THE EFFECT OF WINNING A FIRST-CHOICE SCHOOL ENTRY LOTTERY ON STUDENT PERFORMANCE: EVIDENCE FROM A NATURAL EXPERIMENT[1]

Fang Lai[2]

New York University

(April 2007)

[1] This research has been supported by the Social Science Research Council (SSRC), the Russell Sage Foundation (RSF), and the Berkeley Center of Chinese Studies (CCS). We would like to thank the Education Bureau of the Beijing's Eastern City District, the Institute of Higher Education of Tsinghua University, and the Data Sea Survey Institute for their support while in China.

[2] Corresponding author: fang.lai@nyu.edu; Comments and suggestions are welcome.

1. Introduction

There are heated debates surrounding school choice in the field of Education Policies. In the US, proponents claim that increased choice will force the current school system to improve efficiency (e.g. Hoxby, 2000a). This claim is based on the assumption that academic quality is the major concern of parental school choice. However, studies have shown that parental preferences about schools might be heterogeneous (e.g., Hastings 2005a). In addition, critics have raised concerns that "increased choice will isolate the most disadvantaged students in the worst schools and parents may not be informed enough to make choices in the best interests of their children" (Cullen, Jacob, and Levitt, 2005, 2006; Elacqua, 2006; Ladd, 2002; Teske and Schneider, 2001; Hsieh and Urquiola, 2006). The empirical studies of the effects of various forms of school choice on student performance have found mixed evidence (Angrist, et al., 2002;[3] Howell and Peterson, 2002; Teske and Schneider, 2001; Hsieh and Urquiola, 2006; Rouse, 1998; Cullen, Jacob and Levitt, 2005, 2006; Mayer, et. al., 2002; Krueger and Zhu, 2002). Many of these studies are random assignment experiments with school vouchers or school choice lotteries administered in the U.S. and South America. They estimate the average treatment effect (ATE) on student performance (in most cases measured by test scores) of winning first-choice school entry or gaining a voucher to transfer to a private school. For example, exploiting the randomized lotteries in the school choice program for Chicago public schools (CPS), Cullen, Jacob and Levitt (2006) found no significant difference in academic achievement between lottery winners and losers, but lottery winners did excel in some nonacademic outcomes. This sort of evidence is essential to the resolution of the debates surrounding school choice, as in order to assess school choice, we need to first examine whether student performance is improved by entering the school of one's choice.

In China, as part of the on-going education reform, local governments have replaced the previous merit-based or geographical proximity-based middle school admission procedures with various forms of random assignment of students to different middle schools. In most cases, the random assignment is to some extent conditional on students' school applications. The government hopes to achieve improvement in the equity of educational opportunity while accommodating individual

[3] They found a beneficial effect of winning a private secondary school lottery on student academic performance.

preference of school. In order to decide to what extent or in what manner school choice should be allowed, the same questions concerning various school choice and voucher programs in the U.S. and other parts of the world arise here: Is there any beneficial effect on student academic performance from entering the school of one's choice? If there is any, what are the sources of this effect? Will school choice lead to social stratification due to differential parental school choice? Answers to these questions are important to assess the effect of school choice on educational quality and equity. This paper is the first attempt, as far as we know, to rigorously assess the consequence of school choice on student performance using the Chinese school choice lottery program; yet the results will not only be relevant to the on-going school choice reform in China, but also shed light on school choice programs worldwide.

To answer the above questions, this paper exploits the preference-based random assignment of students to middle schools resulting from the educational reform in Beijing's Eastern City District in 1998. The data set consists of the census data and administrative data on 7000 students who entered middle school in 1999 and graduated from middle school in June 2002, and the survey and administrative data on school characteristics including school facilities and teacher characteristics. 4,937 of these students submitted their middle school application, and then 4,646 of them were randomly selected by the schools, as the schools they applied to were oversubscribed. We estimate the effect of entering one's first-choice school by comparing the lottery winners (i.e., students who were randomly selected into their first-choice school) and lottery losers (i.e., students who were randomly selected out of their first-choice school) within the same lottery of first-choice school.

Although the middle school admission procedure was designed to be a random assignment based on student school applications, there were several departures from random assignment in its implementation. In particular, transferring school after the random assignment incurred significant differences in some important socioeconomic characteristics between lottery winners and losers. Thus, instead of the simple comparison between lottery winners and losers, we first estimate the "probability of being treated" using a rich set of observed covariates, and then conduct a regression analogy of propensity score matching to estimate the local average treatment effects and the differential treatment effects between high-stake lottery takers (i.e. those who chose top-tier schools as their first choice) and low-stake lottery takers. We use various specifications of propensity scores to allow for high flexibility in the control function form, and check the robustness of the estimates by using various specifications in the estimation of propensity scores and 2SLS. Further, we use thick support, i.e. observations with estimated propensity scores between 0.33 and 0.67, to gauge the possible bias in estimation by unobserved individual characteristics. Finally, as test scores are missing for around half of the observations for various reasons, we also use a two-sided truncation model and weighed regressions to gauge this problem.

Results show that entering one's first-choice school does not have significant beneficial effects on the student test scores in the High School Entrance Exam (HSEE) 2002.[4] However, the beneficial effects of entering one's first-choice school are larger for students who applied to the top-tier schools (i.e., taking a high-stake lottery) than those who chose other schools as their first choice (i.e., taking a low-stake lottery). This indicates that entering one's first-choice school does bring more beneficial effects on academic performance for students who were more academically ambitious than those who were not. Moreover, even though academic quality is a major factor in parental school choice in general, parental preferences of schools are heterogeneous to some extent. In particular, students applying for the top-tier schools tend to have stronger academic and socioeconomic backgrounds, indicating sorting in school choice along socioeconomic status, which is also observed by many studies (e.g. Hsieh and Urquiola, 2006). Still further, many of the oversubscribed schools were outperformed by undersubscribed schools in the HSEE 2002 after the re-shuffling of students across schools via randomization. Thus, parents seemed to select schools based on their performance prior to the advent of school choice reform, suggesting that misinformation might lead to inefficient school choice. These are all possible reasons for the overall insignificant effects of entering one's first-choice schools on student performance. Finally, there seems to be neither lottery winning effects nor differential lottery winning effects between high-stake and low-stake lottery takers after controlling for various school characteristics.

[4] Students who entered middle school in 1999 graduated middle school and entered high school in 2002.

The remainder of this paper is organized as follows: Section 2 provides information on the middle school reform, the new middle school admission procedure, and the data set. Section 3 checks the validity of randomization in school assignment and describes the methodology of the "regression analogy" of propensity score matching. Section 4 estimates the overall effects of winning a lottery of first-choice school entry on student academic performance and the difference of these effects between high-stake and low-stake lottery takers, and conducts a robustness check to the estimates. Section 5 explores the sources of the heterogeneous effects of winning a first-choice school entry lottery between high-stake and low-stake lottery takers. Section 6 concludes the paper.

2. Background and Data

2.1. The middle school education reform in the eastern city district

Before 1998, primary school graduates in Beijing's Eastern City District were admitted on a merit basis by public middle schools. The students filled out their middle school applications upon graduating from primary school. Then they took a district-wide uniform Middle School Entrance Exam (MSEE). The oversubscribed middle schools first admitted students who reported these schools as their first choice according to their overall test score on MSEE, recruiting students with highest scores first. Then, if the school could not fulfill its enrollment target by admitting all first-choice applicants, it went to students who chose it as their second choice, and so on, until the school fulfilled its enrollment target. Students whose test score could not enable them to enter any of their choices would be randomly assigned to undersubscribed schools.

This middle school admission mechanism, on the one hand, provided incentives for the middle schools to improve their efficiency and performance, as their performance was essential to students' application decisions and the quality of the students they received.[5] On the other hand, however, schools with better previous performance always got better students in the long run; thus, these schools tended to rely on the high quality of entering students instead of high school quality to keep up their performance and attract high quality applicants. Similarly, rather than making a good effort in improving school

[5] Once they were oversubscribed , the schools were able to pick the students with highest performance.

quality, schools with worse performance were inclined to blame the inferior quality of the entering students for poor performance. As a result, over time, this admission procedure reinforced the existing inequalities in performance across schools without necessarily improving overall school efficiency. In fact, the performance of the middle schools was highly heterogeneous by the time of the education reform in Beijing in 1998, reporting, for example, vastly different performance on the city-wide HSEE, with excellence rates[6] ranging from 55% to 100%.

The sharp gap in performance across middle schools brought about fierce competition in the MSEE among primary school graduates. Introducing such fierce competition and high inequalities in children's access to "quality schools" at such an early age[7] had been considered both unfair and detrimental to the children's physical and psychological development. Thus, for quite some time, the public demanded an equalization of access to quality schools. Because directly equalizing resources across schools would have been very difficult given the existing dramatic heterogeneity across schools and vested interests, and because the government considered that the merit-based selection mechanism puts unhealthy pressure on the children, it decided to first equalize the quality of entering students across schools. As a result, an educational reform was introduced in 1998, introducing radical changes in the middle school admission procedure.

The education bureau abandoned the merit-based selection, and divided the district into school neighborhoods based on primary school affiliation, which did not necessarily reflect household residential proximity, and assigned neighborhood-specific enrollment quotas to each middle school in the district. Each school neighborhood was linked to four to seven middle schools in a way so that all neighborhoods had equal access to the quality schools, and the formerly best middle schools were

[6] The excellence rate is the percentage of students in a middle school with test scores higher than 455 out of 560 on the HSEE. The excellence rate is a good indicator of the chance of being accepted by a good high school, which is, in turn, a strong predictor of access to university education.

[7] The students usually take the district-wide uniform middle school entrance exam at the age of 12, when they graduate from primary school.

usually available to more than one school neighborhood, whereas most low-quality[8] schools were only available to the school neighborhood of proximity.

