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

Yun Pu†

This version: May 29, 2019

Abstract

The deferred acceptance mechanism (DA) replaced the Boston mechanism (BM) in college admissions in China. In this paper, I compare the empirical performance of these two mechanisms in the Chinese context by developing an empirical model and applying it to college admissions in Guangxi, Hebei, and Sichuan provinces. Then, I conduct counterfactuals to empirically compare the BM and DA in these three provinces for given years. I find that not only is the BM superior to the DA in terms of total welfare but also that most students receive lower utility after the switch from the BM to DA.

JEL classification: C51, C78, D47, D61.

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

*I thank Ian Sheldon, Javier Donna, Jason Blevins, Yongyang Cai, and Lixin Ye for advice and guidance. I thank Hengshui High School for providing the data on the number of students achieving each score in Hebei. I am grateful to the participants of the seminars held at the China Meeting of the Econometric Society (Shanghai), the International Association for Applied Econometrics Annual Conference (Montréal), the European Meeting of the Econometrics Society (Cologne), and the Annual Conference of the European Association for Research in Industrial Economics (Athens). I also acknowledge the grant received from the International Association for Applied Econometrics and the allocation of computing time at the Ohio Supercomputer Center. Any errors that remain are solely mine.

†School of Public Finance and Taxation, Southwestern University of Finance and Economics, email: puyun@swufe.edu.cn

1. Introduction

On June 6 – 8 every year, approximately ten million high school seniors take the college entrance exam (Gaokao) in China. For most of these young adults, the exam is the most important 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) that 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. The offices use a variant of either the Boston mechanism (BM, the sequential mechanism or Shunxu Zhiyuan) or the deferred acceptance mechanism (DA, the parallel mechanism or Pingxing Zhiyuan) to assign students to colleges according to their exam-grade-based ranking, their preferences, and the admission quotas.[1]

Under the BM, students' first-choice colleges are critical. If a student is rejected by the first-choice college, she will probably be assigned to a substantially worse college or even have no assignment. Thus, students must pick their first choice carefully, and many complain about the risks in the decision-making process. By contrast, under the DA, if a student is rejected by one of her choices, the next choice will be considered. Since a student's admission is guaranteed only if the exam score is above the cutoff threshold[2] for one of the colleges in the preference list, students feel safer under this mechanism. In addition, the BM has inferior theoretical properties: the DA is Pareto optimum for college admissions in China[3], while the BM is not. Furthermore, the BM is not strategy-proof. Under the BM,

[1]See Appendix C for details.

[2]The cutoff threshold is the score of the lowest-ranked student admitted by a college.

[3]In college admissions in China, all colleges share the same priority for students. This priority is strict.

students strategically select their preferences; hence, the true college preferences cannot be determined directly from the data. Due to these drawbacks, the BM was abandoned in China.[4]

However, the three disadvantages of the BM mentioned above can be called into question. First, both mechanisms assign the same number of students to colleges, and the admission quotas of the colleges remain the same under the two mechanisms. If the preferences of students are sufficiently homogeneous,[5] exchanging assignments between two students would harm one of them. Thus, not all students would benefit from the switch from the BM to DA — at least some of the students will sacrifice other benefits to enjoy the safety of the DA. Second, although the DA is Pareto optimum, it may not yield the highest social welfare when preferences are cardinal, that is, when students are able to not only prefer one college to another but *strongly* prefer one college to another. To the best of my knowledge, the general performance of the DA under cardinal preference has not been studied. Abdulkadiroğlu, Che and Yasuda (2011, 2015) have demonstrated that the DA is not the best mechanism when no priorities exist or when the priorities of students are coarse. Third, although under the BM the true preferences cannot be learned directly from the data, we can still learn the preferences after constructing a model of the mechanism. Therefore, the following questions arise: can we estimate colleges' quality (or more accurately, their attractiveness) by using manipulated preference data under the BM? Can we compare the empirical performance of the BM to that of the DA in China? Can we estimate the welfare loss (or gain) for an

[4]Before 2001, the BM was implemented in all provinces. The DA was first introduced in Hunan province in 2001. By 2012, the BM was applied in only three of thirty-one provinces (Chen and Kesten, 2017).

[5]In China, 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 students should be homogeneous.

individual student when the mechanism switches from the BM to the DA?

In this paper, I develop a model of the BM to recover the true college preferences of students and then perform welfare analysis. To the best of my knowledge, current models in the empirical literature (Agarwal and Somaini, 2018; Calsamiglia, Fu and Güell, 2017; He, 2017; Hwang, 2016) follow coarse school priorities. Such schools do not differentiate between some of the students; thus, many students share the same priority. In this case, the number of priorities is small. For example, Agarwal and Somaini (2018) used only three different priorities for all students. However, in the Chinese case, priorities are strict: colleges differentiate students based on their scores, and each student has a unique rank. In this case, the number of priorities is large. The large number of priorities can make the existing models computationally intractable due to the curse of dimensionality. In addition, the models in the literature require the preference data of individual students (that is, micro data). However, in college admissions in China, such data are highly confidential. For example, Li, Gan and Yang (2010) used such micro data, but they anonymized the names of the provinces due to the sensitivity of the data. The use of micro data also restricts the future reusability of the model because the micro data are not always available. I develop a model that can handle such strict priorities and make estimates using public data: the admission quotas and cutoff thresholds of the colleges.

I estimate the model using public data from the provinces of Guangxi, Hebei, and Sichuan and conduct counterfactuals calculating social and individual welfare under the DA. In these three provinces for the given years, the students apply to colleges after receiving their scores and are assigned to colleges based on the BM. I find that total welfare under the DA is 1.73% – 6.63% lower than that under the BM. The welfare loss from switching from the

BM to the DA is a consequence of the cardinal preferences, given that the DA is the Pareto optimum.[6] I also find that the cutoff thresholds under the BM are looser than those under the DA. Many provinces switching from the BM to DA saw stricter cutoff thresholds after switching.[7] This change reveals the cost of students' enjoying increased safety under the DA: It becomes more difficult to receive admission if one's rank is not competitive. Intuitively, "good" students benefit from increased safety under the DA, while "bad" students suffer from stricter cutoff thresholds. However, how good should a "good" student be? According to my results, only 0.64% – 10.65% of students above the key cutoff threshold[8] benefit from mechanism switching in Round 1 admission; most students suffer from switching.

The rest of this paper is organized as follows, I review the literature in Section 2. I present an example in Section 3. The model to recover true preferences is presented in Section 4. I describe the data in Section 5 and present the results in Section 6. Finally, Section 7 concludes the paper.

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 the DA and that any mechanism, including the DA, is Pareto dominated by the

[6] See Section 3 for an example.

[7] See, e.g., http://news.sohu.com/20140713/n402175085.shtml.

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

top trading cycle (TTC) mechanism. In college admissions in China, a given student has the same priority in any college since colleges rank students based only on their total scores. That is, the DA and TTC are equivalent; therefore, the DA is Pareto efficient and free of justified envy. Ergin and Sönmez (2006) demonstrate that the DA is more efficient than the BM. They provide 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 the DA, students report their preferences truthfully[9]. Hence, students living in region M (or N) will be admitted by school N (or M, respectively). In the BM, students are afraid to report their true preferences and thus do not have the highest priority in their favorite schools. Students 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 the BM is not as efficient as the DA. This example relies heavily on the assumption that 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 the others do[10]. Thus, Ergin and Sönmez (2006) is unable to show that the DA performs better than the BM in college admissions 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

[9]Dubins and Freedman (1981) show that students are unable to improve their utility by lying under the DA.

