Preliminary Draft –
Please Do Not Cite or Distribute without Permission of the Authors
Comments Welcome.

Does selective high school improve student achievement? Effects of Exam Schools in Beijing[1]

Kathryn Anderson

Xue Gong

Kai Hong

Xi Zhang

November 2014

Abstract: We examine the effect of selective public high schools, which admit students solely by pre-existing achievement, on student achievement in Beijing, China, using regression discontinuity and kink designs. We explore the weak instrument issue in the framework of regression discontinuity and kink designs, which is associated with the student potential manipulation under the Boston mechanism and the trade-off between school quality and distance, by tests which are robust to the presence of weak instrument. Despite that selective exam schools have higher peer quality by design, and sometimes are also equipped with more experience teachers and better facilities, we find that selective exam high schools at best have modest effects on test scores, and few schools have positive effects on the probability of taking exams and being qualified for college admission, and in most times we do not find any positive effects. Mediators including peer achievement, student/teacher ratio and the percentage of certificated and experienced teachers show weak explanatory power to those findings. The main results hold for students in both of the science and art tracks. We find similar main results in not only groups of schools divided by selectivity, but also a wide range of individual schools along a large part of the distribution of pre-existing achievement.

Keywords: Economics of Education, Selective Schools, Peer Effects, Regression Kink Design, Weak Instrument

JEL Codes: H52, I20, R10

---

[1] Financial support from Beijing Planning Office of Philosophy and Social Science (12JYC021) is acknowledged by Xue Gong and Xi Zhang. Anderson: Department of Economics, Vanderbilt University. Email: kathryn.anderson@vanderbilt.edu. Gong: Beijing Academy of Educational Sciences. Email: gong.sophie@gmail.com. Hong: Department of Economics, Vanderbilt University. Email: kai.hong@vanderbilt.edu. Zhang: Beijing Academy of Educational Sciences.

1. Introduction

On June 26, 2014 in Beijing, more than 80,000 ninth grade students took the high school entrance exam; the exam was given in almost 3000 classrooms and lasted for three days. Performance on the exam would decide which high school students attend in the fall semester. College admission is also based on an exam taken after high school, and graduation from a higher-ranked college usually indicates that the student will obtain a good job with a higher salary and better working conditions. Parents and students think that graduation from an elite high school with high quality teachers and high-ability peers increases the chances that the student is admitted later to a good college; graduation from an elite high school can therefore be the ticket to a student's successful future career. High schools vary in quality, and the competition to get into elite schools is extremely severe; the minimum requirement for admission to some elite schools is to score as high as 95% on the entrance exam.

The evidence on the effectiveness of elite schools which admit students by exam score is under debate in the United States. A recent article in *Slate* magazine advocates that "super-elite public schools aren't necessary any more" (Salam, 2014). The article looks at Stuyvesant High School in New York City to propose several potential problems at elite high schools; most of these problems are associated with the fierce competition at the schools which encourages languishing among lower-ranked students and cheating on exams. The article suggests that reinvention of the elite school does not help solve those problems, and the best solution is to close such one-size-fit-all elite schools and spread gifted students across a wide range of high schools.

The elite school model is found in countries other than China and the United States, Romania, Trinidad and Tobago, and the United Kingdom have elite schools that admit students based on an admission test score, especially at the high school level. Singapore is another country in which the majority of public high schools are exam schools.

Evaluations of student performance in exam schools produce mixed results. The positive effect on test scores is found in studies by Clarke (2010) in the United Kingdom, Pop-Eleches and Urquila (2013) in Romania and Jackson (2010) in Trinidad and Tobago. Other studies show little to no effect of exam schools on student outcomes, including Abdulkadiroglu et al.(2014) in Boston and New York City, the United States, Dobbie and Fryer (2014) in New York City, the United States and Lucas and Mbiti (2014) in Kenya. The studies on elite schools which admit students by lotteries also show no consensus. The positive effect on test scores is found by Hastings and Weinstein (2008) in Charlotte-Mecklenburg, the United States and little evidence of positive effect on student achievement is found by Zhang (2014) in China and Cullen et al.(2006) in Chicago, the United States.

