Matching Mechanisms, Matching Quality, and Strategic Behavior: Evidence from China

Mingming Ma **Binzhen Wu** **Xiaohan Zhong**

School of Economics and Management

Tsinghua University

Abstract

The school-choice literature has predicted that the Boston (BOS) mechanism is not strategy-proof and can deliver unfair matching outcomes. The serial dictatorship (SD) mechanism, a special case of the top trading cycles (TTC) mechanism, can help address these problems. However, there is little empirical evidence based on field data. This paper fills in the gap by exploiting a rich dataset from a Chinese high-school that provides information on students' reported preference and matching outcomes during the college admissions. We find that given the college entrance exam score, female students are matched with worse colleges than male students under the BOS mechanism. The main reason is that females choose worse colleges as their first-choices when reporting their preferences. These gender differences disappear when we switch from the BOS to the SD mechanism. However, compared with the Boston mechanism, the SD mechanism neither improves the average fairness of matching outcomes, nor reduces the average preference manipulation significantly.

Keywords: College Admission System; Matching Mechanism; Risk Averse; Strategic Behavior; Unfair Matching Outcome

JEL Classification: C78; D61; D78; I28

1 Introduction

China's college admission is one of the largest and most influential matching problems in the world. In 2009, around 10 million high school graduates competed for the slots available in colleges. The overall admission rate was around 60%, yet this rate fell to 40% for universities that offer bachelor degrees, and to 3% for top 70 universities (Educational Statistics Yearbook of China, 2010). This brings up the typical question for multi-to-one matching problems: if it is not possible for all students to go to their most desired colleges, how should space in different colleges be allocated?

The school choice literature usually focuses on three properties of a mechanism to see whether it is socially desirable: fairness (or stability), efficiency, and strategy-proofness. Theories predict that, some mechanisms, such as the Top Trading Cycles (TTC) or Serial Dictatorship (SD) as its special case is strategy-proof, fair and efficient. Another mechanism, the Boston (BOS) Mechanism, is inferior to those two since it is not strategy-proof (not truth-telling) and inefficient, and it is more likely to deliver unfair matching outcomes (Abdulkadiroğlu and Sonmez, 2003; Ergin and Sonmez, 2006; Kesten, 2006, Haeringer and Klijn, 2009). However, there is little empirical evidence based on field data.

We take advantage of the policy changes in China's college admission system to conduct empirical comparisons between the BOS mechanism and the SD mechanism. More specifically, China's college admission system is comprised of two stages. Stage one is a national college entrance exam called the CEE or *gaokao* in Mandarin. Stage two involves a matching mechanism that assigns students to colleges on the basis of students' CEE scores and their reported preferences. In order to improve the system's fairness and reducing manipulation in preference reporting, numerous policy changes have taken place. An important one is the shift from the BOS mechanism to the SD mechanism. In the BOS mechanism, the admission priority is based on reported preference first, then on the CEE score. Each college considers only the students who list the college as their first choice in the first round. Only when there are remaining quotas after the first round will a college consider admitting students who rank the college as their second choice. In contrast, in the SD mechanism, the priority is based first on the CEE score and then on reported preference. Students with higher scores always have priorities among the colleges they list in the preference lists no matter whether they list them as the first-choice.

By exploiting a rich dataset from a high school in Hebei province in China that provide detailed information on students' CEE scores, their preference lists, and matching outcomes, we are able to provide real-life evidence about the fairness of matching outcomes and strategic behavior under the BOS and SD mechanism. Moreover, the policy change in Hebei province, which introduced the SC mechanism to replace the BOS mechanism, makes it possible to compare the BOS mechanism with the SD mechanism and to identify the causal effect of the mechanisms on fairness and preference manipulation.

The results show substantial amount of mismatch under the BOS mechanism, which indicates the unfairness of the mechanism. Besides, preference manipulation seems to be prevalent under the BOS mechanism, because the majority of the students choose a college different from the best college they are eligible for as their first-choice in the preference lists. For the SD mechanism, it is somewhat surprising that we do not find significant improvement on the average fairness of matching. Mismatch is still present for the majority of the students under the SD mechanism. Moreover, given the demographics, there is also no significant difference in the preference manipulation between the BOS and the SD. These findings are consistent with the result of the experimental study by Calsamiglia et al (2007), who find no significant difference in preference manipulation between TTC and BOS mechanism when there are constraints on the number of college choices.

Although the SD mechanism on average is not significantly more desirable than the BOS mechanism, female students are better off while male students are worse off when switching from a BOS mechanism to a SD mechanism. Specifically, under the BOS mechanism, given the CEE score ranks, female students report worse first-choice colleges in the preference list and are admitted into worse colleges than male students. When it comes to the SD mechanism, both gender differences become insignificant. These findings are consistent with the literature on gender difference in behavior economics, where females are shown to be more risk averse and more averse to competition than males (see Croson and Gneezy, 2009 for a review). In our context, the first-choice of college in the preference list is crucial under the BOS mechanism as students may end up with poor-quality schools or even none schools if they fail to be admitted into the first-choice college. As a result, females who are more risk averse tend to be more conservative in preference reporting than males under the BOS mechanism. The high risk of a competitive first choice is reduced largely under the SD mechanism and thus the gender difference is no longer significant.

Our results imply that the policy change from the BOS mechanism to the SD mechanism does not fulfill its initial intention of improving fairness and reducing preference manipulation, although it is effective in reducing the unfairness in relation to gender difference. To our best knowledge, this is the first paper using real data to examine both strategic behavior and welfare properties of matching mechanisms and compare between two most studied matching mechanisms. Previous studies focus on theoretical analyses. Empirical studies are mostly based on lab experiments. This is also the first empirical study evaluating the matching mechanisms in China's college admission system, which is one of the largest matching problems in the world.

The paper is organized as following. Section 2 briefly introduces the institutional background and discusses the theoretical predictions by reviewing the school choice literature. Section 3 describes the data, measurement of true preference, and definition of mechanism regimes. Section 4 analyzes the matching quality of the two mechanisms. Section 5 examines the strategic behavior under the two mechanisms. Section 6 reports robustness checks and section 7 concludes.

2 Institutional Background and Theoretical Predictions

2.1 Institutional Background

The design of college admission mechanisms is an important and widespread concern for millions of families in China, because the college admission is a major determinant for student's future course of life. Admission mechanisms also matter for a country's long-term economic growth, because better matching can improve the allocation efficiency of human resources. The quality of college admission in China also matters for other countries, because of the increasing globalization of the labor and higher education markets.[1]

The college admission system in China is a typical one-sided, multi-to-one matching problem: 1) every school can admit more than one student, but every

[1] The number of Chinese students going abroad to pursue graduate studies has increased by 30% annually since 1999, reaching 0.28 million in 2010 (Educational Statistics Yearbook of China, 2010).

student can only be admitted by one school; 2) the priority order of schools on students is predetermined and known by students. In China, it is well known that schools' priority orders on students are almost entirely determined by the CEE scores. Students with higher test scores are always preferred; hence, all schools with the exceptions of specialist schools essentially have the same priority order on students.[2]

In addition, admission process for each province can be seen as an independent matching problem. The reason is that the enrollment quota of each college for each province is predetermined and publically known before students submit their preference. Moreover, CEE is conducted uniformly within each province. As a result, we can simply use data from one province, which is Hebei province in our case, to evaluate the matching mechanisms.

The admission procedure is comprised of two stages. In stage one, students take the exam CEE and submit their preference lists for colleges. The deadline for preference submission varies by province and time, and can be prior to the exam or post the release of CEE scores. In Hebei province, students submit the preference after they know their scores and ranks. In stage two, a matching mechanism is used to match students with colleges according to students' CEE scores and their reported preferences.

For preference reporting and matching, colleges are categorized into different batches on the basis of college types (Davey, Lian, and Higgins, 2007). Typically, there are 4 batches of colleges with decreasing reputation and quality.[3] Batch 1

[2] Applicants to some specialist programs are screened by additional criteria, such as some art departments (e.g., audition), military and police schools (political screening and physical exam), and some sports programs (tryout). There are policies that allow certain groups of students to enter a university with lower scores, including ethnic minority groups, sports-people, children of army personnel, and disabled applicants. It is also possible for students to be recommended to universities, but the number of students entering through this route is very small (Davey, Lian and Higgins, 2007).

[3] There is another batch, called batch 0. This batch includes mainly art departments, military and police schools and admits students ahead of all other batches. However, since the admission of batch 0 does not follow the regular procedure and may consider qualifications besides scores, such as audition,

comprises elite universities. Only students whose CEE scores are at the top of the score distribution in each province can be admitted into these universities. Batch 2 consists of the remaining public universities. Batch 3 contains mainly private universities that offer bachelor degrees and charge substantial tuition and fees. Batch 4 is composed of vocational colleges. In each batch, there can be sub-batches. For example, in Hebei province, batches 1 and 2 both had two sub-batches from 2006 and 2008. Specifically, the sub-batches were 1A and 1B in batch 1, and 2A and 2B in batch 2. 1A is better than 1B, and 2A is better than 2B.[4] In 2009, there were no such sub-batches for batches 1 and 2.

Students submit their preferences for colleges by batches, and colleges recruit students following the batch sequence.[5] Colleges in better batches have higher priority on admission. Specifically, colleges in batch 1 admit students first. After these colleges finish their admission, colleges in batch 2 are allowed to admit students. Colleges in batch 4 admit last. Within each batch, the sub-batches also admit students sequentially. For example, batch 1A (2A) admits students before batch 1B (2B). Therefore, we have a matching problem for each batch and the matching problems in different batches are sequentially correlated. Given that all students prefer colleges in lower batches, it is in the students' interest to try their best in each batch in a pure strategy equilibrium, as if this batch were the only batch they face.[6]

Each batch applies one matching mechanism to match applicants with colleges in that batch and the matching mechanism can vary across batches. Two matching

political screening, and physical exam, we exclude students who are admitted into batch 0.