[10]This is called the serial dictatorship mechanism.

can prefer one school to another but cannot *strongly* prefer one school to another. These assumptions are not realistic.

Abdulkadiroğlu, Che and Yasuda (2011) considered incomplete information and cardinal preference and assumed that schools have no priority and that although students have the same ordinal preference, their cardinal preferences may be different. They showed the BM performing slightly better than the DA. Abdulkadiroğlu, Che and Yasuda (2015) generalize this idea, illustrating that the DA is not Pareto efficient when the priorities are coarse under complete information and cardinal preference. That is, these papers assume that the priorities are not strict, an assumption that is essential for their results. Unfortunately, priorities in college admissions in China are strict. Therefore, these theoretical papers do not apply to the Chinese case.

My paper is related to four previous papers using structural models to compare the two mechanisms empirically, all of which found that the BM is more efficient than the DA. He (2017) studied Chinese high school admissions and found that students suffer, regardless of whether they are naïve or sophisticated, if others perform as in the data after the switch from the BM to DA. Calsamiglia, Fu and Güell (2017) analyzed public school admissions in Barcelona, Spain using a counterfactual analysis showing that average welfare decreases by 1020 euros when switching from the BM to DA. Agarwal and Somaini (2018) scrutinized public elementary school admissions in Cambridge, MA. They found that the immediate acceptance mechanism (a variant of the BM) performed better than the DM. Hwang (2016) proved that both naïve and sophisticated students follow a simple rule, which is used to partially identify the model. In the empirical application, ex ante welfare was found to be high in the BM.

However, the models in all the empirical papers studying the BM follow the coarse schools' priorities. If the priorities are strict, the number of priorities is large, which makes the existing models impossible to compute due to the curse of dimensionality. Nevertheless, models that present strict priorities are still rare in the theoretical literature. Comparing two mechanisms that fall under strict priorities helps to test the theoretical literature in scenarios outside their assumptions. In addition, the existing models use the maximum log-likelihood or the moment inequality for the estimation and thus rely on individual preference data, which are highly confidential and unavailable in the Chinese case. Thus, departing from the literature, I develop a BM model that can handle strict priorities and does not require micro data; the model requires only the admission quotas and cutoff thresholds of the colleges, which are public knowledge and can be found in the college application guides.

3. Motivating Example

I present an example to illustrate why the welfare loss when switching from the BM to DA is a consequence of cardinal preferences. Suppose that there are two colleges M and N and five students x, y, z, w, and r. M will admit one student, and N will admit three students. x is the top student, y the second, z the third, w the fourth, and r the bottom. M is known to be a better school than N, but the students do not know other students 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 the 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 the DA if the preferences are ordinal;

by contrast, if the preferences are cardinal, I can further assume that y prefers M *slightly*, while z prefers M *strongly*. Under this assumption, the DA is no longer optimum because assigning z to M and y to N will yield higher social welfare. Under the 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 strongly prefers M: 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 the BM, yielding higher social welfare than that under the DA. Intuitively, this result stands: since students are allowed to express their cardinal preferences under the BM but not under the DA, the BM potentially yields better social welfare than does the DA.

This example is consistent with our empirical findings. In the example, the cutoff for M is 3 (at z) under the BM and 2 (at y) under the DA. However, in the results, I find that the cutoff thresholds under the BM are looser than those under the DA. In addition, intuitively, "good" students (such as y) will benefit from the increased safety under the DA, while "bad" students (such as z) will suffer from the stricter cutoff after switching from the BM to DA. In the results, none of the bottom half of the students above the key cutoff benefit from the switch.

4. Model

I propose a model to recover the true preferences of students from the data of cutoffs and admission quotas generated under BM. Suppose that there are L colleges in China. College l

has quota A_l for a given province[11]. Therefore, the college can admit up to A_l students. The students are ranked: student i is the ith-highest-ranked student. I 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. A student is able to observe her own rank and quotas $\boldsymbol{A}$ (but not cutoffs $\boldsymbol{N}$) when submitting the preferences.

Each student has $L+1$ available choices, including the outside option $(\{1, 2, \ldots, L\} \bigcup \{\emptyset\})$. A student will receive $\xi_l + \varepsilon_{il}$ in utility provided the student is admitted by college l or 0 if assigned to the outside option. Here, ξ_l is the mean attractiveness [12] of the college to the students, whereas ε_{il} is student i's private preference in addition to ξ_l. I further assume that ε_{il} is unknown to other students, but its distribution $\mathcal{E}$ remains common knowledge. Moreover, the students submit their major (disciplinary) preferences for each college that they choose. A college then assigns students to majors. Students can reject the assignment, but the rejection may result in their being admitted to much worse colleges. Wu and Zhong (2014) indicates that almost all the students accept the assignment. Therefore, I can separate admission into two stages. In the first stage, the students are assigned to colleges, which are considered as composite goods. In the second stage, the students are assigned to majors. This argument was proposed by Wu and Zhong (2014). In this paper, I only consider the first stage. Any major preference for college l of a student is in ξ_l.

I consider a Bayesian Nash equilibrium. Student i will use a strategy $\sigma_i(\boldsymbol{\xi}, \varepsilon_{il}; \sigma_{-i}, \mathcal{E})$ to maximize the expected return. To simplify the model, I further assume that only the first step of BM is considered. College admission is quite competitive in China, and all

[11]As described in the Appendix C (page 37), I consider the admission process for each province independently, so the model is for one province only. Additionally, I consider only Round 1 admission.

[12]We use "attractiveness" and "quality" interchangeably in this paper. Indeed, the attractiveness is the quality of a college as perceived by the students.

spaces at most colleges are filled in the first step of BM. Then, student i expects to receive $(\xi_l + \varepsilon_{il})\mathbb{P}^a_l(i, \boldsymbol{A}; \mathcal{E}, \sigma_{-i}, \boldsymbol{\xi})$[13] in utility provided college l is chosen, where $\mathbb{P}^a_l(i, \boldsymbol{A})$ is the admission probability of student i for college l. A student will apply a strategy to maximize the expected return. Therefore, student i will choose college l^* if and only if college l^* maximizes $(\xi_l + \varepsilon_{il})\mathbb{P}^a_l(i, \boldsymbol{A})$ and the utility is positive. If student i receives negative utility from choosing all the colleges, he/she will choose the outside option and receive 0 in utility. Consistent with Fack, Grenet and He (2019); He and Magnac (2018), I also introduce cost into the application: students will not consider colleges with very small admission probability. This assumption has little effect on the welfare analysis. Students receive nearly nothing from choosing a very-low-admission-chance (VLAC, hereafter) college. However, this assumption technically removes multiple "numerical equilibriums" for the ease of computation: student i receives very similar utility from a VLAC college and the outside option with a nonzero probability. Therefore, each student i solves a maximization problem,

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

where α is a small positive number. I also solve $\mathbb{P}^a_l(i, \boldsymbol{A})$ recursively.

Lemma 1. $\mathbb{P}^a_l(i, A_l; A_{-l})$[14] $= \mathbb{P}^a_l(i-1, A_l; A_{-l})(1-\mathbb{P}^c_l(i-1, \boldsymbol{A})) + \mathbb{P}^a_l(i-1, A_l-1; A_{-l})\mathbb{P}^c_l(i-1, \boldsymbol{A})$ *for* $i \geq 2$ *and* $A_l \geq 1$*, where* $\mathbb{P}^c_l(i, \boldsymbol{A})$ *is the chance of student* i *choosing college* l*. In addition,* $\mathbb{P}^a_l(1, A_l; A_{-l}) = 1$ *for* $A_l \geq 1$ *and* $\mathbb{P}^a_l(i, 0) = 0$ *for* $i \geq 1$*.*

The proof is in Appendix D.1 on page 40. As an intuitive example, suppose that college l

[13] Hereafter, slightly abusing the notation, I use $\mathbb{P}^a_l(i, \boldsymbol{A})$ instead of $\mathbb{P}^a_l(i, \boldsymbol{A}; \mathcal{E}, \sigma_{-i}, \boldsymbol{\xi})$.

