4-2017

Sorting, School Performance and Quality: Evidence from China

Yang Song
ysong@colgate.edu

Sorting, School Performance and Quality: Evidence from China*

Yang Song†

Department of Economics, Colgate University

November 2016

Abstract

School choice reforms give talented students the option to sort out of low-performing schools but often leave disadvantaged students behind. This study shows how a Chinese city was successful in helping its low-performing schools to catch up by encouraging talented students to sort into these schools. The city identified eleven low-performing middle schools and guaranteed elite high school admission to their top ten-percent graduates. This study documents that the policy improved school performance by 0.19-0.26 standard deviations. Using data on lottery middle school assignment, I further test for potential mechanisms, including strategic sorting and improvement in school value-added.

Keywords: Education Inequality; School Choice; Incentives; Sorting; Peer Effects; Top Ten-Percent Quota Policy.

JEL Classification Numbers: I21, I24, I25, I28.

*I am grateful to Daniel Berkowitz, Thomas Rawski, Werner Troesken and Randall Walsh for their advice and support. I also wish to thank Susan Dynarski, John Bound, Eric Chyn, Lixin Colin Xu, Marc Law, Sebastian Galiani, Paul Glewwe, Tate Twinam, as well as participants at CIERS at the University of Michigan, Xiamen Ronald Coase Institute Workshop, 2nd Summer School of Socioeconomic Inequality at the University of Chicago, the University of Pittsburgh, Colgate University, HKUST, HKBU, and Lingnan University for helpful comments.

†E-mail: ysong@colgate.edu. Address: 13 Oak Drive, Hamilton, NY, 13346.

1 Introduction

School choice policies including vouchers and lotteries have been widely adopted in many countries. The underlying idea is that these policies give students and parents more freedom in choosing schools, and schools are under more pressure to improve quality to attract students. However, these school-choice policies have had mixed results and, somewhat problematically, have in some cases increased sorting and stratification.[1] When given choices, students with low socioeconomic status (SES) are less likely to apply to a good school(Ajayi, 2011; Walters, 2013; Hoxby and Avery, 2013). Theoretical and empirical evidence shows that students with high SES and high ability sort out of low-performing schools, leaving disadvantaged students behind (Epple and Romano, 1998; Levin, 1998; Hsieh and Urquiola, 2006; Galiani et al., 2008; Chakrabarti, 2013; Muralidharan and Sundararaman, 2015).Given the importance of peer composition in the education production function (Epple and Romano, 2011), this sorting pattern may lead to more inequality (Calabrese et al., 2012). How to change sorting given choices so as to close the performance gaps is still an open question.

This study evaluates a policy that provides incentive for high ability students to voluntarily enroll in low-performing schools under a choice-based lottery school assignment system. Changsha, a Chinese provincial capital city with a population of seven million, introduced the top ten-percent quota policy in 2007. The education bureau chose one or two low-performing public middle schools in each district and guaranteed admission to an elite high school for the top ten-percent of 9th grade graduates from each of these schools. This paper answers two questions: Did the top ten-percent quota policy narrow the school performance gap? And if so, what are the underlying mechanisms?

To estimate the policy impact, I employ a generalized difference-in-differences identification strategy with a panel data set of middle school graduation exam performance from 2004 to 2011. Validity of random policy assignment is verified by testing whether various observable pre-policy school characteristics, including school size, school performance, and distance to elite high school, could predict policy treatment status. Results from generalized difference-in-differences regressions show that the policy schools improved their average performances by around 0.2 standard deviations in the middle school graduation exam. Their school ranking also increased by around twelve percentile. Placebo tests were conducted to show that the treatment effects do not generalize on other low-performing schools.

[1]See Hoxby (2000), Rothstein (2007), Figlio and Rouse (2006), Chakrabarti (2008), Rouse et al. (2013), and for mixed results on effects of introducing school choice and increased school competition; see Rouse and Barrow (2009); MacLeod and Urquiola (2013) for reviews. A large literature on the impacts of winning a school lottery or school voucher have also found mixed results in various locations, like Milwaukee, Columbia, New York, Chicago, Charlotte-Mecklenburg and China (Abdulkadiroğlu et al., 2011; Cullen et al., 2006; Deming et al., 2014; Krueger and Zhu, 2004; Rouse and Barrow, 2009; Zhang, 2016).

This positive policy impact could be working through various possible mechanisms, including a composition effect, a tournament effect, and a peer effect.First, a conceptual framework in section 5 illustrates the trade-offs on whether to change enrollment choice from an over-subscribed school to a low-performing policy school. It predicts that above average students who are not at the top of the talent distribution are most likely to switch to low-performing policy schools.[2] Second, competition to place at the top ten-percent may stimulate a higher effort level exerted by students, especially the top-performing ones. Third, with better peer groups and a more active learning environment, it may bring positive spill-over effects on non-switchers: students who chose an over-subscribed school but lost the lottery and students who would have chosen a low-performing school anyway. The first channel redistributes students across schools; the latter two channels increase the value-added at the policy schools.

To tease out the mechanisms, I exploit Changsha's preference-based lottery middle school assignment. Since 1996, the Changsha education bureau has assigned a fixed number of seats in several neighborhood public middle schools to each elementary school every year. A sixth grade student chooses one from the short list of middle schools assigned to his/her elementary school.[3] In cases of over-enrollment, a lottery takes place and randomly assigns winners to their chosen school and losers to the under-subscribed low-performing school, some of which were assigned the top ten-percent quotas. This allows me to analyze changes in students' school choices and compare the outcomes of lottery winners with those of lottery losers to obtain unambiguous results on the value-added gap between policy schools and over-subscribed schools.

Observing the school choices by sixth graders, I compare the baseline performance of students who voluntarily chose the policy schools before and after the policy. I found that sixth graders with high math scores and high SES were more likely to choose a policy school after the policy. In particular, students with high, but not the highest, sixth grade math scores changed their school choice to policy schools after the policy introduction. This result is consistent with predictions from the conceptual framework.

Using lottery assignment as the instrument variable, I estimate the local average treatment effect (LATE) of attending a policy school before and after the introduction of the top ten-percent quota policy. Estimates show that policy school attendance caused a 0.3

[2]Benefit from switching occurs when a student is more likely to get admitted to an elite high school. In other words, it is when a student's probability of making it into the top ten-percent of a low-performing school is higher than that of making it into the top thirty-percent among all students in the city. Cost of switching to a low-performing middle school is having lower peer quality.

[3]In China, elementary school goes from first to sixth grade and middle school goes from seventh to ninth grade.

standard deviation decrease in academic performance of lottery losers before the policy; this value-added gap was closed after the policy. With better peer quality, policy schools may improve their value-added for all students; extra effort to place at the top ten-percent brings a tournament effect only for the high-performing students. To estimate heterogeneous effects, I conduct instrumental variable quantile treatment effect (QTE) analysis (Abadie et al., 2002) to test how the policy changed the distribution of value-added gap. Results shows that before the policy, the value-added gap was negative for most deciles across the distribution, and more so for high-performing students. After the policy, the value-added gaps were closed for most of the deciles, except for the sixtieth and ninetieth percentile. Since students at low quantiles would only be subject to changes in peer effects but not tournament effects, improvements on value-added at low deciles suggest that peer effects are at work. The policy closed not only the performance gap but also the value-added gap between the low-performing policy schools and the over-subscribed schools.

This study has some implications on recent school choice reform. Attending private or charter schools sometimes brings academic and/or nonacademic benefits to the lottery winners.[4] However, school choice reforms may increase sorting and may widen the performance gap (Epple and Romano, 1998; Levin, 1998; Hsieh and Urquiola, 2006; Calabrese et al., 2012; Galiani et al., 2008; Chakrabarti, 2013). Similar to previous studies (Ajayi, 2011; Butler et al., 2013; Hastings et al., 2008; Hoxby and Avery, 2013; Walters, 2013), I find that students with lower SES are less likely to choose an over-subscribed school. Previous research attempted to improve school choice by providing information for parents and students and found that it helps in some cases, but not in others.[5]

Aside from efforts to help disadvantaged students choose and attend better schools, many studies have also looked more directly on how to improve the quality of low-performing schools. Angrist et al. (2013) and Dobbie and Fryer (2013) found that the "No Excuses" model of urban education is the key to charter school effectiveness. In the case of the top ten-percent quota policy, combining school choice with incentives for good students to attend lower-performing schools helps change the sorting patterns and narrow the performance gap. More importantly, instrumental QTE results show that the value-added gap is also closed almost everywhere across the distribution, which suggests that positive peer effects brought on by the composition changes are at work.

[4] A large literature has looked at the effect of attending a chosen school in a lottery setting and has found mixed evidence (Abdulkadiroğlu et al., 2011; Angrist et al., 2002; Dobbie and Fryer, 2011; Angrist et al., 2016; Cullen et al., 2006; Deming, 2011; Deming et al., 2014; Krueger and Zhu, 2004; Peterson et al., 1998; Rouse, 1998; Rouse and Barrow, 2009; Witte, 1997).

[5] Positive effects of providing information were found in some cases, for example Hastings and Weinstein (2008), but not in others (Banerjee et al., 2010; Mizala and Urquiola, 2013).

This study also relates to how relative evaluation changes sorting. A similar policy in the U.S. is the top x-percent rule in Texas, California and Florida, which guarantees flagship state university admission for top x-percent of seniors in all high schools. These policies and Changsha's policy differ in the school level (middle school v.s. high school), in the affected schools (some low-performing schools v.s. all schools) and most importantly, in their purposes. Changsha introduced the top ten-percent quota policy to improve low-performing schools by changing composition and improving school quality collectively, while top x% in the U.S. mainly aims at improving the minority students representation in selective colleges after the affirmative action ban (Long, 2004; Long et al., 2010). Therefore, while sorting emerges in both cases, the change in sorting was unintended in Texas (Cullen et al., 2013), but it was expected and beneficial in the case of Changsha.

Results here on the top ten-percent quota policy complement previous findings on the Texas Ten-Percent Law. Cullen et al. (2013) and this study provide converging evidence that relative evaluation brings different sorting behaviors and improves student composition in low-performing schools. Although the top x-percent rule fails to promote the opportunity of minority groups as well as the affirmative action (Long, 2004; Long et al., 2010), it helps low-performing high schools to improve their performance faster than other schools (Cortes and Zhang, 2011). I also find that in the case of Changsha, policy schools caught up in their performance. Further, this study advances previous studies by estimating the value-added gap before and after the policy, exploiting Changsha's unique lottery school assignment, and shows that the school quality improved as well. To what extent these results would generalize to environments with large variations in racial composition and instructions is uncertain.

The rest of the paper is organized as follows. Section 2 provides the background on the top ten-percent quota policy and choice-based lottery middle school assignment. Section 3 describes data construction, and thus why and how I use the data to conduct the analysis. Section 4 evaluates the policy impact on school performances using a generalized difference-in-differences framework. Section 5 provides a conceptual framework on school choice and exploits the preference-based lottery middle school assignment to tease out changes in composition and in value-added. Section 6 concludes.

2 Policy Background

Although China has experienced rapid economic growth in the past few decades, enlarging inequality has brought pressing social problems. School choice is an especially controversial topic. Differences in education quality and academic performance across schools have been widened, both within and across cities. Students bear heavy pressure to compete for access to good schools and avoid bad schools, starting from a very early age. Lump-sum fees for high quality schools and an increasingly large and expensive tutoring industry put students with low SES in worse situations. To alleviate these problems and improve equal education opportunities, governments from the central to local levels have been experimenting various policies to narrow school quality gap. One of such policies is Changsha's top ten-percent quota policy.

2.1 Top Ten-Percent Quota Policy

Changsha is a provincial capital city in South-central China, with a population of about seven million. While there are several rural districts/counties in the city, the top ten-percent quota policy only relates to the five urban districts.[6] At the elementary school level in these urban districts, there are around 18,000 students in each cohort and about 240 schools; at the middle school level, there are around 20,000 students in each cohort and about 75 schools. Elementary schools run from first until sixth grade, middle schools run from seventh until ninth grade, and high schools run from tenth until twelfth grade. At the end of their ninth grade, students take the Middle School Graduation Exam (MSGE), which determines high school admissions.