[4] The design of the batches is introduced in http://www.gaokao.com/e/20130410/516516dfba99c.shtml.

[5] Depending on the province and the year, students may not submit preferences for all the colleges and all batches at the same time. For example, it is possible that in the first round, students submit preference only for the first-choice colleges in some batches. After the matching results of the first round have been nailed down, the second round starts and students submit preference for the second-choice colleges in these batches or for the first choice colleges in other batches.

[6] The case with mixed-strategies is more complicated because students may choose a strategy that gives a higher probability to choosing a second-batch college and a lower probability to choosing a first-batch college.

mechanisms have been used in China: the BOS and the SD. Details on the mechanisms are introduced in the section 2.2. Before 2003, all provinces applied the BOS mechanism. The SD mechanism rapidly gained popularity in China after its' first introduction in 2003. In 2011, 23 provinces have implemented the SD mechanism. In Hebei province, the SD mechanism has been gradually implemented to replace the BOS mechanism since 2007. As shown in the following table, the change started in some (not all) batches, which is not unusual in China. Particularly, the SD mechanism was first piloted to replace the BOS mechanism in batches higher than 2B (including 2B, 3, and 4) in 2007, then extended to batch 2A in 2008 and all batches in 2009. This means that the BOS mechanism was applied in the remaining batches in these years and all batches in 2006.

<table>
<tr><th>Batch</th><th>2006</th><th>2007</th><th>2008</th><th>2009</th></tr>
<tr><td>1A</td><td>BOS</td><td>BOS</td><td>BOS</td><td rowspan="2">SD</td></tr>
<tr><td>1B</td><td>BOS</td><td>BOS</td><td>BOS</td></tr>
<tr><td>2A</td><td>BOS</td><td>BOS</td><td>SD</td><td rowspan="2">SD</td></tr>
<tr><td>2B</td><td>BOS</td><td>SD</td><td>SD</td></tr>
<tr><td>3</td><td>BOS</td><td>SD</td><td>SD</td><td>SD</td></tr>
<tr><td>4</td><td>BOS</td><td>SD</td><td>SD</td><td>SD</td></tr>
</table>

The number of college choices that students can list can also vary across batches due to the difference in the match mechanism. In Hebei province, the BOS mechanism typically allows for one first choice and one second choice of college in each batch. The SD mechanism generally allows for five choices of colleges in each round.

Besides college choices, students are asked to list their major preferences for each college in the preference list. In Hebei province, there are 6 major choices and one option of "any major" for each college.[7] The admission procedure works as a two-stage game: students are first matched with colleges on the basis of scores and reported college preferences; then, conditioning on their qualification for one college, their major preferences are considered typically on the basis of the BOS mechanism. Thus, colleges can be regarded as composite commodities for students. Students need to incorporate both their major preferences and their admission possibilities over majors when they choose colleges. In another words, the two-stage game will deliver similar outcomes to a one-stage game that considers composite college preference. As a result, we simply focus on the composite college preference.

One distinct feature of the Chinese secondary education system is that high school students have to decide whether to focus on the humanities or the sciences for the rest of their education. Both humanities and science students take Chinese, English, and political science; humanities students take geography, history, and basic mathematics; and science students take physics, chemistry, biology, and advanced mathematics. The CEE also has two sets of exams respectively. Humanities and science students are admitted separately and face different admission quota. College batches or rankings also differ between humanities students and science students.

2.2 Matching Mechanisms, Theoretical Predictions, and Literature

The main difference between the BOS and SD mechanism lies in the importance that the mechanism attaches to students' reported preferences, particularly their first choice of college. Specifically, under the BOS procedure, in the first round of admission, all students' applications are sent to their first-choice colleges in their preference lists, and each college admits the students who ranked the college as their

[7] By choosing the option of "any major," students can get admitted into the preferred college but be assigned to an undesirable major if their CEE score is lower than the eligibility scores of the majors they list but higher than the eligibility scores of other majors in this college.

first choice in order of students' CEE scores. Students who are not admitted in this first round are then considered by their second-choice college in their list. Only when excess quotas remain after the first round will a college extend offers to students who listed it as their second choice. The procedure continues until all colleges have filled their quotas or all colleges listed by every student have been considered.

Under the SD procedure, students with the highest score are considered for admission in the first round, and admitted by their first choice college. In the second round, students with the second highest score are considered for admission. These students are matched with their first choice college as long as these colleges still have spaces available. If not, students' second-choice college is considered. The procedure continues until all students have been considered or all colleges have filled their quotas. The SD mechanism is a special case of the TTC mechanism. When all the colleges have the same preference order over students (the Kesten-acyclic condition), which is true in China, these two mechanisms are equivalent (Abdulkadiroglu and Sonmez, 2003). In addition, in this case, the SD mechanism functions equivalently to the Gale-Shapley (GS) or deferred acceptance (DA) mechanism.

We borrow from the literature for the welfare properties of the BOS and SD mechanism. The focus here is on the fairness (stability) and strategy proofness of the mechanisms. The formal definitions of these properties are as following:

Fairness/Stability: A matching mechanism is fair/stable if, for any student's preferences and score distributions, the mechanism always implements in its Nash equilibrium a matching outcome that is fair/stable. A fair/stable outcome means there are no blocking pairs, in which a student is not accepted by a preferred college, even though this college has vacant slots or recruits students with lower CEE scores.

Strategy-Proofness: A matching mechanism is strategy-proof if, for any student's preferences, truly reporting their preference for schools is a (weakly) dominant strategy for all students.

A fair matching mechanism protects students with high CEE scores in the sense that they will never be worse than students with lower scores. Strategy-proofness helps avoid the unfairness caused by rationality disparities, given that students may have different levels of rationality for strategy manipulation. In addition, strategy-proofness can save students' time on preference manipulation.

For the BOS mechanism, the theoretical literature predicts that it is not strategy-proof (Abdulkadiroğlu and Sonmez, 2003). Yet, in the Chinese context (where the acyclic condition is satisfied because all schools have the same priority order over students), it implements in its Nash Equilibrium (NE) the unique fair matching outcome (Ergin and Sonmez, 2006; Haeringer and Klijn, 2009). However, when the number of players is large, the non-strategy-proofness makes it unpractical to realize the NE and thus the fair outcome, because it is very difficult for every person to correctly predict all other players' strategies and follow the NE strategy.

The incentive to manipulate preference under the BOS mechanism is intuitive. Given the shortage of high-quality colleges, students with very high CEE scores can be assigned to poor-quality colleges or even fail to be admitted by any college if they miss their unique first-choice college.[8] Consequently, students may lie about their preferences by listing a moderately favored college as their first choice to secure a slot in this college. This easily leads to unfair outcomes, in which students with higher CEE scores are not match with better colleges. The difference in risk aversion across demographic groups may result in different preference manipulation and matching

[8] This risk of not being admitted by any college is not trivial among the students whose CEE score is above the eligible score of the admission. In 2006 in Hebei province, 22% of such students are not admitted by any college (http://edu.qq.com/a/20070505/005090.htm). The ratio increased to 33% in 2008 in spite of the college enrollment expansion (http://edu.qiaogu.com/info_314445/).

quality in the equilibrium.[9]

The theoretical literature predicts that when there is no constraint on the number of college choices, the SD mechanism is fair, and strategy-proof, and the matching outcome implemented in NE is unique (Balinski and Sonmez, 1999; Abdulkadiroglu and Sonmez, 2003; and Kesten, 2006). Under the SD mechanism, when it comes to a student's turn to be admitted, the colleges that a student list (in the first round of a batch) essentially are all first choices (for that batch). Whether a college in a student's list admits the student depends on only whether the student's CEE score is higher than the scores of other applicants, not on whether the student literally lists this college as his or her first choice.[10] Therefore, when students can list all colleges, they essentially have no incentive to misreport the preference order over colleges and thus students with higher scores always fare better.

When the number of college choices is limited, the fairness of SD is only attained under very strict conditions (Haeringer and Klijn, 2009). This constraint on choices also gives students incentive to manipulate their preferences to avoid intensively competitive colleges. In Hebei province, the batches that implement the SD mechanism typically allow for 5 choices of colleges in each round. This may explain why the officials in the provincial admission office have publically guided students to be strategic under the SD mechanism:[11]

"One important instruction for preference submission under the SD mechanism: Choose the colleges strategically. List 2 to 3 popular colleges to try luck and 2 to 3

---

[9] The preference manipulation is widely considered as the most important reason for the fluctuations in the number of applications for a given college, as students try to avoid the most-competitive colleges in the last year (http://www.80edu.com/html/gaozhao/gzzt/11217.html).

[10] The order of college choices still matters because when there are multiple colleges having spaces remaining for a student, the student is admitted by the first eligible college in the order of listing. There is another feature in China's system: the number of applications that a college holds can exceed its quotas, by no more than 20%. Yet in practice, colleges are often unwilling to hold more applications than their quota, because this can invoke disputes.

[11] Cite from the notice of "Instructions on filling in application forms in 2007," issued by the admission office of Hebei Province.

unpopular colleges as insurance.”