A student could apply to some or all of the middle schools available in his school neighborhood, and had to rank them in order of preference. These choices were incorporated into a randomizing procedure as follows. First, a computer-generated 10-digit number was randomly assigned to each student. In the first round, within each school neighborhood, the middle schools admitted students who had declared them as their first choice, enrolling students from the lowest numbers, until they reached their neighborhood specific enrollment quota. If there were still slots available after admitting all first-choice applicants, the schools proceeded to the second round, choosing among applicants reporting the school as their second choice in the same manner, and so forth until they met the quota in that school neighborhood. Students who missed their preferred choice in a certain round were transferred to their next best choice in their ranked application. If a student missed all the schools he applied to, he would be randomly assigned to any middle school available to his neighborhood that had not yet filled up its enrollment quota. Thus, conditional on the school neighborhood assignments and student applications, the new enrollment procedure is a random assignment independent of a student's own characteristics and family background. Through this system, students of diverse backgrounds were mixed together and expected to spend their three years of middle school together. The randomization is incomplete in that schools allowed some students to be admitted without randomization if their parents were employed in the school, if the students had received at least a city-level prize in academic or special skill achievements, or if a considerable direct payment to the school was made.

This system is to a certain extent similar to the school choice program implemented in Chicago public schools (CPS) studied by Cullen et al (2005, 2006) and the Charlotte-Mecklenburg school district (CMS) in Hastings et al. (2005ab). Although the students were not explicitly assigned to a "home school," they were aware of the fact that if they missed all their applications, they would end up in the one (or sometimes two) school(s) that was (were) the least popular in their school neighborhood. Due to the preference-based randomization, the heterogeneity of student quality across schools decreases, while

[8] The "quality" referred to by the government here is more relevant to the historical performance of the school than actual "school quality" which cannot be directly observed.

the heterogeneity of the student body within each school increases. Now that the schools cannot solely depend on or blame student quality and related peer effects for school performance, they need to improve efficiency to promote performance and attract students of high quality, provided that school performance is a major consideration in students' school choice.

2.2. The data

A census was conducted in early 2002 by the Education Bureau of Beijing's Eastern City District in all its 27 public middle schools, covering all 7,102 students enrolled in the third year, which is also the last year, of middle school. Dropout and repetition is almost nonexistent in middle schools of this district, and hence our survey population is the population of students who entered middle school in 1999. Students were asked to give their opinions about their study environment, and to answer questions about their attitudes toward school and society. In addition, a questionnaire directed to the students' parents collected information on factors parents considered during school choice, households' wealth, parents' education levels, and some opinions such as parents' expected educational attainment for their children.

To supplement the census data, we collected school administrative data and conducted interviews of around 600 middle school teachers who taught this cohort of students. The school administrative data contain official records on the student middle school applications, school neighborhood affiliation, enrollment information, and primary school attended. It also includes students' primary school graduation test scores in two subjects, Chinese and mathematics, their test scores on all subjects through the six semesters of middle school, and their test scores on the HSEE at the end of their middle school study in 2002. The HSEE is an official city-wide uniform exam and the major mechanism for high school recruitment. However, only students who intend to enter high school take it, and the participation rate was around 70% in 2002. Moreover, we are currently able to get test scores for only two-thirds of the students that took the exam due to some data combination difficulties. Despite these shortcomings, the HSEE has the advantage of being not only uniform across the district, but also graded by one single committee appointed by the Education Bureau. Other uniform exams, though taken by all students, were graded by the schools themselves, introducing possible systematic discrepancy in grading across schools. The teacher data include basic characteristics of the teachers, such as gender, age, quality rank, education, and experience.

2.3. The preference-based randomized process of student assignment to schools and the randomized lottery of first-choice school entry

The district has 27 schools that serve 7,102 students. 2,165 students entered schools without randomization, either because they were transferred from other districts (1,247), or were directly admitted as described above (918). The other 4,937 students went through the random selection process, 367 of whom transferred schools after the random assignment, leaving 4,570 staying with their initial assignment.

There were 15 neighborhoods in 1999,[9] within each of which every student submitted a list of schools in order of preference. It turned out that, of the 27 schools, 15 could accommodate all the students that selected them as their first choice. Among the 4,570 students, the 220 students who chose them were thus directly assigned to their first choice. The remaining 4,350 students reported the other twelve schools as their first choice, all of which had more first-choice applicants than they could accommodate in at least one neighborhood. These most coveted schools thus proceeded to the random selection of their students among these first-choice applicants, and admitted 1,800 students. The 2,550 students who did not get their first-choice school proceeded to their second choice and were selected by the relevant schools in the same manner. Thus, there was random assignment going on in rounds other than the first round, and 556 students were involved in randomization in higher rounds. However, because most randomization happened in the first round, and the first choice of school demonstrates more about parental preference or choice than considerations such as hedging the risk of ending up in the least desirable schools, we will focus on random assignment in the first round and the 4,350 students whose first-choice schools were oversubscribed in the first round and who stayed with their assigned school after the randomization. From now on, unless noted otherwise, "randomization" only refers to

[9] The division in school neighborhoods varies across years.

randomization in the first round, and “oversubscribed schools” refers to schools that were oversubscribed in the first round.

The randomization can be considered as a lottery process. As the random assignment is only within the same school neighborhood, each lottery is a combination of first-choice school in the application sequence and the school neighborhood of the applicant. The student wins the lottery if he is randomly selected into his first-choice school in his school neighborhood. For example, if one of the oversubscribed schools, school No. 1, is available to two school neighborhoods, school neighborhood No. 1 and school neighborhood No. 2, the resulting two lotteries are the lottery of entry to school No. 1 from school neighborhood No. 1, and the lottery of entry to school No. 1 from school neighborhood No. 2, respectively. There are forty-two such lotteries in total.

3. Test of Validity of the Randomization and Regression Analogy of Propensity Score Estimation

3.1. Test of validity of the randomization

In the subsequent analysis of student performance, we will compare students taking the same lottery, i.e., applying to the same first-choice school in the same school neighborhood, arguing that winning or losing the lottery (i.e. whether to enter their first-choice middle school or not) is random. The validity of that analysis requires verification that children randomly selected in or out of a first-choice school within the same lottery are similar. We therefore proceed to test for the validity of the process that took place among the 4,350 students who were subjected to randomization in the first round and stayed with their school assignment,[10] as the same sample was used in subsequent sections to estimate the effect on student academic performance of entering one’s first-choice school.

The indicators used in the tests include two student characteristics (gender and primary school graduation test score), three parental socioeconomic characteristics (parents’ income and education level, and whether they have some connections in the school), and four indices of parents’ attitude toward their children, namely, the index of parents’ declared expectation for the academic achievement of their child, the parents’ opinions of the importance of school in education, the index of parents' lack of involvement

[10] 4,717 students were randomized in the first round, but 367 of them transferred school afterwards, leaving 4,350 students in the sample.

in child's education, and the index of parents' knowledge about their child's study, life, and peer group. Although the attitude variables are measured (or collected by asking the parents to recall their attitudes during the school choice procedure in 1999) during the census which was conducted two-and-a-half years after the school choice procedure, we assume these are good proxies of the parental attitudes when the students entered the middle schools.

Ideally, we should test to see if these characteristics influenced school assignment by looking within each lottery. Yet the lottery-by-lottery test is not appropriate to set up a uniform measure of the validity of the randomization. Moreover, many lottery groups are small,[11] and thus we can gain more power by pooling all the lotteries together and doing an overall test conditional on lottery fixed effects:

$$x_{il} = \alpha + \gamma \cdot winning_i + l_l + \varepsilon_{il} \tag{1}$$

where x_{il} is a characteristic of student *i* taking lottery *l*, l_l are lottery fixed effects, and $winning_i$ is an indicator equal to 1 if the student is selected into the chosen school during the random assignment, and 0 otherwise. The parameter γ measures the average within-lottery difference in the means of characteristics between students winning or losing the first-choice school entry lottery. Perfect randomization would imply a non-significant γ .

Results are reported in Table 1. Compared to those who were randomized out, students who were randomly selected in their first-choice schools have significantly higher parents' income, education, and declared importance attached to school quality. In particular, the difference in parents' incomes is larger than 10% of the mean income, while difference in parents' education is very small in scale, less than 3% of the mean level.

There are several possible reasons for the significance of these characteristics. First, based on the detailed administrative records, we identified 367 students who transferred school after

[11] The lottery group size varies from 26 to 370.

randomization, and 90% of them transferred to oversubscribed schools.[12] While we have no way to get an exhaustive count of the transfer students, we believe that we have identified most of them. The schools would be quite unlikely to accommodate more transfers than this figure, given that they could not drive out the students already admitted through random assignment or through formal nonrandom assignment procedure to make space for more transfers.[13] We exclude these students from the sample because keeping the transferring students in the sample might confound the estimates of winning a lottery on student outcomes. However, for an ideal test of validity of random assignment, the sample should include all students who went through the randomization procedure besides the students who stayed with their initial assignment. Therefore, the tests in this section not only test for the validity of randomization within each lottery per se, but also test for whether the nonrandom attritions via transfers generate significant differences in characteristics between students who were randomly selected in and students who were randomly selected out.

Transferring to other schools after the randomization could only be realized through considerable financial contributions to the coveted schools. Thus, the students who did manage to transfer schools are likely from privileged social backgrounds, and on the upper tail or even outliers in the distribution of socioeconomic backgrounds. By excluding the 367 students who transferred schools after the random assignment, we probably underestimate the average parents' income and education levels of the students who were randomly selected out of a school. In addition, the remaining unidentified transferring students, even though we do not expect there should be many of them for the reasons mentioned above, might also be a source of the significant difference in mean characteristics between lottery winners and losers. The small size of many lottery groups, and the fact that within each lottery group, many more people lost the lottery than won, further diminished the validity of the t-test.

[12] In this case, students participated in the randomization, missed their first choice, but managed to transfer schools afterwards through considerable financial contribution.

[13] After the random assignment procedure, the typical class size of the 12 oversubscribed schools was already around 45. By regulation, only 5 students per class are allowed to enroll after the random assignment enrollment procedure is finished, and on average, there are 5 classes per grade in those oversubscribed schools. It is difficult for schools to exceed this limit by under-reporting, because there was not much physical space left to further increase the class size after it reached 45. Actually, according to the data, the final class size of the oversubscribed schools is 50 on average. Thus, the actual transfer students shouldn't exceed our figures by a large amount.