Middle schools with better past MSGE performances carry better reputations of school quality and attract students with better academic performance and higher socioeconomic status. Large performance gaps intensify sorting by ability across schools. In 2007, Changsha's education bureau initiated and announced the top ten-percent quota policy.[7] One or two low-performing middle schools in each district were chosen to pair with an elite high school. Six middle schools were originally assigned the quota since 2007 and five more were added in 2008. The percentage of elite high school direct admission quota for these policy schools was initially 2% in 2007, then slowly raised to 6%, 8% in 2008 and 2009, and reached the promised 10% since 2010. Altogether, these 11 policy schools have around 3,000 students per cohort, about 14% of the total middle school student population in the city. As of the writing of this paper, the policy is still operating.

For each pair of a low-performing middle school and an elite high school, the top ten-percent 9th grade graduates from the middle school every year are guaranteed to be admitted into the elite high school, without taking the MSGE at the end of 9th grade and competing with all other graduates. Although the top ten-percent ranking method is decided by individual middle schools and varies slightly, they basically use accumulated performance across

[6]The district description of urban and rural here is from a Chinese perspective. Urban areas are more developed and populated and typically have better schools.

[7]The project is called "dui kou zhi sheng" in Chinese pinyin, which literally translates to "pair-wise direct admission".

subjects throughout the three years in middle school. To be eligible to compete for the top ten-percent, students are required to be admitted through the preference-based lottery and attend the school from 7th grade onward. These requirements rule out possibility of late-term transfers.

The main goals of the Top Ten-Percent Policy are to change sorting, narrow performance gap between middle schools and provide better education for students attending the lower-performing middle schools. More equalized performance across middle schools eases the concern of parents and lowers the incentive of sorting. Parents would not worry as much if they send their children to a slightly lower-performing school since peer quality and chances to attend a good high school would now be higher. Students from low SES families who attend a lower-performing school would still get comparable value-added during middle school.

2.2 Preference-based Lottery Middle School Assignment

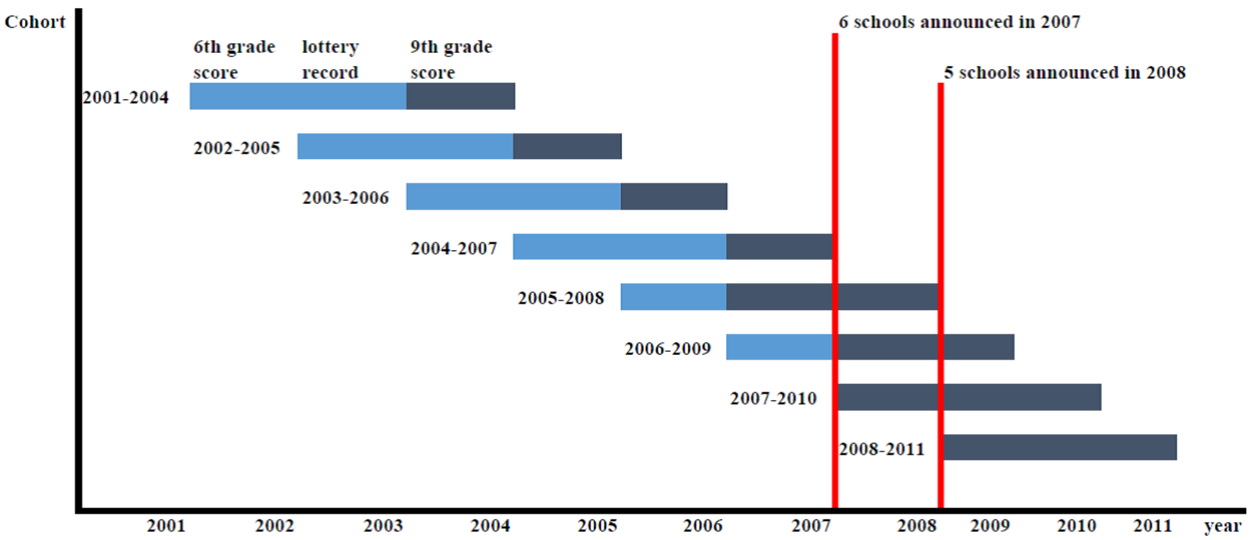
The unique preference-based lottery middle school assignment in Changsha allows me to tease out the mechanisms. Since 1996, Changsha introduced the preference-based lottery middle school assignment. Each elementary school is assigned a fixed number of seats in two or three neighboring middle schools for its graduates (i.e. sixth graders). Each sixth grader can only choose one middle school. If a middle school is over-enrolled from a particular elementary school, a lottery takes place and assigns winners to the chosen middle school and losers to a school that has unfilled seats for that elementary school. Lottery losers will be assigned randomly to one of the under-subscribed middle schools if there are more than one of them.

Before the lottery, some students get pre-admitted to several designated schools with specialized training in art, music, dance, athletics, or foreign languages. Official rules forbid other middle schools to pre-admit students by organizing selection exams. They also require all students to obey the preference-based lottery middle school assignment. I find that students who were pre-admitted or participated in the lottery and chose a high-reputation school have better family background and better academic performance than those who chose a low-performing school.

3 Data

Through the generous cooperation of local officials, I was allowed access to restrictive administrative data sets. Figure 1 illustrates the time line of policy introduction and data availability for different years. Individual level 9th grade Middle School Graduation Exam (MSGE) scores were available for eight cohorts, who entered into middle school through 2001 to 2008 and graduated three years later through 2004 to 2011. For the latter four cohorts, I have the lottery middle school assignment records, which happened in 2005 through 2008. In addition, I collected elementary school graduation exam scores from one school district for the last two cohorts. I also observe students' gender, ethnicity, city residency (hukou) and parental political affiliation for the last two cohorts.

Figure 1: Timeline of Policy Announcement and Data Availability for Different Cohorts



Notes: Data is available on the eight cohorts (2001-2004 cohort means they graduated from elementary schools in 2001 as a sixth grader and finished middle school in 2004 three years later as a ninth grader). For each cohort, the colors of each bar indicate data availability of three records: light blue for "not available" and dark blue for "available"; from left to right these three bars refer to 6th grade score, lottery record and 9th grade score. The red vertical line indicates that six and five policy schools were announced in 2007 and 2008 respectively.

Only observing sixth grade baseline scores for 2007 and 2008 in Yuhua district is the reason why I was restricted to analyze student composition changes in the 2008 policy school in this district. To verify the representativeness of the results, I ran regressions of school average performance as the dependent variable on interaction terms between dummies for policy and for Yuhua district. The coefficient for interaction term is insignificant, which suggests that comparison between policy schools and other schools in Yuhua district is similar with comparisons in other districts. In a separate regression, coefficient for the triple interaction of three dummies (policy, Yuhua district and post-policy) is also insignificant, which indicates that the policy treatment effect is no different in Yuhua district than the other districts.

Analysis on students' school choice reveals sorting patterns. Table 1 shows the summary statistics for 2007-2010 cohort. Pre-admitted students have significantly higher baseline scores and better socioeconomic status than those going through preference-based lottery school assignment. Among those students going through the normal procedure of the preference-based lottery school assignment, students choosing over-subscribed schools have higher baseline scores and better socioeconomic status. This echos with recent literature showing that students from low socioeconomic background are unlikely to choose high-quality schools across many settings (Ajayi, 2011; Butler et al., 2013; Hastings et al., 2008; Hoxby and Avery, 2013).

The main outcome variables in this paper are the 9th grade MSGE scores and elite high school attendance. The exam is high-stakes since its score is the only criteria for high schools to select students.[8] Notice that after the policy, the top ten-percent of students in the policy impacted schools get direct admission to the elite high school without taking the MSGE. Using only the non-missing test scores would under-estimate the policy impact on school performance. Therefore, I imputed the test scores of those students by assigning the highest score observable in that school. Both imputed and non-missing test scores are reported in Table 2. Since 2007, the imputed test scores are slightly higher than the non-missing test scores because of the high scores assigned to the top-x percent students in policy schools.

The MSGE final grade is consist of 6 parts, including Chinese, math, English, social science (history and politics), science (physics and chemistry), integrated subjects (biology, geography, physical education). In 2004 and 2005, final grades were in scores and high schools admissions followed a clear score cutoff. Since 2006, the education bureau changed from actual scores to letter grades A-E, with A being the highest grade and E the lowest.[9] The letter grades are determined by the percentile of student performance in each subject. On average, the cutoff percentiles are as follows: top 23% gets an A, the next 38%, 22% and 11% gets a B, C and D respectively, and the bottom 5% gets a failing grade E. High schools admit students based on letter grades and prefer higher and more balanced grades. To make the grades comparable before and after the grading scale change, I assign letter grades to each student for each subject in 2004 and 2005 by calculating the percentile category they are in, and then adding up a total score.

I use the panel of 9th grade MSGE performance to estimate the impact of the policy on treated middle schools. The number of schools and characteristics of all schools and policy schools are presented in Table 2. Earlier years have more missing data than later years.

[8]Few exceptional students get directly admitted because of the quota policy, athletic or music specialties, or exceptional academic excellence. Proportion of students who get directly admitted through channels other than quota policy did not change.

[9]The transition in grading scale might be the reason why 9th grade scores in 2006 are significantly lower than other years. Since the grading scale change equally influences students in policy schools and other schools, it does not affect the analysis of this study. I tested the effect of grading scale change on students' relative ranking in the city using 2004 and 2005 data and found that students in policy schools do not experience different change in ranking than other schools.

Table 1: Individual Level Data: Summary Statistics 2007

Sample	(1) All	(2) policy schools	(3) pre-admitted	(4) oversubscribed lottery	(5) lottery &policy
Academic performance					
9th grade score	22.73	21.68	24.71	21.60	21.55
normalized 9th grade score	0.758	0.723	0.824	0.720	0.718
% passing grade	0.795	0.763	0.876	0.766	0.767
% academic high school	0.782	0.679	0.935	0.734	0.706
% elite high school	0.332	0.299	0.529	0.201	0.249
non-academic evaluation	19.13	18.82	19.55	18.94	18.97
imputed 9th grade score	22.78	21.93	24.75	21.65	21.69
Family background					
% female	0.476	0.463	0.481	0.469	0.477
% with city hukou	0.713	0.650	0.832	0.655	0.617
% missing hukou status	0.0642	0.0681	0.0478	0.0664	0.0457
father political	0.239	0.0924	0.454	0.131	0.103
% missing father's political	0.332	0.274	0.346	0.272	0.207
mother political	0.119	0.0209	0.256	0.0469	0.0266
% missing mother's political	0.360	0.294	0.388	0.284	0.216
Middle school admission					
pre-admission	0.388	0.00345	1	0	0
over-subscribed lottery	0.286	0.497	0	1	1
policy school	0.138	1	0.00123	0.240	1
Observations	14,699	2,027	5,709	4,199	1,007

Note: Column 2 describes policy school students, column 3 describes students who were pre-admitted, column 4 describes students who chose an over-subscribed school and assigned by lottery; column 5 describes students who were assigned by lottery to a policy school. % academic high school indicates the percentage of 9th grade graduates attending an academic high school; some other graduates attend vocational schools or stop going to schools. Non-academic evaluation consists of teacher and self-rated measures of four abilities, including civics, learning ability, atheistic ability, and practical ability. Imputed 9th grade score is constructed by assigning the highest grade of their cohort to the missing grade of direct admitted students who did not take the exam. Having city hukou means that a student is born in a city and enjoys the public goods of that city; it is often used as a measure of socioeconomic background. Father and mother political is a dummy that equals one if the parent is affiliated with any party; parental party affiliation is an indicator of better family background.