On the whole, the comparison of the BOS and the SD mechanism demands empirical studies using data from the real world. However, only a few empirical studies have been conducted, and most of these studies are based on experimental data which induce students’ true preference by small or modest monetary payoffs (Chen and Sonmez, 2006; Featherstone and Niederle, 2008; Klijn, Pais, and Corsatz, 2010; Pais and Pinter, 2008). The evidence obtained is mixed. For example, Chen and Sonmez (2006) find that the BOS mechanism elicits significantly more preference manipulation than other mechanisms and renders lower fairness. However, Calsamiglia et al (2007) replicates the experiment in Chen and Sonmez (2006) but imposes constraints on the number of colleges to choose. They find no significant difference in preference manipulation and fairness between the SD and BOS mechanism.

Among the few empirical studies that use real data, Abdulkadiroglu et al. (2006) focuses on the adoption of the BOS mechanism in the BOS area. The authors find that students who apply strategic manipulations fare better than students without preference manipulations. Abdulkadiroglu, Pathak, and Roth (2009) simulate and compare the matching outcomes of different mechanisms on the basis of the preferences submitted by students who are about to enter high schools in New York City. However, whether the submitted preference is a good proxy for the real preference remains unclear, especially when preference manipulation is prevalent.

3 Data and Measurements

Our data are from a good-quality high school in a city of Hebei Province in China. The city is a typical middle-income city, with a population of 283 million and

GDP per capita $3,700 in 2008. The dataset provides information on the students graduating from this high school between 2006 and 2008. It has information on each student's preference submitted and admission outcome. Moreover, it contains students' CEE scores, score ranks in the province, and the scores of their last 3 mock exams that help students get familiar with the real CEE. Many individual characteristics, with the exception of family income, are available, including student's gender, age, minority status, *Hukou* status (urban vs. rural), type of class (whether in elite classes or regular classes), and field of study (humanities vs. science). We also collect provincial-level information such as each college's admission quota for Hebei province and the entire distribution of CEE score in the province in each year.

One key difficulty in examining the fairness and strategic behavior of different mechanisms is that the true preference of students for colleges is unobservable. However, the batches or rankings of colleges are obviously one of the most important determinants for students' preferences, particularly in an education system like China where high schools devote almost all their efforts on increasing students' admission rate into good colleges. Therefore, we would like to use college batches or some commonly recognized ranking as a proxy for students' preference orders.

It is common knowledge that colleges in lower batches are better in quality (reputation) and more competitive in the labor market. Therefore, all students prefer to be admitted by colleges in lower batches. However, as described in the institutional background, since admissions in different batches are independent in the sense that they admit at different time (sequentially), there is no fundamental reason for students to lie about their preference on batches. Moreover, since students have significant different preference on batches, it is uncommon to sacrifice their preference on better batches to insure admissions to worse batches. As a whole, the batch classification is not good for testing strategy manipulation. Students incline to manipulate college choices within batch so that they can insure to be admitted by the batch that they are eligible for. Therefore, we need to exploit the ranking information within batches.[12]

Although there are no academic consensus on college ranks, it is widely recognized that the most popular and influential ranking for student is presented in a series of books written by Shulian Wu (Wu, 2006; 2007; 2008; 2009).[13] Wu's books report three kinds of rankings, which classify the colleges into 11 tiers, 125 tiers, and 501 tiers respectively. A college's ranking depends on the field, that is, a college's ranking of the science field can be different from its ranking of the humanities field. The ranking is based on colleges' reputation, academic strength, and students' career prospective. Although students have different preferences over a specific college, it is reasonable to argue that students have a somewhat homogeneous preference order on these college tiers, particularly when a small number of tiers, such as 11 tiers, are considered.

Table A.1 gives the distribution of colleges in each tier in 2007. The best tier is 12, the worst is 1, and 0 represents not admitted by any college. It shows that the top two tiers have only about 13 colleges, the tiers in the middle have more than 30 colleges, and the bottom tiers have more than 60 colleges. This table also shows the distribution of students admitted in each tier in our sample, which indicating that most tiers have more than 50 students.

For the analyses, we exclude students who are admitted by a college with no ranking information.[14] We also exclude students who are not admitted by any

---

[12] The batch classification changed in 2009 in Hebei, which makes longitudinal comparison difficult.

[13] http://www.gaokao.com/e/20091103/4b8bd4851e023.shtml; http://baike.baidu.com/view/163202.htm. We find that student's college ranks are highly correlated with the batches that students are admitted into. The correlation coefficient is 0.84 and significant at 0.1% level. This confirms the reliability of the college ranking to some extent.

[14] 410 students are excluded due to missing information on the college ranking. Most of them (93%)

school.[15] The reason is that most of these students are admitted by some school but choose not to matriculate because they are not satisfied with the matching outcome and may take the CEE again next year. In another word, their matching outcome may not be worse than others. After these exclusions, we have about 1000 observations left. These sample restrictions can induce sample selection bias. We also try the result including them to see how sensitive the results are.

Table 1 gives the summary statistics of the main variables. It shows that more than half of the students (55%) of the students are females. Most of the students (93%) have urban Hukou, which results from the fact that the school locates in the city. About 10% of the students are minorities. The average age is 18.8. Around 29% of the students are in the elite classes that usually has the best teachers in the school. The average standardized score in analogue CEE exams is right skewed after we exclude students who are not admitted or have no information on college ranking. The average score in three mock exams is considered as a better measure for the academic aptitude in high school than the one-shot CEE score. About 67% of the students take the science exam. Our sample does not include repeat exam takers. In the analysis, we use students' ranks of CEE scores among students taking the same type of exam instead of the raw CEE scores because it is the ranks that determine schools' priorities. Lower rank means better performance in the CEE. Students in our sample school on average perform quite well in the CEE because this school is one of the top high schools in the province.

The next three columns of Table 1 show the demographics of students who are affected by different mechanisms. To simplify the explication, we use the SD batches to represent batch 2B or higher batches in 2007 and batch 2 (including 2A and 2B) or

---

are admitted into batch 3.

[15] About 15% of our sample is not admitted by any college. This rate is 19% for those affected by the Boston mechanism, 15% for those affected by the SD mechanism.

higher in 2008, and the BOS batches to represent all the other batches in these two years and all batches in 2006. SD (BOS) batches use SD (BOS) mechanisms to admit students. The matching mechanism that affects a student depends on not only the batch that the student is actually admitted into, but also the batch that the student is likely to be admitted into in a fair match. By a fair match, we means the matching outcome is determined solely by the CEE score and students with higher CEE scores are always admitted by better batches. We define the batch in a fair match as the fair batch or the "fairly-likely-to-be" batch. More specifically, the fair or "fairly-likely-to-be" batch is calculated based on the student's actual CEE score and the fair eligibility score of each batch. The fair eligibility score of each batch is lowest score among all the scores of the students admitted by that batch in a fair match, which is calculated on the basis of the information about the score distribution of all students and the total admission quota of all the colleges in each batch.

Therefore, a reasonable definition of the batches that are likely to affect a student include both the actual batch that the student is admitted into and the fairly-likely-to-be batch. Notice that some students are likely to be affected by both the SD mechanism and the BOS mechanism because they are actually admitted into the Boston batches (SD batches) but they should be admitted into the SD batches (Boston batches) in a fair match. There is no obvious way of dealing with the overlaps. Noticing that there are much fewer students admitted into the SD batches in our sample, in our baseline model, we categorize all the students affected by both mechanisms as students affected by the SD mechanism to maximize the observation number under the SD mechanism. In the robust test, we consider other ways of dealing with the overlaps, such as excluding all these observations or categorize them as students affected by the BOS mechanism.

To summarize, we distinguish two groups of students, one is the students who are likely to be admitted into SD batches and thus affected by the SD mechanism, and the other is the remaining students who are likely to be affected by the BOS mechanism. Column 2 shows the demographics of students affected by the BOS mechanism are quite similar to those of the population. However, these students are better than the population average in terms of academic performance such as CEE scores and scores in mock exams. The proportion of students in the elite classes is also larger. This is not surprising because the BOS batches are on average better than SD batches (in 2007 and 2008).

Column 3 displays the demographics of students affected by the SD mechanism. Since the SD mechanism is only applied in batch 2 or higher batches, the number of observations is much smaller. Compared with students affected by the BOS mechanism (column 2), students affected by the SD mechanism have worse CEE scores and scores in mock exams, and are less likely to be in the elite classes. These students are also more likely to come from rural areas and take the science exam. To compare the mechanisms based on a more homogenous sample, we exclude students who are affected by the BOS mechanism and are likely to be admitted by batch 1 (column 4). We notice that the difference in academic performance between the BOS mechanism and SD mechanism gets much smaller, although still exists. Other demographics also differ. This motivates us to use difference-in-difference regression when comparing the BOS and the SD mechanism.

4 Matching Quality

4.1 Descriptive Analysis

To empirically measure the fairness of a matching outcome for a given CEE score distribution of all students, we first characterize the unique fair matching outcome for the given score distribution and given popularity of colleges. Particularly, we calculate the college tier that each student can be admitted into when the matching outcome is fair, and this tier is labeled as the fair (college) tier. Then we compare the actual tier of college that a student is admitted into (labeled as actual tier) with the fair tier. If any student is admitted by a tier that is not the same as the fair tier, the matching outcome is not fair. Otherwise, the matching outcome is fair.

Given that students' preference is represented by the ranks of colleges, a fair matching result must be that students with higher CEE scores are assigned to colleges with better ranking. As a result, the fair college tier depends on the entire CEE score distribution in the province and the admission quota of each college tier. More specifically, if the top-tier (tier 12) colleges have an admission quota of n and the second-best tier colleges have a quota of m, then the fair tier for students whose CEE score ranks among the top n in the province is tier 12 and the fair tier for students whose CEE score ranks between n+1 and n+m is tier 11. Similarly, we can derive the fair tier for other students. Since we know the quota of each college (and hence each tier) and each student's rank of CEE score in the province, we can calculate the fair tier for every student.

