i-1, k-1, A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A}); k > 0, i \geq 2 \end{cases} \tag{10}$$

If none of the first $i-2$ students chooses the college l and the student $i-1$ does not choose the college l, none of the first $i-1$ students chooses this college. If (1) k of the first $i-2$ students choose the college l and the student $i-1$ does not choose the college l *or* (2) $k-1$ of the first $i-2$ students choose the college l and the student $i-1$ chooses the college l, k of the first $i-1$ students choose this college. These derive Equation 10.

We start from the initial conditions. If the college l does not admit any student, the chance to be admitted is zero. We have

$$\mathbb{P}_l^a(i, 0; A_{-l}) = 0, \forall i \geq 1 \tag{11}$$

In addition, if the college l has a positive quota, it is impossible for the college to reject the top ranked student. We have

$$\mathbb{P}_l^a(1, A_l; A_{-l}) = 1, \forall A_l \geq 1 \tag{12}$$

[13] $\mathbb{P}_l^a(i, A_l; A_{-l}) = \mathbb{P}_l^a(i, \boldsymbol{A})$

For the case that $i \geq 2$ and $A_l \geq 2$, we have

$$\begin{aligned}
\mathbb{P}_l^a(i, A_l; A_{-l}) =& \sum_{k=0}^{A_l-1} \mathbb{P}_l^o(i, k, A_{-l}) \\
=& \sum_{k=1}^{A_l-1} \mathbb{P}_l^o(i, k, A_{-l}) + \mathbb{P}_l^o(i, 0, A_{-l}) \\
=& \sum_{k=1}^{A_l-1} \mathbb{P}_l^o(i-1, k, A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) \\
&+ \sum_{k=0}^{A_l-2} \mathbb{P}_l^o(i-1, k, A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A}) + \mathbb{P}_l^o(i-1, 0, A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^a(i-1, A_l - 1; A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A})
\end{aligned} \tag{13}$$

For the case that $i \geq 2$ and $A_l = 1$, we have

$$\begin{aligned}
\mathbb{P}_l^a(i, A_l; A_{-l}) =& \mathbb{P}_l^o(i, 0, A_{-l}) \\
=& \mathbb{P}_l^o(i-1, 0, A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^a(i-1, A_l - 1; A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A})
\end{aligned} \tag{14}$$

□

B.2. Proof of theorem 1 on page 10

Theorem 1. *For all l, $A_l/N_l - 1/N_l \sum_{i=1}^{N_l} \mathbb{P}_l^c(i, \boldsymbol{A}) \xrightarrow{a.s.} 0$*

Proof. $\mathbb{1}_l(i)$ is an indicator function. It denotes whether the student i chooses the college l as her first choice. If $\mathbb{1}_l(i) = 1$ she chooses the college l. If $\mathbb{1}_l(i) = 0$, she does not choose the college. $\mathbb{1}_l(1)$, $\mathbb{1}_l(2)$,..., $\mathbb{1}_l(N_l)$ are independent random variables because the private preferences (i.e. ε_{il}) are independent. Her probability of choosing the college l is $\mathbb{P}_l^c(i, \boldsymbol{A})$. The mean of $\mathbb{1}_l(i)$ is $1 \times \mathbb{P}_l^c(i, \boldsymbol{A}) + 0 \times (1 - \mathbb{P}_l^c(i, \boldsymbol{A})) = \mathbb{P}_l^c(i, \boldsymbol{A})$. Let $\omega_l(i) = \mathbb{1}_l(i) - \mathbb{P}_l^c(i, \boldsymbol{A})$. Based on the *variance criterion for averages, Kolmogorov* in Corollary 3.22 of Kallenberg (1997), we have

$$\frac{1}{N_l} \sum_{i=1}^{N_l} \omega_l(i) \xrightarrow{a.s.} 0 \tag{15}$$

This equation holds if $\frac{1}{N_l^2}\sum_{i=1}^{N_l}\mathbb{E}\omega_l^2(i) < \infty$. In addition,

$$\begin{aligned}\omega_l^2(i) =&(\mathbb{1}_l(i) - \mathbb{P}_l^c(i,\boldsymbol{A}))^2\\ =&\mathbb{1}_l^2(i) + (\mathbb{P}_l^c(i,\boldsymbol{A}))^2 - 2\mathbb{1}_l(i)\mathbb{P}_l^c(i,\boldsymbol{A})\\ \leq&1+1+2=4\end{aligned} \tag{16}$$

Thus, $\frac{1}{N_l^2}\sum_{i=1}^{N_l}\mathbb{E}\omega_l^2(i) \leq \frac{1}{N_l^2}\sum_{i=1}^{N_l}4 = \frac{4}{N_l} < \infty$. Equation 15 holds.

The cutoff line is N_l. This also means that A_l of the top N_l students choose the college l. $\sum_{i=1}^{N_l}\mathbb{1}_l(i) = A_l$. We have

$$\frac{1}{N_l}\sum_{i=1}^{N_l}\omega_l(i) = \frac{1}{N_l}\sum_{i=1}^{N_l}\mathbb{1}_l(i) - \frac{1}{N_l}\sum_{i=1}^{N_l}\mathbb{P}_l^c(i,\boldsymbol{A}) = \frac{A_l}{N_l} - \frac{1}{N_l}\sum_{i=1}^{N_l}\mathbb{P}_l^c(i,\boldsymbol{A}) \xrightarrow{a.s.} 0 \tag{17}$$

□

B.3. *Proof of theorem 2 on page 11*

Theorem 2 (Uniqueness). *Given that $\boldsymbol{A}$ and $\boldsymbol{\xi}$, $\boldsymbol{N}$ are uniquely generated by $\boldsymbol{\Psi}$ if α is small enough.*

Proof. The relationship among $\boldsymbol{\xi}$, $\boldsymbol{A}$ and $\boldsymbol{N}$ is equivalent to

$$A_l - \Big(\sum_{i=1}^{\lfloor N_l\rfloor}\mathbb{P}_l^c(i,\boldsymbol{A})^{14} + (N_l - \lfloor N_l\rfloor)\mathbb{P}_l^c(\lfloor N_l\rfloor + 1,\boldsymbol{A})\Big) = 0, \forall l \tag{18}$$

and we let $f_l(N_l) = A_l - \Big(\sum_{i=1}^{\lfloor N_l\rfloor}\mathbb{P}_l^c(i,\boldsymbol{A}) + (N_l - \lfloor N_l\rfloor)\mathbb{P}_l^c(\lfloor N_l\rfloor + 1,\boldsymbol{A})\Big)$. When $\alpha = 0$, the students do not rule out a college even if the admission probability is small[15]. In this case, $\mathbb{P}_l^c(i,\boldsymbol{A})$ is strictly positive because the chance of $\varepsilon_{il'} < -\xi_{l'}$ for all $l' \neq \{0,l\}$ does not vanish. If $\varepsilon_{il'} < -\xi_{l'}$ ($\forall l' \neq \{0,l\}$), $\varepsilon_{i0} \leq 0$ and $\varepsilon_{il} > -\xi_l$, the student chooses the college l since she can get positive utility only from this college. Thus, $f_l(N_l)$ is a strictly decreasing function. $f_l(0) = A_l > 0$ and $f_l(\infty) \leq A_l - \infty \times \min_i \mathbb{P}_l^c(i,\boldsymbol{A}) < 0$. There is a unique N_{l0} such that $f_l(N_{l0}) = 0$.

Let $N^* = \lceil \max_l f_l^{-1}(0)\rceil + 1$ where $\lceil\bullet\rceil$ is the smallest integer not smaller than the $\bullet$. We

[14] $\mathbb{P}_l^c(i,\boldsymbol{A}) = \mathbb{P}_l^c(i,\boldsymbol{A},\boldsymbol{\xi})$, which depends on $\boldsymbol{\xi}$

[15] See details in Equation 1 on page 9

also know that $\mathbb{P}_l^a(i, A_l; A_{-l}) \leq \mathbb{P}_l^a(i-1, A_l; A_{-l}) \forall i > 1, A > 0$ because

$$\begin{aligned}
\mathbb{P}_l^a(i, A_l; A_{-l}) =& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^a(i-1, A_l - 1; A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A}) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^c(i-1, \boldsymbol{A}) \sum_{k=0}^{A_l-2} \mathbb{P}_l^o(i-1, k, A_{-l}) \\
\leq& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^c(i-1, \boldsymbol{A}) \sum_{k=0}^{A_l-1} \mathbb{P}_l^o(i-1, k, A_{-l}) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})(1 - \mathbb{P}_l^c(i-1, \boldsymbol{A})) + \mathbb{P}_l^a(i-1, A_l; A_{-l})\mathbb{P}_l^c(i-1, \boldsymbol{A}) \\
=& \mathbb{P}_l^a(i-1, A_l; A_{-l})
\end{aligned} \tag{19}$$

Thus we know that

$$\mathbb{P}_l^a(N^*, A_l; A_{-l}) \leq \mathbb{P}_l^a(i, A_l; A_{-l}) \forall i < N^* \forall l \tag{20}$$

Now we take a positive α such that $\alpha < \min_l \mathbb{P}_l^a(N^*, A_l; A_{-l})$. $\mathbb{P}_l^a(i, A_l; A_{-l})$ and $\mathbb{P}_l^c(i-1, \boldsymbol{A})$ are unchanged for $\forall i < N^*$ due to Equation 20. This also keeps all N_{l0} unchanged. Moreover, $f_l(N_l)$ is a non-increasing function when $\alpha > 0$. From our construction, $f_l(N^*) < 0$ for all l. So N_{l0} is still the unique solution. □