Table 2: Administrative Panel Data Description

Entire Sample							
year	# schools	# students	% no missing score	avrg # stu per school	% female	9th grade score	imputed 9th grade score
2,004	84	24400	90.5%	258.0	51.12%	17.28	17.28
2,005	61	21442	81.3%	343.3	47.15%	17.22	17.22
2,006	82	20941	94.0%	261.52	47.37%	14.88	14.88
2,007	78	20429	88.9%	258.1	47.84%	16.90	16.93
2,008	72	17150	100.0%	235.6	47.21%	17.13	17.20
2,009	73	24984	83.5%	339.2	47.13%	17.19	17.25
2,010	74	25753	87.1%	345.6	46.78%	16.98	17.08
2,011	70	25493	93.3%	341.5	48.44%	17.40	17.51
Policy Schools							
year	# schools	# students	% no missing score	avrg # stu per school	% female	9th grade score	imputed 9th grade score
2,004	8	3081	70.3%	274.6	51.56%	16.86	16.86
2,005	7	1862	88.1%	262.3	48.91%	15.38	15.38
2,006	11	2744	100.0%	233.3	45.63%	13.66	13.66
2,007	11	2691	92.5%	244.6	47.17%	16.09	16.27
2,008	11	2120	100.0%	192.5	48.44%	16.53	17.01
2,009	11	3636	71.4%	330.5	45.48%	17.28	17.69
2,010	11	3494	74.8%	317.6	46.35%	17.30	17.89
2,011	11	2813	91.7%	255.7	49.17%	17.64	18.27

Note: The table presents some summary statistics of the Middle School Graduation Exam administrative test score data set. In 2004 and 2005, one or two districts participated in a different set of exams as experiments for a measurement reform, resulting in some missing test score data for those two years.

Altogether, the eleven policy schools have around 3,000 students per cohort, which is about 14% of all students in the city. Comparing the last two columns across both panels, we can see that policy schools have lower 9th grade scores than other schools across all years.

To evaluate the change in quality of composition, I merge 6th grade baseline records in one district with lottery records for 2007-2010 and 2008-2011 cohorts. The matching rate is higher than 90%.[10] In a separate merge, I match the lottery records in 2005 through 2007 with corresponding MSGE records in 2008 through 2010.[11] The matching rate is higher than 70% across these cohorts.[12] More details on these merges and data set construction can be found in the appendix.

4 Does the Top Ten-Percent Quota Policy work?

4.1 Comparisons of Trends in Performance

To look at the overall trend in the school performances, I plot the MSGE average scores of the treatment group (policy schools) and that of the control group across years. The exercise is done on three samples: all schools, low-performing schools, and a balanced sample. For each sample, the numbers of schools in the treatment and control groups are (11, 98), (11, 37) and (4, 33). The entire panel data set has 11 policy schools and 98 other schools.[13] Low-performing schools are defined as schools with average 9th grade score below average in 2004.[14] The balanced sample includes 4 policy schools and 33 other schools, which have MSGE scores records for 2004 through 2011.

Figure 2 shows three sets of comparisons of normalized 9th grade scores for policy schools and the control schools, before and after the introduction of the policy. Each includes scatter plots of group averages and linear fitted lines. Two vertical lines in each figure indicate the timing of the policy change. Three figures differ in the sample they use to plot the scatter and linear fitted lines as defined in the previous paragraph. In all samples, the pre-trends of

[10] Non-perfect matching rate might due to changing names, typing errors in data, transferring, moving out of the city, etc.

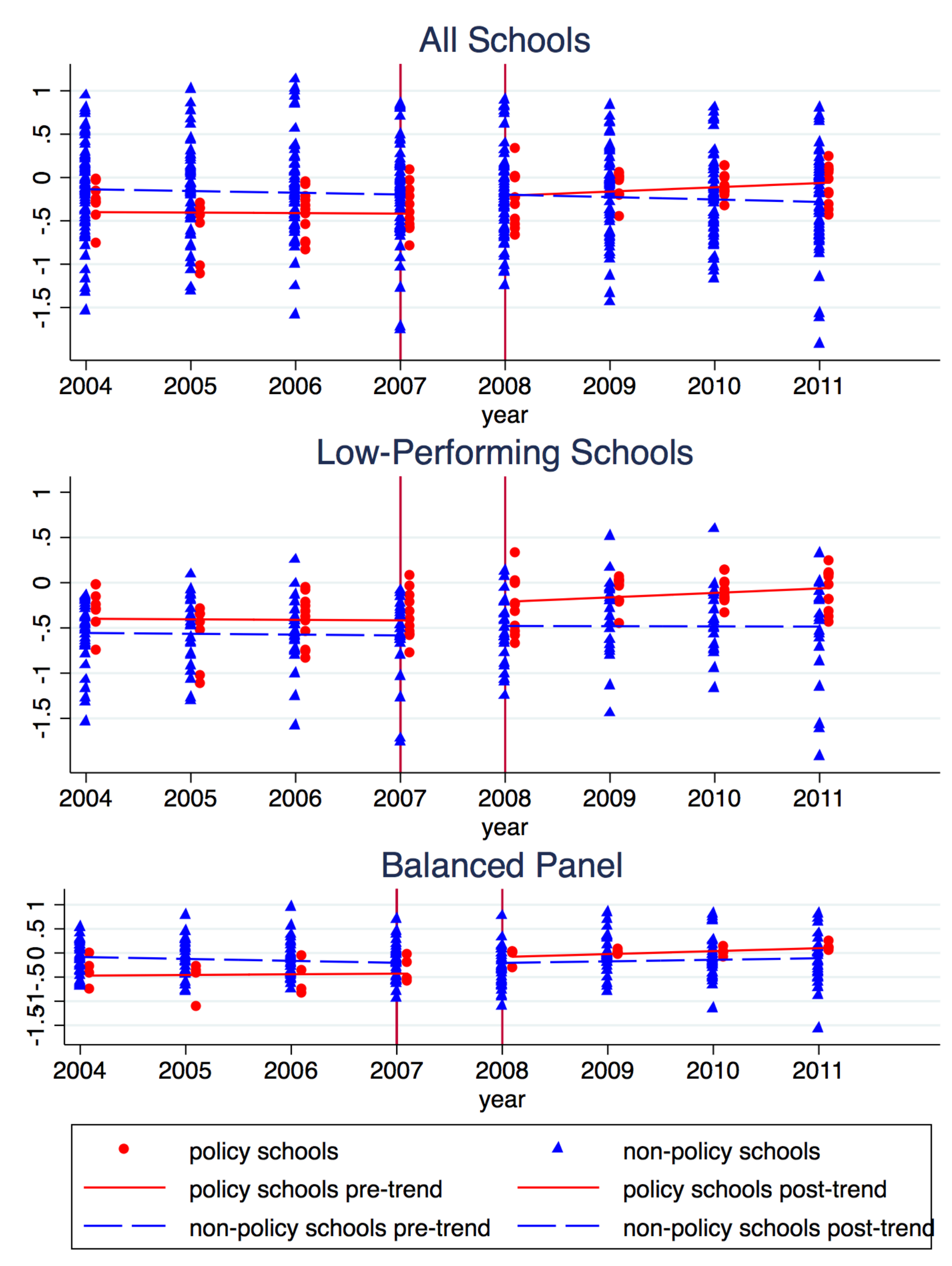
[11] The reason why I do not include 2008-2011 cohort in the lottery analysis is because in 2008 there was a change in the lottery school choice mechanism. Private schools were included in the choice set. Since then, there were two stages of lotteries, first for private schools and then for public schools. One needs to make four choices, one public schools, one private schools, whether to go on to public school lottery if win the private lottery, and whether to go on if lose.

[12] Reasons for non-matching observations include typos in records and students who transferred (in and out of the city).

[13] The sum is larger than the number of schools in any year, because some schools were shut down or merged into other schools while some new schools were opened in later years.

[14] Three policy schools did not have data on MSGE scores in 2004, and I assign them to the low-performing schools sample. The fact that they are low-performing is verified by comparing the performance of these schools with other policy schools.

Figure 2: Normalized 9th Grade Scores: Policy Schools and Other Schools



Notes: Two vertical lines in each figure indicate the timing of the policy change. Three figures differ in the sample they use to plot the scatter and linear fitted lines. The top figure uses the entire sample and compares the average academic performance of policy schools with all other schools from 2004 to 2011. The middle figure uses low-performing schools with average 9th grade score below medium in 2004 and plots their performances across different years. The bottom figure uses the balanced panel. Pre-trends of treatment and control groups are statistically parallel in all three plots.

treatment and control groups are statistically indifferent from each other in all three plots in Figure 2, which verifies the parallel trend assumption needed for a difference-in-differences estimation.

The top figure uses the entire sample. Before the policy, the performance gap between policy schools and other schools is around 0.5 standard deviations; after the policy, the performance gap was gradually closed.[15]

The best control group is not all other schools, but other low-performing schools without receiving the policy treatment. Panel B uses only the low-performing schools to make such comparison. Before the introduction of top ten-percent quota policy in 2007, policy schools' average academic performances were slightly higher than other low-performing schools; both sets of schools performed below average with normalized standardized scores at around -0.6 -0.5 and had an improving trend. After that, the policy schools improved more rapidly and increased their average scores to above average, while the other low-performing schools improved slightly but still had normalized average performance at around -0.5.

The bottom figure uses the balanced sample. Since previous years have more occurrences of missing data, we may worry if data is missing for policy schools when they coincidentally performed well or poorly in that pre-policy year, which would bias the treatment effect upward or downward. The balanced sample only includes 37 schools and may not give us accurate estimates of the impact. Therefore, this sample serves as a robustness check to ensure that the main results are not due to accidental biases from an unbalanced panel. For the schools with no missing data, before 2007 policy schools had a slight improving trend; after 2007 they made a parallel movement upward by around 0.5 standard deviations and ended at around average score.

4.2 Test for Selection in Treatment Status

Not all low-performing schools received the treatment of the top ten-percent quota policy.[16] Elite high schools were only willing to set aside a limited number of seats for unconditional acceptance of top Ten-Percent students from low-performing schools.

To verify that the policy schools were not chosen based on observed characteristics, I run a set of probit regressions to test whether pre-policy characteristics can predict policy treatment status.

$$D[policy]_{st} = \alpha + \mathbf{X_{st}}\beta + \epsilon_{st} \tag{1}$$

[15] Note that the scores are normalized every year, so if policy schools' scores improved over time, the other schools' normalized scores, by design, would fall.

[16] Recall from the policy background section that six schools have been treated since 2007 and five other schools since 2008.

$D[policy]_{st}$ is a dummy variable for policy treatment status. A *policy 2007* school is treated by the Top Ten-Percent Policy starting from 2007; a *policy any* school is any policy school. X_j represents a set of pre-policy characteristics, including normalized MSGE average performances, female percentages, numbers of students, and distance to a policy-targeted elite high school in the district.

Marginal effects are reported in Table 3. Columns (1),(2),(5) and (6) include all schools in the sample; columns (3),(4),(7) and (8) only use low-performing schools, defined by having below-average MSGE performance, in the sample. The two columns on the right are more relevant to the question of whether some were chosen over others while all were low-performing schools.

Panel A reports results on how pre-policy performance, school size and composition predict treatment status. The only statistically significant coefficients are those for MSGE score in columns (1) and (2), which means that schools treated by policy are generally lower-performing ones. What matters is whether the coefficients in columns (3) and (4) are statistically significant, which informs us that, among the low-performing schools, the education bureau did not choose the relatively better schools to benefit from the policy.

In Panel B, I test whether distance to elite high school explains policy treatment status. The data is constructed as documenting distances between pairs of elite high school and low-performing middle schools.[17] Results show that middle schools closer to the elite high schools within the district were as likely to be treated by the policy as the ones further away.

This exercise shows that no evidence was found for policy treatment selection on the observable school characteristics. If the education bureau assigned quotas to the policy schools in 2007 because these schools would grow faster after 2007 for unobserved reasons, regardless of the top ten-percent quota policy, the results would be biased. However, given the objective of equalizing school performance, the education bureau would not have chosen these schools if they were already improving by themselves. Instead, they would have chosen other schools that needed help.

[17]Since distance doesn't change across years, there are less observations in Panel B than in Panel A. For the districts that only one elite high school is targeted to be paired with low-performing middle school, low-performing schools only appear once in the regression sample, with the distance to that elite high school as the explanatory factor and whether it was treated by policy as the outcome variable. For some districts, there are two policy-targeted elite high schools, thus each middle school in those districts appears in the regression sample twice.