We define mismatch as the difference between the actual college tier admitting a student and his/her fair college tier. When the value of mismatch is positive (negative), it means the student is up-matched (down-matched) with a college that is better (worse) than the college in the fair matching. The value of mismatch represents the extent of the mismatching. For the entire population of applicants for college, the number of up-matched should be the same as the number of down-matched. Our sample is, however, not a representative sample of the population, since the high school is one of the best high schools in the city. So the mismatching may not be summarized to zero.

Figure 1 shows the distribution of mismatch for students who are mostly affected by the BOS mechanism. It indicates that more than 68% of the students in our sample are mismatched in the sense that they are not admitted by the fair college tier. About 37% of the students are up-matched and 31% of the students are down-matched, which means that the majority of our mismatched students are better-off under the BOS mechanism.

Table 1 shows that the proportion of mismatching declines under the SD mechanism, but the difference is only 2 percentage points and not significant at 10% level (column 1). Compared with the BOS mechanism, the SD mechanism is less likely to up-match students (44% vs. 48%), and more likely to down-match them (32% vs. 30%). However, these differences are not significant (at 10% level) either.

Table 1 also shows the variations in mismatch across different groups. We first notice that under the BOS mechanism, female students are much more likely to be mismatched than male students. This mainly results from the fact that they are much more likely to be down-matched than male students under this mechanism. The probability of being up-matched is not significantly different between genders. Consistently, the average value of mismatch or the average distance between the actual tier and the fair tier is significantly lower among female students.

In contrast, when it comes to the SD mechanism (Part B of Table 1), the difference between genders is mostly gone. Moreover, the disparity in the gender difference between the BOS mechanism and SD mechanism is not due to the difference in the batch composition of the students. When we restrict the sample for the BOS mechanism to the students who are mostly likely to be admitted in the second or higher batch, females are still significantly more likely to be down-matched than males (Part C of Table 1). One concern here is that females may differ from males in other attributes. We resort to the multivariate regressions in the next section to further examine this issue.

The extent of mismatch does not vary significantly with rural status, ethnic status, age group, and exam type, with the exception that rural students in the second or higher batch tend to be more likely to be mismatch in both mechanisms and minorities in the second or higher batch are more likely to be up-matched under the BOS mechanisms. We also notice that students who have top half scores in the mock exams are less likely to be up-matched and more likely to be down-matched than their counterparts. We find very similar pattern if we compare students in the elite classes with students in the regular class, or compare students with top-half CEE scores with students with bottom half CEE scores. However, these differences need to be explained with caution. By design, it is impossible for the students with the top fair tier to be up-matched and students with bottom fair tier to be down-matched.

4.2 Regression Analysis

4.2.1 Heterogeneous Matching Quality

Since individual characteristics are often correlated, we rely on multivariate regressions to examine how the matching quality depends on students' characteristics. In a fair matching, a student's matching outcome should depend solely on his/her CEE score rank. In another word, for a given CEE score rank, if the matching outcome varies with a student's demographics, it implies the matching is less likely to be fair. Therefore, we design the econometric model as following:

$$Y_{it} = \beta_0 + \beta_1 * X_{it} + \beta_2 * Year_t + \beta_3 * Fair\text{-}tier_{it} + e_{it} \quad (1)$$

Here, Y_{it} is the tier of college that student *i* is admitted into in year *t*. X_{it} is the vector of all individual characteristics, including gender, *Hukou* status, minority status, age, average scores in the three mock exams, the type of the CEE that a student takes (humanities vs. science), and the class dummies. We also control for CEE score rank and its square and cubic terms to allow the score ranks to have nonlinear effect on the admission outcome. *Fair-tier* represents 11 dummies of college tiers that students are eligible for in the fair matching. In a fair matching outcome, students are admitted into colleges in the fair tier. Therefore, after we control for fair-tier dummies, there should be no other variables affecting the actual college tier. If students' demographics affect the actual college tier, the matching outcome is less likely to be fair.

Table 3 reports the results. Since the actual values taken on by the college tier are irrelevant and larger values are assumed to correspond to better outcomes, we apply the Ordered Logit model. The first column considers the entire sample. Here we see that given CEE scores and fair college tiers, female students on average fare worse than male students. Column 2 restricts the sample to students who are most likely to be admitted under the BOS mechanism. The gender difference in preference reporting gets larger and more significant.

Column 3 considers the students who are most likely to be affected by the SD mechanism. Now the gender difference disappears. To further examine the difference between the two matching mechanisms, we pool the sample together, control for the dummy of SD mechanism, and add the interaction term between demographics and SD mechanism (Column 4). The result confirms that these two mechanisms differ significantly in terms of gender gap in college tiers that students are admitted into.

In column 5, we restrict the sample to students who are most likely to be admitted by batch 2 or higher batches and affected by the BOS mechanism, so that the BOS mechanism are more comparable to the SD mechanism in terms of the batch composition and the sample size. Column 6 examines the difference between the BOS mechanism and SD mechanism among these students and confirms the significant difference in the gender gap between these two mechanisms.

For other demographics, column 1 indicates that under the BOS mechanism, urban students and students with top half scores in the mock exams are matched with better colleges than their counterparts (rural students and students with bottom half scores in the mock exams) respectively. The difference disappears when we restrict the sample of the BOS mechanism to the non-first-batch students, which indicates that the effect of BOS mechanism on rural-urban difference or academic-aptitude difference in column 1 is mainly present among students in batch 1.

In addition, column 3 indicates that although gender difference disappears under the SD mechanism, rural-urban difference remains. Moreover, differences between age group and between ethnic groups emerge. These significant variations across demographic groups are somewhat surprising as the SD mechanism is famous for its emphasis on the CEE scores and intention of fair matching. The significant difference between demographic groups for a given CEE score indicates that the SD mechanism is not fair either. It can result from the constraints on the number of college choices in the preference reporting. On the other hand, comparing column 3 and 5 shows that given the CEE score, students most likely admitted under the SD mechanism on average are matched with better college tiers than students most likely admitted under the BOS mechanism.

Another way to check the demographic difference in the matching quality is to directly examine how the mismatch varies with individual characteristics. Table 4 shows that the results are quite similar to the descriptive statistics in Table 2. In particular, females are 9 percentage points more likely to be mismatched than males under the BOS mechanism (column 1). This is mainly because females are 13 percentage points more likely to be down-matched than males under the BOS mechanism (column 2), as the probability of up-matching does not differ between genders (column 3). Accordingly, females on average have significantly lower value of mismatch than males (column 4). These disadvantages of female students are not significant any more in the SD mechanism (columns 5-8). These results still hold when we restrict the sample to the non-first batch students for the BOS mechanism (not reported due to space limit).

4.2.2 Average Matching Quality: Difference-in-Difference Estimation

Although females fare better in the SD mechanism, it is not clear how the SD mechanism has improved the average fairness, measured by the amount of mismatch. To examine this issue, we directly measure the difference in the mismatch between the BOS mechanism and SD mechanism. Since in our data, these two mechanisms are applied to different batches, we apply the difference-in-difference framework to exclude the time-invariant differences between batches. The model is as following:

$$Y_{it} = \beta_0 + \beta_1 * SD\text{-}batch_i + \beta_2 * Post_t + \beta_3 * SD\text{-}batch_i * Post_t + \beta_4 * X_{it} + e_{it} \quad (2)$$

Here, Y_{it} is a measure of matching quality, such as the proportion of mismatched, up-matched, or down-matched students. *SD-batch* represents the batches that adopt the SD mechanism in the post-treatment periods. It is batch 2A or higher batches in 2008 and 2B or higher batches in 2007. *Post* represents a year dummy, which is 1 for post-treatment years and 0 for other years. $SD\text{-}batch_i * Post_t$ is the interaction term between SD batches and post-treatment year dummy. X_{it} is the individual characteristics, including individual demographics (gender, *Hukou* status, minority status, age, average scores in the three mock exams), the type of the CEE exam, and the class dummies, CEE score ranks and it square and cubic terms, and 11 dummies of fair-tiers.

We first apply the model to year 2006 and 2007, where 2006 is the pre-treatment year and 2007 is the post-treatment year. Batch 2B is the SD-batch or the treatment batch, and other batches are the control batches. Notice that the control batches can also be affected by the policy changes because of the spill-over effect, particularly for the students on the margin. However, the influence is much weaker. By excluding the marginal students, we will gauge how important the spill-over effect is. To compare the SD mechanism and BOS mechanism based on a somewhat homogenous sample, we restrict the sample to students who are most likely to be admitted by batch 2 or higher batches. As a result, batch 2A is the control batch.

Table 5 reports the results. The first column does not control for students' demographics. It shows that on average, the SD mechanism is less likely to mismatch students than the BOS mechanism; however, the difference is not significant at the 10% level. In the second column on, we control for students' demographics. The result hardly changes. Columns 3 to 5 indicate that students in our sample school are less likely to be up-matched and more likely to be down-matched and have lower average value of mismatch under the SD mechanism than under the BOS mechanism, but the differences are again not significant. These results indicate that the SD mechanism does not improve the fairness significantly, at least for the students in our sample school.

We can also apply the model to year 2007 and 2008, where 2008 is the post-treatment year, 2007 is pre-treatment year. In this case, batch 2A is the treatment batch, and batch 2B or higher batches is the control batch. The results shown in columns 6 to 9 are similar to the results using year 2006 and 2007, except that the differences in the probability of up-match and the value of mismatch become significant at 10% level.