As a result, γ, measuring the average within-lottery difference in characteristics between lottery winners and losers might be substantially driven up by only a small nonrandom attrition (around 8% of the total sample) of outliers.

In column (2) of Table 1, we thus restrict the test to lotteries that have at least 140 applicants, as results from large lottery groups should be more immune to the influence of outliers. Not only are the parents' income and education no longer significantly different, but also the magnitude of the differences declines substantially, indicating that the drop in significance is not solely due to the decreased precision.

Nevertheless, significant differences in characteristics between lottery winners and losers indicate that we need to control for important individual characteristics in subsequent analysis.

3.2 "Regression analogy" of propensity score estimates

3.2.1. "Regression analogy" versus semi-parametric propensity score matching

The effect of winning a lottery on student performance can be measured by comparing the performance of lottery winners and losers within the same lottery, which, as many existing studies have suggested, can done by regressing the student test scores on an indicator of winning a lottery, controlling for the lottery fixed effects. In our case, however, as some of the lottery winners and losers have significantly different characteristics due to transferring school, the OLS estimate of the effect of entering one's first-choice school (treatment effect) on student outcomes might suffer selection bias. The only possible way for the lottery losers to enter their first-choice school was to transfer to their desired school via significant financial payment (in the "legitimate" name of "donation to school construction"), which requested strong enough parental socioeconomic backgrounds due to the scarcity of slots left in each school for transferring students. Thus, in our case, departures from randomization were driven by more observable characteristics than in many other cases. Moreover, the data set contains unusually rich information of not only the demographic and socioeconomic characteristics but also important attitudinal measures of the student and their parents. An immediate thought of treating this selection bias is thus to estimate the treatment effect controlling for variables highly related to the decision and eligibility of transferring schools, including the variables representing family socioeconomic background and parental care and attitudes about the child's education. However, this approach does not allow for much flexibility in the functional form other than linearly controlling for these covariates and some higher-order terms and interactions of these covariates. We hence implement a "regression analogy" of the propensity score matching method used in non-experimental contexts (e.g. Winship and Morgan, 1999; Dehejia and Wahba, 1999[14]). The propensity score *P(X)* indicates "the probability of being treated," and is estimated using a logit or probit model based on a set of covariate *X*. In semi-parametric propensity score matching, untreated observations and treated observations are non-parametrically matched on the estimated propensity scores, and the treatment effect is estimated by comparing observations in treatment and control groups that have similar propensity scores, generating an unbiased estimate of treatment effect conditioning on observed characteristics. In this paper, instead of various semi-parametric propensity score matching methods as proposed by Rosenbaum and Rubin (1983, 1984), Rubin and Thomas (1996), Smith (1997), etc., we include various specifications of propensity scores in the regression model as controls and estimate the treatment effects on the common support (i.e. the region in the propensity score distribution with enough observations from both the treatment and control group). Formally,

$$y_{il} = \beta_0 + \beta_1 winning_i + \beta_2 (P(X_i) - \overline{P(X_i)}) \cdot winning_i + f(P(X_i)) + l_l + \varepsilon_{il} \tag{2}$$

Where y_{il} is the performance of student *i* who took lottery *l*, and l_l indicates the lottery fixed effects. $f(P(X_i))$ is a flexible function of propensity scores estimated using covariate set X_i, and is included to control for the selection bias due to the observed characteristics; the interaction of the deviation of individual propensity score from the mean propensity score and treatment dummy $(P(X_i) - \overline{P(X_i)}) \cdot winning_i$ is included to capture the heterogeneous treatment effects over the

[14] Their primary strategy was to use stratification and matching on propensity scores in the paper instead of the regression analogy.

propensity score distribution. β_1 thus represents the local average treatment effect (LATE), and β_2 indicates the magnitude and direction of the selection bias.

In addition to all controls we have included in equation (2), we can also include interaction of propensity scores and lottery dummies $P(X_i) \cdot l_l$ to control for heterogeneity in probability of being treated across different lotteries. As semi-parametric propensity matching estimators are usually the weighted average of the estimators computed within each balancing block of the propensity scores (as the observed characteristics are comparable between untreated and treated observations only within each block), we can also include the balancing block dummies (B_b , where b is the index of balancing block) as controls to make this approach closer to the semi-parametric scenario. Thus, the complete format is as below:

$$y_{ilb} = \beta_0 + \beta_1 winning_i + \beta_2 (P(X_i) - \overline{P(X_i)}) \cdot winning_i + f(P(X_i)) + l_l + B_b + l_l \cdot P(X_i) + \varepsilon_{il} \qquad (3)$$

We do not implement the semi-parametric propensity matching estimators, but a "regression analogy" instead for several reasons. First, we want to use the within-lottery group random assignment. Even though there are significant differences in parental characteristics between the treatment (lottery winners) and control group (lottery losers), indicating imperfect randomization, the random assignment procedure should have eliminated much more differences in unobserved characteristics between the treatment and control groups than a typical non-experimental design that many studies using propensity score matching methods have worked on. In addition, the rule of the remaining "selection" after randomization (i.e. transferring school) is also more visible in our case, as it strongly depends on the socioeconomic backgrounds of the parents.[15] As randomization is only within each lottery, we should also consider systematic differences across lottery groups. In this case, semi-parametric propensity score matching requires estimating the propensity scores for each of the 42 lotteries separately (see Dehejia

[15] Parental care is also essential, however, parents cannot make the transferring happen simply by caring about their children.

and Wahba, 2005) and then conducting nonparametric matching using the estimated scores within each lottery. Technically, it is very difficult to find a single set of covariates that both contains enough information and satisfies the balancing conditions (i.e., for each balancing block, the mean propensity scores are the same for lottery winners and losers) for all lottery groups. If we use different sets of covariates to estimate the propensity scores for different lottery groups, it is difficult to justify the comparability of the resulting matching estimators of treatment effect across lottery groups. On the contrary, with the "regression analogy" of propensity score matching, we can control for selection bias and systematic difference across lottery groups by simply including lottery fixed effects, a flexible function of propensity scores, and the interaction of lottery fixed effects and propensity score, and estimate the within-lottery group treatment effect. This is technically more tractable than the nonparametric approach.

Furthermore, we are not only interested in the overall effect of winning the lottery in first-choice school entry, but also the heterogeneous effects of winning the lottery on different groups, especially the groups with different preferences in school choices. Using the "regression analogy" of propensity score matching, we can simply examine the heterogeneous treatment effects by interacting the treatment with the relevant group characteristics. In contrast, the procedure will be more complicated if we use the semi-parametric propensity score matching, as we need to conduct propensity score matching not only within each lottery group, but also within each subgroup of each lottery group, which exacerbates the dilemma we described in the last paragraph.

Last but not least important, the differences in baseline characteristics between lottery winners and losers are likely to be driven by identified and unidentified cases of transferring schools, yet not only within the student's own lottery group or school neighborhood: limited school-transfer quotas in the oversubscribed schools were open to lottery losers from all school neighborhoods. It thus makes intuitive sense to estimate the propensity score using the whole sample instead of each lottery group separately, and to match the propensity score using the whole sample while controlling for lottery fixed effects instead of non-parametrically matching treatment and control cases within each lottery group.

As a result, we have to give up the flexibility in the functional form and efficiency in estimation offered by the semi-parametric propensity score matching, and choose the "regression analogy" of it for practical convenience.

3.2.2. Dealing with violations of assumptions for propensity score matching

There are several problems with this approach of the "regression analogy of propensity score matching." First, misspecification of the control functional form might introduce bias to the estimate. We deal with this problem by trying various forms of $f(P(X_i))$ and see if the estimates are robust.

Second, both the semi-parametric propensity score matching and this regression analogy of it assume that selection into treatment is solely based on observed characteristics; namely, there is no systematic difference in outcomes between the lottery winners and losers had both of them lost the lottery, i.e. ((Conditional Independence Assumption or CIA). Moreover, both methods assume that the random assignment of treatment conditioning on the propensity score $P(X_i)$ is equivalent to the random assignment of treatment conditioning on X_i. Thus, propensity score matching has the advantage of alleviating the "curse of dimensionality" compared to nonparametric matching. However, in practice, these two assumptions might not necessarily hold. To deal with violations of the last assumption, we can additionally control X_i in equation (2) or (3) to gauge the balancing quality of propensity score. As for the selection on the remaining unobserved characteristics after conditioning on $P(X_i)$, Black and Smith (2004) showed that unobservable factors on average play a larger role on the tails of propensity score distribution than on the region near 0.5, and thus we present the estimates on the thick support, i.e. observations with propensity scores falling within the interval of (0.33, 0.67) as they did.

4. The Effects of Winning a First-choice School Lottery on Academic Performance

4.1. The overall effects of entering one's first-choice schools on academic performance

4.1.1. Estimates

We now proceed to the analysis of the impact of school choice on student performance, measured by students' scores on the HSEE. This exam includes five subjects—Chinese, mathematics, English, physics, and chemistry—graded on a scale of 120 points for the first three subjects, 100 points for physics, and 80 points for chemistry. The overall score determines admission to high school. In 2002, 21 public high schools recruited their students using these HSEE results. The recognized top five high schools required a minimum score of 450, while the other high schools admitted students with scores of at least 389. With those thresholds, 71% of the students qualified for high school.

We confine our analysis to the 4,350 students who enrolled through the random assignment process described in the previous section. We exclude from the sample the students who completely avoided the random assignment process, those who transferred schools after the randomization, and those who chose undersubscribed schools as their first choice and thus enrolled in the schools of their choice without going through the randomization. We exclude these students because we cannot rigorously identify the effects of winning the lottery of first-choice school entry on these students from the effects of individual characteristics affecting both their school choice and their school performance.