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

admits one student ($A_l = 1$). Then, I can simplify the equation as

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

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

In Lemma 1, I 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, as is student i's choice. Therefore, this lemma indicates that the equilibrium of this game and the choice of student i are unique almost surely.

Although the strategies represented in the model are difficult for the students, most of the students will still act as the model suggests and will not perform different sophistication for the following two reasons. First, as noted, the college entrance exam is the most important exam in China and, in general, in students' lives. Second, I consider only Round 1 admission in this paper. Fewer than the top 10% of students are considered in this round, and the quality of the colleges involved in this round is better than that in other rounds. Students in this round cherish the opportunities they are presented with, and thus, will try extra hard to perform optimally.

I now want to estimate the college's attractiveness ξ_l. Then, I will know the preferences of the students. The next theorem establishes a mapping from the quota $\boldsymbol{A}$ and the cutoff threshold $\boldsymbol{N}$ to $\boldsymbol{\xi}$

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

The proof is provided in Appendix D.2 on page 41. 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. The law of large numbers can be used to further loosen the assumption of the behaviors of the students: even if some of the students do not perform optimally, their misbehavior will be averaged and will diminish in the large sample. I have two remarks on this theorem. First, $\boldsymbol{\varepsilon}$ are the random variables in the theorem. They constitute the actual choices of the students and, thereafter, A_l. Second, the cutoff N_l being large requires the total number of students admitted to be large, which is equivalent to admission quotas of most colleges being large and/or the number of colleges being large. If the admission quotas are not large for at least some colleges, the BM may not collapse to the DA. In an extreme example, suppose there are T colleges with the same mean attractiveness, where T is a large number. Each college admits one student. For student $T/2$, the admission probability of any college is 50%, and the student may be not admitted by the favorite feasible college. 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; thus, the assumption of a 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 $\boldsymbol{\xi}$. I use the second mapping to estimate the mean attractiveness $\boldsymbol{\xi}$ and therefore reveal the true preferences of students. Notably, $\boldsymbol{N}$ are assumed to be integers in the theorem. If $\boldsymbol{N}$ are not only integers, I extend the theorem to be

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

where $\lfloor N_l \rfloor$ is the largest integer not larger than N_l. This ensures that both mappings will be continuous. The identification, estimation, and counterfactual analysis methods are discussed in Appendix A and B.

5. Data

To recover the mean attractiveness $\boldsymbol{\xi}$ and compare the two mechanisms, I need the data of the quotas $\boldsymbol{A}$ and the cutoff thresholds $\boldsymbol{N}$. I collect data for Guangxi, Hebei, and Sichuan[15] from different sources. In Guangxi, the Guangxi Provincial Academy of Recruitment and Examination (Gvangjsih Cauhswngh Gaujsi Yen in the 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. I calculate the cutoff threshold for each college from the lowest score for students admitted to each college and the number of students with 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 students are eligible to apply to colleges in the second and fourth round, while the choices of major and college are limited in Round 0. Thus, most highly ranked students apply in Round 1. In this paper, I combine Round 1 and Round 1 for college prep into one round and study this round only. I 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" ("Quan-

[15] I plot the locations of the three provinces in Figure 1 on page 14.

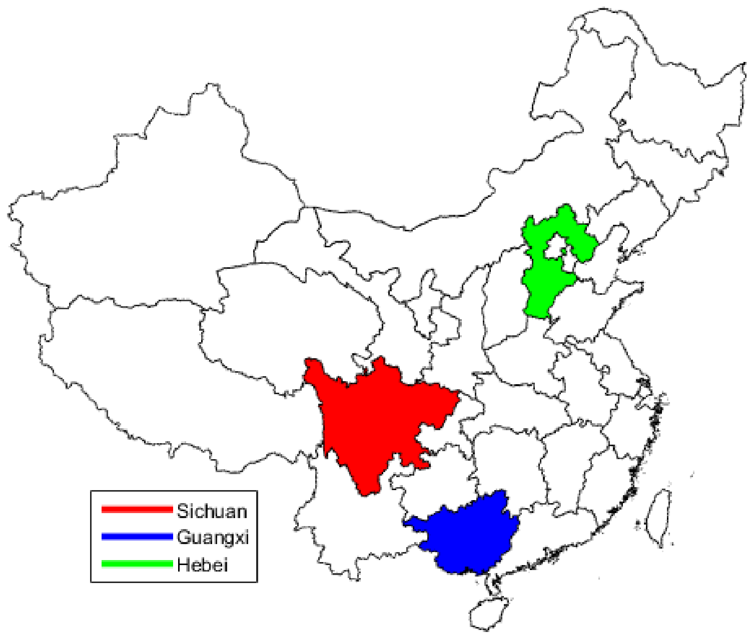


Fig. 1. Location of the Three Chinese Provinces

guo 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, I could not locate page 166 of these statistics for science major students, so I 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. Therefore, I assume that these two colleges did not admit students for the 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, I address it by again assuming that this college did not admit science major students in 2007. Since Xi'an International Studies University in fact admitted only five science majors from Hebei that year, this assumption does not significantly affect the results. I 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, I analyze only 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 thresholds stricter than that college (i.e., $\sum_{j=1}^{l} A_{(j)}$), leading to nonexistent results (i.e., $\mathbf{\Phi^{-1}} = \varnothing$) based on Theorem 2. This may be caused by errors in the data and/or my simplifications. To solve this issue, I 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, I 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 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 I corrected the score to 618 from 518. In the admission process, the first two rounds are Round 0 and Round 1; I consider only Round 1 in this paper.

The student placement office in each province conducts admissions for science and arts majors separately; thus, I 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.

6. Empirical Results

The results are similar for all three provinces; thus, I report only the results for science majors in Guangxi for 2008 (the other results are reported in Table I – IV on Page 22 – 24). For the BM model, Figure 2 on page 17 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. I emphasize that attractiveness is the average preference of the students; if a given student chooses a college, she must receive nonnegative utility from it because she receives zero from the outside option. A student may be interested only in some small number of colleges (such as high-ranking ones) and thus may receive positive utility only from these colleges. Thus, the average preference (or attractiveness) is negative for most colleges.