5 Strategic Behavior in Preference Reporting

This section examines students' strategic behavior in preference reporting to explain the findings on matching quality. Since 84% of students are admitted by their first-choice college in our sample, we focus on the students' first-choice colleges in the batches that they are likely to be admitted into. [16]

5.1 Descriptive Analysis

To empirically measure students' strategic behavior for a given score distribution, we first need to characterize students' true preference on the colleges that they are eligible for. By assuming that students have a homogeneous preference order on the college tiers (the higher tier, the better), we only need to determine the best college tier that a student can be admitted into. A reasonable benchmark for calculating this "best tier" is to assume that students with higher CEE scores are assigned to colleges with better ranking.

Then we compare the fair tier with the tier of the first-choice college in the batch that a student is most likely to be admitted into. If a student chooses a college tier different from the best tier, the student is more likely to be strategic. Similar to the analyses for matching quality, we define *misreport* as the difference between the first-choice college tier and the best college tier that the student can be admitted into. When the value of *misreport* is positive (negative), it means the student up-report (down-report) his or her preference. Notice that

[16] This rate being admitted by the first-choice college is 95% for students affected by the Boston mechanism, but 39% for students affected by the SD mechanism.

although we use the word of “strategic”, the difference between the first-choice tier and the best tier can result from students’ bias in estimating the best tier. Positive (negative) difference can be due to the upward-bias (downward bias) of students’ estimation.

Figure 2 gives the distribution of misreport for students who are likely to be affected by the BOS mechanism. It shows the majority of the students misreport their preference. 38% of the students choose a college with a higher tier than their best tier in the fair matching and 31% of the students choose a lower tier than their best tier. The average value of misreport is around 0.

For the heterogeneity of preference reporting, we first notice that under the BOS mechanism, female students are much more likely to misreport than male students. This mainly result from the fact that they are much more likely to down report their first-choice than male students under this mechanism. The probability of up-reporting is not significantly different between genders. Consistently, the average value of misreport is significantly lower among female students. These are all consistent with the pattern of matching outcomes.

Also consistent with the results of the matching outcome, the gender difference in preference reporting disappears under the SD mechanism (Part B of Table 6). Also, the disparity in the gender difference in preference reporting between the BOS mechanism and SD mechanism does not seem to result from the difference in the batch composition of the students (Panel C of Table 6).

The extent of misreport also varies with rural status, ethnic status, age group, and exam type, but the pattern is more mixed. Different from the results for matching outcomes, students who are in the elite classes, have top half academic aptitude, or have top half CEE scores are more likely to up-report and less likely to be down-report than their counterparts. These indicate that they may be more confident than their counterparts.

5.2 Regression Analysis

5.2.1 Heterogeneous Strategic Behavior

For the regression analysis on how the strategic behavior changes with students' characteristics, we first apply the regression model represented in equation (1). The dependent variable Y_{it} now is the tier of the first-choice college of the batch that student i is most likely to be admitted into. The covariates are the same as those in the regression on matching outcomes.

Table 7 reports the results based on the Ordered Logit model. The first column considers the entire sample and indicates that given CEE scores and best college tiers, female students on average choose a college with lower tier as their first-choice than male students. Column 2 restricts the sample to students who are most likely to be admitted under the BOS mechanism. The gender difference in the matching quality gets larger and more significant. Column 3 considers the students who are most likely to be affected by the SD mechanism. Now the gender difference disappears. Column 4 confirms that the gender disparity in preference reporting is significantly different between the BOS mechanism and the SD mechanism. These results still hold when we restrict the sample to students who are most likely to be admitted by batch 2 (column 5 and 6).

The gender difference in preference reporting and its variation between mechanisms are completely consistent with the results on matching quality. It means females are admitted to worse colleges than males because they choose worse colleges as their first-choice. Why females choose worse colleges is beyond this paper.

However, as indicated by the literature on behavior economics, female students are more risk averse and more averse to competition (see Croson and Gneezy, 2009 for a review). As a result, when the first-choice is crucial and a highly competitive college is risky, females are more likely to give up the highly competitive school and secure a slot in a school which they have better chance to be admitted into. In another word, females are much more likely to down report their preference and thus to be down matched with colleges under the BOS mechanism. When it comes to the SD mechanism, the first-choice is not that crucial and the matching is based on score first. Therefore, females can try good schools without losing the slot in schools that is less competitive. As a result, the difference between males and females is largely gone.

We also observe variations in preference reporting in other dimensions, for example, students with higher academic aptitude tend to choose better colleges under the SD mechanism (significant at 10% level). However, we have not had a good explanation for it.

Table 8 directly examines how the misreport varies with individual characteristics. The first four columns show the results for the BOS mechanism. Consistent with Table 7, we find that under the BOS mechanism, females are 4 percentage points more likely to downward report their first-choice college than males (column 1). Slightly different from the descriptive statistics, here females are less likely to up report their first-choice college (column 2). As a result, the difference in the possibility of misreport is not significant between males and females (column 3). The next four columns show the results for the SD mechanism. The gender differences are all insignificant. Moreover, the change in the gender difference is not due to the difference in batch composition between these two mechanisms, as the change remains if we restrict the sample to batch 2 for the BOS mechanism.

5.2.2 Average Preference Manipulation: Difference-in-Difference Estimation

We apply the regression model represented in equation 2 to examine the average difference in preference reporting between the BOS mechanism and SD mechanism. The results are in Table 9. Similar to the analyses on matching outcome, the first five columns apply the difference-in-difference model to year 2006 and 2007 and restrict the sample to students who are most likely to be admitted by batch 2.

The theory predicts that students will be truth-telling under the SD mechanism but strategic under the BOS mechanism. However, we do not find empirical evidence for this theoretical prediction. On average, there is no systematic difference in preference reporting between the SD mechanism and BOS mechanism. The results based on year 2007 and 2008 also confirm this finding (column 5 to 9). This finding is consistent with the results for matching outcome, where we do not find systematic difference either. This might result from the fact that in the practice, there is limited number of college choices when reporting the preference.

6 Robustness Check

To examine whether our results are sensitive to the definition of college tier, we try two other rankings of colleges. One has 127 categories, and the other has 502 categories. The results for matching outcome and preference reporting are both shown in Table 10. The results are quite similar to those in the baseline model.

Humanities students and science students face different admission quota, are admitted separately, and have different preference on colleges. Controlling for the dummy of exam type may not be enough to control for all the differences. As a result, we run the regressions for humanities and science students separately. The results are shown in Table 11, which tells that the conclusions still hold.

The regression method imposes assumption on the function form. To relax the assumption, we apply the 5 nearest-neighbor matching method. More specifically, to estimate the gender difference in matching quality, we try to find 5 matches or male students who are most similar to a female student, matched on *year* (exact match), *score rank*(exact match), *rural*, *minority*, *elite class*, *age* and *scores in the analogue CEE*. Table 12 reports the result. Again, we find similar gender gap in the tier of college that admits students and the first-choice college under the BOS mechanism. The differences are no longer significant under the SD mechanism.

7 Conclusion

We employ the data from a high school in Hebei province of China to empirically examine the fairness and strategic behavior of two matching mechanisms: BOS mechanism and SD mechanism. Theoretical studies and experimental studies have shown that the BOS mechanism is not strategy-proof and can deliver unfair matching outcomes. In contrast, the SD mechanism, which corresponds to the TTC or SD mechanism, is strategy-proof and delivers fair outcomes.

Our findings confirm that the BOS mechanism is unfair as a lot students are matched with colleges other than the colleges in the fair matching. Also, the reported first-choice college is different from the best feasible college for the majority of the students, which is evidence of preference manipulation. However, for the SD mechanism, we do not find significant improvement in the average fairness of matching outcome and reduction in preference manipulation compared with the BOS mechanism.

Moreover, we identify significant gender differences in matching outcomes and preference reporting under the BOS mechanism. Specifically, females are matched with worse colleges than males, because they list worse colleges as their first-choice when reporting preference. These gender differences are no longer significant under the SD mechanism. This implies that given CEE scores, female students fare better under the SD mechanisms. These findings can be explained by the argument that female students are more risk averse than male students.

The reforms on matching mechanisms in China lean toward mechanisms that try to reduce preference manipulation and emphasize CEE scores more. Our findings cast doubt on the effectiveness of the policy changes that shift from the BOS mechanism to TTC or SD mechanisms. The findings about the fairness also echo the challenges raised by many students. Hou et al. (2009) conduct a survey of college students in Shanghai and finds that introducing SD options is favored by only slightly more than half of the students. Many students think the policy protects students with high CEE scores more than students with high academic aptitude. Other criticisms on the reform include scores becoming the only standard for admission and the tendency of the reform to drive students toward placing excessive importance on university ranking, but little on their own professional interests (Yang, 2009).

There are several caveats to interpreting our results. First, our data come from a top high school in the city. The results may not be directly generalized to other types of schools. Second, the term of preference manipulation or strategic behavior needs to be explained with caution. We assume students all know the best college that their CEE scores are eligible for. As a result, the difference between the first-choice and the best college is defined as preference manipulation. Yet, some students may not have complete or precise information on the best college they can be admitted into. Also, students may not be perfectly rational in real life. All these issues warrant further research.

References

[1] Abdulkadiroglu, Atila, and Tayfun Sonmez, School Choice: A Mechanism Design Approach, *American Economic Review*, 93:3 (Jun., 2003), pp. 729-747.

[2] Abdulkadiroglu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sonmez, Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access, mimeo, May 2006.