Due to the missing information on student test scores, the observation number drops to 2,360, and we will deal with this missing data problem in later sections. We first regress student test scores in all subjects on the dummy of whether students enter their first-choice school (winning the lottery) or not (losing the lottery), conditional on lottery dummies as randomization only took place within the same lottery group. The regression model is:

$$y_{ilsc} = \alpha + \delta_0 winning_i + l_l + v_{sc} + \varepsilon_{ilsc} \tag{4}$$

where y_{ilsc} is the score obtained by student *i* who took lottery *l* and enrolled in school *s* and class *c*. $winning_i$ is a dummy variable equal to 1 if student *i* won the lottery and entered his or her first-choice school, v_{sc} is a random effect for class *c* in school *s*, and ε_{ilsc} the unobserved student heterogeneity factor. We include a classroom random effect to capture potential correlation between test scores within a class due to unobserved classroom characteristics, such as the head-teacher characteristics.[16] Summary results of equation (4) in Table 2 show that the impact of entering one's first-choice school is insignificant in determining the overall test scores and test scores in individual subjects except for English. Entering one's first-choice school significantly increases English score by 3.13.

To control for the departures from randomization mentioned in the previous section, the subsequent columns of Table 2 report results controlling for various student individual characteristics and various specifications of control functions using the estimated propensity scores (i.e. the "regression analogy" of propensity score matching) as described in equation (2) or (3). The results are shown in Panels A, B, and C of Table 3. All regressions from columns (1) to (6) include the estimated propensity score and the interaction of deviation from the mean propensity score and the treatment dummy to control for both "the probability of being treated" and the heterogeneous treatment effects over the distribution of "the probability of being treated." Column (2) additionally controls for the quintic polynomial of the estimated propensity score; column (3) includes all terms in column (2), and also controls for the balancing block dummies and the interaction of propensity score and lottery dummies; besides all terms in column (3), column (4) additionally controls for the covariates used in estimating the propensity score to gauge the balancing quality of the propensity scores. Columns (5) and (6) estimate the same models as columns (3) and (4), respectively, except that they use the thick support with propensity score falling within the interval of (0.33, 0.67) to alleviate the possible bias due to unobserved characteristics. Regressions in all other columns are estimated on the common support.

The covariates used to estimate propensity score are different across different panels. Panel A only uses parents' income, parents' average years of education, the age of each parent, student gender,

[16] Clustering at the school level also give similar estimates and standard errors.

and student initial test scores, as they are the most relevant characteristics in determining the possibility of transferring to their first-choice school if the students lost the lotteries. Propensity scores in Panel B are estimated using not only the covariates in Panel A but also the quadratic terms of these covariates. Propensity scores in Panel C are estimated using a larger set of covariates: in addition to the covariates used in Panel A, the propensity score estimation in Panel C also includes the cubic polynomial of covariates used in Panel A, each parent's age, squared and cubic terms of each parent's age, dummies of the primary school the student graduated from, the interaction of primary school dummies and the primary school test score, whether the student had won an official award during primary school, father's professional rank,[17] mother's professional rank, parents' ideal of the child's final educational attainment, the parental opinion of the importance of school on child's education, index of parents' lack of involvement in child's education and life, whether the family is a single-parent family, whether the family had a study room for the student, payment for middle school entry (e.g. fees for transferring school), squared term of the payment for school entry, and cubic terms for the school entry payment, the interaction of father's professional rank and parents' income, the interaction of father's professional rank and parents' average years of education, the interaction of parents' income and parental ideal of the student final educational attainment, and the interaction of parents' average years of education and the parental ideal of the student's final educational attainment. Using this larger set of the covariates substantially drops the observations on the common support due to the fact that fewer cases have complete information on all of the above variables; as a result, only 1,337 observations can be used in the final regressions.

To balance the completeness of covariates included and the number of observations remaining on the common support, we also estimate the propensity scores using different sets of covariates. The results of these different specifications are not significantly different from the results reported in Table 3 (and thus not reported). We can see that even though the estimates of the effects of first-choice school lottery winning (i.e. the treatment effects) are weakly significant at 0.1 level in Panel B, the significance is not robust to the inclusion of additional covariates used in propensity score estimation. Moreover, the

[17] Parents' professional ranks are based on the self-reported titles/ranks of parents in their respective profession.

estimates of the treatment effects are mostly insignificant when using the larger set of covariates in propensity score estimation, and the signs of the coefficients are also mixed. These results indicate that there are no systematic overall beneficial effects of entering one's first-choice school.

4.1.2. Preferences in school choices and possible reasons for insignificant effects of entering one's first-choice school on student performance

At first sight, the results are not intuitive that students did not benefit academically from entering a school that was both popular and of their choice. The first possible reason has been raised by various studies (e.g. Hastings et al., 2005ab), i.e., the parents might have heterogeneous demands for schools. For example, via school choice, some parents pursue improvement in academic performance, while others might just seek convenience in transit. In this case, entering one's first-choice school might not necessarily bring about academic improvement. In fact, we find that parents did care about the academic merit of the school. The eleven schools oversubscribed in the first round (for simplicity, we call them Type A schools in the subsequent paragraphs) in all school neighborhoods ranked No. 1 to No. 11 among all 27 public middle schools in HSEE in 1999.[18] The schools that could admit all students who chose them as first choice but were oversubscribed in some higher round (denoted as Type B schools) in general had better rank in HSEE 1999 than the schools undersubscribed in all rounds (for simplicity, we call them Type C schools). Moreover, only 220 out of 4,570 students chose Type B or C school as their first choice. Still further, as shown in Table 4, around 34% of the parents directly listed the previous performance and reputation of schools among the three most important factors in school choice decision, and 73% listed teacher quality, which was highly relevant to school academic performance. Nonetheless, we did observe a certain level of heterogeneity in parental preferences of schools, as 29% of parents chose distance from home to school as one of the three most important considerations in school choice, and 56% of parents chose school spirit and discipline.

A second possible reason is that even if the parents chose schools based on academic merits, they might not attain what they expected because of misinformation of the school performance after

[18] The cohort taking HSEE in 1999 were recruited before the reform and thus under the merit-based system.

reshuffling the students across schools via the randomization process. In Figure 1, we plot the change in the rank of a school in the HSEE between 1999 and 2002 against its initial rank in 1999, with 1 indicating best performance and 27 indicating worst performance. We use the rank of the school rather than its average score to avoid issues of comparability of exams in different years. The three school types defined earlier (A, B, and C) are marked with different symbols. We find a negative correlation between performance in 1999 and subsequent change in ranks between 1999 and 2002, except that the several top schools stayed on top. Furthermore, many previously worse schools outperformed those previously better schools. The reversion in trend seems to be too much to be explained by mean-reversion alone. This suggests that the former placement process by which students were competitively selected by the best school, as well as the resulting higher peer quality, accounted for a large part in the superior performance of these schools. Under the newly implemented randomization process, while the relative performance among schools had dramatically changed due to the reshuffling of students, parents evidently still chose schools according to previous performance. Thus, even if the parents chose the school based on academic considerations, they might not obtain what they had desired with their school choices. This indicates the importance of good information on school performance to school choice.

4.2. Heterogeneity in the effects of winning the lottery of first-choice school entry on test scores between high-stake lottery takers and low-stake lottery takers

4.2.1. Basic estimation

Considering that parents might have heterogeneous preferences and misconceptions about school quality, in examining the genuine effectiveness of school choice on academic improvement, it is instructive to look at whether entering one’s first-choice school has larger effects on academic performance for students who value academic performance in their school choice more than other students. Moreover, we need to remove the confounding influence from misconceptions about school academic quality based on school performance previous to the reform. Thus, we examine whether the effects of entering one’s first-choice school on academic performance are significantly different between students who chose one of the four top-tier schools as their first choice (i.e. the high-stake lottery takers) and those who did not (i.e. low-stake lottery takers). The top tier school titles are granted by the government via strict and comprehensive evaluation of the school, and the top tier schools also had many privileges in public funding allocation previous to the reform. Thus, the superior performance of the top tier schools was less likely to be solely boosted by the superior quality of entering students and the corresponding peer effects, and there should be less misinformation about the quality of these schools.

Table 5 compares the characteristics of high-stake and low-stake lottery takers. In general, the high-stake lottery takers, besides being more ambitious in academic achievement, which is demonstrated by higher median parental expectation of the student's final educational attainment, have stronger academic and family background: they have significantly higher average standardized primary school test scores, higher average parents' monthly income and years of education; the median father's professional rank for high-stake takers exceeds the median of the low-stake takers by one level. We thus include the dummy of high-stake takers in the estimation to control for difference in individual characteristics between the two types of lottery takers. More importantly, the difference in academic and socioeconomic backgrounds between these two types of lottery takers indicates that the post-reform admission procedure might sort students into different schools according to academic performance and socioeconomic status, and students with better academic and socioeconomic backgrounds are more likely to apply for and thus end up in better schools. This might lead to social stratification across schools, which might be further exacerbated by the fact that a few students with strong academic and socioeconomic backgrounds were able to avoid the randomization or to transfer schools after the randomization.

To estimate the heterogeneous effects of winning a first-choice school lottery and student performance and examine whether students with higher academic ambitions obtain more academic improvements by winning the lottery than students with lower academic ambitions, we use the model below:

$$y_{ilsc} = \alpha + \delta_0 winning_i + \delta_1 winning_i \times stake_i + \eta \cdot stake_i + l_l + \varepsilon_{ilsc} \quad (5)$$

Where $stake_i$ indicates the level of the stake student *i* took (1 if high-stake, and 0 if low-stake), l_l indicates the lottery fixed effects, and ε_{ilsc} is the error term allowing correlations in test scores of students within the same school *s* and the same class *c*.[19] By including the interaction of $winning_i$ and $stake_i$, δ_0 indicates the effects of winning the lottery on low-stake lottery takers, and δ_1 indicates the difference in the effects of winning the lottery between high-stake and low-stake lottery takers. The term $\eta \cdot stake_i$ captures the systematic difference in individual characteristics between the two types of lottery takers so that the estimate of δ_1 is not confounded by difference in individual characteristics between the two types of lottery takers, which is rather sharp according to Table 5. The results are reported in the first column of Table 6. It shows that winning a first-choice school lottery does not have any significant effect on the overall test scores for low-stake lottery takers. However, this effect is significantly stronger for high-stake lottery winners with an increase of around 16 points, which is around 0.3 of a standard deviation of the overall test scores of the sample.