B.4. Proof of theorem 3 on page 11

Theorem 3 (Existence)**.** *Rank N_l from the smallest to the largest as $N_{(1)}$, $N_{(2)}$, ..., $N_{(L)}$. When α is small enough and given $\boldsymbol{A}$ and $\boldsymbol{N}$, (1) $\boldsymbol{\Psi^{-1}} = \varnothing$ if $\exists l$ $N_{(l)} \leq \sum_{j=1}^{l} A_{(j)}$, where $A_{(j)}$ is the admission quota of the college with the cutoff threshold $N_{(j)}$; (2) $\boldsymbol{\Psi^{-1}} \neq \varnothing$ if $N_l > \sum_{j=1}^{L} A_j$ for all l.*

Proof. If $\exists l$ $N_{(l)} \leq \sum_{j=1}^{l} A_{(j)}$, we add $f_{(1)}(N_{(1)})$[16], $f_{(2)}(N_{(2)})$, ..., $f_{(l)}(N_{(l)})$ up,

$$\begin{aligned}
\sum_{j=1}^{l} f_{(j)}(N_{(j)}) =& \sum_{j=1}^{l} A_{(j)} - \sum_{j=1}^{l} \Big(\sum_{i=1}^{\lfloor N_{(j)} \rfloor} \mathbb{P}_{(j)}^c(i, \boldsymbol{A}) + (N_{(j)} - \lfloor N_{(j)} \rfloor)\mathbb{P}_{(j)}^c(\lfloor N_{(j)} \rfloor + 1, \boldsymbol{A}) \Big) \\
\geq& \sum_{j=1}^{l} A_{(j)} - \sum_{j=1}^{l} \Big(\sum_{i=1}^{\lfloor N_{(l)} \rfloor} \mathbb{P}_{(j)}^c(i, \boldsymbol{A}) + (N_{(l)} - \lfloor N_{(l)} \rfloor)\mathbb{P}_{(j)}^c(\lfloor N_{(l)} \rfloor + 1, \boldsymbol{A}) \Big) \\
=& \sum_{j=1}^{l} A_{(j)} - \Big(\sum_{i=1}^{\lfloor N_{(l)} \rfloor} \sum_{j=1}^{l} \mathbb{P}_{(j)}^c(i, \boldsymbol{A}) + (N_{(l)} - \lfloor N_{(l)} \rfloor) \sum_{j=1}^{l} \mathbb{P}_{(j)}^c(\lfloor N_{(l)} \rfloor + 1, \boldsymbol{A}) \Big)
\end{aligned} \tag{21}$$

The chance of choosing the outside option is always strictly positive. When $\varepsilon_{il} < -\xi_l$ ($\forall l \neq 0$) and $\varepsilon_{i0} \geq 0$, the student chooses the outside option because she can get non-negative utility

[16] $f_l(\bullet)$ is defined in Appendix B.3 on page 34

only from this option. Therefore, $\sum_{j=1}^{l} \mathbb{P}^c_l(\bullet, \boldsymbol{A}) < 1$. We have

$$\begin{aligned}\sum_{j=1}^{l} f_{(j)}(N_{(j)}) \geq & \sum_{j=1}^{l} A_{(j)} - \Big(\sum_{i=1}^{\lfloor N_{(l)} \rfloor} \sum_{j=1}^{l} \mathbb{P}^c_{(j)}(i, \boldsymbol{A}) + (N_{(l)} - \lfloor N_{(l)} \rfloor) \sum_{j=1}^{l} \mathbb{P}^c_{(j)}(\lfloor N_{(l)} \rfloor + 1, \boldsymbol{A}) \Big) \\ > & \sum_{j=1}^{l} A_{(j)} - \Big(\sum_{i=1}^{\lfloor N_{(l)} \rfloor} 1 + (N_{(l)} - \lfloor N_{(l)} \rfloor)1 \Big) \\ = & \sum_{j=1}^{l} A_{(j)} - N_{(l)} \geq 0 \end{aligned} \tag{22}$$

Equation 22 indicates $\exists j \leq l \ f_{(j)}(N_{(j)}) > 0$. No $\boldsymbol{\xi}$ can be the solution of these functions. $\boldsymbol{\Psi^{-1}} = \varnothing$.

If $\forall l \ N_l > \sum_{j=1}^{L} A_j$, we apply the Poincaré-Bohl theorem. For the convenience, I copy the theorem from Fonda and Gidoni (2016):

Theorem 2 (Poincaré-Bohl)**.** *Assume that Ω is an open bounded subset of $\mathbb{R}^N$, with $0 \in \Omega$, and that $f : \overline{\Omega} \to \mathbb{R}^N$ is a continuous function such that*

$$f(x) \notin \{\beta^{17} x : \beta > 0\}, \textit{ for every } x \in \partial\Omega.$$

Then, there is an $\bar{x} \in \overline{\Omega}$ such that $f(\bar{x}) = 0$(Fonda and Gidoni, 2016, Page 4)

$\boldsymbol{\Psi}(\boldsymbol{\xi}, \boldsymbol{A})$ is a continuous function when α is small enough[18]. We consider that $\boldsymbol{\xi}$ is in the big ball (i.e. $\sum_l \xi_l^2 \leq R^2$). We would like to show $\boldsymbol{\Psi}(\boldsymbol{\xi}, \boldsymbol{A}) - \boldsymbol{N} \notin \{\beta\boldsymbol{\xi} : \beta > 0\}$ where $\boldsymbol{\xi}$ satisfies $\sum_l \xi_l^2 = R^2$. Then, according to the Poincaré-Bohl theorem, there is at least a $\boldsymbol{\xi_0}$ in the ball satisfying $\boldsymbol{\Psi}(\boldsymbol{\xi_0}, \boldsymbol{A}) - \boldsymbol{N} = 0$.

When $\boldsymbol{\xi}$ satisfies $\sum_l \xi_l^2 = R^2$, there is at least one $|\xi_l| \geq R/\sqrt{L}$. We divide this into two cases: (1) $\exists l \ \xi_l \geq R/\sqrt{L}$ and (2) $\nexists l \ \xi_l \geq R/\sqrt{L}$

Case 1 ($\exists l \ \xi_l \geq R/\sqrt{L}$)**.**

Divide the real line into $L+2$ sections: $(-\infty, 0)$, $[0, R^{1/L}/\sqrt{L})$, $[R^{1/L}/\sqrt{L}, R^{2/L}/\sqrt{L})$, ..., $[R^{(L-1)/L}/\sqrt{L}, R/\sqrt{L})$, $[R/\sqrt{L}, \infty)$. There are L colleges. ξ_l are in these sections. At least one section except $(-\infty, 0)$ does not have ξ_l. We suppose this section being $[R^{k/L}/\sqrt{L}, R^{(k+1)/L}/\sqrt{L})$. M out of L ξ_l are not smaller than $R^{(k+1)/L}/\sqrt{L}$ while others are smaller than $R^{k/L}/\sqrt{L}$. This procedure distinguishes the big ξ_l ($\xi_l \geq R^{(k+1)/L}/\sqrt{L}$) and the small ξ_l ($\xi_l < R^{k/L}/\sqrt{L}$). There is at least one big ξ_l since $\exists l \ \xi_l \geq R/\sqrt{L}$.

$\widehat{N}_l$ denotes the result of $\Psi_l(\boldsymbol{\xi}, \boldsymbol{A})$ while N_l is the real cutoff line of the college l. Consider these M colleges. We rank $\widehat{N}_l$ of these colleges from the smallest to the largest as $\widehat{N}_{\{1\}}$, $\widehat{N}_{\{2\}}$, ..., $\widehat{N}_{\{M\}}$. For any college m among these M colleges, we have

[17] It is α in Fonda and Gidoni (2016). We change it to β because we have defined α in our paper.

[18] See details in Theorem 2 on page 11

$$
\begin{aligned}
0 =&A_{\{m\}} - \Big(\sum_{i=1}^{\lfloor \widehat{N}_{\{m\}} \rfloor} \mathbb{P}^c_{\{m\}}(i, \boldsymbol{A}) + (\widehat{N}_{\{m\}} - \lfloor \widehat{N}_{\{m\}} \rfloor) \mathbb{P}^c_{\{m\}}(\lfloor \widehat{N}_{\{m\}} \rfloor + 1, \boldsymbol{A}) \Big) \\
\leq &A_{\{m\}} - \Big(\sum_{i=1}^{\lfloor \widehat{N}_{\{1\}} \rfloor} \mathbb{P}^c_{\{m\}}(i, \boldsymbol{A}) + (\widehat{N}_{\{1\}} - \lfloor \widehat{N}_{\{1\}} \rfloor) \mathbb{P}^c_{\{m\}}(\lfloor \widehat{N}_{\{1\}} \rfloor + 1, \boldsymbol{A}) \Big)
\end{aligned} \tag{23}
$$

We add all these inequalities up. We get

$$
\sum_{m=1}^{M} A_{\{m\}} - \Big(\sum_{i=1}^{\lfloor \widehat{N}_{\{1\}} \rfloor} \sum_{m=1}^{M} \mathbb{P}^c_{\{m\}}(i, \boldsymbol{A}) + (\widehat{N}_{\{1\}} - \lfloor \widehat{N}_{\{1\}} \rfloor) \sum_{m=1}^{M} \mathbb{P}^c_{\{m\}}(\lfloor \widehat{N}_{\{1\}} \rfloor + 1, \boldsymbol{A}) \Big) \geq 0 \tag{24}
$$

Then, we show that $\sum_{m=1}^{M} \mathbb{P}^a_{\{m\}}(i, \boldsymbol{A})$ does not vanish for $i \leq \sum_{m=1}^{M} A_{\{m\}}$. Since i is finite, $\sum_{m=1}^{M} \mathbb{P}^a_{\{m\}}(i, \boldsymbol{A}) \to 0$ requires at least $A_{\{1\}}$ out of top $i-1$ students almost surely choose the college (1); at least $A_{\{2\}}$ of these students almost surely choose the college (2); ...; and at least $A_{\{M\}}$ of these students almost surely choose the college $\{M\}$. This contradicts with $i \leq \sum_{m=1}^{M} A_{\{m\}}$. Thus $\mathbb{P}^a_{\{m\}}(i, \boldsymbol{A}) > 0$ for at least one m, in which one of them is denoted by m^*.