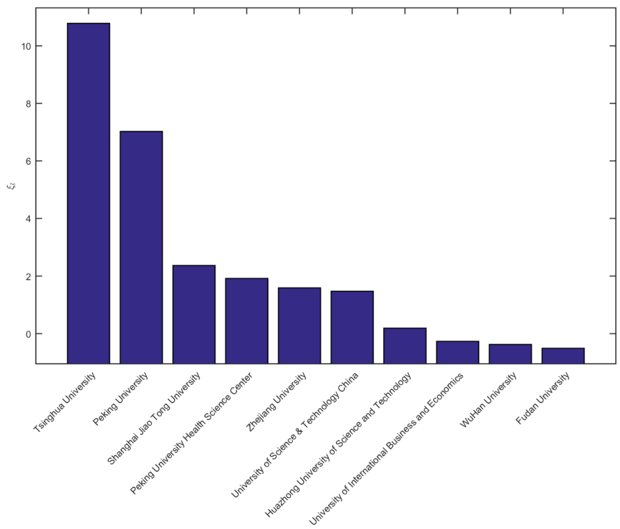


Fig. 2. Attractiveness ξ_l of Top 10 Universities for Science Major Students in Guangxi for 2008

Then, I simulate and compare the BM and DA based on estimated attractiveness. The cumulative welfare change when switching from the BM to DA is reported in Figure 3, on page 18, and the individual change is reported in Figure 4, on page 19. In these two graphs, the x-axis is the rank of a student, and the y-axis is the average utility change for all students ranked slightly better than the student (Figure 3) or the utility change of this student (Figure 4). 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. However, Figure 5 on page 20 indicates that only 129 (0.9%) students in fact benefit from the switch, while 14241 students are above the key cutoff threshold. All of these students are among the top 197 students, as shown in Figure 4. Social welfare would increase after the switch only if fewer than 335 students existed, as suggested in Figure 3.

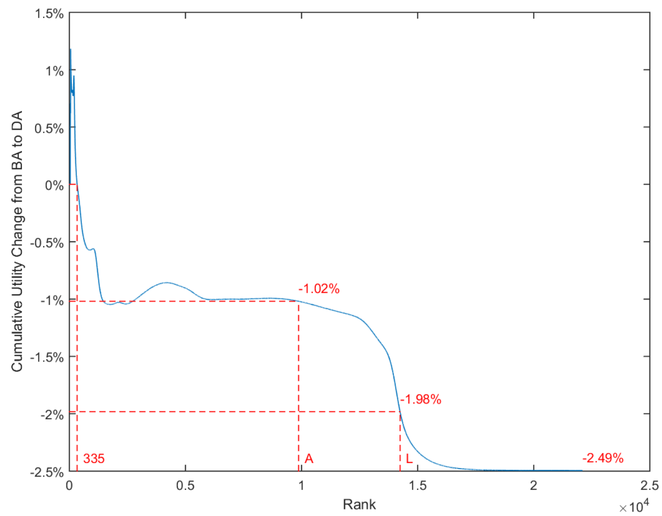


Fig. 3. Cumulative Welfare Change When Switching from the BM to DA for Science Major Students in Guangxi for 2008

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

In Figure 6, on page 21, the cutoff thresholds of all colleges become stricter after the switch. This is the reason for the welfare loss: a student who can be admitted under the BM

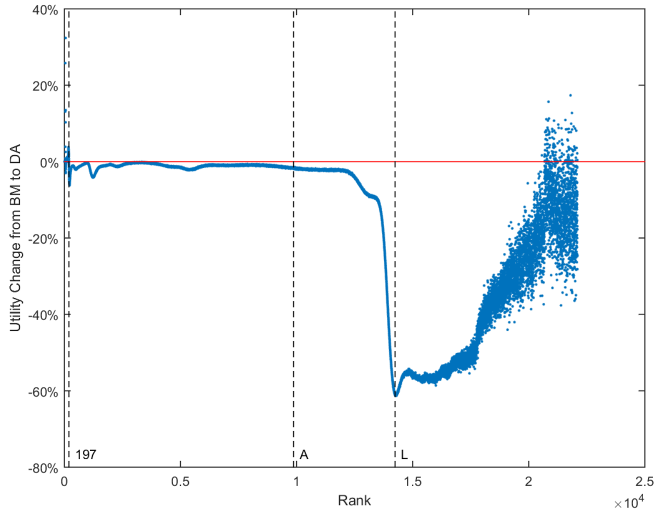


Fig. 4. Individual Welfare Change When Switching from the BM to DA for Science Major Students in Guangxi for 2008

may be rejected under the DA. As shown in Figures 3 and 4, students who receive scores slightly 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 the BM but not under the DA because of the stricter cutoff, so they receive much lower utility under the DA.

In Figure 4, 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 the simulation increases when utility is close to zero, which contributes to the noisiness of the results.

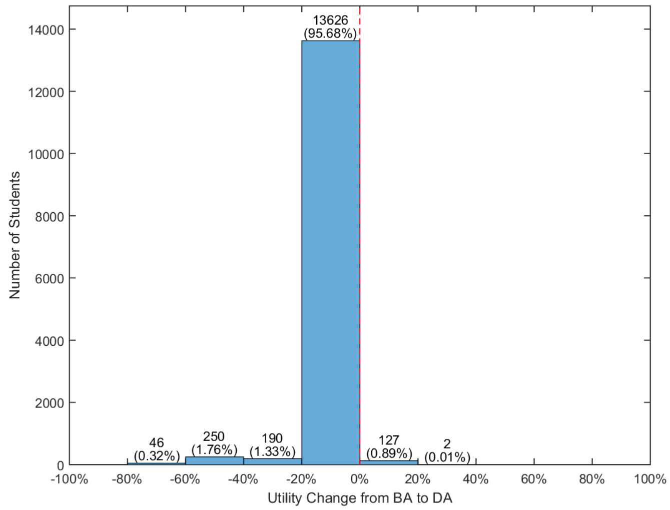


Fig. 5. Histogram of Welfare Change When Switching from the BM to DA of Science Major Students above the Key Cutoff Threshold (Top 14241 Students) in Guangxi for 2008

7. Conclusion

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

This paper makes the following contributions. First and most importantly, this paper shows that from a social welfare perspective, the BM is a better approach than the DA to conducting college admissions in China. Historically, the BM was implemented in all the provinces of the Chinese mainland, whereas currently, the DA is employed by most provinces. The results indicate that this switch from the BM to DA has been costly: the total welfare

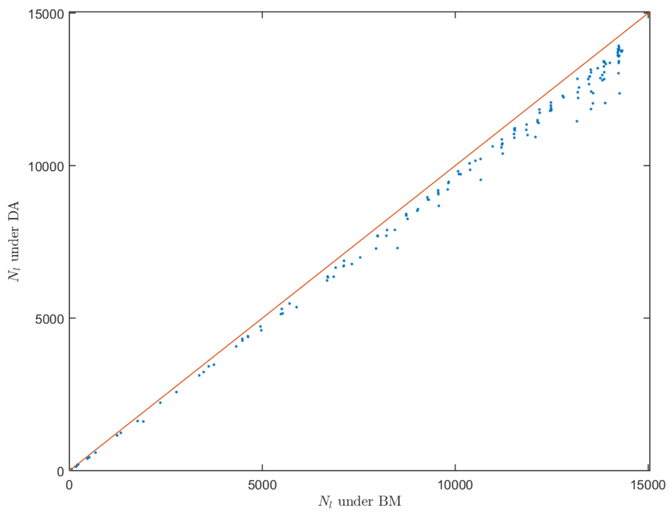


Fig. 6. Cutoff Threshold of Each University in the BM vs. DA for Science Major Students in Guangxi for 2008

of students has decreased 1.73% – 6.63% due to the switch. On the Chinese mainland, approximately 150 colleges 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 quality is the same. Further, if I assume that one key college is worth 1 billion dollars, the switch costs 2.595 – 9.945 billion dollars.

Second, to the best of my knowledge, the existing literature considers only the coarse priorities scenario for comparing the two mechanisms under cardinal preferences. This paper indicates that the BM performs better than the DA not only in the coarse priorities scenario, as discussed in the literature, but also in the strict priorities scenario, as described in the Chinese case.

Third, departing from the literature, my model does not need micro data on the submitted preferences of individual students; instead, the model requires only the admission quota and

Table I: 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
		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.01523$		$1/L\|\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\|_1 = 0.00327$

cutoff threshold for each college. Micro data are difficult to obtain and may be restricted for privacy reasons; my use of public data makes the results easier to replicate and makes the model potentially more widely usable.