4.2.2. Estimation using various specifications of the estimated propensity scores

Table 6 also shows estimates of δ_0 and δ_1 controlling for various functional forms of the estimated propensity scores. In addition to the basic model in equation (5), all regressions from columns (1) to (6) include the estimated propensity scores, and the interaction of the deviation of individual propensity score from the mean propensity score and the winning status. In addition, column (2) controls for the quintic polynomial of the estimated propensity scores; column (3) controls for the quintic polynomial of the estimated propensity score, the dummies of balancing block, and the interaction of propensity score and lottery dummies; column (4) includes all controls in column (3), and additionally

[19] We have also tried allowing the error terms to be correlated within school, and the estimates and standard errors did not change significantly.

controls for the covariates used in estimating the propensity scores to gauge the balancing quality. In order to alleviate the possible bias due to unobserved confounding factors, columns (5) and (6) estimate the same models as in columns (3) and (4), respectively, using the thick support with propensity score falling within the interval of (0.33, 0.67). We also tried other functions of the estimated propensity scores, but the results were similar to those in this table.

The effects of winning a lottery on the overall test score are either insignificant or negatively significant for low-stake takers, and the estimate ranges from -2.54 to -16.19. Winning the lottery has significantly improved test scores for high-stake lottery takers compared to low-stake lottery takers at least at the 5% level. This differential treatment effect is smaller in magnitude yet still significant when controlling for the covariates used in propensity score estimation (column (4)). On the thick support, after controlling for the covariates used in propensity score estimation, compared to a decrease of 16.19 points for low-stake takers (0.3 of a standard deviation of the overall test scores), by winning the lottery, high-take lottery takers gained 22.15 points (0.4 of a standard deviation of the overall scores), which indicates a roughly six-point (0.1 of a standard deviation of the overall scores) net gain in test scores by winning a high-stake lottery. Interestingly, the estimates on thick support are more significant than estimates over the whole common support region in the propensity score distribution and the estimates from models without controlling for various terms of propensity scores. There might be several possible reasons. First, the differential treatment effects between the two types of lottery takers might be more significant and larger on the middle section of the propensity score distribution. It is hard to find theoretical evidence in the literature to support this explanation, though. Second, the measurement error in the characteristics of observations in the tails of the propensity score distribution might lead to underestimate of the actual effects. Third, as the bias due to the unobserved characteristics should be lower in the thick support than in the tails of the propensity score distribution, this actually indicates that the unobserved characteristics might not lead to an overestimate of the treatment effects and the differential treatment effects between the two types of lottery takers. Instead, neglecting the unobserved might even lead to an underestimate of the significance and magnitude of the relevant coefficients.

The generally negative effects of winning a lottery on the test scores of the low-stake lottery takers become significant on the thick support. This pattern might indicate that, once they lost some high quality students due to random assignment, except for the top-tier schools, the oversubscribed schools may not necessarily be a better match for the students than the undersubscribed schools, due to the post-reform convergence in performance shown in Figure 1. As a result, being randomly assigned to one's first choice might bring some negative effects on student performance as a result of misconception of the school's academic quality.

4.2.3. Additional robustness check

4.2.3.1. Using different set of covariates and different models in propensity score estimation

We have tried various sets of covariates and different specifications for propensity score estimation and found that the estimates of the treatment effects and the differences in these effects between the two types of lottery takers are robust. We show some selected results in Table 7. All regressions control for a quintic polynomial of the estimated propensity scores, the interaction of winning a first-choice school entry and the deviation of individual propensity score from the mean level, the interaction of propensity score and lottery dummy, lottery fixed effects, the balancing block fixed effects, and the covariates used in propensity score estimation. The first column shows the results in column (4) of Table 6. Column (1) shows estimates using propensity scores estimated by a subset of the larger set of covariates, namely, the parents' average income and education, the age of each parent, student gender, primary school affiliation, and initial test scores. By using this subset, the number of observations in the common support increases from 1,337 (when using the larger set) to 2,124. The squared terms of the covariates are added in the propensity score estimation model, and the results of lottery winning effects are shown in column (2). Columns (3) and (4) correspond to columns (1) and (2), respectively, except that observations on thick supports are used instead of observations on the common support. The estimates of the treatment effects and differential treatment effects are robust to these different formulas. The estimates are also robust whether the propensity scores are estimated using the probit model or the logit model. (Table 7 only shows results with the logit model).

4.2.3.2. 2SLS estimation using predicted initial assignment as instrument variable

We have identified 367 students who transferred schools after randomization and excluded them from the sample. Even though the actual number of school transfers after randomization might be unlikely to be much larger than this figure due to limited space for transferring students in each school, we cannot rule out the possibility of unidentified transferring students remaining in the sample. In this section, we will check if the estimates are robust to the existence of unidentified transferring students.

For the 367 students who transferred schools after randomization, 68 of them transferred to the school of their first choice. The initial winning status of these 68 students should be lottery losers,[20] while they are observed as lottery winners. Thus we introduce the 367 transferring students back into the sample, and the initial winning status is employed as instrument for their observed winning status inferred from their school application sequences and actual school assignment.

Furthermore, we use the parents' self-reported entry payment for middle school admission in the census data to detect possible remaining under-reported transferring students. Even though explicitly asked about the amount of expenditure spent on activities such as school transfer payment, some parents seemed to mistake this expense for the various middle school payments such as tuition payment and book expenditure. Thus, the responses to this question are quite noisy. Moreover, nearly 1/3 of the parents did not respond, possibly either because of unwillingness to respond or considering the questions irrelevant to them as they did not have this payment. Despite these, if the parents reported a high amount of payment and their child ended up in their first-choice schools, it is reasonable for us to suspect a school transfer case. In particular, it is difficult to imagine that a typical public middle school entry payment should exceed 6,000RMB or even 10,000RMB if the students entered via random assignment.

[20] We have also used a different approach, which assigns the predicted initial winning status as losing and the observed wining status of the lottery as winning if the students finally transferred to any Type A school. These transferring students were most likely to transfer because they had lost the lottery and were faced with a Type B or C school. They might not be able to transfer to their first-choice school due to limited space availability; however, their ultimate goal might just be to enter a Type A school instead of the very school they reported as their first choice—thus it is more appropriate to assign these students as the observed lottery winners than otherwise. The results using this approach are similar to those using the approach presented in the paper.

Ideally, the student initial winning/losing lottery status should be used as an instrument to the observed lottery winning status in a 2SLS estimation of the treatment effects and the differential treatment effects between the two types of lottery takers. As we do not directly observe the actual student initial school assignment, especially for students suspected of transferring schools given abnormally high school entry payment, we have to make some prediction. Thus, we construct a rough measure of the initial lottery winning status using the information of entry payment: if the school entry payment exceeds a certain threshold and the student entered his or her first-choice school (i.e. observed as a lottery winner), we would classify her as a case of unidentified school transfer, and assign her initial winning status as a lottery loser. The initial winning status is the same as the observed status if the student was not enrolled in her first-choice school or if her reported entry payment is below the threshold. This rough measure of initial winning status is used as an instrument for the actual lottery winning status to estimate the treatment effects and differential treatment effects between the two types of lottery takers.

Because of the roughness of this instrument and the arbitrariness in setting the cutoff point of the school entry payment to predict the school transfer and initial winning status, we do not expect a precise estimate from this approach. Nonetheless, this approach might be indicative of the existence and direction of the bias due to unidentified transfers. Table 8 reports the 2SLS estimates using school entry payment thresholds of 0, 6000, and 10000, respectively. The odd-number columns assume that parents who did not respond to the payment question had their children stay with their initial assignment. The even-number columns estimate the same models as the corresponding odd-number columns, except for assuming that observed lottery winners whose parents did not report the school entry payment lost the lottery initially. The 2SLS estimates of the differential effects of entering one's first-choice schools using various school entry payment cutoff points are consistent with previous estimates, except that the estimates are insignificant in the even-number columns. Given that the instrument is just a rough prediction of the initial winning status, and that the instrumental variable estimator is less efficient than the OLS estimator, this does not compromise the robustness of the results.

Besides the basic formulas of 2SLS controlling for the lottery dummies, to alleviate the confounding influences discussed in previous sections, we additionally control for various flexible functions of the estimated propensity score, balancing blocks, covariates and the relevant interactions included in models in previous sections, and the results did not change (not reported in Table 8).

4.2.3.3. Missing observations

A final concern is that only 2,360 of the 4,350 students have HSEE scores, which substantially decreases the number of observations in the regressions. There are three major reasons for missing HSEE scores: first, the students did not expect to successfully pass the threshold of high school entrance and thus did not take the HSEE; second, the students' persistent excellent performance exempted them from the HSEE; third, the data center was unable to merge some students' test scores with the administrative and census data for various reasons, such as typographical errors in the student names in the database. Unfortunately, with the available information, we are unable to distinguish non-attendance from data entry errors, and thus we will treat them as a joint problem.

We find that on average, within each selection channel, students who won the first-choice school lottery were 3% less likely to have missing HSEE scores than students who lost, and this difference is marginally significant at the 5% level (p-value equal to 0.049). Thus, missing HSEE scores might not be random.

We conduct several tests to explore whether the nonrandom patterns of missing HSEE scores compromise the estimates and the results are presented in Table 9. First, as shown in column (1), controlling for the percentage of students with missing HSEE scores, their average semester score relative to the school average, and the square terms of these two variables do not significantly affect the estimates of the effects of winning a lottery or the differential treatment effects on high-stake and low-stake lottery takers. Second, we weight each observation in the regression models by the inverse of the predicted probability of non-missing HSEE scores (i.e. the probability of being included into the sample) to correct for the sampling bias introduced by missing scores. The probability of non-missing HSEE scores are predicted using polynomials of students' semester scores over the five semesters, individual and parental characteristics such as student gender, primary and middle school dummies,[21] primary school test scores, and parents' income and years of education. Column (2) shows that the results from weighted regression are consistent with the original results. The results are also robust to various specifications of the model that predicts the sampling weights.