One reason for the mixed results is that the elite exam school is not a homogenous educational experience; many characteristics related with achievement production can affect academic outcomes. In most cases exam schools admit higher ability students ; Students are influenced by their peers, and in most cases the peer effect in education indicates that students learn more if they interact with smarter peers (Hoxby and Weingarth, 2006; Lavy et al., 2010). However, studies of exam schools find positive peer effects in some cases (Jackson, 2013, Trinidad and Tobago) and no peer effects in others (Abdulkadiroglu et al., 2014, the United States).

High school quality is another possible characteristic of exam schools, although it is difficult to precisely measure quality. Some literature shows that school quality and teacher qualification have positive effects on student performance (Lai et al., 2011, China). Since school quality is multi-dimensional, it is difficult to isolate the features of exam and other schools that have the most significant impacts on student performance.

One difficulty in measuring the effect of exam schools is the admission rule itself. Students admitted by higher-ranked schools perform better than other students on the admission test. Performance on the test reflects how well students did in school before the test, and is positively correlated with student ability. Students in higher-ranked schools are more likely to obtain higher scores on exams even if they were enrolled in a lower-ranked school because these students, on average, are more able. The crucial empirical problem in the evaluation of the effectiveness of elite exam schools is modelling this selection by ability. Many recent studies implement the regression discontinuity design (RDD) as a solution to the selection problem.

Our paper evaluates the effect of model schools in Beijing on academic performance using the RDD strategy. Each school has a minimum test score for admission. None of the cutoffs are deterministic; students can attend an exam school with scores below the relative cutoff, and students with scores above the cutoff can attend another lower-ranked school. The structure for admission into high schools fits the fuzzy RDD setting[2]. The forcing variable by which students are assigned to schools is their score on the entrance exam to high schools (SEEH); the main outcome is the score on the entrance exam to colleges (SEEC).

Our study contributes to the existing literatures in several ways. We not only examine the effect of several of the most selective schools, which we call model schools, but also examine all of the other (regular) schools in a large area of China. Both types of schools follow the same rules to admit students, but regular and model schools have different admission cutoff points. We

[2] Students can attend a higher-ranked school because they are able to obtain extra scores if certain requirements are satisfied, such as having minority race or being the child of a martyr. Students choose a lower-ranked school rather than higher-ranked schools for which they are eligible because when students report preference they do not list those higher-ranked schools before the school they actually attend.

examine a broader range of exam schools along the distribution of student achievement. This provides us stronger external validity. We also try to link the effect of exam schools with the admission mechanism. The manipulation that students strategically choose not to report their preferences truthfully leads to weak change in the treatment status around the cutoff, which makes it hard to estimate reliable effects of model schools under the context of weak instruments. We solve this problem by using the regression kink design (RKD) to take into account the potential kink rather than discontinuity in the possibility of treatment at the cutoff.

Our paper extends the literature by studying selective exam schools in a developing country -- China. We explore the effects of exam schools in different tracks by subject. Finally, our paper examines whether there are heterogeneous effects by gender and parental education and occupation, and we provide insights into whether selective exam schools decrease the gender gap in achievement and achievement variation by parental background.

2. Analytical Framework

2.1 Education Production Function

Educational outcomes are produced with inputs through a production function. A typical educational production function is described in Hanushek (1979) and presented in (1) below.

$$A_{it} = f\left(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i, e_i^{(t)}\right) \qquad (1)$$

where A_{it} is the achievement of student i at time t; $B_i^{(t)}$, $P_i^{(t)}$ and $S_i^{(t)}$ are accumulated family background characteristics, peer group characteristics and school inputs for student i at time t; I_i is the student's innate ability, which does not change over time; and $e_i^{(t)}$ denotes unobserved variables.

t and t' denote the times when the SEEC and SEEH are measured respectively. At time t' we have a similar educational production function, and we can use achievement at t' to capture the effect of all inputs accumulated until t' . We rewrite Equation (1) as follows:

$$A_{it} = f\left(A_{it'}, B_i^{(t-t')}, P_i^{(t-t')}, S_i^{(t-t')}, e_i^{(t-t')}\right) \qquad (2)$$

If the changes in schools, which include peer group characteristics and school inputs, are due to attendance in different high schools H_i, then the production function for SEEC becomes:

$$A_{it} = f\left(A_{it'}, B_i^{(t-t')}, H_i, e_i^{(t-t')}\right) \qquad (3)$$

If there are no systematic differences in changes in family background characteristics and unobserved inputs between t and t' between schools and the production function is linear, we have the following expectation of SEEC, where γ is the effect of attending school H on SEEC:

$$E(A_{it}) = \alpha + \beta A_{it'} + \gamma H_i \qquad (4)$$

2.2 Exam School Applications and Admissions

The school assignment system of Daxin District is a variation of the Boston mechanism. Boston mechanism and its variations have been used in many cities in the United States, such as Boston, Chicago, and Denver, and in other countries like the United Kingdom. This mechanism is simple to implement and low cost, but it is manipulable; students have an incentive to misreport their preferences over schools (Pathak and Sonmez 2008, 2013; Abdulkadiroglu and Sonmez, 2003; He, 2012).