Table 3: Test for Policy School Treatment Status Selection

Panel A. Whether school performance and size predict policy treatment

	(1)	(2)	(3)	(4)
Sample	All Schools		Low Performing Schools	
VARIABLES	policy2007	policyany	policy2007	policyany
normalized MSGE score	-0.0884***	-0.105***	-0.0809	0.0655
	(0.0316)	(0.0335)	(0.0655)	(0.0768)
number of students	1.06e-05	3.31e-05	1.99e-06	0.000177
	(0.000103)	(0.000109)	(0.000202)	(0.000233)
percent female	0.415	0.308	0.641	0.286
	(0.254)	(0.301)	(0.465)	(0.562)
Obs (School by year)	208	272	110	144
Pseudo R2	0.129	0.0571	0.0348	0.0150

Panel B. Whether distance to the district elite high school predict policy treatment

	(5)	(6)	(7)	(8)
Sample	All Schools		Low Performing Schools	
VARIABLES	policy2007	policyany	policy2007	policyany
distance to elite high school	0.0141	0.0139	0.0241	0.0150
	(0.0263)	(0.0390)	(0.0450)	(0.0606)
distance squared	-0.00174	-0.00340	-0.00308	-0.00502
	(0.00242)	(0.00396)	(0.00404)	(0.00589)
Obs (Middle-High School Pairs)	130	130	71	71
Pseudo R2	0.0240	0.0495	0.0309	0.0786

This table reports probit regression results on whether pre-policy school characteristics can predict whether a school will be treated by the policy. The sample used for column (1)-(3) is all observations in the year 2004, 2005 and 2006; the sample used for column (4)-(6) is low-performing schools in those three years. Marginal effects and standard errors in parentheses are reported, *** p<0.01, ** p<0.05, * p<0.1. Other than having a lower performance than city average, all other specifications fail to predict policy treatment status.

4.3 Difference-in-Differences Estimation

The pre-trends of treatment and control groups are statistically parallel in all three plots in Figure 2, which verifies the key assumption needed for a difference-in-differences estimation. In addition, pre-policy school characteristics fail to predict treatment status of a school, which indicates that policy schools were not chosen based on these observable characteristics and can be deemed as an exogenous shock. To obtain a magnitude of the impact, I first use the following difference-in-differences specification:

$$Y_{it} = \alpha_i + \delta_t + \beta D(policy)_i * D(post)_t + \epsilon_{it} \tag{2}$$

Y_{it} stands for school i's performance measures, including normalized average 9th grade score, normalized ranking and percentage of students attending elite high schools in year t. Normalized average 9th grade score has a mean of 0 and a standard deviation of 1 for every year. Normalized ranking is constructed by dividing the school ranking (the higher the better) with total number of schools that year. I add school fixed effects α_i and year fixed effects δ_t to control for differences in exams and cohorts across years and fixed differences between schools. The coefficient of interest is β, which is the impact of the policy on the outcome variables Y_{it}. The dummy indicator $D(policy)_i * D(post)_t$ takes value of one when school i had the quota in year t. For the six policy schools assigned the quota in 2007, this dummy is one for year 2007 onward; for the five policy schools assigned the quota in 2008, it is one for year 2008 onward; for all other schools, it is zero for all years. Following Bertrand et al. (2004), ϵ_{it} is clustered at the school level to avoid over-rejection caused by serial correlated error terms.

I also use two alternative specifications, one replaces year fixed effect dummies with a linear time trend and the other replaces them with school-specific time trends. Estimated treatment effects do not vary much with year fixed effects or a linear time trend, but decrease slightly in magnitude and significance level with school-specific time trends. Note that such decreases are expected in a short panel, where school-specific time trends overlap with the variation in our variable of interest, postXpolicy. For similar reasons explained in the previous subsection, I conduct the analysis for three samples.

Results are shown in Table 4. Panel A, B, and C use all schools, low-performing schools, and a balanced panel, respectively, as the sample of analysis. Estimates using the balanced panel are the largest and those using only low-performing schools are the smallest. First, Panel A shows that, compared with all other schools, policy schools experienced 0.3 standard deviations of improvement in 9th grade school average scores when controlling for year fixed effects or a time trend. The estimated treatment effect is smaller, 0.185 standard deviations, when controlling for school-specific time trends. Their school ranking also rose by around 13 percentile compared to other schools, which would be moving up 10 ranks among 80 schools.

Panel B focuses on the low-performing schools and compares the policy schools with other low-performing schools. It shows that policy schools improved 0.256 standard deviations in 9th grade school average scores when controlling for year fixed effects. Since this is a short panel, school-specific time trends are highly correlated with the variable of interest (postXpolicy). Nevertheless, column (3) in Panel B shows that even when controlling for school-specific time trends, the estimate remains statistically significant at the ten percent level, although drops slightly to 0.185 standard deviations. On the other hand, the normalized ranking of policy schools doesn't seem to improve significantly when compared with other low-performing schools.

Panel C serves as a robustness check for whether the results are driven by schools with incomplete information, which were either closed down, newly opened, or took a different set of exams because of measurement reform experiment. For this set of regressions, I keep only 37 schools that remain in the data set throughout the eight years, among which there are four out of 11 policy schools.[18] Results are even more pronounced with this balanced panel than with previous two samples.

While the interacted dummy variable in Equation 2 captures the average treatment effect on the policy schools, it assumes a one-time shift in the performances instead of gradual changes, especially considering that the policy slowly increased the quota percentage during the first three years. As mentioned in the background section, the percentage of elite high school direct admission quota for policy schools varied across time, from the initial 2% in 2007, slowly raised to 6%, 8% in 2008 and 2009, and reached the promised 10% since 2010. The following alternative specification uses policy dosage instead of binary indicators and looks at whether the effect is responsive to the treatment intensity. $Policydosage_{it}$ equals zero for all schools before the policy started. It then takes the value of the effective quota percentage for policy schools and zero for all other schools. For example, it equals 2 for policy schools in 2007 and 10 for policy schools in 2010.

$$Y_{it} = \alpha_i + \delta_t + \beta policydosage_{it} + \epsilon_{it} \tag{3}$$

Table 5 shows that in all specifications, policy dosage has a positive and significant impact on school performance. Similar with patterns in Table 4, estimates using balanced panel are the largest and those using only low-performing schools are slightly smaller. Compared with all other schools, policy schools improved by 0.04 standard deviation for each quota

[18] The other seven policy schools had one year of missing score either in 2004 or in 2005 because of measurement reform experiment.

Table 4: Difference-in-differences: Treatment Effect on School Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Depend. Variable	Normalized 9th grade score			Normalized Ranking		
Panel A. All Schools						
postXpolicy	0.307***	0.314***	0.185*	0.130**	0.127**	0.0553
	(0.0858)	(0.0815)	(0.0939)	(0.0590)	(0.0569)	(0.0776)
Obs	597	597	597	597	597	597
R-squared	0.111	0.062	0.904	0.067	0.039	0.863
Number of schools	109	109	109	109	109	109
Panel B. Low-performing Schools						
postXpolicy	0.256**	0.273***	0.185*	0.0704	0.0758	0.0553
	(0.0987)	(0.0886)	(0.0929)	(0.0632)	(0.0589)	(0.0768)
Obs	287	287	287	287	287	287
R-squared	0.151	0.108	0.853	0.139	0.122	0.758
Number of schools	48	48	48	48	48	48
Panel C. Balanced sample						
postXpolicy	0.477***	0.472***	0.312**	0.287***	0.278***	0.187*
	(0.134)	(0.131)	(0.128)	(0.0824)	(0.0806)	(0.109)
Obs	296	296	296	296	296	296
R-squared	0.191	0.108	0.826	0.163	0.104	0.806
Number of schools	37	37	37	37	37	37
School FE	Y	Y	Y	Y	Y	Y
Year FE	Y	N	Y	N	Y	N
Time Trend	N	Y	Y	N	Y	Y
School Specific Time Trend	N	N	Y	N	N	Y

Notes: Robust standard errors clustered at the school level in parentheses. Significance level indicated by *** p<0.01, ** p< 0.05, * p< 0.1. The time span of the school panel data ranges from 2004 to 2011. "postXpolicy" equals one for policy schools during the years after the initial effective year. Each column has dependent variable listed on top row and control variables indicated at the bottom four rows. Panel B takes school with lower than average score in 2004; Panel C takes schools with observed performance in all eight years.

percentage treatment they receive. Because the policy dosage variable is highly correlated with the number of years treated, this result could be also come from the increase in the policy impact as time pass by, as students have more time to change their sorting behavior and work hard for the direct admission quota. Either way, this set of results show that the policy impact on school performance was gradual instead of a one-time shift.

4.4 Placebo Test

To verify that treatment effects only exist for the policy schools instead of for any low-performing schools, I run a falsification test of Equation 2 on the panel data without policy schools, and assign a placebo treatment status to other low-performing schools. Coefficients of the interaction dummies are reported in the Appendix Table A4. Placebo treatment effects are statistically insignificant in all except one specification: outcome variable being normalized ranking and controls including school and year fixed effects. This effect goes away after including a linear time trend, which indicates that the only statistically significant placebo treatment effect is because of variable construction. Normalized school ranking is constructed as ranking divided by total number of schools that year, and total number of schools decreases from 2004 to 2011 overall. Looking back at actual treatment effect estimates in Table 4, including the linear time trend did not change the results. Therefore, the falsification test suggests that treatment effects only apply to the policy schools.

4.5 Change in Distributions of Academic Performances

I plot the 9th grade score distributions for policy schools and non-policy schools before and after 2007 in Figure 3a. Solid lines are for before 2007 and dotted lines are for after 2007; red lines are for policy schools and black ones are for non-policy schools. To see the change in distributions of policy schools' normalized 9th grade scores, I compare the dotted and solid red lines. The Kolmogorov-Smirnov test rejects that two distributions are the same, indicates that the distribution significantly moved to the right and students in the policy schools improved their MSGE performance. A smaller proportion of students get a normalized score one standard deviation below mean and more get a score around the average. On the other hand, there was not a significant change in the distributions of non-policy schools before and after 2007.

To quantify the change in distributions, I run quantile regressions with the following specification:

$$Y_{it} = \alpha_i^p + \delta_t^p + \beta^p D(policy)_i * D(post)_t + \epsilon_{it}^p \tag{4}$$

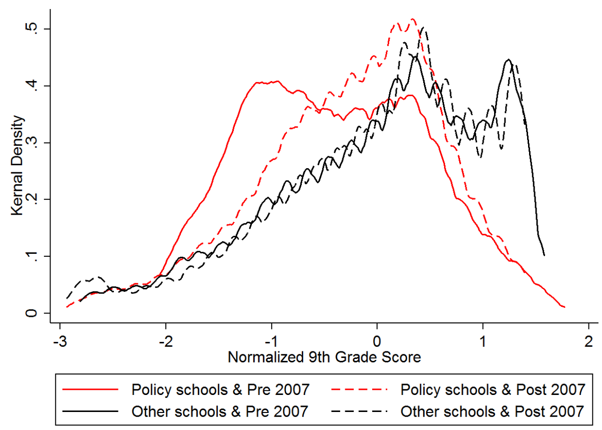
where $0 < p < 1$ indicates the proportion of the population having scores below the pth

Table 5: Policy Dosage Effect on School Performance and Ranking

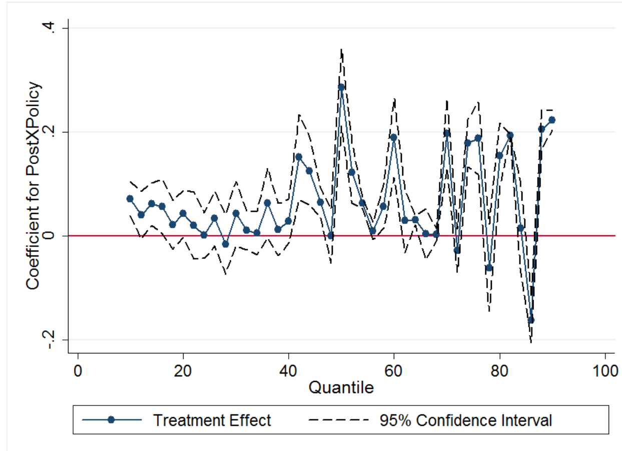
	(1)	(2)	(3)	(4)	(5)	(6)
Depend.Variable	Normalized 9th grade score			Normalized Ranking		
	Panel A. All Schools					
policy dosage	0.0383***	0.0411***	0.0570***	0.0165**	0.0176**	0.0251*
	(0.0104)	(0.00994)	(0.0165)	(0.00702)	(0.00683)	(0.0149)
Observations	597	597	597	597	597	597
R-squared	0.116	0.072	0.905	0.070	0.047	0.864
Number of schools	109	109	109	109	109	109
	Panel B. Low-performing Schools					
policy dosage	0.0335***	0.0365***	0.0570***	0.00959	0.0110	0.0251*
	(0.0120)	(0.0113)	(0.0163)	(0.00753)	(0.00730)	(0.0147)
Observations	287	287	287	287	287	287
R-squared	0.163	0.121	0.857	0.143	0.128	0.762
Number of schools	48	48	48	48	48	48
	Panel C. Balanced sample					
policy dosage	0.0564***	0.0583***	0.0701***	0.0337***	0.0345***	0.0434*
	(0.0158)	(0.0155)	(0.0255)	(0.00966)	(0.00954)	(0.0233)
Observations	296	296	296	296	296	296
R-squared	0.190	0.114	0.827	0.161	0.110	0.807
Number of schools	37	37	37	37	37	37
School FE	Y	Y	Y	Y	Y	Y
Year FE	Y	N	N	Y	N	N
Time Trend	N	Y	Y	N	Y	Y
School Specific Time Trend	N	N	Y	N	N	Y

Notes: Robust standard errors clustered at the school level in parentheses. Significance level indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The time span of the school panel data ranges from 2004 to 2011. "Policy dosage" equals zero for all school year combination where there were no quota policy and equals the percentage of admission quota for policy schools after the policy started. Each coefficient comes from a separate regression with dependent variable listed in the top row and control variables indicated in the bottom four rows. Panel B takes school with lower than average score in 2004; Panel C takes schools with observed performance in all eight years.