Finally, we estimate a two-sided truncation regression models. If the students with missing HSEE scores had average semester scores lower than 65 points,[22] we assume that they did not expect to meet the high school entrance threshold and skipped the HSEE, and assign these observations as left-censored at 389 points (threshold of high school entrance). On the other hand, if the students with missing HSEE scores had average semester scores higher than 85 points,[23] we assume that their persistent excellent performance exempted them from HSEE, and assign these observations as "right-censored" at 450 points (threshold of top high school entrance). We use multiple imputation to impute the values of the remaining missing HSEE test scores using the semester test scores and other characteristics, assuming these scores were missed by data input/combination error, which should cause missing at random. By combining results from five imputed data sets following the approach introduced in Allison (2002), the final results are calculated and presented in column (3) of Table 9. The number of observation increased to 3,671, with 849 "left-censored" observations, 339 "right-censored", and 2,483 uncensored observations. The coefficient estimates are consistent to the original results, and robust to other reasonable cut-off points for the truncation and different multiple imputation models. The results are consistent with the original ones. We have also estimated the above models controlling various functions of the estimated propensity scores described in previous sections, and obtained similar results (not reported). To summarize, no significant evidence of selection bias due to missing HSEE scores are found.

[21] Middle school dummies are included to control for both the systematic school effects on HSEE attendance and the systematic difference in grading the semester exams across schools.

[22] 60 is the threshold of passing the semester exam, and we use 65 because a passing grade might not be enough for entering high school. The long existing four categories in the semester exam are: excellence: 85 and above; good: 75-84; pass: 60-74; fail: below 60.

[23] 85 is the threshold of excellence in the semester exam for most schools.

5. Sources of the Differential Effects of Winning a First-choice School Lottery on Test Scores between High-stake and Low-stake Lottery Takers

Why is there such a significant difference in the effect of entering one's first-choice school on test scores between the high-stake and low-stake lottery takers? First, the fixed effects of stake level should have captured the average systematic difference in individual characteristics between high-stake and low-stake lottery takers, and thus general difference in unobserved individual characteristics between these two types of lottery takers alone cannot be the source of the differential treatment effects. Second, high-stake and low-stake takers applied to different middle schools, and the difference in school quality might have driven the differential effect of entering one's first-choice school between these two types of lottery takers. Third, between high-stake and low-stake lottery takers, there might be some systematic difference in the idiosyncratic benefits of winning a lottery, such as better quality of student-school match conditioning on winning.

To explore these various sources, we include various school characteristics as additional controls in the model to see if the interaction of winning a lottery and the stake level is still significant. We cannot use the school fixed effects as they are perfectly correlated with the lottery winning status. Three sets of school characteristics variables are included. The first set of variables include characteristics of school facility, such as teacher-student ratio, the number of students per library, the number of students per science lab or computer lab in the school, the number of students per classroom, the number of students per acre of playground area, and the squared terms of these variables. The second set of the variables include the peer characteristics, namely, the average initial test score of the peer students in the same school, gender ratio of the peer students, the proportion of award-winning students for the same cohort, the peer students' average parents' years of education and income, and the proportions of parents in each of the five tiers of professional rank, together with the within-school standard deviation (except that we use the inter-quartile range for parent income) of these peer characteristics. The last set of variables include teacher characteristics averaged at school level, including the proportion of teachers of quality rank II, the proportion of teachers with quality rank III and above,[24] the proportion of teachers holding at least a university degree, the proportion of teachers securing a higher degree via teacher training program, and the average years of teaching.

The results are reported in Table 10. Panel A reports estimates without controlling for the school characteristics. Column (1) of Panel B reports estimates controlling for the school characteristics. Column (2) additionally controls for the estimated propensity scores and the interaction of winning a lottery and the deviation of individual propensity score from the mean level. Column (3) controls for all the terms in columns (2), as well as the quintic polynomial of the estimated propensity score, balancing block dummies, and the interaction of the estimated propensity scores and the lottery dummies. Columns (4) and (5) correspond to columns (2) and (3), respectively, except that they use the thick support to alleviate the lingering selection bias due to unobserved characteristics. The propensity scores are estimated using the aforementioned larger set of covariates.

In all cases, the differential treatment effects between low-stake and high-stake lottery takers are insignificant. Across all different model specifications, the teacher characteristics are the only set of variables consistently jointly significant at the 0.05 level. The peer characteristics are significant in some cases, but not in all cases, while in none of the regressions the school facility characteristics are jointly significant. We change the set of covariates used in propensity score estimation and the specifications of the model and get similar results.

When we include each set of school characteristics one at a time (not reported here), both the significance of winning a lottery and the additional effect of winning a high-stake lottery drop with the inclusion of teacher characteristics. With many model specifications, the significance of the coefficients also drops with the inclusion of school facility or peer characteristics, yet the school facility characteristics are often jointly insignificant. Of course, we cannot observe all characteristics of each of these three aspects of school characteristics, and thus cannot fully identify the genuine effects of teacher characteristics, school facility, and peers. Nevertheless, the differential treatment effects between high-stake and low-stake lottery takers become insignificant after including these observable school characteristics in the model. This indicates that the differential treatment effects between the two types

[24] The teacher's quality rank is an official evaluation of the comprehensive quality of a teacher, and it has four levels.

of lottery takers are not driven by the idiosyncratic benefits or losses of school choice in relation to the differential characteristics of these two types of lottery takers (e.g., discrepancy in the quality of student-school match) that are uncorrelated to these observed school characteristics.

6. Conclusions

This paper exploits the preference-based random assignment of students across schools in Beijing's Eastern City District to estimate the effects of entering the school of one's first choice on student academic achievement, and the heterogeneity in these effects among students with different academic ambitions by comparing winners and losers of the randomized lottery of first-choice school entry, and the differential effects of winning the lottery for high-stake and low-stake lottery takers. A regression analogy of propensity score matching is employed to control for a rich set of observed characteristics that were highly relevant to student school-transferring decisions after the initial random assignment and the possible resulting selection bias; the remaining bias due to unobserved factors is alleviated by conducting the estimation on the thick support containing observations with propensity scores between 0.33 and 0.67.

Results show that although the overall effects of winning a lottery to enter one's first-choice school are not significant, the students who applied to one of the top-tier schools benefit more from winning the lottery than students who applied to other schools. This indicates that the benefit in academic achievement from entering the school of one's choice is higher for students putting higher value on the academic achievement. These results are robust to various specifications of propensity score estimation, different control functional forms using the estimated propensity scores, 2SLS estimation to deal with potential unidentified school transfers, and missing HSEE scores. Finally, the differential academic benefits of winning the lottery between the two types of lottery takers are mostly explained by the differences in school facility, teacher characteristics, and the characteristics of the schoolmates in the same cohort. There seems to be no significant remaining idiosyncratic effects of school choice (such as student-school matching quality) orthogonal to these observed school characteristics.

Both heterogeneity in parental school choice and misconceptions about school academic quality caused by the sharp convergence of school performance after the reform might contribute to the insignificant overall effects of winning the lottery on student performance, which provides side evidence of the importance of adequate information about schools to a school choice program.

The new admission system is beneficial in that after the reshuffling of the qualities of entering students, the genuine school quality is more observable to the public (as over time, the misconceptions about the school quality will not last long), and the school can no longer rely on high student quality to keep up its good performance or explain away its poor performance using poor student quality as an excuse. As a result, they have higher incentives to improve the efficiency in order to attract high-quality students, which might finally lead to an improvement of the overall academic performance, especially when the parents do value academic achievement in school choices. However, the analysis also raises concerns about this new middle school admission system. First, as there is still space for students to avoid random assignment, this new system protects the welfare of rich students with poor academic performance as they can pay to enter their desired middle schools as they did under the old merit-based system; entering their desired schools might be even less costly than before as they do not have to pay if they win the lottery. On the contrary, the new system thwarts the prospect of poor students with good performance as they now have to count on luck instead of their performance to gain entry to their desired schools, and their performance will not help if they lose the lottery. This might even hurt their incentives to work hard in primary school and thus lower their performance. Moreover, the students with weak academic and socioeconomic backgrounds are less likely to apply to top-tier schools, possibly because the high-stake lottery is more risky and involves higher opportunity cost. As a result, the current middle school admission system might create social stratification across schools even though the initial intent of the reform is to reduce the overall inequality of access to quality schools. Thus, relevant policies are urgently needed to regulate enrollment via procedures other than randomization, or create approaches to help students with disadvantaged socioeconomic backgrounds to enter their desired schools.

References

Allison, Paul D., 2002. "Miss Data: Quantitative Applications in the Social Sciences." *Sage Publications, Inc.*

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *The American Economic Review*, 2002, 92(5), 1535-1558.

Black, Dan A., and Jeffrey A. Smith, 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." *Journal of Econometrics*, 121(2004), 99-124.

Cullen Berry J., Brian A. Jacob, and Steven D. Levitt, 2005. "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." *Journal of Public Economics*, 89(2005), 729-760.

Cullen Berry J., Brian A. Jacob, and Steven D. Levitt, 2006. "The Effect of School Choice on Participants: Evidence from Randomized Lotteries." *Econometrica*, 74(5), 1191-1230.

Dehejia, Rajeev and Sadek Wahba, 1999. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association*, 94(448), 1053-1062.

Dehejia, Rajeev and Sadek Wahba, 2005. "Practical Propensity Score Matching: A Reply to Smith and Todd", *Journal-of-Econometrics***,** 125(1-2): 355-64.