[3] Abdulkadiroglu, Atila, Parag A. Pathak, and Alvin E. Roth, Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match, *American Economic Review*, 99:5 (Dec., 2009), pp. 1954–1978.

[4] Balinski, Michel, and Tayfun Sonmez, A Tale of Two Mechanisms: Student Placement, *Journal of Economic Theory*, 1999(84), pp. 73-94.

[5] Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn, Constrained School Choice: An Experimental Study, mimeo, June 2009.

[6] Chen, Yan, and Tayfun Sonmez, School Choice: An Experimental Study, *Journal of Economic Theory*, 127(2006), pp. 202-231.

[7] Davey, Gareth, Chuan De Lian, and Louise Higgins, The University Entrance Examination system in China, Journal of Further and Higher Education, Nov. 2007 (31:4), pp. 385-396.

[8] Ehlers, Lars, and Jordi Masso, Matching Markets under (In)complete Information, mimeo, Jan., 2007.

[9] Ergin, Haluk I, Efficient Resource Allocation on the Basis of Priorities, *Econometrica*, 70:6 (Nov., 2002), pp. 2489-2497

[10] Ergin, Haluk, and Tayfun Sonmez, Games of School Choice under the Boston Mechanism, *Journal of Public Economics*, 90(2006), pp. 215-237.

[11] Featherstone, Clayton, and Muriel Niederle, Ex Ante Efficiency in School

Choice Mechanisms: An Experimental Investigation, mimeo, Dec. 2008.

[12] Haeringer, Guillaume, and Flip Klijn, Constrained School Choice, *Journal of Economic Theory*, 144(2009), pp. 1921-1947.

[13] Hou, Dingkai, Meijiao Zhang, and Xiaona Li. 2009. "The Effect of the SD Option Mechanism in Shanghai." *Research in Education Development (in Chinese)*, 7: 27–31.

[14] Kesten, Onur, On Two Competing Mechanisms for Priority-based Allocation Problems, *Journal of Economic Theory*, 127(2006), pp. 155-171.

[15] Klijn, Flip, Joana Pais, and Marc Vorsatz, Preference Intensities and Risk Aversion in School Choice: A Laboratory Experiment, mimeo, Apr. 2010.

[16] Li, Hongbin and Binzhen Wu, Education Inequality in China: Evidence from College Application and Admission, working paper, 2011.

[17] Pais, Joana, and Agnes Pinter, School Choice and Information: An Experimental Study on Matching Mechanisms, *Games and Economic Behavior*, 2008 (64), pp. 303-328.

[18] Pathak, Parag A., and Tayfun Sonmez, Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism, mimeo, May 2008.

[19] Wu, Shulian, "Guide for Preference Submission," various years (2006, 20007, 2008), Beijing, Chinese statistical publisher.

[20] Yang, Deguang. 2009. "Comments on the Fact that Famous Schools Admit All High Score Students and Analyses of SD Options." *Peking University Education Review (in Chinese)*, 7: 179–182.

Figure 1. Distribution of Mismatch: the Boston Mechanism

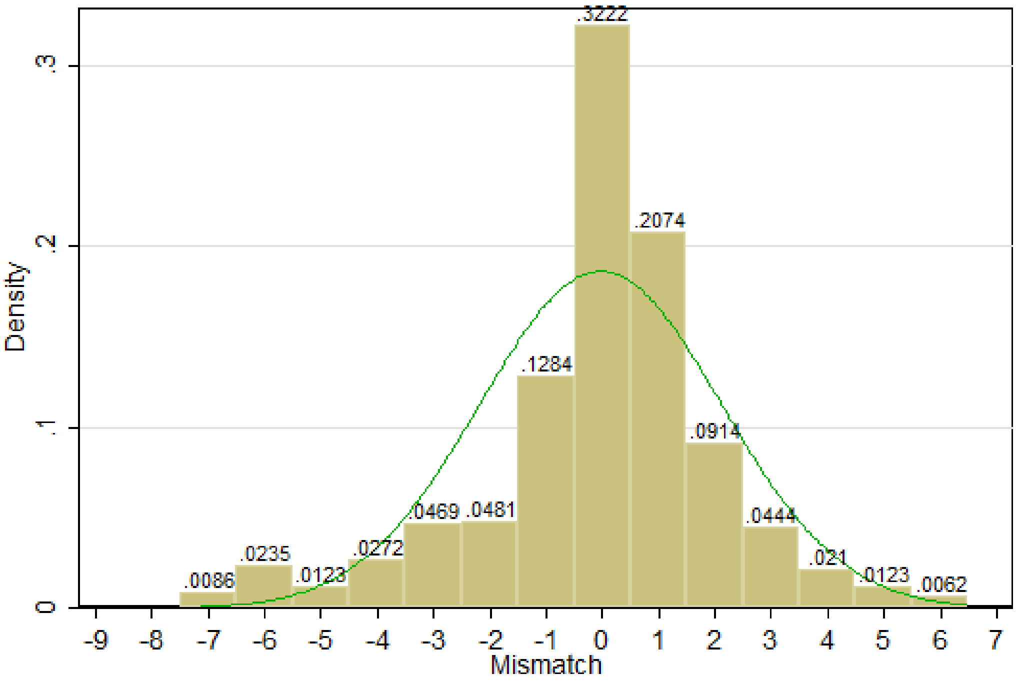

Figure 2. Distribution of Misreport: the Boston Mechanism

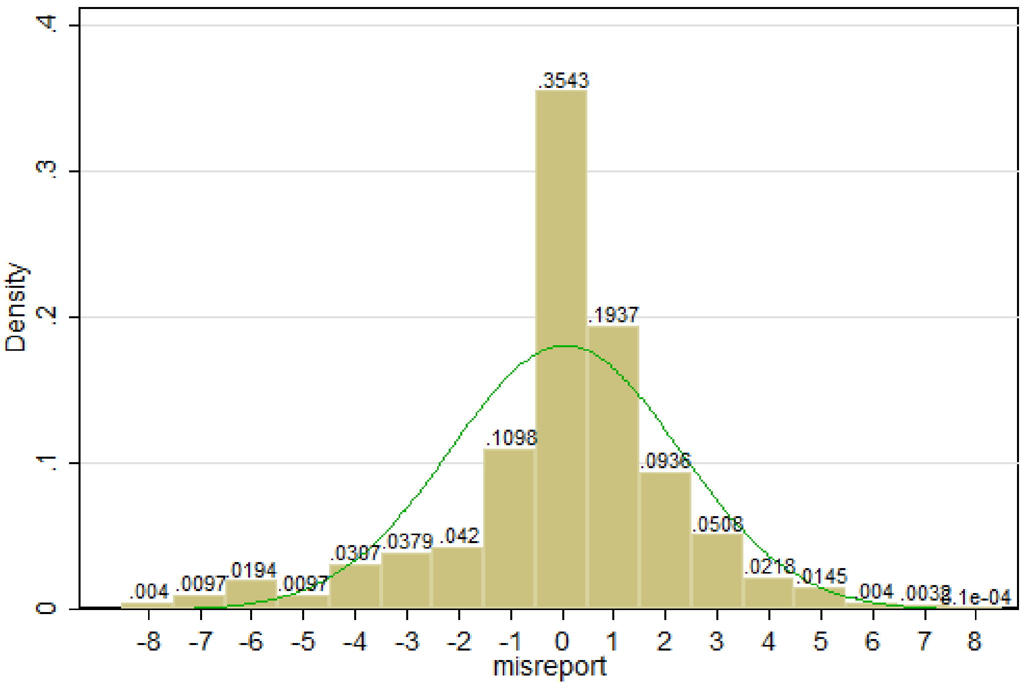

Table 1. Descriptive Statistics

Variable	All observations	All batches	Students likely to be admitted by batch 2	
		BOS mechanism	SD mechanism	BOS mechanism
	(1)	(2)	(3)	(4)
Female	0.55 (0.50)	0.55	0.56	0.62
Rural	0.07 (0.26)	0.07	0.10	0.03
Minority	0.10 (0.30)	0.10	0.11	0.07
Age	18.76 (0.56)	18.73	18.84	18.82
Elite class	0.29 (0.45)	0.33	0.10	0.07
Scores in 3 mock exams	0.45 (0.61)	0.53	0.14	0.29
Science	0.67 (0.47)	0.66	0.73	0.66
CEE score ranks	21.15 (32.44)	18.44	31.63	25.31
Observation	1031	819	212	191

Note: standard deviation in the parenthesis in the first column.

Table 2. Descriptive Statistics for Matching Quality: Mismatch

		Total	Difference between groups					
			Female -Male	Rural -Urban	Minority -Han	Old -Young	Science - Arts	Scores in analogue CEE: top half - bottom half
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
BOS mechanism, all batches	Mismatched	0.68	0.09***	-0.01	-0.03	0.02	-0.04	-0.03
	Up-match	0.37	-0.01	-0.07	0.01	0.03	0.02	-0.12***
	Down-match	0.31	0.10***	0.06	-0.04	-0.01	-0.06	0.09**
	Value of mismatch	-0.16	-0.35**	-0.34	0.08	-0.02	-0.02	-0.66***
SD mechanism, second batch	Mismatched	0.76	-0.01	-0.20*	0.11	0.07	0.02	-0.01
	Up-match	0.44	0.07	-0.02	0.04	-0.03	-0.10	-0.18**
	Down-match	0.32	-0.07	-0.18	0.07	0.09	0.12	0.17**
	Value of mismatch	0.39	0.26	-0.38	0.84	-0.06	-0.30	-1.17***
BOS mechanism, second batch	Mismatched	0.78	0.09*	-0.29*	0.23**	0.10	-0.07	0.02
	Up-match	0.48	-0.06	-0.16	0.40***	0.11	-0.08	-0.15**
	Down-match	0.3	0.15**	-0.14	-0.17	-0.01	0.01	0.16**
	Value of mismatch	0.64	-0.69*	-0.15	1.93***	0.32	-0.33	-0.72**

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. Column 2 gives the difference between for example female student and male students. It is obtained by regression mismatch measures on the dummy of female students.