Recall $\xi_{\{m^*\}}$ being one of the big ξ_l. $(\xi_{\{m^*\}} + \varepsilon_{i(m^*)})\mathbb{P}^a_{\{m^*\}}(i, \boldsymbol{A})$ is almost surely larger than those with the small ξ_l when $R \to \infty$. Thus, student i will choose college $\{m^*\}$ or other colleges with big ξ_l. This leads $\sum_{m=1}^{M} \mathbb{P}^c_{\{m\}}(i, \boldsymbol{A}) \to 1$ for $\forall i \leq \sum_{m=1}^{M} A_{\{m\}}$. If $\sum_{m=1}^{M} \mathbb{P}^c_{\{m\}}(\lfloor \widehat{N}_{\{1\}} \rfloor + 1, \boldsymbol{A}) > 0$, then $\widehat{N}_{\{1\}} \leq \sum_{m=1}^{M} A_{\{m\}} + \iota$ almost surely when $R \to \infty$, where ι is a arbitrary small positive number. Then, $\widehat{N}_{\{1\}} - N_{\{1\}} < 0$ when ι is small enough. This leads no $\beta > 0$ satisfying $\widehat{N}_{\{1\}} - N_{\{1\}} = \beta \xi_{\{1\}}$.

Unfortunately, $\sum_{m=1}^{M} \mathbb{P}^c_{\{m\}}(\lfloor \widehat{N}_{\{1\}} \rfloor + 1, \boldsymbol{A})$ may converge to zero when $R \to \infty$. For example, among the top $\sum_{m=1}^{M} A_{\{m\}}$ students, $A_{\{1\}}$ students choose college $\{1\}$ almost surely; $A_{\{2\}}$ students choose college $\{2\}$ almost surely; ... $A_{\{M\}}$ students choose college $\{M\}$ almost surely. Student i ($i > \sum_{m=1}^{M} A_{\{m\}}$) will not choose the M colleges almost surely because she realizes the low probability of her to be admitted. In this case, $\widehat{N}_{\{1\}}$ may be much larger. To solve this problem, we play the same trick as in Equation 23 for all colleges. We get

$$
\sum_{l=1}^{L} A_l - \Big(\sum_{i=1}^{\lfloor \widehat{N}_{(1)} \rfloor} \sum_{l=1}^{L} \mathbb{P}^c_l(i, \boldsymbol{A}) + (\widehat{N}_{(1)} - \lfloor \widehat{N}_{(1)} \rfloor) \sum_{l=1}^{L} \mathbb{P}^c_l(\lfloor \widehat{N}_{(1)} \rfloor + 1, \boldsymbol{A}) \Big) \geq 0 \tag{25}
$$

where $\widehat{N}_{(1)}$ is the smallest $\widehat{N}$. There are two circumstances: (1) $\sum\limits_{l=1}^{L} \mathbb{P}_l^c(i, \boldsymbol{A}) \to 1$ for all $i \le \sum\limits_{l=1}^{L} A_l$ when $R \to \infty$ or (2) $\exists i \le \sum\limits_{l=1}^{L} A_l$ such that $\sum\limits_{l=1}^{L} \mathbb{P}_l^c(i, \boldsymbol{A}) < 1$ when $R \to \infty$.

Let us consider the first circumstance. Students shall not almost surely choose the outside option when $R \to \infty$, because the chance of $\varepsilon_{i0} \le 0$ does not vanish. This ensures $\sum\limits_{l=1}^{L} \mathbb{P}_l^c(\lfloor \widehat{N}_{(1)} \rfloor + 1, \boldsymbol{A}) > 0$. We have $\widehat{N}_{(1)} \le \sum\limits_{l=1}^{L} A_l + \iota$. In addition, $\xi_{(1)} > 0$. If $\xi_{(1)} \le 0$, $\mathbb{P}_{(1)}^c(i, \boldsymbol{A}) > 0 \Longrightarrow \mathbb{P}_0^c(i, \boldsymbol{A}) > 0$. If a student may choose the college (1), she also has a positive probability to choose the outside option. This is impossible in this circumstance because $\sum\limits_{l=1}^{L} \mathbb{P}_l^c(i, \boldsymbol{A}) \to 1$. If no student i ($i \le \sum\limits_{l=1}^{L} A_l$) may choose the college (1), $\widehat{N}_{(1)} \not\le \sum\limits_{l=1}^{L} A_l + \iota$. This leads to a contradiction. When $R \to \infty$ and ι is small enough, $\widehat{N}_{(1)} - N_{(1)} < 0$ as $N_{(1)} > \sum\limits_{l=1}^{L} A_l$. No $\beta > 0$ satisfies $\beta \xi_{(1)} = \widehat{N}_{(1)} - N_{(1)}$.

Now consider the second circumstance. If a student $i^{\triangle}$ may choose the outside option, she may also choose the colleges with the big ξ_l. $\sum\limits_{m=1}^{M} \mathbb{P}_{\{m\}}^c(i^{\triangle}, \boldsymbol{A}) > 0$. In addition, $\sum\limits_{m=1}^{M} \mathbb{P}_{\{m\}}^c(i, \boldsymbol{A}) \to 1$ for $\forall i \le \sum\limits_{m=1}^{M} A_{\{m\}}$, so $i^{\triangle} > \sum\limits_{m=1}^{M} A_{\{m\}}$. $\widehat{N}_{\{1\}} < i^{\triangle} < \sum\limits_{l=1}^{L} A_l < N_{\{1\}}$. The big ξ_l are all larger than zero. This leads no $\beta > 0$ satisfying $\beta \xi_{\{1\}} = \widehat{N}_{\{1\}} - N_{\{1\}}$.

Case 2 ($\nexists l \; \xi_l \ge R/\sqrt{L}$)**.**

If $\nexists l \; \xi_l \ge R/\sqrt{L}$, then $\exists l^{\triangledown}$ such that $\xi_l^{\triangledown} \le -R/\sqrt{L}$. When R becomes big, the student would like to choose the outside option rather than the school $l^{\triangledown}$. Thus, $\widehat{N}_{l^{\triangledown}}$ can be arbitrarily large if we choose a large R. We choose a R such that $\widehat{N}_{l^{\triangledown}} - N_{l^{\triangledown}} > 0$. This leads no $\beta > 0$ satisfying $\beta \xi_{l^{\triangledown}} = \widehat{N}_{l^{\triangledown}} - N_{l^{\triangledown}}$.

All in all, the conditions of the Poincaré-Bohl theorem are satisfied. We have $\boldsymbol{\Psi^{-1}} \neq \varnothing$ if $N_l > \sum_{j=1}^{L} A_j$ for all l

$\square$

B.5. *Proof of Theorem 4 and Corollary 1 on page 11*

Theorem 4 (Non-Uniqueness)**.** $\boldsymbol{\Psi^{-1}}$ *may be multi-valued.*

Corollary 1. *For college l, its cutoff threshold N_l may increase (i.e. be looser) when it becomes more attractive.*

Proof. We prove the theorem and the corollary by raising an example. Suppose that there are two colleges: college l and college l'. Each college would like to admit one student ($A_l = A_{l'} = 1$). For simplification, α is small. $\xi_l \ge \xi_{l'} > 0.5$. The support of ε_{i0} is $(-\infty, \gamma)$, while ε_{i1} and ε_{i2} are exponential distributed. γ is a small positive number. This ensures that the top students do not consider the outside option.

Af first, we calculate the choice probabilities for the top students[19] given the admission probabilities. The utility that a student can get from the college l is $(\xi_l + \varepsilon_{il})\mathbb{P}_l^a$[20]. The utility that she can get from the college l' is $(\xi_{l'} + \varepsilon_{il'})\mathbb{P}_{l'}^a$. She chooses the college l if and only if $(\xi_l + \varepsilon_{il})\mathbb{P}_l^a \geq (\xi_{l'} + \varepsilon_{il'})\mathbb{P}_{l'}^a$. We can get the choice probability

$$\mathbb{P}_l^c(i, \boldsymbol{A}) = \int\int_{(\xi_l+\varepsilon_{il})\mathbb{P}_l^a \geq (\xi_{l'}+\varepsilon_{il'})\mathbb{P}_{l'}^a} \exp(-\varepsilon_{il} - \varepsilon_{il'}) d\varepsilon_{il} d\varepsilon_{il'} \tag{26}$$

This can be divided into two cases.