In the exam school choice problem, there are a number of students who are willing to be assigned one seat at a high school. The set of students is $= \{i_1, \dots, i_n\}$, and the set of high schools is $S = \{s_1, \dots, s_m\}$. Each school has a capacity $C = \{c_1, \dots c_m\}$. Each student reports a strict preference ordering over schools. The set of preferences is $P = \{p_1, \dots p_n\}$, where p_i is a function such that $p_i: \{1, \dots, m\} \rightarrow \{s_1, \dots s_m\}$. Students take the EEH and obtain a score SEEH.

The set of SEEH for all n students is $A = \{A_1, \ldots, A_n\}$. Each school also has a strict preference ranking of all students. The set of school preferences is $Q = \{q_1, \ldots q_m\}$, where q_s is a function such that $q_s: \{1, \ldots, n\} \rightarrow \{i_1, \ldots i_n\}$.

The outcome of the mechanism is determined in several rounds as follows:

Round 1. Schools only consider the students who listed them as the first choice. Each school admits those students by SEEH until there is no student left who has listed it as the first choice or its capacity is fulfilled.

Round 2. Schools with available seats consider the unassigned students who have listed them as the second choice. Each school admits those students by the order of SEEH until there is no student left who has listed it as the second choice or its capacity is fulfilled.

Round k. Schools whose capacity is not fulfilled admit unassigned students who have listed them as the k-th choice by SEEH.

For school s_m the cutoff for admission is determined as $d_m = A_l$, where A_l refers to SEEH of student l who is either the c_m-th student ranked by SEEH or the last student admitted in round k. In both cases, student l has the lowest SEEH among those who are assigned a seat at school m.

Since each student can list at most eight schools, it is possible that after round eight some schools still having unfilled seats. These schools can contact any unassigned students to see whether they would like to attend this school. After all schools are filled, the students who are still unassigned have to try other options, such as attending private or professional schools or leaving school for the labor market.

In this admission mechanism, parents are likely to avoid top ranking their truly most preferred exam school for their child; if their child's SEEH is not high enough for that school, her chance of being admitted by the second choice is greatly diminished because her second choice school can only admit her if it still has unfilled seats after it admits all students who list it as the first choice. The level of manipulation depends on various characteristics of parents and children (Lai et al., 2009; Pais and Pinter, 2008). On average one can expect that with all of the other factors fixed, parents who expect that their child's SEEH will be around the cutoff are more likely to manipulate; they will top rank a school with lower cutoff for safety.

2.3 Regression Discontinuity and Kink Design

Equation (5) models the SEEC outcome:

$$A_{itj} = \alpha + \beta A_{it'j} + \gamma H_{ij} + e_{itj}, \qquad (5)$$

where A_{itj}, $A_{it'j}$, H_{ij}, α, β and γ are described in 3.1 and the subscript j indicates school j. e_{itj} is the error term that includes all of the unobserved factors which are correlated with SEEC. If the high school assignment is uncorrelated with the error term, an OLS regression gives us consistent and unbiased estimation of the effect of attending school H_j on SEEC.

High school assignment is not exogenous (section 3.2). To account for this endogeneity, we implement RDD and RKD with SEEH as the forcing variable. We assume that the unobserved factors in the error term in (5) are uncorrelated with high school assignment at the corresponding cutoffs. Therefore, we estimate the following equation:

$$A_{ijt} = \alpha + \beta f\left(A_{ijt'}\right) + \gamma H_i + e_{ijt}, \qquad (6)$$