Figure 3: Distribution Changes: General Pattern and Performance Gap



(a) Changes in Scores



(b) Change in Performance Gap: Quantile Regression Estimates

Notes: The left figure plots distributions of individual level normalized 9th grade scores for policy school students and non-policy school students before and after 2007. The right figure plots estimates of the interaction term postXpolicy in quantile regressions from 10th percentile to 90th percentile with a 2 percentile increment.

percentile. Estimates for different quantiles are plotted in Figure 3b, with 95% confidence intervals around them. The estimates are very volatile, especially for the top percentiles. This may be due to the discrete nature of translated letter grades. For the lower end of the distribution, there is not much change in the performance; for the middle part, we see positive changes for some quantiles and insignificant changes for others; for the top end, estimates are very volatile. The quantile regression results show a noisy picture of changes in the distribution of performance gap and do not offer insights on changes in selection and in school quality.

5 Underlying Mechanisms

5.1 A Conceptual Framework

I document in the previous section that the gap between the policy impacted low-performing middle schools and the others has been shrinking after the policy. The top ten-percent quota policy might work through three interacting channels. First, it increases the chance to be admitted into a top high school for students attending the policy schools. Therefore it encourages 6th grade elementary school graduates, especially the high performing ones, to choose the policy impacted schools. I refer to this as a *composition effect*. Secondly, if the first channel exists, having better peers may bring a positive *peer effect* to students attending the policy schools (Sacerdote, 2001; Zimmerman, 2003; Ding and Lehrer, 2007; Sacerdote, 2011). Thirdly, competition for the top ten-percent within these schools may also help the schools improve their average performance. I refer to this as a *tournament effect*. Top-performing students in these schools could be affected by both peer effects and a tournament effect; while low-performing students are most likely affected by peer effects but not the tournament effect, since they know that placing at the top ten-percent is unlikely.

To understand what drives the impact, I modify the theoretical framework from Cullen et al. (2013) to help illustrate the mechanisms and motivate the empirical tests for changes in composition. For simplicity, I abstract away from heterogeneous neighborhood characteristics, transportation cost and tuition and assume they are identical. This assumption is plausible because of the following reasons: the public middle schools one can choose from are nearby and in the same district; public transportation is cheap and convenient; this city has rather low crime across all districts; there is little ethnic variation across neighborhoods; public middle school tuition is regulated (around USD50 per semester).

The decision of school choice by parents is driven by the expected impact schools will have on their children's future prospects. In China, returns to elite college attendance is high (Li et al., 2012) and college entrance exam is fiercely competitive. Therefore, I set the objective of school choice to maximize college entrance exam performance, Y_{ij}. It depends on student's innate ability, a_i, learning progress before entering high school, y_{ij}, and probability of attending an elite high school, $prob_{ij}$. I assume learning progress in middle school increases with one's ability a_i, peer quality q_j, and school characteristics and learning atmosphere γ_j. Probability of attending an elite high school is included because elite high schools have better school quality and peers, and a significantly higher percentages of students attending an elite college.

$$max_j \;\; Y_{ij} = Y_{ij}\{a_i,\; y_{ij}(a_i,\; \text{peer}q_j,\; \gamma_j\},\; prob_{ij}(y_{ij})) \;\; \text{for } j = 1, 2...n \tag{5}$$

Without loss of generality, assume that families face two public middle school choices, $j = 1, 2$, and school 1 has a better peer quality $q_1 > q_2$, and $y_{i1} > y_{i2}$. The policy guarantees school 2 elite high school admission for top ten-percent and changes $p_{i2}(y_{i2}, \kappa_{i2})$, where κ_{i2} indicates the probability of placing at top ten-percent in school 2.

$$max_j \;\; Y_{ij} = Y_{ij}\{a_i,\; y_{ij}(a_i,\; \text{peer } q_j,\; \gamma_j\},\; prob_{ij}(y_{ij}, \kappa_{ij})) \;\; \text{for } j = 1, 2...n \tag{6}$$

Now, for a set of students, the policy changes the comparison between probability of attending an elite high school in school 1 and school 2, $prob_{i1}(y_{i1}) < prob_{i2}(y_{i2}, \kappa_{i2})$. These students switch to school 2 when the benefit from higher probability of attending an elite high school is larger than the cost of having a lower peer quality, $\frac{\partial Y_{ij}}{\partial prob_{ij}} > \frac{\partial Y_{ij}}{\partial y_{ij}} \frac{\partial y_{ij}}{\partial q_j}$.

On the other hand, for highest performing students, the probability of getting into an elite high school is $prob_{ij}(y_{ij}) = 1$. Because they can get into an elite high school with or without top-ten percent policy, they have no incentive to bear the costs of downgrading to school 2 with lower peer quality. For lowest performing students, the probability of getting into an elite high school is $prob_{ij}(y_{ij}) = 0$ regardless of their school choice, because they cannot outperform 65% in the city MSGE, nor can they outperform 90% in low-performing schools for the top ten-percent quota policy.

To sum up, the introduction of top ten-percent quota policy changes $prob_{ij}$ for some students i and policy schools, and it alters some middle to high performing students' school choices. This affects policy schools in two dimensions: change in peer composition q_j and change in value-added y_{ij} caused by change in peer composition q_j and competitive learning atmosphere γ_j.

5.2 Using Lottery Records to Tease Out Mechanisms

Empirically, the lottery middle school assignment provides the opportunity to separately look at changes in composition and in value-added. Recall that each elementary school is assigned a certain number of seats to around three nearby middle schools. Each 6th grader chooses only one middle school. When a middle school gets oversubscribed in that elementary school, a lottery randomly determines school assignment.

Policy schools rarely gets oversubscribed in any elementary schools. Therefore, there are only two types of 6th graders who attend policy middle schools: those who voluntarily enroll in policy schools, and those who choose some other school but lose the lottery and get randomly assigned to a policy school. The composition effect mainly captures changes in the ability of the first type, "voluntary enrollees". I analyze the baseline scores of sixth graders who voluntarily chose a policy school and test if there is evidence of strategic switching after the policy by high-performing students.

For the second type, "lottery losers", they did not strategically change their school choice after the policy, yet they still may benefit from attending a policy school compared to previous cohorts, because now they have better peers. I exploit the random lottery assignment to evaluate differences in value-added between the policy schools and the chosen schools, before and after the policy, to see how the value-added gap was changed by the policy-induced peer composition change. High-performing lottery losers may benefit more from good peers and a tournament incentive than lower-performing lottery losers, therefore I conducted instrumental quantile treatment effect analysis to detect differential treatment effects along the distribution.

5.3 Change in Composition

The top ten-percent quota policy increases the expected return of attending the policy schools, changes the trade-off of school choice, and therefore incentivizes some sixth graders to strategically choose the policy schools. As illustrated in the conceptual framework, switching to a policy school involves trade-offs. A student may benefit from the expectation that he/she will be in the top ten-percent of the graduating class, but at the same time may suffer from a lower peer quality. Therefore, it is unclear whether students would respond to the policy by changing their school choice.

The composition effect has two dimensions, quantity and quality. Higher percentages of sixth graders may voluntarily choose the policy; among those sixth graders who choose policy schools, the average baseline performance may be higher than before. First, to test the quantity dimension of demand change, I use four years of lottery choice records from 2005 to 2008 and compute the percentages of students choosing the policy-impacted schools for each elementary school. I run the following regression to check if there was a significant increase in the percentages of students choosing the policy schools.

$$Perc(S_{jt}) = \theta_j + \beta D(post)_t + \epsilon_{jt} \tag{7}$$

$Perc(S_{jt})$ is the percentage of 6th graders choosing policy school in elementary school j in year t; θ_j stands for elementary school fixed effect; $D(post)_t$ is a dummy for post policy years. β is the coefficient of interest, which indicates how many more students, in percentage of their elementary schools, chose policy schools after the change. Table 6 shows that the percentages of students who chose policy schools did not significantly change after the policy announcement. Therefore, there was not a significant increase in the quantity of students voluntarily enrolled in policy schools.

To test the change in composition quality, I analyze the incoming students' sixth grade scores. Since I only have the sixth grade scores for one district in 2007-2010 and 2008-2011 cohorts, I compare the change in scores of elementary students choosing the policy school announced in 2008 for that district. To verify the generalizability of results from this district, I compare the differences in school characteristics and policy impacts of this district and other districts. Results show that the comparison between policy and non-policy schools and policy impacts are not statistically different from those of other districts. The conditional logistic regression with elementary school fixed effects is specified as below:

$$Pr(S_i = 1 | \mathbf{x_i}, D(post)_t) = F(\alpha_i + \mathbf{x_i}\beta_1 + \beta_2 D(post)_t + D(post)_t * \mathbf{x_i}\beta_3) \tag{8}$$

where F is the cumulative logistic distribution, $F(z) = \frac{exp(z)}{1+exp(z)}$. S_i equals to 1 if student i

Table 6: Composition effect
Change in Percentages of Sixth Graders Choosing Policy Schools

	(1)	(2)	(3)	(4)
Sample	All Elementary schools		Elementary schools with Non-Zero Percentage Choosing Policy Schools	
Dependent Variable	% Choosing 2007 Policy Schools	% Choosing Any Policy School	% of Choosing 2007 Policy Schools	% of Choosing Any Policy School
After 2007	-0.0131		-0.0103	
	[0.0107]		[0.0203]	
After 2008		-0.00745		-0.00618
		[0.0133]		[0.0170]
Elementary School FE	Y	Y	Y	Y
Observations	979	979	243	430
R-squared	0.617	0.722	0.931	0.907

Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1.
Sample includes 2005-2008 four cohorts of elementary school graduates' school choice. *After 2007* equals to 1 if the year is after 2007 and *After 2008* equals to 1 if the year is after 2008.

Table 7: Are elementary students with high ability and high SES more likely to choose a policy school after the policy?

Dependent Variable	Indicator for choosing a policy school				
	(1)	(2)	(3)	(4)	(5)
postXmath	0.049**				
	(0.0174)				
postXreading		0.0217			
		(.0297)			
postXfather politics			0.0055**		
			(0.00203)		
postXmother politics				0.0095**	
				(0.0045)	
postXfemale					0.017
					(0.015)
Observations	4,151	4,151	2,746	2,499	4,570

Each column reports the marginal effect of the interaction term in a fixed effect logit regression controlling for elementary school fixed effects, corresponding to Equation 8. Robust standard errors were allowed to clustered at the elementary school level. Significance level is indicated by *** p<0.01, ** p<0.05, * p<0.1. Father and mother political statuses are dummy variables for any parental political affiliation, which are proxies for high socioeconomic status. The table shows positive and statistically significant changes in family background and baseline academic performance for students who voluntarily enroll in policy schools.

Table 8: Conditional logit estimates of choosing a policy school: Who Are Switching?