Case 1 ($\mathbb{P}_l^a \xi_l < \mathbb{P}_{l'}^a \xi_{l'}$).

$$\begin{aligned} \mathbb{P}_l^c(i, \boldsymbol{A}) &= \int_0^\infty \int_{(\xi_{l'}+\varepsilon_{il'})\mathbb{P}_{l'}^a/\mathbb{P}_l^a - \xi_l}^\infty \exp(-\varepsilon_{il} - \varepsilon_{il'}) d\varepsilon_{il} d\varepsilon_{il'} \\ &= \int_0^\infty \exp(-\varepsilon_{il'}) \exp(-(\xi_{l'} + \varepsilon_{il'})\mathbb{P}_{l'}^a/\mathbb{P}_l^a + \xi_l) d\varepsilon_{il'} \\ &= \frac{\mathbb{P}_l^a}{\mathbb{P}_l^a + \mathbb{P}_{l'}^a} \exp(\frac{1}{\mathbb{P}_l^a}(\mathbb{P}_l^a \xi_l - \mathbb{P}_{l'}^a \xi_{l'})) \end{aligned} \tag{27}$$

Case 2 ($\mathbb{P}_l^a \xi_l \geq \mathbb{P}_{l'}^a \xi_{l'}$).

$$\begin{aligned} \mathbb{P}_l^c(i, \boldsymbol{A}) &= \int_0^{\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a - \xi_{l'}} \int_0^\infty \exp(-\varepsilon_{il} - \varepsilon_{il'}) d\varepsilon_{il} d\varepsilon_{il'} \\ &\quad + \int_{\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a - \xi_{l'}}^\infty \int_{(\xi_{l'}+\varepsilon_{il'})\mathbb{P}_{l'}^a/\mathbb{P}_l^a - \xi_l}^\infty \exp(-\varepsilon_{il} - \varepsilon_{il'}) d\varepsilon_{il} d\varepsilon_{il'} \\ &= \int_0^{\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a - \xi_{l'}} \exp(-\varepsilon_{il'}) d\varepsilon_{il'} \\ &\quad + \int_{\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a - \xi_{l'}}^\infty \exp(-\varepsilon_{il'}) \exp(-(\xi_{l'} + \varepsilon_{il'})\mathbb{P}_{l'}^a/\mathbb{P}_l^a + \xi_l) d\varepsilon_{il'} \\ &= 1 - \exp(-\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a + \xi_{l'}) + \frac{\mathbb{P}_l^a}{\mathbb{P}_l^a + \mathbb{P}_{l'}^a} \exp(-\mathbb{P}_l^a \xi_l/\mathbb{P}_{l'}^a + \xi_{l'}) \\ &= 1 - \frac{\mathbb{P}_{l'}^a}{\mathbb{P}_l^a + \mathbb{P}_{l'}^a} \exp(-\frac{1}{\mathbb{P}_{l'}^a}(\mathbb{P}_l^a \xi_l - \mathbb{P}_{l'}^a \xi_{l'})) \end{aligned} \tag{28}$$

For the student 1, her admission probability is one for both colleges. We can get her choice probabilities from Equation 27 and Equation 28.

[19] Thus we do not consider the outside option.

[20] $\mathbb{P}_l^a = \mathbb{P}_l^a(i, \boldsymbol{A})$ for abbreviation.

$$\begin{aligned}\mathbb{P}^c_l(1,\boldsymbol{A}) =& 1-\frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)\\ \mathbb{P}^c_{l'}(1,\boldsymbol{A}) =& \frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)\end{aligned} \tag{29}$$

Then we calculate the admission probabilities for the student 2.

$$\begin{aligned}\mathbb{P}^a_l(2,1;1) =& \mathbb{P}^a_l(1,1;1)(1-\mathbb{P}^c_l(1,\boldsymbol{A}))+\mathbb{P}^a_l(1,0;1)\mathbb{P}^c_l(1,\boldsymbol{A})\\ =& \frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)\end{aligned} \tag{30}$$

$$\begin{aligned}\mathbb{P}^a_{l'}(2,1;1) =& \mathbb{P}^a_{l'}(1,1;1)(1-\mathbb{P}^c_{l'}(1,\boldsymbol{A}))+\mathbb{P}^a_{l'}(1,0;1)\mathbb{P}^c_{l'}(1,\boldsymbol{A})\\ =& 1-\frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)\end{aligned} \tag{31}$$

In addition, we have

$$\exp\left(\xi_l-\xi_{l'}\right)-\frac{\xi_l+\xi_{l'}}{2\xi_{l'}}\geq 0 \tag{32}$$

because the equality holds when $\xi_l=\xi_{l'}$ and its first derivative with respect to ξ_l is $\exp\left(\xi_l-\xi_{l'}\right)-\dfrac{1}{2\xi_{l'}}$, being larger than zero based on our assumption.

Equation 32 leads to $\mathbb{P}^a_l(2,\boldsymbol{A})\xi_l\leq\mathbb{P}^a_{l'}(2,\boldsymbol{A})\xi_{l'}$. We then have

$$\begin{aligned}\mathbb{P}^c_l(2,\boldsymbol{A}) =& \frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)\exp\left(\xi_l-\xi_{l'}\frac{1-\frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)}{\frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)}\right)\\ =& \frac{1}{2}\exp\Big(2\xi_{l'}\big(1-\exp\left(\xi_l-\xi_{l'}\right)\big)\Big)\end{aligned} \tag{33}$$

We add the probabilities of choosing the college l up for the two students. We get

$$\begin{aligned}g(\xi_l,\xi_{l'}) =& \mathbb{P}^c_l(1,\boldsymbol{A})+\mathbb{P}^c_l(2,\boldsymbol{A})\\ =& 1-\frac{1}{2}\exp\left(-(\xi_l-\xi_{l'})\right)+\frac{1}{2}\exp\Big(2\xi_{l'}\big(1-\exp\left(\xi_l-\xi_{l'}\right)\big)\Big)\end{aligned} \tag{34}$$

If we take $\xi_l=\xi_{l'}$, $g(\xi_l,\xi_{l'})=1$. Thus, $\Psi_l=2$ based on our definition. We take the derivative of $g(\xi_l,\xi_{l'})$ with respect to ξ_l.

$$\frac{\partial g(\xi_l,\xi_{l'})}{\partial\xi_l}=\frac{1}{2}\exp\Big(-\left(\xi_l-\xi_{l'}\right)\Big)-\xi_{l'}\exp\Big(2\xi_{l'}\big(1-\exp\left(\xi_l-\xi_{l'}\right)\big)+\xi_l-\xi_{l'}\Big) \tag{35}$$

When $\xi_l=\xi_{l'}$, Equation 35 can be simplified,

$$\frac{\partial g(\xi_l, \xi_{l'})}{\partial \xi_l} = \frac{1}{2} - \xi_{l'} \tag{36}$$

Given $\xi_{l'} > 0.5$, $\frac{\partial g(\xi_l,\xi_{l'})}{\partial \xi_l} < 0$. Ψ_l shall increase when ξ_l slightly increases. This proves the Corollary 1. When the college l becomes more attractive, its cutoff line is higher (i.e. it admits worse students.). The reason is that the student 2 is afraid to apply the college l. For the student 1, she is more willing to apply the college l if the college is more attractive. However, the student 2 observes this change of the willingness. She understands her admission probability becoming lower. Therefore, she is more likely to apply the college l', the safer option.

For Theorem 4, we note that $g(\xi_l, \xi_{l'}) \equiv 1$ for any $\xi_l = \xi_{l'}$. Thus $\Psi_l = 2$. Still when $\xi_l = \xi_{l'}$, $\mathbb{P}^c_{l'}(1, \boldsymbol{A}) + \mathbb{P}^c_{l'}(2, \boldsymbol{A}) = 1 - \mathbb{P}^c_l(1, \boldsymbol{A}) + 1 - \mathbb{P}^c_l(2, \boldsymbol{A}) = 1$, which leads to $\Psi_{l'} = 2$. Therefore, $\boldsymbol{\Psi}(\{\xi_l, \xi_l\}, \boldsymbol{A})$ generate the same result for $\forall \xi_l$. $\boldsymbol{\Psi^{-1}}$ may be multi-valued. $\square$

B.6. Proof of Theorem 5 and Theorem 6 on page 15

Theorem 5. *Based on* $\boldsymbol{\Phi^{-1}}$

$$\xi_{(l)} = \log \frac{\left(1 + \sum_{k=l+1}^{L} \exp\left(\xi_{(k)}\right)\right) A_l}{N_{(l)} - \sum_{k=1}^{l} A_{(k)}} \tag{5}$$

for all $l < L$ *and*

$$\xi_{(L)} = \log \frac{A_{(L)}}{N_{(L)} - \sum_{k=1}^{L} A_{(k)}} \tag{6}$$

Theorem 6. *Rank* $A_l / \exp(\xi_l)$ *from smallest to largest as* $A_{(1)}/\exp(\xi_{(1)})$, $A_{(2)}/\exp(\xi_{(2)})$, *...,* $A_{(L)}/\exp(\xi_{(L)})$. *Based on* $\boldsymbol{\Phi}$

$$N_{(l)} = \sum_{k=1}^{l} A_{(k)} + \frac{1 + \sum_{k=l+1}^{L} \exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l)}\right)} A_{(l)} \tag{7}$$

for all $l < L$ *and*

$$N_{(L)} = \sum_{k=1}^{L} A_{(k)} + \frac{1}{\exp\left(\xi_{(L)}\right)} A_{(L)} \tag{8}$$

Proof. First, we shall prove Theorem 5. We know $\boldsymbol{N}$ and $\boldsymbol{A}$. We want to get $\boldsymbol{\xi}$. Equation 4 tells us the relationship.