Attendance at high school H_j is positively correlated with the indicator of whether the SEEH is higher or equal to the cutoff for that school, but this relationship is not deterministic; students may manipulate the assignment mechanism by proposing and attending a school with a lower SEEH cutoff, and it is possible for students with SEEH lower than the cutoff to attend the corresponding high school, as long as certain extra requirements are satisfied. Taking such fuzziness into account, we estimate two types of treatment effect. First, we replace the school attendance variable in equation (6) with an indicator of whether the SEEH is higher or equal to the cutoff and get equation (7):

$$A_{ijt} = \alpha + \beta f\left(A_{ijt'}\right) + \gamma 1\left\{A_{ijt'} \geq c_j\right\}_i + e_{ijt}, \qquad (7)$$

Equation (7) identifies the intent-to-treat (ITT) effect of high school; it estimates the effect of high school eligibility rather than the effect of attending high school. Second, we estimate a fuzzy RDD model in (8) using the indicator $1\left\{A_{ijt'} \geq c_j\right\}_i$ as an instrument.

$$\begin{cases} H_i = a + f\left(A_{ijt'}\right) + b1\left\{A_{ijt'} \geq c_j\right\}_i + u_i \\ A_{ijt} = \alpha + \beta f\left(A_{ijt'}\right) + \gamma H_i + e_{ijt}, \end{cases} \qquad (8)$$

Equation (8) identifies the local average treatment effect (LATE) of attending a high school for compliers at the cutoff who would attend the high school when the SEEH exceeded the cutoff and would not attend the high school when the SEEH was less than the cutoff.

For the RKD, we also use a function of SEEH to control for endogeneity and estimate the following equation:

$$A_{ijt} = \alpha + \beta f\left(A_{ijt'}\right) + \theta f\left(A_{ijt'}\right)1\left\{A_{ijt'} \geq c_j\right\}_i + e_{ijt}, \qquad (9)$$

Equation (9) identifies the ITT effect on high school attendance as the coefficient (θ) on the interaction term. Similarly, the equation for fuzzy RKD is described in equation (10):

$$\begin{cases} H_i = a + f\left(A_{ijt'}\right) + b1\left\{A_{ijt'} \geq c_j\right\}_i + cf\left(A_{ijt'}\right)1\left\{A_{ijt'} \geq c_j\right\}_i + u_i \\ A_{ijt} = \alpha + \beta f\left(A_{ijt'}\right) + \gamma H_i + e_{ijt}, \end{cases} \qquad (10)$$

Equation (10) identifies the average treatment effect of the probability of attending high school for the compliers defined in a similar way as in Equation (8).

3. Data

3.1 Descriptive Statistics

The data analyzed include administrative data on student demographics and outcomes from the Daxin District of Beijing. Daxin District used to be a county in the Beijing metropolitan area and became a suburban district in 2001 with the expansion of Beijing City. Information on school level is also reported. The data include 11 high schools and 3868 grade 12 students surveyed in 2008. Those students took the entrance exam for high schools (EEH) in 2005 and the entrance exam for colleges (EEC) in 2008. Not all of the students took the EEC; 1.2% of the students in model schools and 10.2% in regular schools did not take the EEC. Students self-choose whether to take the EEC. Later we show that such self-selection is not a threat to our RDD (RKD) analysis. Two out of 11 high schools are model schools - No.1 and Xinghua High schools. The other nine are regular schools.

Students enter high school by a system of admission rules and take either the science or art track. The admission and track systems are explained in detail in Appendix A. Scores on both exams, as well as student characteristics, family background, and school characteristics, are collected for students in both school types and tracks. The descriptive statistics on these characteristics by model schools are shown in Tables 1 and 2, and detailed descriptive statistics by individual schools are in appendix Tables A1, A2 and A3.

Table 1: Descriptive Statistics by School Types and Tracks

	Art			**Science**		
	Model School	**Regular School**	**T test**	**Model School**	**Regular School**	**T test**
Exam Score						
SEEC	0.72 (0.77)	-0.38 (0.90)	1.10*** (0.04)	0.81 (0.80)	-0.22 (0.92)	1.03*** (0.06)
SEEH	1.01 (0.46)	-0.21 (0.79)	1.22*** (0.03)	0.82 (0.50)	-0.52 (1.02)	1.34*** (0.06)
Student Characteristics						
Male	0.50	0.47	0.03 (0.02)	0.66	0.68	-0.02
Age	18.56 (0.660)	18.76 (0.724)	-0.20*** (0.03)	18.58 (0.67)	18.76 (0.70)	-0.18*** (0.05)
Parental Backgrounds						

College Father	0.19	0.18	0.00 (0.02)	0.27	0.19	0.08*** (0.03)
College Mother	0.22	0.15	0.07*** (0.02)	0.19	0.17	0.02 (0.03)
Farmer Father	0.50	0.47	0.04 (0.02)	0.40	0.48	-0.08** (0.04)
Farmer Mother	0.50	0.47	0.04 (0.02)	0.50	0.43	0.07* (0.04)
No. of Obs	782	1495		278	1015	

Note: SEEC and SEEH are standardized test scores with the mean of zero. Sample means for the 2008 cohort. No. of Obs is the number of observations with at least one non-missing value for the variables listed (or the number of observations with non-missing values for SEEH which is the forcing variable in the RDD).