In summary, in this paper, I find that the BM outperforms the DA in college admissions in China. However, is BM the best possible mechanism? Future studies may want to propose new mechanisms that are better than BM or, conversely, to prove that the BM yields the highest welfare in college admissions in China.

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

Ranking	Arts Majors for 2005 Name	ξ_l	Science Majors for 2005 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 III: Top Ten Colleges in terms of ξ_l in Sichuan

	Arts Majors for 2006		Science Majors for 2006	
Ranking	Name	ξ_l	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
		$1/L\Vert\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\Vert_1 = 0.00044$		$1/L\Vert\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\Vert_1 = 0.00097$
	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
		$1/L\Vert\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\Vert_1 = 0.00024$		$1/L\Vert\boldsymbol{\Psi}(\boldsymbol{\xi}^*, \boldsymbol{A}) - \boldsymbol{N_0}\Vert_1 = 0.00059$

Table IV: 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
Sichuan	2007	Science	541 (2.09%)	1962 (7.59%)	2098 (8.12%)	-3.27%	-3.58%	20948	25839

[a] G: the number of students who benefit from the switch and the proportion of these students in all students above the key cutoff threshold in parentheses (i.e. Figure 5 on page 20).
I: the last student in terms of rank who benefits from the switch and is above the key cutoff threshold; the proportion of the rank in the key cutoff in parentheses (i.e., the first vertical line from the left in Figure 4 on page 19).
Breakeven: the maximum number of students where social welfare may increase after the switch; the percentage of this number in the key cutoff in parentheses (i.e. the first vertical line from the left in Figure 3 on page 18).
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.

References

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

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

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

Agarwal, Nikhil, and Paulo Somaini. 2018. "Demand analysis using strategic reports: An application to a school choice mechanism." *Econometrica*, 86(2): 391–444.

Arnosti, Nick, Ramesh Johari, and Yash Kanoria. 2018. "Managing congestion in matching markets." Available at SSRN: `https://ssrn.com/abstract=2427960` or `http://dx.doi.org/10.2139/ssrn.2427960`.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica*, 841–890.

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2017. "Structural estimation of a model of school choices: The Boston mechanism vs. its alternatives." Unpublished manuscript. Available online at: `https://www.ssc.wisc.edu/~cfu/Barcelona.pdf`.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. "Constrained school choice: An experimental study." *American Economic Review*, 100(4): 1860–74.

Chen, Yan, and Onur Kesten. 2016. "Chinese college admissions and school choice reforms: An experimental study." Unpublished manuscript. Available online at: http://yanchen.people.si.umich.edu/papers/Chen_Kesten_Experiment_2016_03.pdf.

Chen, Yan, and Onur Kesten. 2017. "Chinese college admissions and school choice reforms: A theoretical analysis." *Journal of Political Economy*, 125(1): 99–139.

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

Dubins, Lester E, and David A Freedman. 1981. "Machiavelli and the Gale-Shapley algorithm." *The American Mathematical Monthly*, 88(7): 485–494.

Ergin, Haluk, and Tayfun Sönmez. 2006. "Games of school choice under the Boston mechanism." *Journal of Public Economics*, 90(1): 215–237.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019. "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions." *American Economic Review*, 109(4): 1486–1529.

Fonda, Alessandro, and Paolo Gidoni. 2016. "Generalizing the Poincaré–Miranda theorem: The avoiding cones condition." *Annali di Matematica Pura ed Applicata (1923 –)*, 195(4): 1347–1371.

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

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

He, Yinghua. 2017. "Gaming the Boston school choice mechanism in Beijing." Unpublished manuscript. Available online at: https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/by/he/bm.pdf.

He, Yinghua, and Thierry Magnac. 2018. "A Pigouvian Approach to Congestion in Matching Markets." IZA Discussion Paper No. 11967. Available at SSRN: https://ssrn.com/abstract=3301721.

Hwang, Il Myoung. 2016. "A robust redesign of high school match." Unpublished manuscript. Available online at: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxzYW1pbG15b3VuZ2h3YW5nfGd4OjZlN2Y0ZDVhODY5NzkOODU.

Jones, Todd R, and Michael S Kofoed. 2017. "Do peers influence occupational preferences? Evidence from randomly-assigned peer groups at West Point." Unpublished manuscript. Available online at: http://www.toddrjones.com/files/JonesKofoed.pdf.

Kallenberg, Olav. 1997. *Foundations of Modern Probability.*

Li, Feng, Li Gan, and Xiaoling Yang. 2010. "Research on application timing and admission mechanism in aational college entrance exam." *Educational Research*, 10.

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

Pathak, Parag A, and Tayfun Sönmez. 2013. "School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation." *The American Economic Review*, 103(1): 80–106.

Sönmez, Tayfun, and Tobias B Switzer. 2013. "Matching with (branch-of-choice) contracts at the United States military academy." *Econometrica*, 81(2): 451–488.

Wu, Binzhen, and Xiaohan Zhong. 2014. "Matching mechanisms and matching quality: Evidence from a top university in China." *Games and Economic Behavior*, 84: 196–215.

Zhu, Min. 2014. "College admissions in China: A mechanism design perspective." *China Economic Review*, 30: 618–631.

Online Appendix (Not For Publication)

Appendix A. Discussions on Identification

In this section, I consider a more general form of the utility function, where the outside option is ε_{i0} instead of 0. In the following context, I will prove the similarity of the BM and DA when admission quotas $\boldsymbol{A}$ are large. With the help of the DA model, I will show the identification of $\boldsymbol{\xi}$ under the BM. First, the DA is equivalent to the serial dictatorship mechanism in college admissions in China because a student has the same priority in any college (Abdulkadiroglu and Sönmez, 2003). Therefore, in the 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 the DA, the student is assigned to l_3. Any mechanism assigning students to the best available colleges is equivalent to the DA.

Let me construct an imaginary mechanism. This mechanism is the same as the 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 colleges will reject them based on their ranking. Thus, the students will list the best available college as their first choice, and they will be admitted by these colleges. Thus, this imaginary mechanism is equivalent to the DA, and I can use this imaginary mechanism to study the 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). I emphasize that this assumption still holds given freely available information on the cutoff thresholds. To see why, I 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, $\mathbb{P}^c_l(i,\boldsymbol{A}|\boldsymbol{N})\mathbb{P}(\boldsymbol{N}) = \mathbb{P}^c_l(i,\boldsymbol{A},\boldsymbol{N}) = \mathbb{P}^c_l((i,\boldsymbol{A}|\boldsymbol{N}),\boldsymbol{N})$. $\mathbb{1}_l(i|\boldsymbol{N})$ and $\boldsymbol{N}$ are independent, so $\mathbb{1}_l(i|\boldsymbol{N})$ is independent with $\mathbb{1}_l(-i|\boldsymbol{N})$. Intuitively, I may provide fake cutoff thresholds $\boldsymbol{N}^f$ to the students; in such a situation, student 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, I 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., $\forall l\ \mathbb{P}^a_l(i,\boldsymbol{A}) = 1$). Thus, these students compare the utility from each of L colleges, as well as from the outside option, and choose the best one. Student $N_{(1)}+1$ to student $N_{(2)}$ will be rejected by college (1) but accepted by other colleges (i.e., $\forall l \neq (1)\ \mathbb{P}^a_l(i,\boldsymbol{A}) = 1$ and $\mathbb{P}^a_{(1)}(i,\boldsymbol{A}) = 0$). Then, the students compare the utility from each of the $L-1$ colleges and 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., $\forall l \neq (L)\ \mathbb{P}^a_l(i,\boldsymbol{A}) = 0$ and $\mathbb{P}^a_{(L)}(i,\boldsymbol{A}) = 1$); thus, these students compare the utility from college (L) and from the outside option. The remaining students will be rejected by all colleges (i.e., $\forall l\ \mathbb{P}^a_l(i,\boldsymbol{A}) = 0$) and fully understand this fact;

thus, these students choose the outside option.