Table 3. Matching Quality under Different Mechanisms: Ordered Logit model

	Dependent variable: Actual college tier					
	All observations	BOS mechanism, all batches	SD mechanism, batch 2	BOS vs. SD	BOS mechanism, batch 2	BOS vs. SD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.29**	-0.39***	-0.03	-0.40***	-0.75*	-0.67**
	(0.12)	(0.13)	(0.31)	(0.13)	(0.38)	(0.33)
Rural	-0.36**	-0.43**	-0.53*	-0.44**	-1.30	-0.96
	(0.16)	(0.21)	(0.27)	(0.21)	(0.87)	(0.67)
Minority	0.21	0.09	0.99*	0.11	0.70	0.79
	(0.22)	(0.26)	(0.52)	(0.27)	(0.78)	(0.72)
Age	-0.01	0.08	-0.42*	0.08	0.46	0.34
	(0.10)	(0.11)	(0.23)	(0.11)	(0.31)	(0.29)
Scores in	0.30	0.42**	-0.32	0.55***	0.18	-0.27
analogue CEE	(0.19)	(0.21)	(0.47)	(0.21)	(0.55)	(0.68)
Science exam	0.25	0.10	-0.95	0.18	-0.99	-1.45
	(0.46)	(0.59)	(2.37)	(0.47)	(2.58)	(1.32)
SD mechanism				5.81		11.50*
				(4.08)		(6.72)
Female *				0.51**		0.68*
SD mechanism				(0.26)		(0.41)
Rural *				0.21		0.51
SD mechanism				(0.29)		(0.69)
Minority *				0.54		0.18
SD mechanism				(0.46)		(0.86)
Age *				-0.36*		-0.70*
SD mechanism				(0.22)		(0.36)
Academic aptitude*				-0.91**		-0.23
SD mechanism				(0.36)		(0.60)
Observations	1,031	819	212	1,031	191	403

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies.

Table 4. Mismatch under Different Mechanisms

	BOS mechanism, all batches				SD mechanism, batch 2			
Dependent variable:	Mismatched	Down-matched	Up-matched	Value of mismatch	Mismatched	Down-matched	Up-matched	Value of mismatch
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.09**	0.13***	-0.04	-0.40***	0.01	0.03	-0.07	-0.15
	(0.04)	(0.04)	(0.04)	(0.14)	(0.06)	(0.11)	(0.09)	(0.32)
Rural	-0.05	0.09	-0.13**	-0.46**	-0.27**	-0.21	-0.22**	-0.49
	(0.07)	(0.07)	(0.06)	(0.22)	(0.12)	(0.13)	(0.11)	(0.30)
Minority	0.04	-0.01	0.06	0.13	0.07	0.10	-0.00	0.62
	(0.06)	(0.06)	(0.06)	(0.23)	(0.09)	(0.15)	(0.13)	(0.55)
Age	0.04	-0.01	0.06*	0.09	0.08	0.30***	-0.07	-0.48*
	(0.03)	(0.03)	(0.03)	(0.12)	(0.06)	(0.10)	(0.08)	(0.25)
Scores in	0.05	-0.08	0.15**	0.27	-0.09	0.25	-0.24	-0.25
analogue CEE	(0.07)	(0.08)	(0.07)	(0.24)	(0.13)	(0.24)	(0.20)	(0.68)
Science exam	0.07	-0.00	0.07	-0.11		-0.96***	0.51	-2.04
	(0.14)	(0.15)	(0.13)	(0.62)		(0.06)	(0.55)	(2.77)
Observations	817	735	784	819	194	163	199	201

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies. Columns 1-3 and 5-8 apply the Probit model and report marginal effect. Columns 4 and 8 apply the Ordered Logit model. The coefficient for the dummy of science exam is missing in column 5 because it is dropped due to collinearity. The number of observations is different across columns, because the Probit estimation drops some observations when the prediction is perfect.

Table 5. Average Matching Quality for Batch 2: BOS mechanism vs. SD Mechanism

	Sample year: 2006-2007					Sample year: 2007-2008			
Dependent variable:	Mismatched	Mismatched	Up-matched	Down-matched	Value of mismatch	Mismatched	Up-matched	Down-matched	Value of mismatch
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SD batch *	-0.21	-0.20	-0.22	0.34	-0.43	-0.23	-0.37**	0.06	-0.92*
Post treatment	(0.15)	(0.15)	(0.15)	(0.21)	(0.50)	(0.14)	(0.15)	(0.22)	(0.54)
SD batch	0.15	0.10	-0.10	0.04	-0.96**	0.01	0.42***	-0.64***	1.76***
	(0.11)	(0.11)	(0.13)	(0.15)	(0.46)	(0.10)	(0.10)	(0.12)	(0.43)
Post-treatment	0.14	0.16*	0.08	0.04	-0.39	0.15	0.40***	-0.17	0.94**
	(0.09)	(0.09)	(0.12)	(0.12)	(0.40)	(0.12)	(0.15)	(0.22)	(0.47)
Science exam				0.33	-2.35	-0.03	0.11	-0.52	-2.30
				(0.59)	(1.85)	(0.56)	(0.87)	(0.83)	(1.84)
Female		0.10	-0.10	0.26***	-0.39	0.03	-0.05	0.08	-0.18
		(0.07)	(0.08)	(0.08)	(0.26)	(0.06)	(0.09)	(0.09)	(0.25)
Rural		-0.51**	-0.29*	0.00	-0.64	-0.34**	-0.24**	-0.14	-0.42
		(0.22)	(0.15)	(0.25)	(0.61)	(0.14)	(0.12)	(0.13)	(0.40)
Minority		0.23***	0.31**	0.09	1.06**	0.07	0.07	0.03	0.77
		(0.04)	(0.12)	(0.18)	(0.51)	(0.09)	(0.14)	(0.15)	(0.47)
Age		0.10*	0.12*	0.04	0.19	0.05	-0.12	0.20**	-0.58***
		(0.05)	(0.07)	(0.08)	(0.23)	(0.05)	(0.08)	(0.08)	(0.22)
Scores in		0.01	0.04	0.01	-0.08	-0.05	-0.20*	0.18	-0.86**
Analogue CEE		(0.09)	(0.11)	(0.14)	(0.37)	(0.08)	(0.12)	(0.13)	(0.34)
Observations	221	221	243	185	250	251	256	219	265

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies. Columns 1-3 and 5-8 apply the Probit model and report marginal effect. Columns 4 and 8 apply the Ordered Logit model. The coefficient for the dummy of science exam is missing in columns 2 and 3 because it is dropped due to collinearity. The number of observations is different across columns, because the Probit estimation drops some observations when the prediction is perfect.

Table 6. Descriptive Statistics for Preference Manipulation: Misreport

		Total	Difference between groups					
			Female -Male	Rural -Urban	Minority -Han	Old -Young	Science - Arts	Scores in mock exams: top half - bottom half
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
BOS mechanism, all batches	Misreported	0.65	0.06**	-0.04	-0.04	-0.04	-0.08***	0.06**
	Up-report	0.38	-0.01	-0.08	-0.09	-0.01	0.03	-0.07**
	Down-report	0.27	0.07***	0.04	0.05	-0.02	-0.11***	0.13***
	Value of misreport	0.00	-0.29**	-0.46**	-0.33	-0.08	0.29**	-0.83***
SD mechanism, second batch	Misreported	0.8	-0.01	0.00	-0.24***	0.08	0.02	0.08
	Up-report	0.54	0.02	-0.07	-0.12	0.12	-0.05	-0.16
	Down-report	0.26	-0.02	0.07	-0.12	-0.04	0.07	0.24**
	Value of misreport	1.00	0.24	-0.17	-0.36	0.85**	-0.78**	-0.85
BOS mechanism, second batch	Misreported	0.78	0.09*	-0.03	-0.12	0.04	-0.05	0.23**
	Up-report	0.53	-0.02	0.03	-0.09	0.04	-0.02	-0.21
	Down-report	0.25	0.11**	-0.06	-0.03	-0.00	-0.03	0.44***
	Value of misreport	0.81	-0.50*	0.79	-0.38	0.16	-0.02	-2.41***

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively

Table 7. Strategic Behavior under Different Mechanisms: Ordered Logit model

	Dependent variable: tier of first-choice college					
	All observations	BOS mechanism, all batches	SD mechanism, batch 2	BOS vs. SD	BOS mechanism, batch 2	BOS vs. SD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.19***	-0.27***	0.03	-0.29***	-0.30*	-0.29*
	(0.07)	(0.08)	(0.16)	(0.08)	(0.16)	(0.15)
Rural	-0.10	-0.21	0.10	-0.20	-0.33	-0.33
	(0.12)	(0.14)	(0.26)	(0.14)	(0.25)	(0.25)
Minority	-0.13	-0.20	0.02	-0.19	0.12	0.07
	(0.14)	(0.17)	(0.26)	(0.18)	(0.37)	(0.37)
Age	-0.03	-0.03	-0.01	-0.04	0.08	0.07
	(0.06)	(0.06)	(0.14)	(0.07)	(0.13)	(0.13)
Scores in	0.24**	0.17	0.36*	0.31***	0.04	0.10
analogue CEE	(0.10)	(0.11)	(0.21)	(0.11)	(0.19)	(0.18)
Science exam	-0.41	-0.55	0.05	-0.58	-0.83	-0.59
	(0.43)	(0.48)	(0.93)	(0.45)	(0.66)	(0.52)
SD mechanism				-1.20		0.50
				(2.50)		(3.47)
Female *				0.38**		0.35*
SD mechanism				(0.15)		(0.21)
Rural *				0.31		0.39
SD mechanism				(0.26)		(0.33)
Minority *				0.25		-0.04
SD mechanism				(0.27)		(0.44)
Age *				0.10		-0.05
SD mechanism				(0.13)		(0.18)
Academic aptitude*				-0.30*		0.18
SD mechanism				(0.16)		(0.22)
Observations	1,031	819	212	1,031	191	403

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies.