Table 2: Descriptive Statistics of Schools

	Model School	**Regular School**	**T test**
Urban	1	0.56	0.44 (0.39)
Number of Students Enrolled	1894 (554.4)	1085 (410.3)	809** (335.1)
Number of Teachers	154.5 (2.1)	94.8 (47.1)	59.72 (34.70)
Student/Teacher Ratio	12.29 (3.76)	11.95 (1.61)	0.34 (1.54)
Percentage of Teachers with Advanced Certificate	0.42 (0.01)	0.23 (0.14)	0.19 (0.11)
Percentage of Teachers Younger than Age 35	0.43 (0.03)	0.64 (0.14)	-0.21* (0.11)
Minimum Score of SEEH for Admission in 2005	482 (11.31)	424.1 (36.99)	57.89* (27.43)
No. of Obs.	2	9	

Students in the model schools in both tracks do better on average on both the SEEC and SEEH. The performance gap between model and regular schools on the SEEC is positive in both tracks. Students in the model schools are slightly younger. In general students in the model schools come from families with more advantaged social status; their parents are more likely to hold college degrees. There is no significant difference between students in model schools and regular schools in gender and parental probability of agricultural occupation.

The two types of schools are quite different in their characteristics. Both of the model schools are located in urban areas, but only 55 percent of regular schools are urban. More students are enrolled and more teachers are employed in model schools than in regular schools. The quality of teachers is higher in the model schools; the teachers are more experienced, and a higher percentage of them has an advanced certificate. As a consequence, the model schools are popular among high-achieving students. This leads to a much higher minimum SEEH for enrollment as well as higher-quality peers in the model schools.

3.2 Validity Test

Before implementing the RDD analysis, we first check the validity of the RDD by testing whether, at the threshold, there is a discontinuity in student frequency, background variables or policy characteristics related to school choice.

3.2.1 Density Test

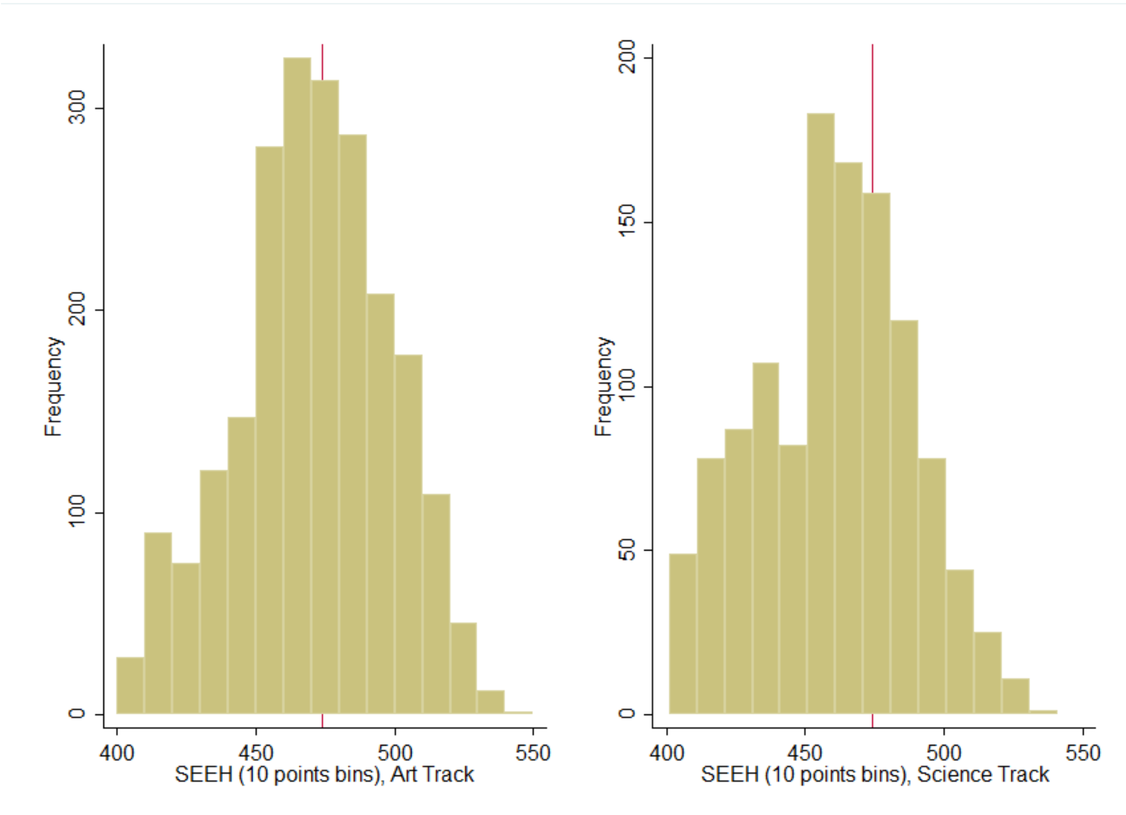
RDD produces unbiased estimates of the effects of high school on subsequent student achievement, if there is imperfect manipulation of the school choice and assignment. Perfect manipulation is unlikely at the cutoff; the cutoff is unknown to anyone[3], and nobody can perfectly control it. When people grade the SEEH exam, they do not know whose paper they are grading because the name is sealed. Moreover, there are many graders. A few cheating graders, if any, should not lead to systematic manipulation. Nevertheless, we still need to formally test for manipulation. We first examine whether there is any evidence of manipulation by displaying the distribution of students observed in our data. According to McCrary (2008), discontinuity in the density around the threshold indicates the risk of endogenous sorting which violates the RDD assumptions.