$$A_{(l)} = N_{(1)}\frac{\exp\left(\xi_{(l)}\right)}{1+\sum\limits_{k=1}^{L}\exp\left(\xi_{(k)}\right)}+\left(N_{(2)}-N_{(1)}\right)\frac{\exp\left(\xi_{(l)}\right)}{1+\sum\limits_{k=2}^{L}\exp\left(\xi_{(k)}\right)}+\ldots+\left(N_{(l)}-N_{(l-1)}\right)\frac{\exp\left(\xi_{(l)}\right)}{1+\sum\limits_{k=l}^{L}\exp\left(\xi_{(k)}\right)} \tag{37}$$

$$A_{(l+1)} = N_{(1)}\frac{\exp\left(\xi_{(l+1)}\right)}{1+\sum\limits_{k=1}^{L}\exp\left(\xi_{(k)}\right)}+\left(N_{(2)}-N_{(1)}\right)\frac{\exp\left(\xi_{(l+1)}\right)}{1+\sum\limits_{k=2}^{L}\exp\left(\xi_{(k)}\right)}+\ldots+\left(N_{(l+1)}-N_{(l)}\right)\frac{\exp\left(\xi_{(l+1)}\right)}{k+\sum\limits_{k=l+1}^{L}\exp\left(\xi_{(k)}\right)} \tag{38}$$

where Equation 37 and Equation 38 are two lines of Equation 4. We multiply Equation 37 by $\exp\left(\xi_{l+1}\right)/\exp\left(\xi_{l}\right)$ and substitute the result into Equation 38. We get

$$A_{(l+1)} = A_{(l)}\frac{\exp\left(\xi_{(l+1)}\right)}{\exp\left(\xi_{(l)}\right)} + \left(N_{(l+1)}-N_{(l)}\right)\frac{\exp\left(\xi_{(l+1)}\right)}{1+\sum\limits_{k=l+1}^{L}\exp\left(\xi_{(k)}\right)} \tag{39}$$

Arrange it. We get

$$N_{(l+1)} - N_{(l)} = A_{(l+1)} + A_{(l+1)}\frac{1+\sum\limits_{k=l+2}^{L}\exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l+1)}\right)} - A_{(l)}\frac{1+\sum\limits_{k=l+1}^{L}\exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l)}\right)} \tag{40}$$

where we define $\sum\limits_{k=L+1}^{L}\exp\left(\xi_{(k)}\right)=0$. Sum $N_{(2)}-N_{(1)}$, $N_{(3)}-N_{(2)}$, ..., $N_{(l+1)}-N_{(l)}$ up. We get

$$N_{(l+1)} = \sum_{k=1}^{l+1} A_{(k)} + \frac{1+\sum\limits_{k=l+2}^{L}\exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l+1)}\right)}A_{(l+1)} \tag{41}$$

Arrange it. We get

$$\xi_{(l+1)} = \log\frac{\left(1+\sum\limits_{k=l+2}^{L}\exp\left(\xi_{(k)}\right)\right)A_{l+1}}{N_{(l+1)}-\sum\limits_{k=1}^{l+1}A_{(k)}} \tag{42}$$

Theorem 5 has been proven.

Now let us look at Theorem 6. If we know how to map l to (l), the proof has been completed in Equation 41. However, we need to know $\boldsymbol{N}$ to generate the mapping from l to (l). $N_{(l)}$ is the lth smallest value in $\boldsymbol{N}$. We only know $\boldsymbol{A}$ and $\boldsymbol{\xi}$. We shall prove $A_l/\exp\left(\xi_l\right)$ generating the same mapping from l to (l).

At first, we show the existence of $\boldsymbol{N}$.[21] We rank $A_l/\exp\left(\xi_l\right)$ from the smallest to the largest as $A_{[1]}/\exp\left(\xi_{[1]}\right)$, $A_{[2]}/\exp\left(\xi_{[2]}\right)$, ..., $A_{[L]}/\exp\left(\xi_{[L]}\right)$. We have

[21] We emphasize that Theorem 2 is not valid since $\mathbb{P}_l^a$ depends on $\boldsymbol{N}$ in the simplified model.

$$
\begin{aligned}
&\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq \frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &(1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &A_{[l+1]} + (1+\sum_{k=l+2}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &\sum_{k=1}^{l+1} A_{[k]} + (1+\sum_{k=l+2}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq \sum_{k=1}^{l} A_{[k]} + (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})}
\end{aligned}
\tag{43}
$$

If we let

$$
N_{[l]} = \sum_{k=1}^{l} A_{[k]} + (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})} \tag{44}
$$

We have $N_{[l+1]} \geq N_{[l]}$. $N_{[l]}$ is one set of solution. $\boldsymbol{N}$ exists.

Then, we show the uniqueness of $\boldsymbol{N}$. If we have $A_l/\exp(\xi_l) > A_{l'}/\exp(\xi_{l'}) \Longleftrightarrow N_l > N_{l'}$ and $A_l/\exp(\xi_l) = A_{l'}/\exp(\xi_{l'}) \Longleftrightarrow N_l = N_{l'}$, $N_{[l]}$ is the unique set of solution, as the mapping from l to $[l]$ and that from l to (l) are equivalent. If $\boldsymbol{N}$ is not unique, we have another set of N_l such that $\exists l, l'$, $N_l < N_{l'}$, $A_l/\exp(\xi_l) \geq A_{l'}/\exp(\xi_{l'})$ or $\exists l, l'$, $N_l = N_{l'}$, $A_l/\exp(\xi_l) \neq A_{l'}/\exp(\xi_{l'})$. In either case, the order of N_l is different from the order of $A_l/\exp(\xi_l)$.

Case 1 ($\exists l, l'$, $N_l < N_{l'}$, $A_l/\exp(\xi_l) \geq A_{l'}/\exp(\xi_{l'})$)**.**

We rank $\boldsymbol{N}$ from the smallest to the largest as $N_{(1)}$, $N_{(2)}$, ..., $N_{(L)}$. $l = (m)$ and $l' = (m')$. Since $N_l < N_{l'}$, $m < m'$. From Equation 43, we have $A_{(m)}/\exp(\xi_{(m)}) \leq A_{(m+1)}/\exp(\xi_{(m+1)})$ $\leq A_{(m+2)}/\exp(\xi_{(m+2)}) \ldots \leq A_{(m')}/\exp(\xi_{(m')})$. If all equalities hold, $N_{(m)} = N_{(m+1)} \ldots = N_{(m')}$. This contradicts with our assumption. Thus we have $A_{(m)}/\exp(\xi_{(m)}) < A_{(m')}/\exp(\xi_{(m')})$. This contradicts with $A_l/\exp(\xi_l) \geq A_{l'}/\exp(\xi_{l'})$.

Case 2 ($\exists l, l'$, $N_l = N_{l'}$, $A_l/\exp(\xi_l) \neq A_{l'}/\exp(\xi_{l'})$)**.**

We apply the same strategy in Case 1. $N_{(m)} = N_{(m')}$ indicates $A_{(m)}/\exp(\xi_{(m)}) = A_{(m+1)}/\exp(\xi_{(m+1)})$ $\ldots = A_{(m')}/\exp(\xi_{(m')})$. This contradicts with $A_l/\exp(\xi_l) \neq A_{l'}/\exp(\xi_{l'})$.

Therefore, the mapping $l \to (l)$ generated from $A_l/\exp(\xi_l)$ and the one from N_l are equivalent. $\boldsymbol{N}$ is unique. □

B.7. Proof of Theorem 7 on page 16

Theorem 7. *$\Psi_l/A_l = \Phi_l/A_l$ and $\Psi_l^{-1} = \Phi_l^{-1}$ assuming (1) $N_l \to \infty$ and $A_l/N_l > 0$ for $\forall l$; (2) $N_l/N_{l'}$ is finite for any l and l'; (3) α is small enough; and (4) ε_{il} of Ψ_l has the same distribution with that of Φ_l.*

Proof. We consider the full model. We present another representation of $\mathbb{P}_l^a(i, \boldsymbol{A})$

$$\mathbb{P}_l^a(i, \boldsymbol{A}) = \mathbb{P}(\sum_{j=1}^{i-1} \mathbb{1}_l(j) < A_l) \tag{45}$$

The student i can be accepted by the college l if and only if fewer than A_l among top $i-1$ students choose the college l. $\mathbb{P}_l^a(i, \boldsymbol{A}) = 1$ for $i \leq A_l$. We only need to consider $i > A_l$. We have

$$\frac{1}{i-1} \sum_{j=1}^{i-1} \mathbb{1}_l(j) - \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) \xrightarrow{a.s.} 0 \tag{46}$$

The proof is the same as that of Theorem 1. $A_l/N_l > 0$, so $i \to \infty$ when $N_l \to \infty$ and $i > A_l$. We get

$$\begin{aligned}
&\mathbb{P}(\frac{1}{i-1} \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \frac{N_l}{i-1}\nu < \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbb{1}_l(j) < \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \frac{N_l}{i-1}\nu) = 1 \\
\Longleftrightarrow &\mathbb{P}(\frac{1}{N_l} \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu < \frac{1}{N_l} \sum_{j=1}^{i-1} \mathbb{1}_l(j) < \frac{1}{N_l} \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu) = 1 \\
\Longleftrightarrow &\mathbb{P}(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l < \sum_{j=1}^{i-1} \mathbb{1}_l(j) < \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l) = 1
\end{aligned} \tag{47}$$

where ν is an arbitrary small positive number when N_l is large enough. Using the same logic, we also have

$$\mathbb{P}(\sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l < A_l < \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l) = 1 \tag{48}$$

Combining Equation 45, Equation 47 and Equation 48, we have

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l \leq \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l\Big) \leq \mathbb{P}_l^a(i, \boldsymbol{A}) \leq \mathbb{P}\Big(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l \geq \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l\Big) \\
\Longleftrightarrow &\mathbb{P}\Big(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) \leq -2\nu N_l\Big) \leq \mathbb{P}_l^a(i, \boldsymbol{A}) \leq \mathbb{P}\Big(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) \geq 2\nu N_l\Big)
\end{aligned} \tag{49}$$