Dependent Variable	Indicator for choosing a policy school	
	(1)	(2)
6th grade score of	math	reading+math
post X top1%	-0.381	-0.159
post X top5%	0.877	0.096
post X top10%	1.341**	1.580**
post X top20%	1.130**	0.702
post X top30%	0.794	0.635
post X top40%	1.344**	0.925
post X top50%	0.589	1.611***
Elementary School FE	Y	Y
Observations	4,570	4,570

Robust standard errors in parentheses, clustered at the elementary school level. Significance level:*** p<0.01, ** p<0.05, * p<0.1

This table reports coefficients from fixed-effects logit regressions. The coefficients can not be interpreted as magnitudes of changes in likelihood of choosing a policy school. Significance level and signs of the coefficients provides information on who are switching. Sample used is merged lottery records with sixth grade scores from one district in 2007 and 2008. In 2008, one middle school was impacted by the policy in this district. The dependent variable equals to one if a sixth grader chose the 2008 policy school. "post" equals to 1 for the year of 2008. "top x%" equals to 1 if a sixth grade's score is in the top x% in their cohort, all "top x%" categories are mutually exclusive and all regressions include dummies for each category.

chose a 2008 policy school in year t; α_i stands for elementary school fixed effect; $D(post)_{it}$ is a dummy for year 2008 and $\mathbf{x_{it}}$ is a vector of student i's 6th grade math score and reading (Chinese) score. The coefficient of interest is β_3, which tells us how the ability of incoming students in the policy schools changed after the policy was announced.

Table 7 shows the regression results for Equation 8. Marginal effects on the interaction terms between post policy and student characteristics are reported. The interaction terms of normalized sixth grade math score and dummies for parental political affiliations are highly significant and positive. It indicates that students who have better math score and better socioeconomic background are more likely to choose policy schools after the introduction of the ten-percent quota policy.

Although we now know that students with higher baseline math scores switch to choose policy schools, we do not know if it is because of a heavy downgrade by a few top students, or a small downgrade by many medium-ranking students. By analyzing the trade-off faced by these students, the highest-performing students have less incentive to downgrade than the second-tier-ranking students because the highest-performing students are confident about getting elite high school admission, even without the policy guarantee. Therefore, they have no benefit and no incentive to pay the cost of having lower quality peers by switching to a policy school.

To test this hypothesis and understand the characteristics of the students who switch, I use the same regression specification as Equation 8 and replace the actual baseline score with the percentile category dummies that are mutually exclusive. For example, if the top 5% takes value of 1, it means that a student has a baseline score between 1% to 5%. Estimates in Table 8 confirm the hypothesis that the highest-performing 6th graders were not more likely to switch, while sixth graders with above-average math scores showed statistically significant switching patterns.

5.4 Change in Value-added

In the previous subsection, I show that the top ten-percent quota policy attracts students with better math scores to voluntarily enroll in the treated lower-performing schools. One of the concerns is that the policy may have improved school average performance only by redistributing students, without changing the school quality at all. After the policy, there are three types of students in the policy schools: strategic switchers, students who would have chosen the policy schools anyway and students who choose an over-subscribed school but lose the lottery and get randomly assigned to a policy school. The previous subsection analyzes school choices by the first two types of students; this subsection takes the last type of students, lottery losers, and compare them with the lottery winners to estimate value-added gaps between policy schools and over-subscribed schools.

Change in value-added can come from several channels. First, policy schools may bring higher value-added to the cohorts entering middle school after the policy than previous cohorts, because of better peers attracted by the policy. In addition, competition to place at the top ten-percent of the graduating class encourages students to exert more effort. Especially since policy schools determine ranking by three-year accumulated performance, students need to work consistently throughout the three years. High-performing students may benefit more from tournament incentive than lower-performing ones.

Different from school performance, school quality is usually difficult to measure because of endogenous selection. Better students often sort into schools with better reputation, which makes it hard to disentangle whether the higher performance in these schools comes from incoming students' ability or school quality. The random lottery assignment allows me to use it as an instrument to evaluate differences in value added with the Local Average Treatment Effect (LATE) model by Imbens and Angrist (1994). These LATE estimates provide measurements of the school quality gap before and after the policy introduction and therefore enable us to see changes in value-added.

The instrumental analysis uses the sample of students who chose an over-subscribed school and randomly assigned to their choice schools or an under-subscribed policy school. Let $Y_i(1)$ be student i's potential test score if she attends a policy school, and let $Y_i(0)$ be her test score if she attends her choice school. D_i indicates the "treatment", policy school attendance, and Z_i is an indicator for lottery outcome. Let $D_i(1)$ and $D_i(0)$ denote potential treatment status as a function of Z_i. The following assumptions are needed for LATE framework:

1. Independence and Exclusion Restriction: $(Y_i(1); Y_i(0); D_i(1); D_i(0))$ is independent of Z_i.

2. Nontrivial First Stage: $Pr(Z_i) = E[D_i l Z_i]$.

3. Monotonicity: $D_i(1) > D_i(0)$ for all i.

The first assumption requires that lotteries are random and do not affect test scores through any channel but policy school attendance. The second assumption requires that lottery losers are more likely to attend policy schools on average. Monotonicity assumption requires that winning the lottery does not encourage any student to attend a policy school instead of the choice school. All three assumptions are satisfied in this study's sample.

First, to verify the random lottery school assignment, I use a probit model to regress students' pre-lottery characteristics on their lottery outcomes. If the lotteries are random, pre-lottery characteristics should not be able to predict the lottery outcome. I include a group of dummy variables to control for the lottery choice and the elementary school attended, since lotteries happen at the elementary school level. The regression equation is

$$Z_{ic} = \alpha_c + \alpha_1 X_i + \epsilon_{ic} \tag{9}$$

Z_{ic} equals 1 if the student i lost the lottery. X_i represents pre-lottery characteristics including gender, city residency (hukou), parental political status, Chinese and math scores in elementary school graduation exam. Parental political status and elementary school graduation exam scores are only available for 2007, not for 2005 or 2006.

Table 9 reports the marginal effects of X_i and verifies the lottery randomness for 2005-2008 and 2007-2010 cohorts. Pre-lottery characteristics cannot predict lottery outcomes for these two cohorts. Conditional on taking the lottery, winning the lottery is an exogenous event that sends students who made the same lottery choice into different middle schools in these two years. This gives us a device to peel away the endogenous school choice problem and compare the value added of the policy impacted ones with the not impacted ones. Results don't change if I put more explanatory variables, for example sixth grade scores and family background, in 2007-2010 cohort. I do not include 2006-2009 cohort in the 2SLS and instrumental quantile treatment effect analysis due to missing data.

This finding is of central importance for this paper because one of the concerns for the policy is that it may have improved school average performance only by redistributing students and may have not improved the school quality of these low-performing schools at all. The analysis here shows that before the policy, among the students who chose an over-subscribed school, lottery losers who were randomly assigned to a low-performing policy school performed worse than their elementary school classmates who won the lottery and assigned to an over-subscribed school in the middle school graduation exams. After the policy, their average outcomes were about the same. To verify the predictive power of losing a lottery on attending a policy school, I run the first stage probit regression, controlling for lottery choice group fixed effects and available student characteristics. Table 10 reports the marginal effects and the Pseudo R squared, which indicates that using lottery outcomes to instrument for policy school attendance is nontrivial.

$$D_i = \kappa_c + \alpha_1 Z_i + X_i \alpha_2 + \mu_{ic} \tag{10}$$

After verifying the assumptions needed for the LATE framework, students can be divided into three types: always takers, who attend regardless of the lottery outcomes ($D_i(1) = D_i(0) = 1$), never takers, who never attend policy schools ($D_i(1) = D_i(0) = 0$), and compliers, who are

Table 9: Lottery Randomness Verification

Dependent Variable: Winning a Lottery			
Year	2005	2007	2007
female	0.0254	-0.00228	0.0115
	(0.0180)	(0.0151)	(0.0175)
hukou		-0.0208	0.0157
		(0.0176)	(0.0204)
father political status			0.0381
			(0.0296)
mother political status			0.0606
			(0.0466)
Lottery fixed effects	Y	Y	Y
Observations	2,558	2,747	2,747

Robust standard errors in parentheses clustered at the lottery level, *** p<0.01, ** p<0.05. Lottery fixed effects are interacted dummy variables with elementary schools and middle school choice. "Hukou" means whether the student has city residency or not. "Father (mother) political status" indicates whether the student's parent has a party affiliation or not.

induced to attend by receiving offers ($D_i(1) > D_i(0)$). The instrumental variables methods can consistently estimate LATE, the average treatment effect for compliers (Imbens and Angrist, 1994):

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0)] \tag{11}$$

I use a two-stage least squares (2SLS) method to estimate LATE. The regression specification is

$$Y_i = \theta_c + \beta_1 D_i + X_i\beta_2 + \epsilon_{ic} \tag{12}$$

where Y_i is the normalized middle school graduation exam score for student i, D_i is a dummy variable indicating policy school attendance, and X_i is a set of elementary school and lottery choice indicators and student characteristics. The first stage is specified in Equation 10.

Table 11 reports the results of the 2SLS. The observations are less than the total number of students participating in school lottery assignment, because only lotteries involving a policy school have variations in first stage outcome, i.e. policy school attendance. The comparison between 2SLS estimates for two cohorts shows that the gap between value-added by policy schools and oversubscribed schools were closed by the top ten-percent quota

Table 10: First Stage: Use losing a lottery to instrument for policy school attendance

Dependent Variable: attending a policy school		
Cohort	2005-2008	2007-2010
losing a lottery	0.244***	0.393***
	(0.0237)	(0.0116)
female	0.0245	0.0001
	(0.0246)	(0.0160)
hukou		0.0202
		(0.0192)
Lottery fixed effects	Y	Y
Observations	1,130	2,066
Pseudo R2	0.249	0.361

Robust standard errors in parentheses clustered at the lottery level, *** p<0.01, ** p<0.05, * p<0.1. Lottery fixed effects are interacted dummy variables with elementary schools and middle school choice. hukou means whether the student has city residency or not.

Table 11: Average Change in School Quality Gap: 2SLS Results

Dependent Variable: Normalized 9th grade score		
	(1)	(2)
Cohort	2005-2008	2007-2010
attending a policy school	-0.302***	-0.0805
	(0.0641)	(0.0495)
Obs	1,130	2,066

Robust standard errors in parentheses, *** p<0.01. This table reports results of 2SLS regressions on two cohorts, before and after policy respectively. Each coefficient shows the average school value-added gap between policy schools and over-subscribed schools.

policy.

As discussed at the beginning of this subsection, there are several channels that the policy could have helped close the value-added gap. Peer effects could help improve value-added to all students; a tournament effect could increase value-added to high-performing students who have a chance to compete for the top-ten percent quota. Therefore, if we observe an value-added improvement to the low-performing students, that would be evidence that peer effects were at work.

The LATE masks the heterogeneous treatment effects across students with different academic performance. To see whether peer effects were at work, it is important to estimate the treatment effect across the distribution. Here I use instrumental Quantile Treatment Effect analysis (Abadie et al., 2002) to analyze the gaps in distributions of value-added between policy schools and chosen middle schools. Similar with 2SLS, this exercise is carried out for two cohorts, one before the policy and one after the policy.

The instrumental QTE is "an Abadie-type weighting estimator of the causal effect of treatment on quantiles for compliers" (Angrist and Pischke, 2008). The relationship between the QTE estimator and quantile regression is analogous to that between 2SLS and OLS. The set up for QTE estimation is described as following. A scalar outcome variable Y is students' normalized ninth grade middle school graduation scores. D is a binary treatment indicator for policy school attendance, and Z is a binary instrument for losing a lottery. X stands for a set of dummies for elementary school and school choice, and other student characteristics.