The distribution of students by SEEH is presented in Figure 1. The figure shows an approximately normal distribution peaking at the 50 percent threshold in both tracks, providing

[3] They can obtain information from the cutoffs in previous years and predict the current cutoff. There is always prediction error, and on average it is quite unlikely for us to have a substantial amount of students who perfectly predict the cutoff.

no evidence of endogenous sorting. This picture by itself is not enough assurance that there is no manipulation. We perform McCrary's density test. The discontinuity estimates are 0.11 with standard error of 0.09 for the art track and -0.07 with a standard error of 0.11 for the science track. The test results provide no evidence of endogenous sorting across the model school cutoffs.

Figure 1: Density of Students by SEEH: Model School



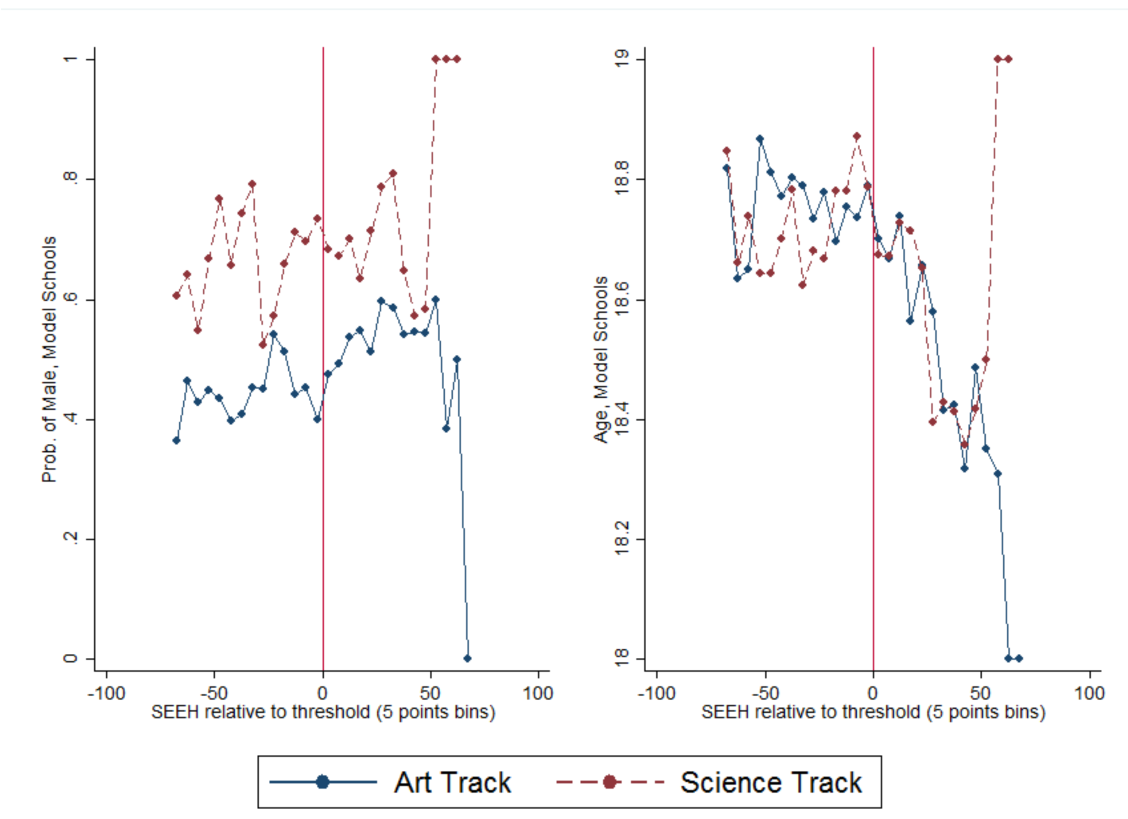
Note: Sample includes all students in the corresponding track in Daxing District of Beijing in 2008. SEEH is censored at 400 from the left.