$\mathbb{P}_l^c(j, \boldsymbol{A})$ has a positive lower bound when $\mathbb{P}_l^a(j, \boldsymbol{A}) \geq \alpha$. The chance of $\varepsilon_{il'} < -\xi_{l'}$ for all $l' \neq \{0, l\}$ does not vanish. If $\varepsilon_{il'} < -\xi_l'$ ($\forall l' \neq \{0, l\}$), $\varepsilon_{i0} < 0$ and $\varepsilon_{il} > -\xi_l$, the student chooses the college l since she can get positive utility only from this college. $\kappa > 0$ denotes this lower bound. $\mathbb{P}_l^c(j, \boldsymbol{A}) \geq \kappa$. If $i \leq N_l + 1 - \lceil 2\nu N_l/\kappa \rceil$,

$$
\begin{aligned}
&\mathbb{P}_l^a(i,\boldsymbol{A})\\
\geq&\mathbb{P}\Big(\sum_{j=1}^{i-1}\mathbb{P}_l^c(j,\boldsymbol{A})-\sum_{j=1}^{N_l}\mathbb{P}_l^c(j,\boldsymbol{A})\leq -2\nu N_l\Big)\\
=&\mathbb{P}\Big(-\sum_{j=i}^{N_l}\mathbb{P}_l^c(j,\boldsymbol{A})\leq -2\nu N_l\Big)\\
\geq&\mathbb{P}\Big(-(N_l-i+1)\kappa\leq -2\nu N_l\Big)=1
\end{aligned}
\tag{50}
$$

Likewise, when $i \geq N_l+1+\lceil 2\nu N_l/\kappa\rceil$, $\mathbb{P}_l^a(i,\boldsymbol{A})=0$ if $\mathbb{P}_l^a(i-1,\boldsymbol{A})\geq\alpha$. If $\mathbb{P}_l^a(i-1,\boldsymbol{A})<\alpha$, $\mathbb{P}_l^a(i,\boldsymbol{A})<\alpha$ because $\mathbb{P}_l^a(i,\boldsymbol{A})\leq\mathbb{P}_l^a(i-1,\boldsymbol{A})$. In both cases, $\mathbb{P}_l^a(i,\boldsymbol{A})<\alpha$. The student i does not consider the college l. ω denotes a small positive number such that $1+\lceil 2\nu N_l/\kappa\rceil<\omega N_l$ for all l. ν can be an arbitrary small positive number when N_l is large. So can ω.

Being consistent with the simplified model, $\varepsilon_{il}+\gamma$ is i.i.d Extreme Value Type 1 distributed. For the college l, we do not count the choice probabilities of the students with $0\leq\mathbb{P}_l^a(i,\boldsymbol{A})<1$. This causes that a college admits fewer students at the cutoff line N_l. Mathematically we have

$$
\begin{aligned}
\frac{A_{(l)}}{N_{(l)}}=&\frac{1}{N_{(l)}}\sum_{i=1}^{N_{(l)}}\mathbb{P}_{(l)}^c(i,\boldsymbol{A})\\
\geq&\frac{1}{N_{(l)}}\Big(N_{(1)}(1-\omega)\frac{\exp(\xi_{(l)})}{1+\sum_{k=1}^{L}\exp(\xi_{(k)})}\Big)\\
&+\frac{1}{N_{(l)}}\Big(\big(N_{(2)}(1-\omega)-N_{(1)}(1+\omega)\big)_+\frac{\exp(\xi_{(l)})}{1+\sum_{k=2}^{L}\exp(\xi_{(k)})}\Big)\\
&\ldots+\frac{1}{N_{(l)}}\Big(\big(N_{(l)}(1-\omega)-N_{(l-1)}(1+\omega)\big)_+\frac{\exp(\xi_{(l)})}{1+\sum_{k=l}^{L}\exp(\xi_{(k)})}\Big)\\
=&\frac{A_{(l)}}{N_{(l)}}\Big(\frac{N_{(1)}}{A_{(l)}}(1-\omega)\frac{\exp(\xi_{(l)})}{1+\sum_{k=1}^{L}\exp(\xi_{(k)})}\Big)\\
&+\frac{A_{(l)}}{N_{(l)}}\Big(\big(\frac{N_{(2)}}{A_{(l)}}(1-\omega)-\frac{N_{(1)}}{A_{(l)}}(1+\omega)\big)_+\frac{\exp(\xi_{(l)})}{1+\sum_{k=2}^{L}\exp(\xi_{(k)})}\Big)\\
&\ldots+\frac{A_{(l)}}{N_{(l)}}\Big(\big(\frac{N_{(l)}}{A_{(l)}}(1-\omega)-\frac{N_{(l-1)}}{A_{(l)}}(1+\omega)\big)_+\frac{\exp(\xi_{(l)})}{1+\sum_{k=l}^{L}\exp(\xi_{(k)})}\Big)
\end{aligned}
\tag{51}
$$

where the first equality is used in the full model. $(X)_+ = X$ if $X > 0$ while $(X)_+ = 0$ if $X \leq 0$. We ignore the marginal students for a college l, whose $\alpha \leq \mathbb{P}^a_l(i, \boldsymbol{A}) < 1$. The non-marginal students have definite beliefs. They act as they do in the simplified model. Likewise, we count the choice probability of a student with $\alpha \leq \mathbb{P}^a_l(i, \boldsymbol{A}) < 1$ as 1 for the college l. This causes that a college admits more students at the cutoff line N_l. Mathematically we have

$$\begin{aligned}
\frac{A_{(l)}}{N_{(l)}} =& \frac{1}{N_{(l)}} \sum_{i=1}^{N_{(l)}} \mathbb{P}^c_{(l)}(i, \boldsymbol{A}) \\
\leq & \frac{1}{N_{(l)}} \Big(N_{(1)}(1-\omega) \frac{\exp(\xi_{(l)})}{1 + \sum_{k=1}^{L} \exp(\xi_{(k)})} + 2\omega N_{(1)} \Big) \\
& + \frac{1}{N_{(l)}} \Big(\big(N_{(2)}(1-\omega) - N_{(1)}(1+\omega) \big)_+ \frac{\exp(\xi_{(l)})}{1 + \sum_{k=2}^{L} \exp(\xi_{(k)})} + 2\omega N_{(2)} \Big) \\
& \ldots + \frac{1}{N_{(l)}} \Big(\big(N_{(l)}(1-\omega) - N_{(l-1)}(1+\omega) \big)_+ \frac{\exp(\xi_{(l)})}{1 + \sum_{k=l}^{L} \exp(\xi_{(k)})} + \omega N_{(l)} \Big) \\
=& \frac{A_{(l)}}{N_{(l)}} \Big(\frac{N_{(1)}}{A_{(l)}}(1-\omega) \frac{\exp(\xi_{(l)})}{1 + \sum_{k=1}^{L} \exp(\xi_{(k)})} + 2\omega \frac{N_{(1)}}{A_{(l)}} \Big) \\
& + \frac{A_{(l)}}{N_{(l)}} \Big(\big(\frac{N_{(2)}}{A_{(l)}}(1-\omega) - \frac{N_{(1)}}{A_{(l)}}(1+\omega) \big)_+ \frac{\exp(\xi_{(l)})}{1 + \sum_{k=2}^{L} \exp(\xi_{(k)})} + 2\omega \frac{N_{(2)}}{A_{(l)}} \Big) \\
& \ldots + \frac{A_{(l)}}{N_{(l)}} \Big(\big(\frac{N_{(l)}}{A_{(l)}}(1-\omega) - \frac{N_{(l-1)}}{A_{(l)}}(1+\omega) \big)_+ \frac{\exp(\xi_{(l)})}{1 + \sum_{k=l}^{L} \exp(\xi_{(k)})} + \omega \frac{N_{(l)}}{A_{(l)}} \Big)
\end{aligned} \tag{52}$$

In the two equations, $N_{(l')}/A_{(l)}$ is finite due to our assumptions. ω can be arbitrarily small when $\boldsymbol{N}$ is large enough. When $\omega \to 0$, the full model collapses to the simplified model. This completes the proof.