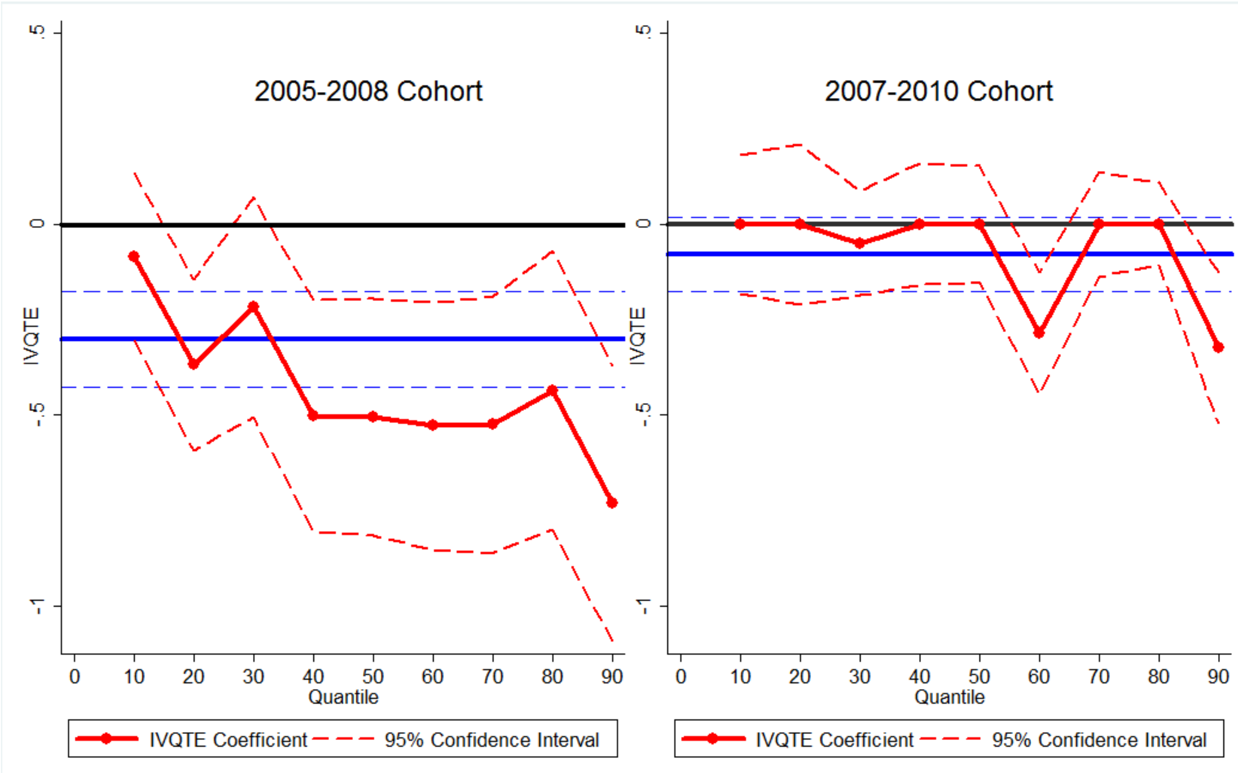
$$Q_\theta(Y_i|X_i, D_i, D_{1i} > D_{0i}) = \alpha_\theta D_i + X_i' \beta_\theta \tag{13}$$

where $Q_\theta(Y_i|X_i, D_i, D_{1i} > D_{0i})$ denotes the θ-quantile of 9th grade score conditional on control variables X_i and policy school attendance D_i for compliers.

The instrumental QTE estimation results for Equation 13 are plotted in Figure 4 for 2005-2008 and 2007-2010 cohorts. Looking at the 2005-2008 cohort, we see a larger value-added gap for high-performing students. This is consistent with previous findings on nonlinear peer effects in Chinese secondary schools by Ding and Lehrer (2007). High-performing students benefit more from attending over-subscribed schools, which provide better peer quality. For lottery losers who chose another middle school and were randomly assigned to a policy school in 2005, before the ten-percent quota policy, the gap between the distribution of 9th grade graduation scores and the lottery winners' distribution was significantly negative for seven out of nine deciles. After the policy, however, most estimates are insignificantly different from zero, which indicates that policy schools and over-subscribed schools then have similar value-added. For 60th and 90th percentiles, the value-added by policy schools for 2007-2010 cohort were still lower than over-subscribed schools, but less so than the 2005-2008 cohort.

Tournament effect is unlikely for the lower-achieving students, because they are far from the threshold of top ten-percent and have no incentive to insert extra effort to respond to the policy. Thus, the narrowed value-added gap for lower range of the distribution suggests that there were positive peer effects. High-performing students also benefited from the policy in terms of value-added, which could be a mixture of tournament and peer effects. In fact, if

Figure 4: Changes in Value-added Gaps across Quantiles



Notes: These two figures plot estimates of 2SLS and instrumental quantile regressions from 10 percentile to 90 percentile with a 10 percentile increment for two cohorts, one before the ten-percent quota policy and one after the policy. 2SLS estimates and their 95% confidence intervals are shown by the horizontal lines in blue; those for IVQTE analysis are in red. These estimates indicate the value-added gap between a policy school and an over-subscribed school at different parts of distribution. Before the policy, seven out of nine IVQTE estimates were significantly negative, which indicates that value-added by policy schools was significantly lower than the oversubscribed schools. After the policy, all coeficcients are indifferent from zero or less negative than before. This shows that the policy narrowed the value-added gap between the low-performing policy schools and the over-subscribed schools all across the distribution.

we move the 2005-2008 cohort value-added estimates upward and compare that with 2007-2010 cohort's estimates, the top thirty percentiles were moved upward a little more than the other deciles, which may be evidence for tournament effect, or may come from a nonlinear peer effect. The instrumental QTE analysis does not provide conclusive evidence for the tournament effect, but supports that peer effects are likely at work.

Finally, results here are derived from analyzing students who chose an over-subscribed school and comparing the value-added gap experienced by pre-policy lottery losers and post-policy lottery losers. The school quality change experienced by those who voluntarily enrolled in the policy schools (hereafter voluntary enrollees) may be different. Given that voluntary enrollees have lower baseline performance than those who choose an over-subscribed school, they may benefit less from the peer quality improvement after policy-induced strategic sorting because high-achieving students usually benefit more from having high quality peers. However, as shown in the IVQTE analysis, even the lower distribution also experienced a significant improvement in value-added. In addition, the voluntary enrollees may experience higher value-added than lottery loser enrollees because of some unobservable benefit that prompted them to select into the policy schools. Therefore, the value-added improvement experienced by voluntary enrollees is probably comparable, if not higher, than the results shown above.

6 Discussion and Conclusion

This paper evaluates a current education policy in Changsha that aims at improving low performing middle schools. By guaranteeing seats in elite high schools to the top ten-percent of students attending the policy schools, it seeks to improve the desirability of these schools, attract better incoming students, bring positive peer effects and encourage students to compete for the top ten-percent. I document that the policy helped narrow the gap between low-performing policy schools and the other schools. Sixth graders with better math scores and socioeconomic status are more likely to voluntarily enroll in the policy schools, which improved the incoming students' quality. Findings here on changes in sorting patterns complement previous research on top x-percent policies in the U.S. (Cullen et al., 2013) that students do respond to the incentive of relative grading. In addition, this study uses instrumental quantile treatment effect estimates and further shows that school value-added also improved across different deciles. It suggests that the top ten-percent quota policy was successful in equalizing the performance and school quality between the low-performing schools and over-subscribed schools.

Although this study shows that the value-added changed, it cannot perfectly disentangle different channels. Teacher and parent behaviors may changed in response to this policy as well, in addition to peer effects and tournament effects. Therefore, the observed improvement in school value-added may partially due to teacher and parental morale. However, due to data limitation on parental involvement and teacher effort, this study cannot directly test those channels. Despite of this limitation, results here advances previous findings on top-x percent policy by estimating the the value-added gap before and after the policy introduction, given the unique lottery school assignment system. To what extent this result would apply to other contexts hinges on how much the switchers interact with the other students. Further research is needed to understand whether strategic sorting induced by top x-percent policy in the United States also brought positive spillover effects and potentially changes in racial attitudes.[19]

The nature of standardized test scores prevents us from having an absolute measure of how student performance in the whole city changed after the policy.[20] Instead of improving low-performing schools while keeping over-subscribed schools the same, the previous performance gap and value-added gap may have been largely driven by peer composition, and the top ten-percent quota policy brought changes in sorting patterns which closed both gaps. However, it is unlikely that school quality of the over-subscribed schools declined as much as the gain in school quality of policy schools, since the switchers do not seem to be the top students and probably do not affect overall peer quality in the over-subscribed schools. Consistent with findings in Cullen et al. (2013), the magnitude of sorting was small compared to the overall size of student population, as shown in subsection 5.3. In addition, policy schools were still mostly under-subscribed, suggesting that the policy did not completely flip the compositions of low-performing schools and over-subscribed schools.

The paper provides implications on how a government mandate on school admission process influences sorting behavior, student outcomes, and school outcomes. This study, along with previous studies, find that students with low socioeconomic status are less likely to choose a good school, which suggests that school-choice program is not a panacea and may not suffice to close the income achievement gap. Policies that change sorting patterns and improve low-performing schools' quality need further exploration.

References

Abadie, Alberto, Joshua Angrist, and Guido Imbens, "Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings," *Econometrica*, 2002, *70* (1), 91–117.

Abdulkadiroğlu, Atila, Joshua D Angrist, Susan M Dynarski, Thomas J Kane, and Parag A Pathak, "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots," *Quarterly Journal of Economics*, 2011, *126* (2), 699–748.

Ajayi, Kehinde F, "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," *Working Paper*, 2011.

[19] Carrell et al. (2015) find that exposure to more intergroup contact changes white male students' racial attitudes toward African Americans.

[20] A possible way to measure general equilibrium effects is using the standardized provincial college admission exam, as this would allow comparison between Changsha and other cities. However, there is a trend for students from other cities moving into Changsha, the provincial capital, to secure a better high school education. Therefore, even if Changsha's college admission exam performance improves, we cannot disentangle between the positive general equilibrium effects by the top ten-percent quota policy and the sorting effects of students moving into Changsha's high schools from other cities.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2008.

_ , **Parag A Pathak, and Christopher R Walters**, "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 2013, *5* (4), 1–27.

_ , **Sarah R Cohodes, Susan M Dynarski, Parag A Pathak, and Christopher R Walters**, "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics*, 2016, *34* (2), 377–415.

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer, "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment," *American Economic Review*, 2002, pp. 1535–1558.

Banerjee, Abhijit V, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani, "Pitfalls of Participatory Programs: Evidence from a randomized evaluation in education in India," *American Economic Journal: Economic Policy*, 2010, pp. 1–30.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly journal of economics*, 2004, *119* (1), 249–275.

Butler, JS, Douglas A Carr, Eugenia F Toma, and Ron Zimmer, "Choice in a world of new school types," *Journal of Policy Analysis and Management*, 2013, *32* (4), 785–806.

Calabrese, Stephen M, Dennis N Epple, and Richard E Romano, "Inefficiencies from metropolitan political and fiscal decentralization: Failures of Tiebout competition," *Review of Economic Studies*, 2012, *79* (3), 1081–1111.

Carrell, Scott E, Mark Hoekstra, and James E West, "The impact of intergroup contact on racial attitudes and revealed preferences," Technical Report, National Bureau of Economic Research 2015.

Chakrabarti, Rajashri, "Can increasing private school participation and monetary loss in a voucher program affect public school performance? Evidence from Milwaukee," *Journal of Public Economics*, 2008, *92* (5), 1371–1393.

_ , "Do vouchers lead to sorting under random private school selection? Evidence from the Milwaukee voucher program," *Economics of Education Review*, 2013, *34*, 191–218.

Cortes, Kalena E and Lei Zhang, "The Incentive Effects of the Top 10% Plan," *Working Paper*, 2011.

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

— , **Mark C Long, and Randall Reback**, "Jockeying for position: Strategic high school choice under Texas' top ten percent plan," *Journal of Public Economics*, 2013, *97*, 32–48.

Deming, David J, "Better schools, less crime?," *Quarterly Journal of Economics*, 2011, p. qjr036.

— , **Justine S Hastings, Thomas J Kane, and Douglas O Staiger**, "School Choice, School Quality, and Postsecondary Attainment," *American Economic Review*, 2014, *104* (3), 991–1013.

Ding, Weili and Steven F Lehrer, "Do Peers Affect Student Achievement in China's Secondary Schools?," *Review of Economics and Statistics*, 2007, *89* (2), 300–312.

Dobbie, Will and Roland G Fryer, "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 2011, pp. 158–187.

— **and** — , "Getting beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 2013, *5* (4), 28–60.

Epple, Dennis and Richard E Romano, "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 1998, pp. 33–62.