We also perform the density test for all of the individual schools. The results are presented in appendix Table A4. For the two model schools we have balanced density. For some other schools we have a discontinuity in density at the cutoff, but we can still compare the results from these schools to results from the other schools to add robustness to our main results.

3.2.2 Validity Test of Background Variables

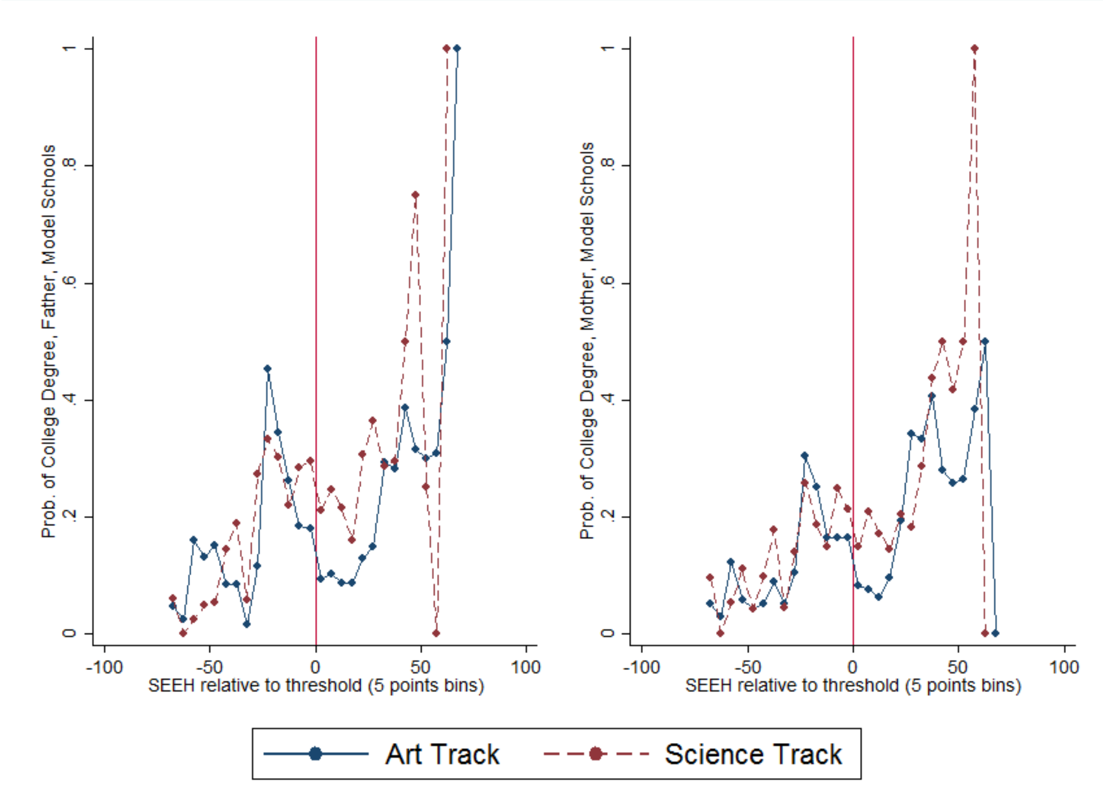
The internal validity of the RD design requires that no relevant variables other than the treatment jump at the cutoff. If some background variables also jump at the cutoff with the treatment status, we are not able to distinguish their effects from the treatment effect. Figure 3 to Figure 5 present graphic results with respect to the validity check of background variables. There are no apparent discontinuities at the cutoff for observed covariates, and in fact many of the covariates show a random-walk pattern with the SEEH. Appendix Table A5 shows the details of validity tests of the background variables for each exam school. In most cases we do not have unbalancedness beyond what would be expected by random chance. As a consequence we do not control for covariates in the model. We still analyze schools with significant unbalanced background variables to add robustness to our results.