The general expression of this model is tedious. I consider $\varepsilon_{il}+\gamma$ to be i.i.d. extreme value type 1 distributed, where $\gamma = 0.5772156649...$ is the Euler constant. This assumption ensures that the mean of ε_{il} is zero. The proof of other types of distributions of private information is similar. From Theorem 1, I obtain the relationship between the average behavior and mean behavior.

$$
\begin{aligned}
&\frac{A_{(1)}}{N_{(1)}} - \frac{\exp\left(\xi_{(1)}\right)}{1+\sum_{l=1}^{L}\exp\left(\xi_{(l)}\right)} \xrightarrow{a.s.} 0 \\
&\frac{A_{(2)}}{N_{(2)}} - \frac{N_{(1)}}{N_{(2)}}\frac{\exp\left(\xi_{(2)}\right)}{1+\sum_{l=1}^{L}\exp\left(\xi_{(l)}\right)} - \frac{N_{(2)}-N_{(1)}}{N_{(2)}}\frac{\exp\left(\xi_{(2)}\right)}{1+\sum_{l=2}^{L}\exp\left(\xi_{(l)}\right)} \xrightarrow{a.s.} 0 \\
&\cdots \\
&\frac{A_{(L)}}{N_{(L)}} - \frac{N_{(1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\sum_{l=1}^{L}\exp\left(\xi_{(l)}\right)} - \frac{N_{(2)}-N_{(1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\sum_{l=2}^{L}\exp\left(\xi_{(l)}\right)} - \dots - \frac{N_{(L)}-N_{(L-1)}}{N_{(L)}}\frac{\exp\left(\xi_{(L)}\right)}{1+\exp\left(\xi_{(L)}\right)} \xrightarrow{a.s.} 0
\end{aligned}
\tag{2}
$$

where A_l/N_l is the average behavior, and the other terms are the mean behavior. Similar to the BM, the DA model relates $\boldsymbol{\xi}$, $\boldsymbol{A}$ and $\boldsymbol{N}$ by Equation 2. $\boldsymbol{\Phi} : \boldsymbol{\xi}, \boldsymbol{A} \to \boldsymbol{N}$ 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 2. *Based on* $\boldsymbol{\Phi^{-1}}$

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

for all $l < L$ *and*

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

Theorem 3. *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* $\boldsymbol{\Phi}$

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

for all $l < L$ and

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

Proof. First, I shall prove Theorem 2. I know $\boldsymbol{N}$ and $\boldsymbol{A}$. I want to obtain $\boldsymbol{\xi}$. Equation 2 tells us the relationship.

$$A_{(l)} = N_{(1)} \frac{\exp\left(\xi_{(l)}\right)}{1 + \sum_{k=1}^{L} \exp\left(\xi_{(k)}\right)} + (N_{(2)} - N_{(1)}) \frac{\exp\left(\xi_{(l)}\right)}{1 + \sum_{k=2}^{L} \exp\left(\xi_{(k)}\right)} + \ldots + (N_{(l)} - N_{(l-1)}) \frac{\exp\left(\xi_{(l)}\right)}{1 + \sum_{k=l}^{L} \exp\left(\xi_{(k)}\right)} \tag{3}$$

$$A_{(l+1)} = N_{(1)} \frac{\exp\left(\xi_{(l+1)}\right)}{1 + \sum_{k=1}^{L} \exp\left(\xi_{(k)}\right)} + (N_{(2)} - N_{(1)}) \frac{\exp\left(\xi_{(l+1)}\right)}{1 + \sum_{k=2}^{L} \exp\left(\xi_{(k)}\right)} + \ldots + (N_{(l+1)} - N_{(l)}) \frac{\exp\left(\xi_{(l+1)}\right)}{k + \sum_{k=l+1}^{L} \exp\left(\xi_{(k)}\right)} \tag{4}$$

where Equation 3 and Equation 4 are two lines of Equation 2. I multiply Equation 3 by $\exp\left(\xi_{l+1}\right) / \exp\left(\xi_{l}\right)$ and substitute the result into Equation 4. I obtain

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

Rearranging the equation yields

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

where I define $\sum_{k=L+1}^{L} \exp\left(\xi_{(k)}\right) = 0$. Summing $N_{(2)} - N_{(1)}$, $N_{(3)} - N_{(2)}$, ..., $N_{(l+1)} - N_{(l)}$ yields

$$N_{(l+1)} = \sum_{k=1}^{l+1} A_{(k)} + \frac{1 + \sum_{k=l+2}^{L} \exp\left(\xi_{(k)}\right)}{\exp\left(\xi_{(l+1)}\right)} A_{(l+1)} \tag{5}$$

Rearrangement then gives

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

Theorem 2 has been proven.

Now let us consdier Theorem 3. If I know how to map l to (l), the proof has been completed in Equation 5. However, I need to know $\boldsymbol{N}$ to generate the mapping from l to (l). $N_{(l)}$ is the lth smallest value in $\boldsymbol{N}$. I only know $\boldsymbol{A}$ and $\boldsymbol{\xi}$. I shall prove $A_l / \exp\left(\xi_l\right)$ generating the same mapping from l to (l).

First, I show the existence of $\boldsymbol{N}$. I rank $A_l/\exp(\xi_l)$ from the smallest to the largest as $A_{[1]}/\exp(\xi_{[1]})$, $A_{[2]}/\exp(\xi_{[2]})$, ..., $A_{[L]}/\exp(\xi_{[L]})$. I have

$$\begin{aligned}
&\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq \frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &(1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &A_{[l+1]} + (1+\sum_{k=l+2}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})} \\
\Longleftrightarrow &\sum_{k=1}^{l+1}A_{[k]} + (1+\sum_{k=l+2}^{L}\exp(\xi_{[k]}))\frac{A_{[l+1]}}{\exp(\xi_{[l+1]})} \geq \sum_{k=1}^{l}A_{[k]} + (1+\sum_{k=l+1}^{L}\exp(\xi_{[k]}))\frac{A_{[l]}}{\exp(\xi_{[l]})}
\end{aligned} \tag{6}$$

If I let

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

I have $N_{[l+1]} \geq N_{[l]}$. $N_{[l]}$ is one set of solutions, and $\boldsymbol{N}$ exists.

Then, I show the uniqueness of $\boldsymbol{N}$. If I 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 solutions as the mapping from l to $[l]$ and that from l to (l) are equivalent. If $\boldsymbol{N}$ is not unique, I 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 ($\exists l, l'$, $N_l < N_{l'}$, $A_l/\exp(\xi_l) \geq A_{l'}/\exp(\xi_{l'})$)**.**

I rank $\boldsymbol{N}$ from 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 6, I 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 my assumption. Thus, I have $A_{(m)}/\exp(\xi_{(m)}) < A_{(m')}/\exp(\xi_{(m')})$. This contradicts $A_l/\exp(\xi_l) \geq A_{l'}/\exp(\xi_{l'})$.

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

I apply the same strategy as 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 $A_l/\exp(\xi_l) \neq A_{l'}/\exp(\xi_{l'})$.