□

Appendix C. Graphs

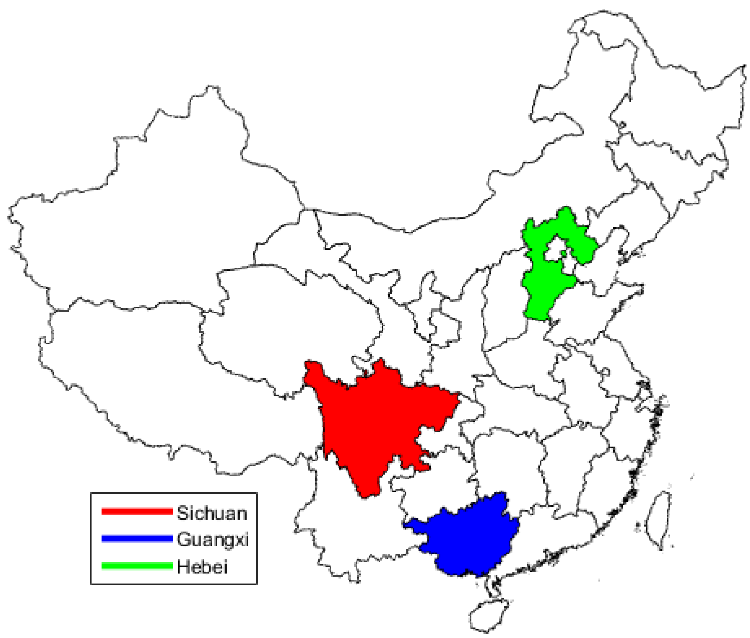
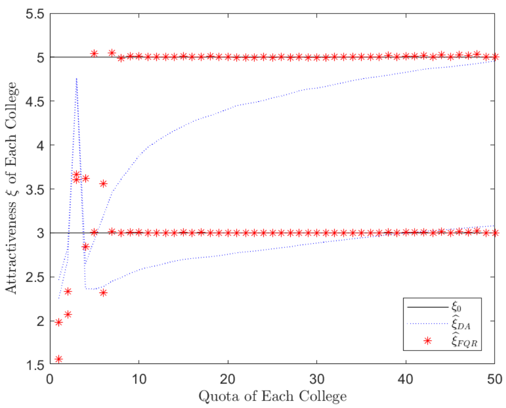
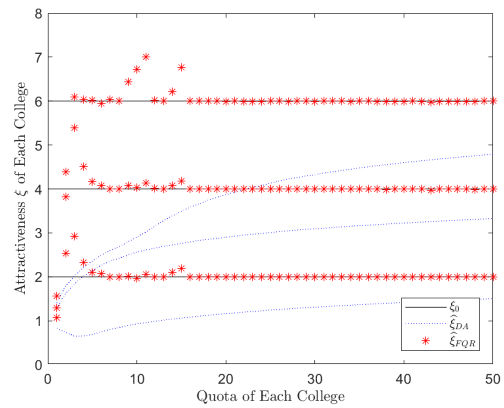


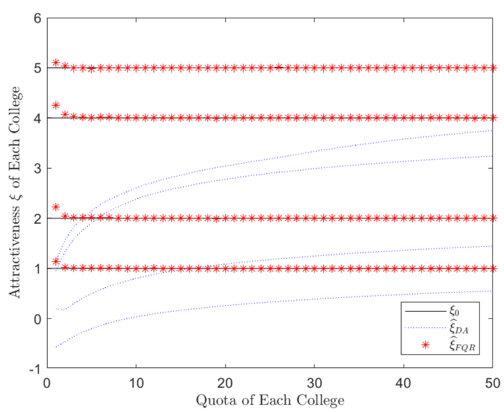
Fig. 1. Location of Three Chinese Provinces



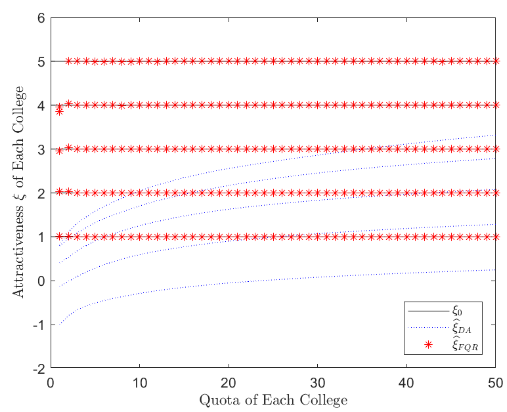
(a) Two Colleges



(b) Three Colleges



(c) Four Colleges



(d) Five Colleges

Fig. 2. DA Model ($\mathbf{\Phi^{-1}}$) vs. FQR for Different Quotas in A Two-, Three-, Four-, or Five-College World

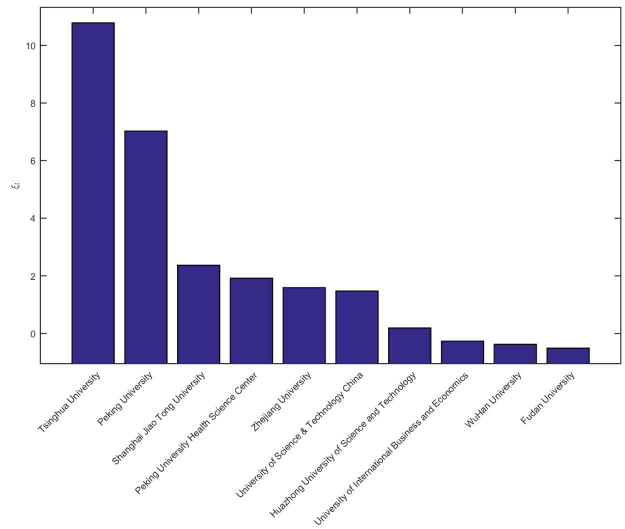


Fig. 3. Attractiveness ξ_l of Top 10 Universities for Science Majors Students in Guangxi for 2008

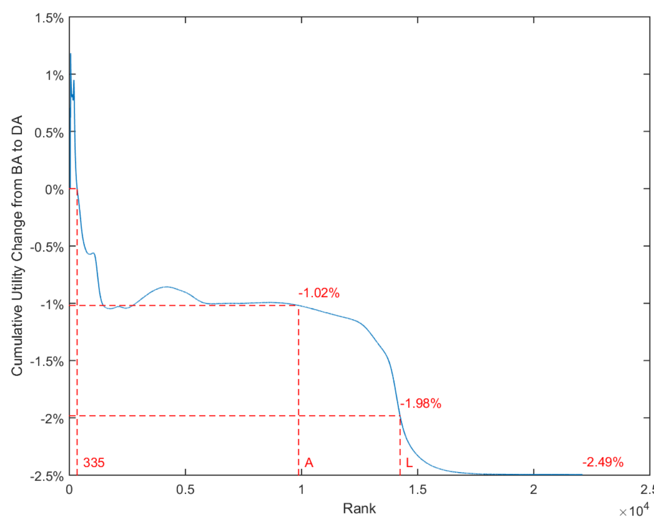


Fig. 4. Cumulative Welfare Change from BM to DA for Science Majors Students in Guangxi for 2008

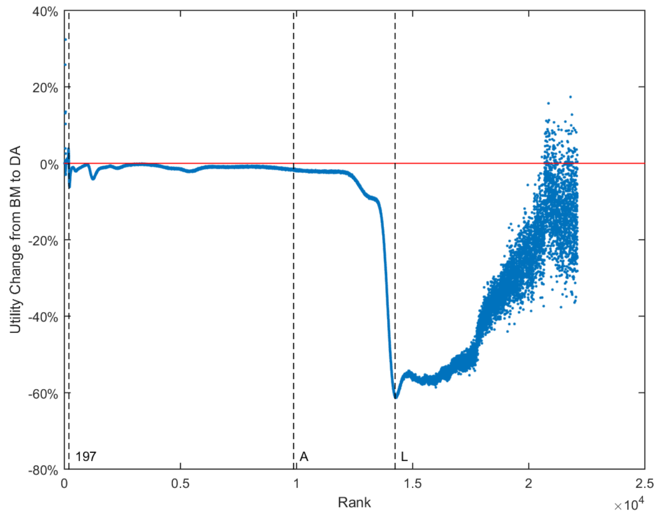


Fig. 5. Individual Welfare Change from BM to DA for Science Majors Students in Guangxi for 2008

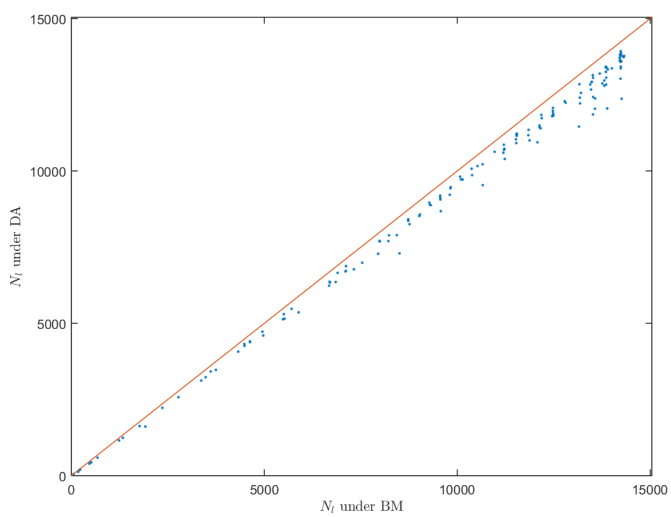


Fig. 6. Cutoff Threshold of Each University in BM vs. DA for Science Majors Students in Guangxi for 2008

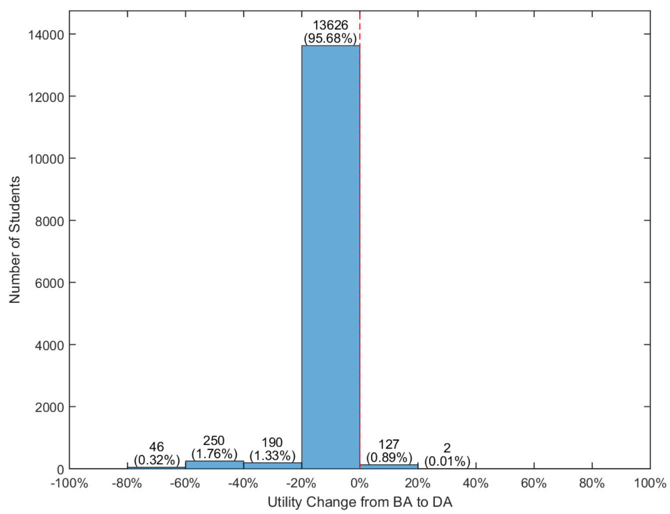


Fig. 7. Histogram of Welfare Change from BM to DA of Science Majors Students above Key Cutoff Threshold (Top 14241 Students) in Guangxi for 2008