Table 8.Misreport under Different Mechanisms

	BOS mechanism, all batches				SD mechanism, batch 2			
Dependent variable:	Down-reported	Up-reported	Misreported	Value of misreport	Down-reported	Up-reported	Misreported	Value of misreport
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.04**	-0.03*	0.00	-0.40***	0.03	-0.00	0.03	-0.10
	(0.02)	(0.02)	(0.02)	(0.10)	(0.03)	(0.04)	(0.04)	(0.17)
Rural	0.03	-0.04	-0.01	-0.14	-0.06	0.01	-0.06	-0.06
	(0.03)	(0.03)	(0.03)	(0.18)	(0.05)	(0.07)	(0.07)	(0.28)
Minority	0.02	-0.03	-0.01	-0.37*	0.07	-0.01	0.06	-0.18
	(0.03)	(0.03)	(0.03)	(0.19)	(0.05)	(0.06)	(0.06)	(0.29)
Age	-0.00	-0.00	-0.01	-0.06	0.01	0.04	0.05	0.00
	(0.01)	(0.01)	(0.02)	(0.08)	(0.02)	(0.03)	(0.03)	(0.15)
Scores in	-0.02	0.06**	0.04	0.16	-0.05	0.08	0.03	0.33
analogue CEE	(0.02)	(0.02)	(0.03)	(0.13)	(0.04)	(0.05)	(0.05)	(0.21)
Science exam	0.13	-0.13	-0.01	-0.80	0.41***	-0.28*	0.13	0.11
	(0.15)	(0.12)	(0.19)	(0.65)	(0.13)	(0.15)	(0.18)	(0.91)
Observations	2,289	2,289	2,289	2,289	584	584	584	584

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies. Columns 1-3 and 5-8 apply the Probit model and report marginal effect. Columns 4 and 8 apply the Ordered Logit model.

Table 9. Average Matching Quality for Batch 2: BOS mechanism vs. SD Mechanism

	Sample year: 2006-2007					Sample year: 2007-2008			
Dependent variables:	Misreported	Misreported	Up-reported	Down-reported	Value of misreport	Misreported	Up-reported	Down-reported	Value of misreport
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SD batch *	0.00	0.00	-0.03	0.09	0.00	-0.02	-0.01	0.00	0.31
Post treatment	(0.06)	(0.06)	(0.08)	(0.11)	(0.25)	(0.06)	(0.08)	(0.11)	(0.26)
SD batch	-0.05	-0.04	-0.25***	0.30***	-1.37***	0.05	0.32***	-0.42***	1.54***
	(0.05)	(0.05)	(0.06)	(0.08)	(0.20)	(0.04)	(0.05)	(0.07)	(0.19)
Post-treatment	0.02	0.04	0.00	0.07	-0.00	0.01	0.05	-0.09	0.13
	(0.04)	(0.04)	(0.06)	(0.07)	(0.18)	(0.04)	(0.06)	(0.09)	(0.20)
Science exam	0.26	0.23	-0.41	-0.85***	-0.80	1.00***	0.81***	-0.85***	-1.32
	(0.29)	(0.30)	(0.34)	(0.18)	(0.96)	(0.00)	(0.01)	(0.27)	(1.41)
Female		0.04	-0.01	0.11**	-0.18	0.00	-0.02	0.06	-0.16
		(0.03)	(0.04)	(0.06)	(0.13)	(0.03)	(0.04)	(0.06)	(0.14)
Rural		0.05	0.02	0.07	-0.37	-0.06	-0.07	0.01	-0.12
		(0.06)	(0.10)	(0.13)	(0.27)	(0.06)	(0.08)	(0.11)	(0.24)
Minority		-0.05	-0.09	0.03	0.13	0.05	-0.01	0.09	0.13
		(0.06)	(0.07)	(0.12)	(0.26)	(0.05)	(0.08)	(0.10)	(0.24)
Age		-0.00	0.03	-0.05	0.09	0.02	0.03	-0.01	-0.06
		(0.03)	(0.04)	(0.05)	(0.11)	(0.03)	(0.04)	(0.05)	(0.12)
Scores in		0.09**	0.13**	0.01	0.21	0.04	0.11**	-0.05	0.22
Analogue CEE		(0.04)	(0.05)	(0.08)	(0.16)	(0.04)	(0.05)	(0.07)	(0.17)
Observations	845	845	849	523	861	806	815	574	830

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies. Columns 1-3 and 5-8 apply the Probit model and report marginal effect. Columns 4 and 8 apply the Ordered Logit model.

Table 10. Robustness Check: Different Measures of College Ranking, OLS regression

	The second college ranking: 127 categories				The third ranking of college: 502 categories			
Dependent variable	College tier admitted		Tier of first-choice college		College tier admitted		Tier of first-choice college	
	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-5.34***	1.27	-2.39**	0.92	-16.01***	-34.47	-10.80**	-5.90
	(1.71)	(3.05)	(0.96)	(1.92)	(8.07)	(24.44)	(4.74)	(12.17)
Rural	-0.69	5.72**	-2.29	3.02	-3.19	-88.02***	-3.35	-10.54
	(3.12)	(2.82)	(1.79)	(2.60)	(14.36)	(30.40)	(8.78)	(20.00)
Minority	0.78	5.45*	-3.63**	0.15	1.72	75.89*	-28.92***	3.36
	(3.09)	(3.29)	(1.84)	(2.76)	(13.74)	(44.19)	(8.78)	(21.82)
Age	-0.14	-2.08	-0.80	1.04	4.08	-25.38	-2.50	4.32
	(1.36)	(2.25)	(0.78)	(1.58)	(6.34)	(21.25)	(3.94)	(11.13)
Scores in analogue CEE	4.60	-0.10	1.41	6.28**	-15.24	8.11	0.66	20.96
	(3.33)	(6.34)	(1.23)	(2.50)	(15.20)	(46.63)	(7.08)	(15.30)
Science exam	-7.58	66.76**	-10.34***	-9.25	87.04***		-14.18	-232.14***
	(6.86)	(26.56)	(5.04)	(8.64)	(27.54)		(52.70)	(62.35)
Observations	818	201	2,285	584	829	205	2,320	600
Adj. R-squared	0.622	0.193	0.643	0.158	0.582	0.240	0.642	0.222

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies. The dummy for science exam type misses in column 6 because it is dropped due to collinearity.

Table 11. Robustness Check: Humanities vs. Science Exam

	Humanities exam				Science exam			
Dependent variable	College tier admitted		Tier of first-choice college		College tier admitted		Tier of first-choice college	
	Matching quality		Preference manipulation		Matching quality		Preference manipulation	
	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism	BOS mechanism	SD mechanism
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.55**	0.31	-0.35**	0.12	-0.41**	0.03	-0.25***	0.06
	(0.24)	(1.12)	(0.16)	(0.45)	(0.16)	(0.35)	(0.09)	(0.18)
Rural	-0.26	0.18	-0.38	-0.17	-0.40	-0.87**	-0.11	0.21
	(0.49)	(0.63)	(0.27)	(0.47)	(0.25)	(0.44)	(0.17)	(0.35)
Minority	0.11	2.01	-0.05	-0.20	0.21	0.42	-0.29	0.00
	(0.49)	(1.29)	(0.26)	(0.52)	(0.34)	(0.63)	(0.23)	(0.34)
Age	0.32	-0.88	-0.04	0.08	-0.04	-0.34	-0.06	-0.05
	(0.22)	(0.56)	(0.13)	(0.34)	(0.12)	(0.27)	(0.07)	(0.16)
Scores in analogue CEE	1.11*	-0.07	0.44*	-0.08	-0.00	-1.12	0.08	0.33
	(0.60)	(1.74)	(0.27)	(0.49)	(0.31)	(0.80)	(0.13)	(0.26)
Observations	282	58	745	177	537	154	1,572	482

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. All regressions control for 11 dummies for students' fair tier in the fair matching, CEE score rank, its square and cubic terms, and classes dummies.

Appendix

Table A.1 Number of Colleges and Students for Each College Tiers or Batches

	Number of colleges		Number of students	
	Humanities	Science	Actual tier	Fair tier
Tier	(1)	(2)	(3)	(4)
Not admitted			184	6
1	281	329	85	281
2	81	150	133	128
3	62	64	69	58
4	43	44	55	54
5	44	48	68	52
6	43	43	89	88
7	42	45	66	103
8	36	40	151	149
9	35	32	112	137
10	24	18	124	110
11	5	7	44	50
12	7	6	35	34
Total	703	826	1215	1250
Batch				
1A	92	116	372	360
1B	41	53	232	338
2A	162	186	248	166
2B	203	246	155	275
3	209	222	398	425
4			36	55
Not admitted			184	6
Total	707	823	1625	1625