Elacqua, Gregory, Mark Schneider, and Jack Buckley, 2006. "School Choice in Chile: Is It Class or the Classroom?" *Journal of Policy Analysis and Management*, 25(3), 577-601.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2005a. "Parental Preference and School Competition: Evidence from a Public School Choice Program," Yale University Department of Economics Working Paper.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2005b. "Evaluating a School Choice Lottery: Preferences, Choices, and Heterogeneous Treatment Effects," Manuscript, Yale University.

Howell, William. G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson, 2002. "School Vouchers and Academic Performance: Results from Three Randomized Field Trials." *Journal of Policy Analysis and Management*, 21(2), 191-217.

Hoxby, 2000a. "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90, 1209-1238.

Hoxby, Caroline. M., 2002. "School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?)." NBER Working Paper 8873.

Hsieh, Chang-Tai and Miguel Urquiola, 2006. "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program." *Journal of Public Economics*, 90(2006), 1477-1503.

Krueger, Alan B. and Pei Zhu "Another Look at the New York City School Voucher Experiment." *National Bureau of Economic Research Working Paper* No. 9418, December 2002.

Ladd, H., 2002. "School Vouchers: a Critical View." *Journal of Economic Perspectives* 16, 3-24.

Fang Lai, Elisabeth Sadoulet, and Alain de Janvry, 2006. "The Effect of School and Teacher Quality on Student Performance: Using a Natural Experiment from the Middle School Reforms in Beijing." Manuscript.

Mayer, Daniel P., Paul E. Peterson, David E. Myers, Christina Clark Tuttle, and William G. Howell, "School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program," Washington, DC: Mathematica Policy Research, Inc., 2002.

Rosenbum PR. and Rubin DB, 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 76, 41-55.

Rosenbaum PR. and Rubin DB, 1984. "Reducing bias in observational studies using subclassification on the propensity score." *J. Am. Stat Assoc*, 79, 516-524.

Rubin DB, Thomas N., 1996. "Matching Using Estimated Propensity Scores: Relating Theory to Practice." *Biometrics*, 52, 249-265.

Smith HL, 1997. "Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies." *Sociology Methodology*, 27, 325-353.

Rouse, Cecilia. E., 1998. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *The Quarterly Journal of Economics*, 113(2), 553-602.

Teske, Paul and Mark Schneider, 2001. "What Research Can Tell Policymakers about School Choice?" *Journal of Policy Analysis and Management*, 20(4), 609-631.

Winship Christopher and Stephen Morgan, 1999. "The Estimates of Causal Effects from Observational Data." *Annual Review of Sociology*, 25 (1999), 659-706.

Figure 1. Change in the school performance rank between 1999-02

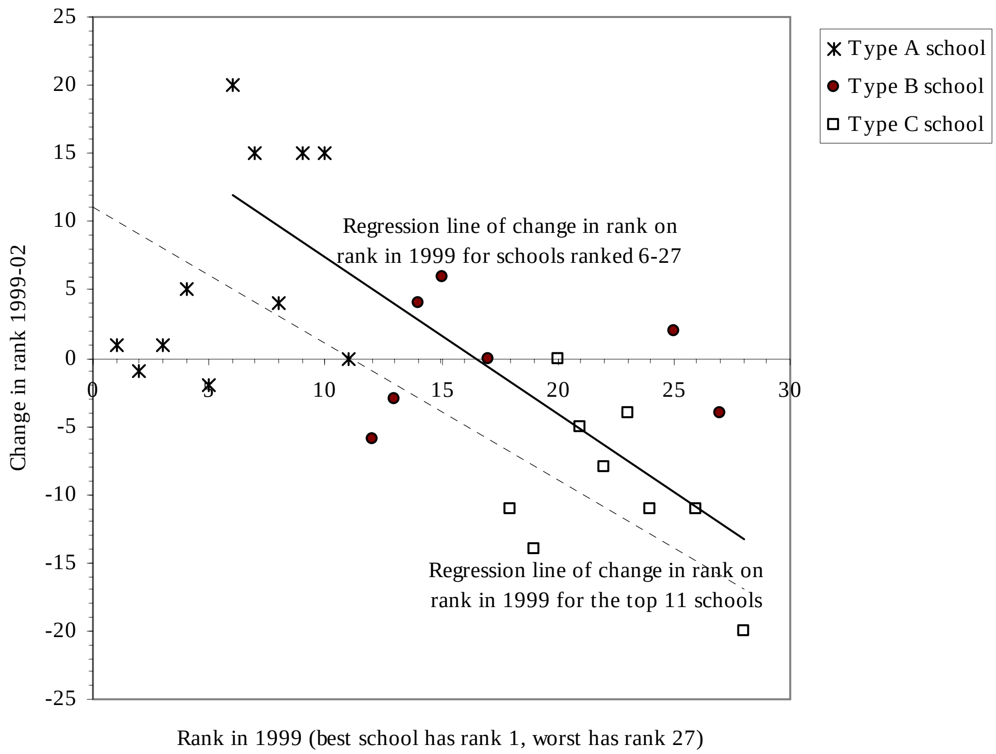

Table 1. Test of the validity of the randomization

	Winners vs. losers All students (1)	Winners vs. losers lottery size no less than 140 students (2)
Child's characteristics		
Female (0/1)		
Randomized in	-0.012	-0.033
	[0.47]	[0.16]
Constant	0.53	0.54
Observations	4350	1975
Standardized primary school test score (-1 to 1)		
Randomized in	0.001	-0.023
	[0.98]	[0.61]
Constant	-0.17	-0.16
Observations	4347	1973
Parents and family characteristics		
Parents' income (Rmb)		
Randomized in	179.029***	91.332
	[0.00]	[0.14]
Constant	1,416	1,455
Observations	4015	1,799
Parents' education level (1 to 7)		
Randomized in	0.122***	0.056*
	[0.00]	[0.07]
Constant	4.28	4.31
Observations	4156	1874
Family has relative in the school (0/1)		
Randomized in	0.009	0.015
	[0.15]	[0.13]
Constant	0.036	0.034
Observations	3353	1506
Parents' attitude		
Parents' expectation on child final education level (1 to 7)		
Randomized in	0.051	0.001
	[0.18]	[0.98]
Constant	4.30	4.32
Observations	4120	1863
Importance of school in education (1 to 5)		
Randomized in	0.111***	0.075**
	[0.00]	[0.05]
Constant	4.43	4.45
Observations	4021	1831.00
Index of parents' lack of involvement in child's education (1 to 5)		
Randomized in	-0.01	-0.063*
	[0.67]	[0.06]
Constant	0.35	0.38
Observations	4207	1903.00
Index of parents' knowledge about their child's study, life, and peer group (1 to 5)		
Randomized in	-0.01	-0.005
	[0.84]	[0.87]
Constant	3.94	3.94
Observations	3965	1789

p-value in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%. Not reported for the constant term.
All regressions have controlled for lottery fixed effects
Column (2) only includes the lotteries with at least 140 applicants.
Expected education level for child are: 1 - middle school, 2 - high school, 3 - professional college, 4- university, 5- master degree, 6- doctor degree

Table 2. The overall effect of winning a first school entry lottery on test scores in HSEE 2002

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Chinese	Math	Physics	Chemistry	English
winning lottery	4.25	0.03	0.99	0.33	-0.18	3.13**
	[0.26]	[0.95]	[0.39]	[0.65]	[0.82]	[0.02]
oberservation	2360	2360	2360	2360	2360	2360
chi square	103.04	66.73	86.44	76.72	82.38	107.4

p-values are reported in the brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3 . The estimates of the effect of winning a first-choice school lottery on student overall test score in HSEE using regresion analogy of propensity score matching

		(1)	(2)	(3)	(4)	(5)	(6)
	without controlling for propensity scores	Panel A					
		propensity score estimated by parent's income, parents' averge years of education, the age of each parent, student gender, and student initial test score					
winning lottery	4.24	5.58	5.45	5.79	2.89	5.75	2.88
	[-3.10 - 11.58]	[0.12]	[0.13]	[0.10]	[0.37]	[0.11]	[0.37]
oberservation	2360	2123	2123	2123	2123	2121	2121
chi square	103.02	117.3	124.41	161.33	366.79	160.48	364.84
		Panel B					
		propensity score estimated by all covariates used in panel A and the squared terms of these covariates excepted gender					
winning lottery		6.17*	6.10*	6.24*	2.69	6.23*	2.81
		[0.09]	[0.09]	[0.08]	[0.38]	[0.08]	[0.36]
oberservation		2124	2124	2124	2124	2106	2106
chi square		126.42	131.11	182.33	440.19	189.53	448.31
		Panel C					
		propensity score estimated by a larger set of covariates than those used in Panel A and B					
winning lottery		3.02	2.95	2.18	-3.43	-2.16	-10.73***
		[0.46]	[0.47]	[0.58]	[0.19]	[0.66]	[0.00]
oberservation		1337	1337	1337	1337	823	823
chi square		87.32	88.51	156.61	720.32	151.02	623.34

95% confidence intervals in the brackets in the first column, and p-values in all other brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%
The first column indicates the regression without controlling for various terms of propensity score.
All regressions control for the lottery dummies (the interaction of school neighborhood and first-choice school).
Column (1) controls for propensity score and the interaction between winning a lottery and the deviation from the mean propensity score.
Column (2) includes all controls in column (1) and a quintic polynomial form of propensity score.
Column (3) includes all controls in column (2), the balancing block dummies, and the interaction between lottery dummy and propensity score.
Column (4) includes all controls in column (3), and the covariates used in estimating the propensity scores.
Columns (5) and (6) are the same as columns (3) and (4), respectively, except that they use the "thick support", i.e. observations with propensity score within (0.33,0.67).