Appendix D. Tables

Table 1: Top Ten Colleges in terms of ξ_l in Guangxi

Ranking	Arts Majors for 2006: Name	ξ_l	Science Majors for 2006: Name	ξ_l
1	Peking University	19.796	Tsinghua University	2.530
2	Fudan University	5.682	Shanghai Jiao Tong University	1.837
3	Renmin University of China	4.303	University of Science & Technology China	1.510
4	University of International Business and Economics	1.866	Peking University	1.173
5	City University Hong Kong	1.266	Nanjing University	0.809
6	China University of Political Science and Law	0.823	Guangxi University	-0.920
7	Beijing Normal University	0.333	Xi'an Jiao Tong University	-0.950
8	Sun Yat-sen University	0.205	Sun Yat-sen University	-0.963
9	Wuhan University	0.117	Central South University	-1.041
10	Nanjing University	-0.117	Central University of Finance and Economics	-1.174
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00745$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00194$
	Arts Majors for 2007		Science Majors for 2007	
1	Peking University	6.124	Peking University	9.502
2	Renmin University of China	1.271	Tsinghua University	7.812
3	Fudan University	0.560	University of Science & Technology China	0.557
4	Nanjing University	0.472	Zhejiang University	0.144
5	Beijing Foreign Studies University	-0.199	Guangxi University	-0.645
6	Zhongnan University of Economics and Law	-0.379	Nanjing University	-1.067
7	Guangxi University	-0.480	Beihang University	-1.233
8	Guangxi Normal University	-0.688	Guangxi Medical University	-1.287
9	Wuhan University	-0.877	Hunan University	-1.322
10	Sun Yat-sen University	-1.059	Nankai University	-1.357
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00152$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00560$
	Arts Majors for 2008		Science Majors for 2008	
1	Tsinghua University	30.256	Tsinghua University	10.782
2	Peking University	18.050	Peking University	7.024
3	Renmin University of China	3.713	Shanghai Jiao Tong University	2.367
4	Nanjing University	2.188	Peking University Health Science Center	1.916
5	University of International Business and Economics	1.362	Zhejiang University	1.589
6	Sun Yat-sen University	1.289	University of Science & Technology China	1.471
7	China University of Political Science and Law	0.552	Huazhong University of Science and Technology	0.189
8	Wuhan University	0.360	University of International Business and Economics	-0.268
9	Nankai University	0.264	Wuhan University	-0.379
10	Zhongnan University of Economics and Law	-0.104	Fudan University	-0.510
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.01523$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00327$

Table 2: Top Ten Colleges in terms of ξ_l in Hebei

	Arts Majors for 2005		Science Majors for 2005	
Ranking	Name	ξ_l	Name	ξ_l
1	Peking University	3.634	Tsinghua University	1.848
2	Fudan University	1.167	Peking University	1.067
3	Zhejiang University	-0.109	Zhejiang University	0.459
4	University of International Business and Economics	-0.185	Tianjin University	0.096
5	Wuhan University	-0.439	Peking University Health Science Center	-0.070
6	Nankai University	-0.440	University of Science & Technology China	-1.238
7	China University of Political Science and Law	-0.454	University of Science and Technology Beijing	-1.313
8	Tsinghua University	-0.515	Harbin Institute of Technology (Harbin)	-1.474
9	Beijing Normal University	-0.576	Huazhong University of Science and Technology	-1.484
10	Renmin University of China	-0.814	Beijing Jiaotong University	-1.490
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00077$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00162$
	Arts Majors for 2006		Science Majors for 2006	
1	Peking University	5.901	Tsinghua University	6.926
2	Tsinghua University	4.255	Peking University	4.849
3	Renmin University of China	3.408	Beihang University	0.687
4	Zhejiang University	0.138	Zhejiang University	0.510
5	Nankai University	0.066	University of Science & Technology China	-0.227
6	Nanjing University	-0.523	Xi'an Jiao Tong University	-0.969
7	Xiamen University	-0.694	Harbin Institute of Technology (Harbin)	-1.399
8	University of International Business and Economics	-1.126	Dalian University of Technology	-1.476
9	Jilin University	-1.281	Nanjing University	-1.586
10	Zhongnan University of Economics and Law	-1.406	Xi'an Electronic and Science University	-1.663
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00067$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00174$
	Arts Majors for 2007		Science Majors for 2007	
1	Peking University	8.101	Tsinghua University	8.920
2	Tsinghua University	3.779	Peking University	5.421
3	Renmin University of China	2.809	Shanghai Jiao Tong University	2.344
4	Fudan University	1.372	Peking University Health Science Center	2.041
5	Zhejiang University	-0.030	Beihang University	1.966
6	Central University of Finance and Economics	-0.379	Fudan University	0.450
7	Nanjing University	-0.396	Zhejiang University	0.356
8	China University of Political Science and Law	-0.460	Xi'an Jiao Tong University	0.281
9	Nankai University	-0.689	Nanjing University	0.188
10	Beijing Foreign Studies University	-0.693	Nankai University	-0.184
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00255$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00203$

Table 3: Top Ten Colleges in terms of ξ_l in Sichuan

Ranking	Arts Majors for 2006: Name	ξ_l	Science Majors for 2006: Name	ξ_l
1	Tsinghua University	3.489	Peking University Health Science Center	6.255
2	Renmin University of China	2.864	Peking University	5.422
3	Fudan University	1.322	Zhejiang University	4.426
4	Sichuan University	0.053	University of Science & Technology China	3.681
5	Southwestern University of Finance and Economics	-0.046	Fudan University	3.346
6	Nanjing University	-0.284	Shanghai Jiao Tong University	2.660
7	Wuhan University	-0.742	Beihang University	1.363
8	Zhejiang University	-0.793	Nanjing University	1.221
9	Tongji University	-1.009	Beijing University of Posts and Telecommunications	1.164
10	Nankai University	-1.248	Shanghai University of Finance and Economics	0.711
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00044$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00097$
	Arts Majors for 2007		Science Majors for 2007	
1	Fudan University	2.830	Peking University	8.208
2	University of International Business and Economics	0.931	Fudan University	8.206
3	Sichuan University	0.850	Peking University Health Science Center	6.472
4	Beijing Foreign Studies University	0.816	Shanghai Jiao Tong University	6.290
5	Southwestern University of Finance and Economics	0.791	Zhejiang University	5.091
6	Nanjing University	0.582	University of Science & Technology China	4.353
7	Zhejiang University	0.491	Renmin University of China	3.381
8	China University of Political Science and Law	0.403	Tongji University	3.156
9	Nankai University	0.196	Nanjing University	2.887
10	Shanghai University of Finance and Economics	-0.719	Beihang University	2.471
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00024$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00059$

Table 4: BM vs. DA

Province	Year	Major	G	I	Breakeven	Loss (L)	Loss	A	L[a]
Guangxi	2006	Arts	241 (6.75%)	308 (8.63%)	546 (15.30%)	-5.87%	-6.63%	2524	3570
Guangxi	2006	Science	232 (1.77%)	1049 (8.01%)	306 (2.34%)	-2.87%	-3.26%	8960	13098
Guangxi	2007	Arts	65 (1.59%)	112 (2.75%)	172 (4.22%)	-2.45%	-3.50%	2898	4077
Guangxi	2007	Science	1506 (10.65%)	4793 (33.91%)	278 (1.97%)	-1.50%	-2.00%	9794	14135
Guangxi	2008	Arts	140 (3.13%)	211 (4.72%)	278 (6.22%)	-3.67%	-4.37%	2967	4468
Guangxi	2008	Science	129 (0.91%)	197 (1.38%)	335 (2.35%)	-1.98%	-2.49%	9875	14242
Hebei	2005	Arts	63 (1.15%)	226 (4.12%)	291 (5.31%)	-4.31%	-4.34%	1743	5480
Hebei	2005	Science	430 (1.86%)	3020 (13.05%)	379 (1.64%)	-1.93%	-1.94%	9476	23145
Hebei	2006	Arts	36 (0.64%)	60 (1.06%)	96 (1.70%)	-4.13%	-4.16%	1839	5656
Hebei	2006	Science	210 (0.88%)	5125 (21.47%)	551 (2.31%)	-1.69%	-1.73%	9187	23866
Hebei	2007	Arts	39 (0.68%)	65 (1.13%)	105 (1.82%)	-3.94%	-3.98%	1722	5764
Hebei	2007	Science	236 (0.93%)	544 (2.14%)	846 (3.33%)	-3.52%	-3.52%	8947	25437
Sichuan	2006	Arts	62 (1.31%)	2090 (44.15%)	155 (3.27%)	-2.86%	-3.77%	3452	4734
Sichuan	2006	Science	307 (1.19%)	731 (2.84%)	1293 (5.03%)	-3.02%	-3.39%	20672	25715
Sichuan	2007	Arts	300 (6.57%)	454 (9.94%)	784 (17.16%)	-4.03%	-4.95%	3488	4569

[a] G: the number of students who benefit from the switch and the percentage of these students in the all students above the key cutoff threshold in the parentheses (i.e. Figure 7 on page 51).
I: the last student in terms of rank who benefits from the switch and is above the key cutoff threshold; the percentage of the rank in the key cutoff in the parentheses (i.e. the first vertical line from the left in Figure 5 on page 50).
Breakeven: the maximum number of students where the social welfare may increase after the switch; the percentage of this number in the key cutoff in the parentheses (i.e. the first vertical line from the left in Figure 4 on page 49).
Loss (L) and Loss: the welfare loss of the students above the key cutoff and that of all the students.
A: the sum of the quotas of all colleges.
L: the key cutoff threshold.