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

The two theorems generate a one-to-one mapping between $\boldsymbol{\xi}$ and $\boldsymbol{N}$ given $\boldsymbol{A}$. The calculation is simple. However, what is the relationship between the BM and DA?

Theorem 4. $\Psi_l/A_l = \Phi_l/A_l$ *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 sufficiently small; and (4)* ε_{il} *of* Ψ_l *has the same distribution as that of* Φ_l*.*

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

$$\mathbb{P}_l^a(i, \boldsymbol{A}) = \mathbb{P}(\sum_{j=1}^{i-1} \mathbb{1}_l(j) < A_l) \tag{7}$$

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

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

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 $i > A_l$. I obtain

$$\begin{aligned}
&\mathbb{P}(\frac{1}{i-1}\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \frac{N_l}{i-1}\nu < \frac{1}{i-1}\sum_{j=1}^{i-1} \mathbb{1}_l(j) < \frac{1}{i-1}\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \frac{N_l}{i-1}\nu) = 1 \\
\Longleftrightarrow &\mathbb{P}(\frac{1}{N_l}\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu < \frac{1}{N_l}\sum_{j=1}^{i-1} \mathbb{1}_l(j) < \frac{1}{N_l}\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu) = 1 \\
\Longleftrightarrow &\mathbb{P}(\sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l < \sum_{j=1}^{i-1} \mathbb{1}_l(j) < \sum_{j=1}^{i-1} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l) = 1
\end{aligned} \tag{8}$$

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

$$\mathbb{P}(\sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) - \nu N_l < A_l < \sum_{j=1}^{N_l} \mathbb{P}_l^c(j, \boldsymbol{A}) + \nu N_l) = 1 \tag{9}$$

Combining Equation 7, Equation 8 and Equation 9 yields

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

$\mathbb{P}_l^c(j, \boldsymbol{A})$ 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 college l since she can receive positive utility only from this college. $\kappa > 0$ denotes this lower bound. $\mathbb{P}_l^c(j, \boldsymbol{A}) \geq \kappa$. If $i \leq N_l + 1 - \lceil 2\nu N_l/\kappa \rceil$,

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

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, $\mathbb{P}_l^a(i, \boldsymbol{A}) < \alpha$. Student i does not consider 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, as can ω.

Consistent with the simplified model, $\varepsilon_{il} + \gamma$ is i.i.d. extreme value type 1 distributed. For college l, I do not count the choice probabilities of the students with $0 \leq \mathbb{P}_l^a(i, \boldsymbol{A}) < 1$; thus, a college admits fewer students at the cutoff line N_l. Mathematically I have

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

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

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

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

□

This theorem indicates that the two models are equivalent when $\boldsymbol{A}$ is large. Intuitively, two types of knowledge are 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 students are guessing. Thus, the BM collapses to the DA. Combining the similarity of the BM and DA and the identification of $\boldsymbol{\xi}$ under the DA, I have the identification of $\boldsymbol{\xi}$ under the BM. I 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 admission, the admitted student is ranked 3000. In this case, $A_l = 1$, which is small, while $N_l = 3000$, which is large. This scenario is common in college admissions in China because the numbers of students and colleges are both large. Therefore, $\boldsymbol{A}$'s being large is more difficult to satisfy. Nonetheless, the two models are similar even if $\boldsymbol{A}$ is not large, based on the theorem. I employ this similarity to estimate the BM with the assistance of the DA.

Appendix B. Estimation, Simulation, and Counterfactual Analysis

B.1. Estimation

I 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 the BM, while $\varepsilon_{i0} + \gamma$ is extreme value type 1 distributed in the closed-form solution of the DA. Thus, the specifications of the two models are not exactly identical. Furthermore, the purpose of the paper is to compare the BM with the DA, so it is unreasonable to assume that the two models are equivalent in the beginning. Instead, I propose a finite quota remedy (FQR) procedure to estimate $\boldsymbol{\Psi^{-1}}$.

Step 1: Obtain $\hat{\boldsymbol{\xi}} = \boldsymbol{\Phi^{-1}}(\boldsymbol{N}, \boldsymbol{A})$. 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: Obtain $\hat{\boldsymbol{N}} = \boldsymbol{\Psi}(\hat{\boldsymbol{\xi}}, \boldsymbol{A})$. I 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: Obtain $\boldsymbol{\xi}_\zeta = \boldsymbol{\Phi^{-1}}(\boldsymbol{N}_\zeta, \boldsymbol{A})$

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 $\boldsymbol{N} - \hat{\boldsymbol{N}}$ via 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, I do not need perfect optimization; in practice, I randomly select 11 ζs within $[-1, 1]$ and choose the best one as ζ^*. The procedure stops after 100 iterations. I record $\boldsymbol{\xi}_{\zeta^*}$ in each iteration, and the best one is denoted by $\boldsymbol{\xi}^*$[16].

Finally, I explain why I am unable to directly use contraction mapping in the estimation. In Appendix I of Berry, Levinsohn and Pakes (1995), the proof of the contraction mapping

[16] In the empirical analysis, one iteration needs 11 cores of an Intel Xeon x5650 CPU to run approximately 20 – 100 minutes depending on the number of colleges and the admission quotas; that is, the optimization requires 367 – 1833 core-hours.

requires the same sign for all $\partial \log(\Psi_l)/\partial \xi_{l'}$, where $l \neq l'$. However, the empirical results indicate that this assumption fails in all cases.

B.2. Simulation

In this section, I will study the performance of FQR and compare it with $\mathbf{\Phi^{-1}}$, which uses the DA model directly to approximate the BM. In Figure 7(a) on page 38, I 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., $A_1 = A_2$). I generate their cutoff thresholds using the BM for the different quotas (i.e., $A_1 = A_2 = 1, 2, \ldots, 50$); then, I use the cutoff thresholds $\boldsymbol{N}$ and the quotas $\boldsymbol{A}$ to estimate the attractiveness $\boldsymbol{\xi}$ by either $\mathbf{\Phi^{-1}}$ or FQR.

I find the FQR outperforms the DA model. The FQR works well even when the quotas are small (e.g., $A_1 = A_2 = 5$), while $\mathbf{\Phi^{-1}}$ does not perform well when $\boldsymbol{A}$ are small. The model works better when colleges admit more students, which coincides with Theorem 4. If I suppose that there are three, four, or five colleges instead of two colleges, the results are similar to Figure 7(a), as reported in Figure 7(b) – 7(d). Therefore, even if the identification of $\boldsymbol{\xi}$ is not guaranteed under small admission quotas, I can still obtain good estimation results based on the identification under large admission quotas and the FQR.