Table 4. Summary of parent's major considerations when making school choice decisions

	% of parents who rank it as No. 1 consideration	% of parents who rank it as one of the first three considerations
Students' own ability compared to the school quality	17.7	33.25
Hedging the risk of entering bad schools	4.62	12.78
Chance of entering the school of one's choice	4.62	18.60
Indicator of knowing someone working in the school	3.76	10.16
Distance from home	10.38	29.21
The spirit and discipline of the school	15.92	56.09
The quality of possible peers	0.74	7.57
School infrstructure and facilities	0.47	9.27
Teacher quality	33.29	72.89
Envionment around the school	0.64	10.75
School reputation and previous performance	7.59	34.02
Suggestions from others	0.27	0.94
Total # of observation	4046	4054

Table 5. Summary statistics of the characteristics of high-stake lottery takers and low-stake lottery takers

		mean	standard deviation	median	inter-quartile range	# of observations
Parents' average income						
	high-stake	1543.18	1665.41	1250.00	1100.00	1292
	low-stake	1303.10	1082.52	1000.00	750.00	3038
Parents' average years of education						
	high-stake	13.07	2.59	12.00	1.50	1345
	low-stake	12.53	2.25	12.00	1.50	3140
Father's professional rank (1 indicates lowest rank, and 5 indicates highest rank)						
	high-stake	2.92	1.15	3.00	2.00	1277
	low-stake	2.68	1.12	2.00	2.00	2983
Mother's professional rank (1 indicates lowest rank, and 5 indicates highest rank)						
	high-stake	2.63	1.11	2.00	2.00	1301
	low-stake	2.42	1.05	2.00	1.00	3015
Parents' ideal of child's educational attainment						
	high-stake	4.59	1.07	5.00	1.00	1317
	low-stake	4.19	1.19	4.00	1.00	3128
Student initial test score						
	high-stake	0.21	0.88	0.43	0.91	1392
	low-stake	-0.09	1.02	0.14	1.11	3287

Parents' ideal of child's educational attainments are: 1- middle school; 1 - middle school, 2 - high school, 3 - professional college, 4 - university, 5 - master degree, 6 - doctor degree and higher

Parents' professional rank are based on the titles/ranks of parents in their respective profession.

Table 6. The effect of winning a high-stake lottery vs. low-stake lottery on the overall test score

		(1)	(2)	(3)	(4)	(5)	(6)
winning low-stake	-0.44	-2.54	-2.62	-3.58	-7.63**	-8.76	-16.19***
	[-8.65 - 7.78]	[0.58]	[0.57]	[0.42]	[0.01]	[0.11]	[0.00]
winning high-stake - winning low-stake	16.31**	19.67**	19.76**	20.94***	15.54***	25.12***	22.15***
	[2.55 - 30.06]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.00]
winning*deviation from mean propensity score		3.16	0.13	-2.25	11.27	11.31	-14.04
		[0.86]	[0.99]	[0.91]	[0.52]	[0.80]	[0.71]
Observations	2360	1337	1337	1337	1337	823	823
chi square	108.36	93.58	94.82	163.73	727.13	157.71	632.15

95% confidence intervals in the brackets in the first column, and p-values in all other brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

The first column indicates the regression without controlling for various terms of propensity score.

All regressions control for the lottery dummies (the interaction of school neighborhood and first-choice school) and the dummies of stake status (high/low).

Column (1) controls for propensity score and the interaction between winning a lottery and the deviation of individual propensity score from the mean propensity score.

Column (2) includes all controls in column (1) and a quintic polynomial form of propensity score.

Column (3) includes all controls in column (2), the balacing block dummies and the interaction between lottery dummy and propensity score.

Column (4) includes all controls in column (3), and the covariates used in estimating the propensity scores.

Column (5) and (6) are the same as column (3) and (4), respectively, except that they use the "thick support", i.e. observations with propensity score within (0.33,0.67).

Table 7. Robustness check with estimates using different specifications in propensity score estimation

	original estimate	(1)	(2)	(3)	(4)
winning low-stake	-7.63**	-2.49	-2.13	-2.5	-1.75
	[0.01]	[0.49]	[0.54]	[0.49]	[0.62]
winning high-stake - winning low-stake	15.54***	18.62***	16.70***	18.58***	15.73***
	[0.01]	[0.00]	[0.01]	[0.00]	[0.01]
winning*deviation from mean propensity score	11.27	-34.65	-30.71	-34.73	-29.53
	[0.52]	[0.52]	[0.48]	[0.52]	[0.51]
Observations	1337	2123	2124	2121	2106
chi square	727.13	376.34	448.42	373.7	455.1

The first column reports the results of column (4) in the Table (6).

Column (1) estimates the same model as column (4) in Table (6), except that the propensity score is estimated by parent's income, parents' years of education, the age of each parent, student gender, student primary school affiliations, and student initial test score.

Column (2) estimates the same model as column (4) in the Table (6), except that the propensity score is estimated by parent's income, parents' years of education, the age of each parent, student gender, student primary school affiliation, student initial test score, and squared terms of all these covariates except gender and primary school dummies.

Columns (3) and (4) correspond to columns (1) and (2), respectively, except that they only use observations on the thick support, i.e. propensity score within (0.33, 0.67).

All regressions control for a quintic polynomial of the estimated propensity scores, the interaction of winning a first-choice school entry and the deviation of individual propensity score from the mean level, the interaction of propensity score and lottery dummy, lottery fixed effects, the balancing block fixed effects, and the covariates used in propensity score estimation the lottery fixed effects and the fixed effects of the stake of the lottery (high/low).

p-values are reported in the brakets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8. 2SLS regressions using the imputed initial assignment (winning lottery/losing lottery) as instrument.

	(1)	(2)	(3)	(4)	(5)	(6)
winning low-stake	-2.95	10.44	-1.21	9.3	-1.71	8.04
	[0.70]	[0.42]	[0.83]	[0.33]	[0.74]	[0.37]
winning high-stake - winning low-stake	21.37*	24.44	15.63*	14.56	16.71**	16.25
	[0.05]	[0.18]	[0.07]	[0.30]	[0.04]	[0.23]
Observations	2397	2397	2397	2397	2397	2397
chi square	110.18	109.16	109.48	109.16	110.27	109.56

p values in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

The odd number columns assume that the observations with missing information of entry payment stayed with their initial assignment.

The even number columns assume that the observations with missing information of entry payment lost the lottery in the initial assignment.

Column (1) and (2) assume all observations with entry payment bigger than 0 as losing the lottery in the initial assignment.

Column (3) and (4) assume all observations with entry payment bigger than 6000 as losing the lottery in the initial assignment.

Column (5) and (6) assume all observations with entry payment bigger than 10000 as losing the lottery in the initial assignment.

All regressions control for the lottery fixed effects and the fixed effects of the stake of the lottery (high/low).

Table 9. Robustness check to missing High School Entrance Exam (HSEE) scores

	Original estimates	(1) Control	(2) Weighted regression	(3) Two-sided truncation model
winning low-stake	-0.44	3.63	-13.15	-0.09
	[-8.65 - 7.78]	[0.28]	[0.11]	[0.97]
winning high-stake - winning low-stake	16.31**	21.07**	24.13**	13.54***
	[2.55 - 30.06]	[0.02]	[0.01]	[0.00]
Number of observations	2360	2294	2360	3671

95% confidence intervals in the brackets in the first column, and p-values in all other brackets.
* significant at 10%; ** significant at 5%; *** significant at 1%
The first column shows OLS results with missing test scores.
All regressions control for the lottery dummies (the interaction of school neighborhood and first-choice school) and the dummies of stake status (high/low).
Column (1) controls for the percentage of missing score in each school and the ratio of the mean semester test score of students with missing HSEE and that of students with HSEE observed, as well as the squared terms of these variables.
Column (3) weights each observation by the inverse of the predicted probability of taking the HSEE.
Column (4) uses the two-sided truncation model: for all students with missing HSEE, those with average semester score equal to or lower than 65 points are left-censored at 389.5 points, those with average semester score equal to or higher than 85 points are right-censored at 450.5, and all other scores are imputed using multiple imputation.

Table 10 . Sources of the effect of entering one's first-choice school - regressions controlling for various school characteristics

	A	B				
		(1)	(2)	(3)	(4)	(5)
	without control	controlling for school facility characteristics, characteristics of peer students, and school average teacher's characteristics				
winning low-stake	-0.45	19.18	-7.53	-10.42	-44.22	-41.59
	[-8.67 - 7.76]	[0.23]	[0.67]	[0.57]	[0.11]	[0.14]
winning high-stake - winning low-stake	16.32**	-14.51	-32.99	-40.45	-76.55	-97.96
	[2.57 - 30.07]	[0.75]	[0.50]	[0.43]	[0.25]	[0.16]
winning*deviation from mean propensity score			19.47	24.84	76.83*	80.15*
			[0.32]	[0.23]	[0.05]	[0.08]
Observations	2360	2040	1163	1163	715	715
chi square statistic of the whole model	108.35	138.16	691.95	742.94	577.55	635.62
chi square statistics of school facility characteristics		10.79	10.05	11.99	11.54	13.46
p-value of the joint significance of school facility characteristics		0.37	0.44	0.29	0.32	0.2
chi square statistics of the peer characteristics		8.27	11.5**	10.8*	6.79	7.28
p-value of the joint significance of the peer characteristics		0.14	0.04	0.06	0.24	0.2
chi square statistics of the teacher characteristics		11.99**	13.56**	14.35**	15.7**	16.9***
p-value of the joint significance of teacher characteristics		0.03	0.02	0.01	0.01	0

95% confidence intervals in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Panel A presents estimates without controlling for the various school characteristics.

Panel B:

Column (1) indicates estimates from the original regression controlling for the school facility, peer student characteristics, and the school average teacher characteristics.

Column (2) indicates the propensity score regression controlling for covariates indicated in the corresponding panel, propensity score, and the interaction of winning a lottery and the deviation of individual propensity score from the mean propensity scores.

Column (3) indicates the propensity score regression controlling for covariates indicated in the corresponding panel, a quintic polynomial of the propensity score, the interaction of winning a lottery and the deviation of individual propensity score, balancing block dummies, and the interaction of propensity score and lottery dummies.

Columns (4)-(5) each correspond to columns (2)-(3), respectively, except that they only use the thick support region of propensity score distribution (within the interval of (0.33, 0.67)).

All regressions include the fixed effects of the stake of the lottery (high/low) and the lottery fixed effects.

The propensity score is estimated by the larger set of covariates.