— **and Richard Romano**, "Peer effects in education: A survey of the theory and evidence," *Handbook of social economics*, 2011, *1* (11), 1053–1163.

Figlio, David N and Cecilia Elena Rouse, "Do accountability and voucher threats improve low-performing schools?," *Journal of Public Economics*, 2006, *90* (1), 239–255.

Galiani, Sebastian, Paul Gertler, and Ernesto Schargrodsky, "School decentralization: Helping the good get better, but leaving the poor behind," *Journal of Public Economics*, 2008, *92* (10), 2106–2120.

Hastings, Justine S and Jeffrey M Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments*," *Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Hastings, Justine, Thomas Kane, and Douglas Staiger, "Heterogeneous preferences and the efficacy of public school choice," *unpublished working paper*, 2008.

Hoxby, Caroline and Christopher Avery, "The Missing " One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 2013, *2013* (1), 1–65.

Hoxby, Caroline M, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, pp. 1209–1238.

Hsieh, Chang-Tai and Miguel Urquiola, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 2006, *90* (8), 1477–1503.

Imbens, Guido W and Joshua D Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, pp. 467–475.

Krueger, Alan B and Pei Zhu, "Inefficiency, Subsample Selection Bias, and Nonrobustness A Response to Paul E. Peterson and William G. Howell," *American Behavioral Scientist*, 2004, *47* (5), 718–728.

Levin, Henry M, "Educational vouchers: Effectiveness, choice, and costs," *Journal of Policy Analysis and Management*, 1998, *17* (3), 373–392.

Li, Hongbin, Lingsheng Meng, Xinzheng Shi, and Binzhen Wu, "Does attending elite colleges pay in China?," *Journal of Comparative Economics*, 2012, *40* (1), 78–88.

Long, Mark C, "Race and College Admissions: An Alternative to Affirmative Action?," *Review of Economics and Statistics*, 2004, *86* (4), 1020–1033.

_ **, Victor Saenz, and Marta Tienda**, "Policy Transparency and College Enrollment: Did the Texas Top Ten Percent Law Broaden Access to the Public Flagships?," *The ANNALS of the American Academy of Political and Social Science*, 2010, *627* (1), 82–105.

MacLeod, W Bentley and Miguel Urquiola, "Competition and Educational Productivity: Incentives Writ Large," in Paul Glewwe, ed., *Education Policy in Developing Countries*, University of Chicago Press, 2013, chapter 7, pp. 243–284.

Mizala, Alejandra and Miguel Urquiola, "School markets: The impact of information approximating schools' effectiveness," *Journal of Development Economics*, 2013, *103* (C), 313–335.

Muralidharan, Karthik and Venkatesh Sundararaman, "The Aggregate Effect of School Choice: Evidence from a two-stage experiment in India," *The Quarterly Journal of Economics*, 2015, *130* (3), 1011–1066.

Peterson, Paul E, David Myers, and William G Howell, "An Evaluation of the New York City School Choice Scholarships Program: The First Year.," 1998.

Rothstein, Jesse, "Does Competition among Public Schools Benefit Students and Taxpayers? Comment," *American Economic Review*, 2007, pp. 2026–2037.

Rouse, Cecilia E, "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*, 1998, pp. 553–602.

_ **and Lisa Barrow**, "School vouchers and student achievement: Recent evidence and remaining questions," *Annual Review of Economics*, 2009, *1* (1), 17–42.

Rouse, Cecilia Elena, Jane Hannaway, Dan Goldhaber, and David Figlio, "Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure," *American Economic Journal: Economic Policy*, 2013, *5* (2), 251–81.

Sacerdote, Bruce, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 2001, pp. 681–704.

_ , "Peer effects in education: How might they work, how big are they and how much do we know thus far?," *Handbook of the Economics of Education*, 2011, *3*, 249–277.

Walters, Christopher, "A structural model of charter school choice and academic achievement," *Unpublished working paper, UC-Berkeley*, 2013.

Witte, John F, "Achievement effects of the Milwaukee voucher program," in "American Economics Association Annual Meeting, New Orleans" 1997, pp. 4–6.

Zhang, Hongliang, "Identification of Local Average Treatment Effects under Imperfect Matching with an Application to Chinese Elite Schools," *Journal of Public Economics*, 2016, *142*, 56–82.

Zimmerman, David J, "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and Statistics*, 2003, *85* (1), 9–23.

Appendix A: Merging Data Set

The original data sets for 2005-2008 and 2006-2009 cohorts were in two pieces and for 2007-2010 and 2008-2011 cohorts were in three pieces, with the information and number of observations listed in the following table. Note that lottery data set B and junior high graduation data set C contain information of all students in the city, while primary school scores in data set A for 2007-2010 cohort is only for one district and 2008-2011 cohort for two districts, one of which is the same as 2007-2010 cohort.

For the composition analysis, I merge data set A and B for 2007-2010 and 2008-2011 cohort for one shared district. Since there is not much time lag between when A and B were collected, i.e. when students graduated from elementary school, the matching rates are high. I only drop few duplicates and the matching rate is around 95%. The unmatched may move to another city or because of mis-typed names that cannot identify by pronunciation of names.

For the instrumental QTE analysis, I merge data set B and C for three cohorts using name, gender, birth date and middle school, the linkage rate is lower. Possible reasons include mis-typed names, incomplete information on birth date, noncompliance of the lottery assignment and transfer. In order to link as many students' record as possible, I gradually relax the criteria of matching.

After each stage, I take out the matched observations and use the remaining unmatched observations in both data sets to do the next stage of matching. Stage 1 gives us the most reliable matches. Stage 2 captures people whose names were mistyped. Stage 3 captures people whose birth date information is inaccurate. Stage 4 and 5 captures students transferred to another middle school.[21] In total, the matching rate is around 80%.

[21] We would expect that the transferred middle school should be, on average, of higher quality than the original one. There are of course other reasons causing transfer, such as moving and transferring to school closer to home.

Table A1. Description of Data Sets

Cohort 2007-2010	
Data Set A: One district 6th grade scores	Name, elementary school, 6th grade Chinese, math scores
Data Set B: City lottery record	Name, elementary school, gender, birth date, class, admission channel, lottery choice and outcome, middle school admitted, parents' political status, hukou, parents' occupation, address, hometown, ethnicity, political status
Data Set C: City 9th grade record	Name, gender, birth date, middle school attended, middle school graduation score, high school admitted.
Cohort 2006-2009	
Data Set B: City lottery record	Name, elementary school, gender, birth date, class, admission channel, lottery choice and outcome, middle school admitted, hukou, address, hometown, ethnicity
Data Set C: City 9th grade record	Name, gender, birth date, middle school attended, middle school graduation score, high school admitted.
Cohort 2005-2008	
Data Set B: City lottery record	Name, elementary school, gender, admission channel, lottery choice and outcome, middle school admitted
Data Set C: City 9th grade record	Name, gender, birth date, middle school attended, middle school graduation score.

Appendix B: More Summary Statistics

The following three tables present us with the descriptive statistics of the data for three cohorts: 2005-2008, 2006-2009 and 2007-2010. I divide all students into three groups: pre-admitted, noncompetitive lottery takers and competitive lottery takers. As we can see, pre-admitted students have higher middle school graduation scores and better family background (in terms of father and mother political status and hukou possession,[22] attend better schools and get higher scores in junior high graduation exams.

Table A2. Individual Level Data: Summary Statistics 2005

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	policy	preadmission	lottery	lottery&policy	lottery&nonpolicy
9th grade score	22.39	20.79	24.61	20.58	19.80	20.72
Non-academic Evaluation	13.20	12.89	13.80	12.77	12.83	12.76
Higher than 90 percentile	0.442	0.0582	0.464	0.388	0.0531	0.0632
Normalized 9th grade score	0.746	0.693	0.820	0.686	0.660	0.691
female	0.475	0.469	0.477	0.457	0.463	0.456
transfer	0.200	0.220	0.0730	0.371	0.369	0.372
policy	0.0994	1	0.00608	0.155	1	0
preadmission	0.385	0.0264	1	0	0	0
lottery	0.240	0.365	0	1	1	1
winlottery	0.480	0.399		0.480	0.399	0.494
obs	12,964	1,289	7,467	4,653	471	2,563

Note: Column 2 describes policy school students, column 3 describes students who were pre-admitted, column 4 describes students who chose an over-subscribed school and assigned by lottery; column 5 describes students who were assigned by lottery to a policy school; column 6 describes students who were assigned by lottery to a non-policy school. Non-academic evaluation is consist of teacher and self-rated measures of four abilities, including civics, learning ability, atheistic ability, and practical ability.

[22] Hukou equals to 1 if the student has the residency record of Changsha. In China, residency record is very important because it gives you access to many benefits in the city, including health care, pension insurance and employment advantages.

Table A3. Individual Level Data: Summary Statistics 2006

Sample	(1) All	(2) policy	(3) preadmission	(4) lottery	(5) lottery&policy	(6) lottery&nonpolicy
9th grade score	22.67	21.68	24.57	21.75	21.90	21.72
Normalized 9th grade score	0.756	0.723	0.819	0.725	0.730	0.724
Academic high school	0.837	0.760	0.951	0.785	0.786	0.785
Elite high school	0.369	0.295	0.575	0.224	0.288	0.208
Non-academic Evaluation	14.59	14.11	15.11	14.37	14.22	14.41
Imputed 9th grade score	22.72	22.06	24.57	21.85	22.23	21.75
female	0.469	0.456	0.489	0.460	0.461	0.460
hukou	0.778	0.754	0.862	0.776	0.752	0.784
missing hukou	0.174	0.0355	0.233	0.103	0.0383	0.121
preadmission	0.387	0.000480	1	0	0	0
policy	0.125	1	0.000155	0.220	1	0
winlottery	0.546	0.197		0.546	0.197	0.645
transfer	0.0910	0.101	0.0881	0.0852	0.120	0.0755
obs	16,665	2,082	6,446	5,694	1,254	4,440

Note: Column 2 describes policy school students, column 3 describes students who were pre-admitted, column 4 describes students who chose an over-subscribed school and assigned by lottery; column 5 describes students who were assigned by lottery to a policy school; column 6 describes students who were assigned by lottery to a non-policy school. % academic high school indicates the percentage of 9th grade graduates attending an academic high school; some other graduates attend vocational schools or stop schooling. Non-academic evaluation is consist of teacher and self-rated measures of four abilities, including civics, learning ability, atheistic ability, and practical ability. Imputed 9th grade score is constructed by assigning the highest grade of their cohort to the missing grade of direct admitted students who did not take the exam. Having city hukou means that a student is born in city and enjoys the public goods of that city; it is often used as a measure of socioeconomic background.

Table A4. Falsification Test for Treatment Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Depend.Variable	Normalized 9th grade score		Normalized school ranking		% elite high school	
Panel A. All Schools						
postXlow-performing	0.0734	-0.0311	0.0883***	0.0393	0.0185	0.0171
	(0.0957)	(0.0873)	(0.0309)	(0.0282)	(0.0161)	(0.0151)
Observations	441	441	441	441	441	441
R-squared	0.062	0.006	0.071	0.010	0.130	0.027
Number of schools	80	80	80	80	80	80
Panel B. Balanced sample						
postXlow-performing	-0.0183	-0.156	0.0540	-0.00873	-0.0101	-0.0113
	(0.156)	(0.145)	(0.0501)	(0.0465)	(0.0269)	(0.0255)
Observations	160	160	160	160	160	160
R-squared	0.158	0.022	0.143	0.006	0.217	0.056
Number of schools	20	20	20	20	20	20
School FE	Y	Y	Y	Y	Y	Y
Year FE	Y	N	Y	N	Y	N
Time Trend	N	Y	N	Y	N	Y

This table reports falsification check for treatment effects. Instead of estimating treatment effects on policy schools reported in Table 4, it tests whether the post-2007 improvement holds true for any low-performing school. Each cell reports the coefficient of "postXlow-performing", which equals to 1 for low-performing schools (defined by below average in 2004) after 2007. The sample ranges from 2004 to 2011 and drops all the policy schools. Each column has dependent variable listed on top row and control variables indicated at the bottom three rows. Panel B takes schools with observed performance in all eight years.Standard errors in parentheses. Significance level indicated by *** p<0.01, ** p<0.05, * p<0.1.