B.3. Counterfactural Analysis

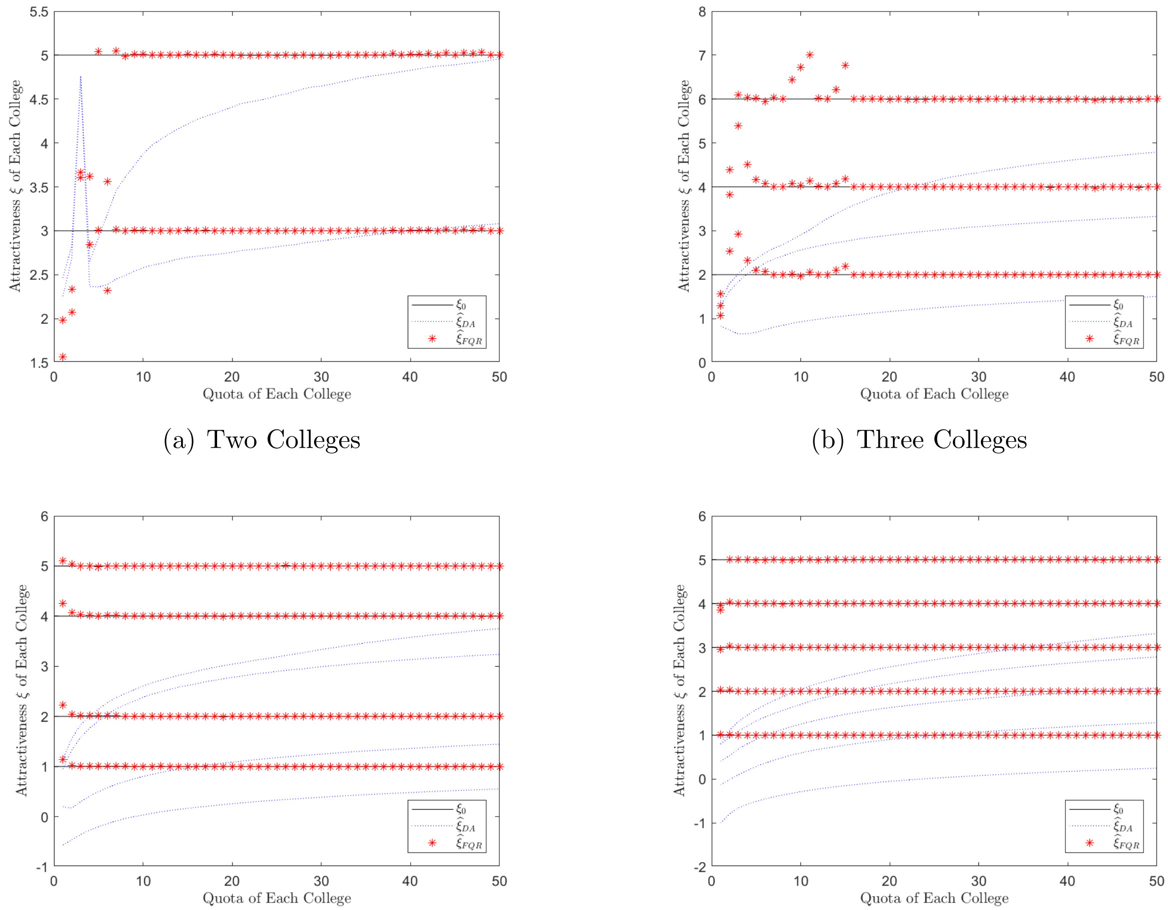
I compare the utility of students under the BM and DA. The expected utility of each student under the BM is calculated from Equation 1. Under the DA, each student will be assigned to the best available college. Therefore, in each simulation, I generate the preferences of students from the estimated mean attractiveness $\boldsymbol{\xi}$ and the distribution of private preferences $\mathcal{E}$. Then, I assign the students to their most preferred available college. I simulate this process 20,000 times.

Appendix C. Boston Mechanism, Deferred Acceptance Mechanism, Serial Dictatorship Mechanism and Their Variants in College Admissions in China

C.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[17]. 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.

[17] 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.



(a) Two Colleges

(b) Three Colleges

(c) Four Colleges

(d) Five Colleges

Fig. 7. DA Model ($\mathbf{\Phi^{-1}}$) vs. FQR for Different Quotas in A Two-, Three-, Four-, or Five-College World

Step k College l has quota A_l^k. The student placement office sends each college a list containing all students who were rejected in Step $k-1$ and chose 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.

C.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 were rejected in Step $k-1$ and chose 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 confirmed admitted.

C.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 $A_{l_1^1}^2 = A_{l_1^1} - 1$.

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 students (that is $A_{l_k^1}^k = 0$), the student is 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 has sent the information of all the students to the colleges. The deferred acceptance mechanism (DA) is equivalent to the serial dictatorship mechanism (SD) in the Chinese setting because in both mechanisms, the students are assigned to the best available college based on their preference lists. I will use the DA and SD interchangeably in this paper, but it is the SD that is actually implemented in college admissions in China.

C.4. Variants in College Admissions in China

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

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

The mechanism also depends on the province and can be either the Boston mechanism (BM), DA, or a mixture of these two. For example, some provinces apply the 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 the BM. I consider a mechanism as the BM if the mechanism applies the BM in its first step.

Appendix D. Proofs

D.1. Proof of lemma 1 on page 10

Lemma 1. $\mathbb{P}^a_l(i, A_l; A_{-l})$[18] $= \mathbb{P}^a_l(i-1, A_l; A_{-l})(1-\mathbb{P}^c_l(i-1, \boldsymbol{A}))+\mathbb{P}^a_l(i-1, A_l-1; A_{-l})\mathbb{P}^c_l(i-1, \boldsymbol{A})$ *for* $i \geq 2$ *and* $A_l \geq 1$*, where* $\mathbb{P}^c_l(i, \boldsymbol{A})$ *is the chance of student* i *choosing college* l*. In addition,* $\mathbb{P}^a_l(1, A_l; A_{-l}) = 1$ *for* $A_l \geq 1$ *and* $\mathbb{P}^a_l(i, 0) = 0$ *for* $i \geq 1$*.*

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

$$\mathbb{P}^a_l(i, A_l; A_{-l}) = \sum_{k=0}^{A_l-1} \mathbb{P}^o_l(i, k, A_{-l}), A_l \geq 1 \tag{10}$$

If college l admits student i, the quota must not be filled by the first $i-1$ students. This case can be broken down into cases where the first $i-1$ students take zero slots, one slot, ..., $A_l - 1$ slots, which results in Equation 10.

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

$$\mathbb{P}^o_l(i, k, A_{-l}) = \begin{cases} \mathbb{P}^o_l(i-1, k, A_{-l})(1-\mathbb{P}^c_l(i-1, \boldsymbol{A})); k = 0, i \geq 2 \\ \mathbb{P}^o_l(i-1, k, A_{-l})(1-\mathbb{P}^c_l(i-1, \boldsymbol{A})) + \mathbb{P}^o_l(i-1, k-1, A_{-l})\mathbb{P}^c_l(i-1, \boldsymbol{A}); k > 0, i \geq 2 \end{cases} \tag{11}$$

If none of the first $i-2$ students chooses college l and student $i-1$ does not choose college l, none of the first $i-1$ students choose this college. If (1) k of the first $i-2$ students choose college l and student $i-1$ does not choose college l *or* (2) $k-1$ of the first $i-2$ students choose college l and student $i-1$ chooses college l, k of the first $i-1$ students choose this college. These results produce Equation 11.

I start from the initial conditions. If college l does not admit any students, the chance of being admitted is zero. I have

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

In addition, if college l has a positive quota, the college cannot reject the top-ranked student. I have

$$\mathbb{P}^a_l(1, A_l; A_{-l}) = 1, \forall A_l \geq 1$$

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

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

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

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

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

□

D.2. Proof of theorem 1 on page 11

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

Proof. $\mathbb{1}_l(i)$ is an indicator function that denotes whether student i chooses college l as her first choice. If $\mathbb{1}_l(i) = 1$, she chooses 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 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), I have

$$\frac{1}{N_l} \sum_{i=1}^{N_l} \omega_l(i) \xrightarrow{a.s.} 0 \tag{12}$$

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

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

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 12 holds.

The cutoff line is N_l. Thus, A_l of the top N_l students choose college l. $\sum_{i=1}^{N_l} \mathbb{1}_l(i) = A_l$. I have

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

□