Figure 3: Graphic Checks of Gender and Age, Model Schools



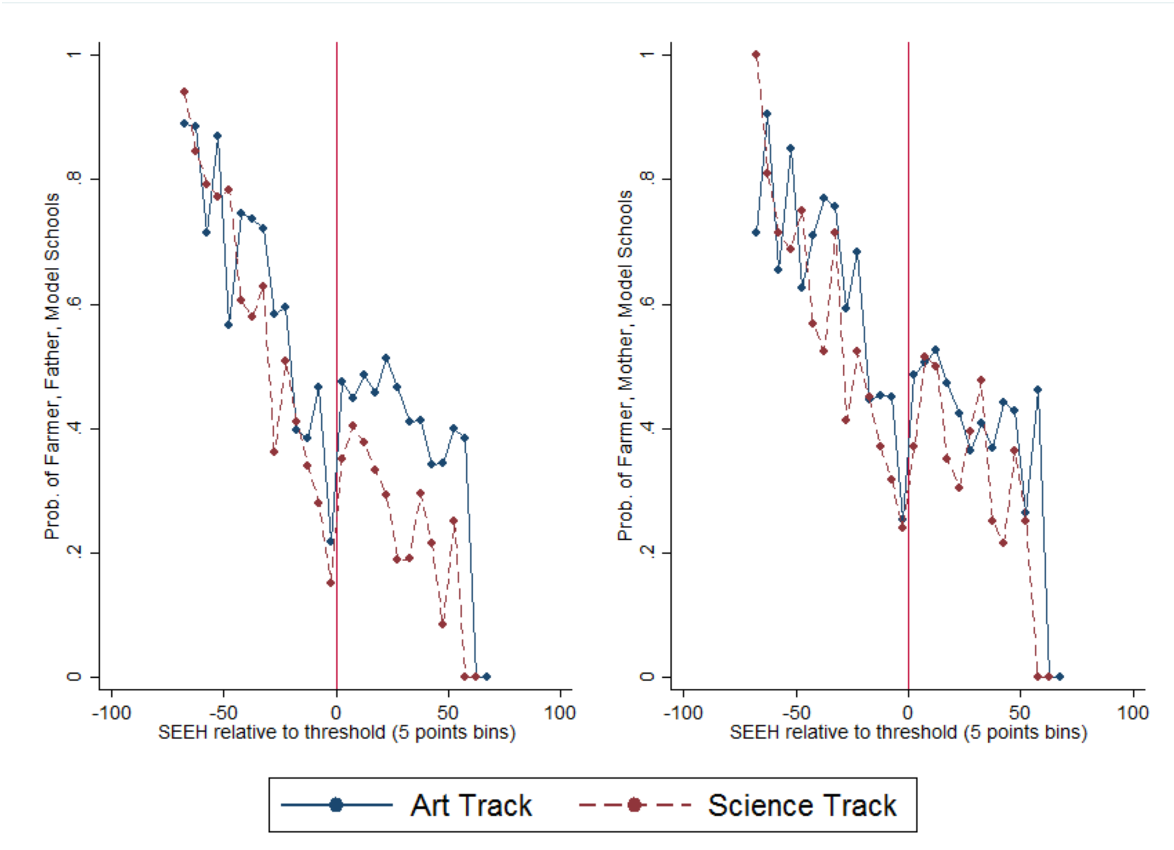
Note: The figure presents the raw SEEH in 5 point wide bins. The probability of being male and age are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin.

Figure 4: Graphic Check of Probability that Parents Have a College Degree, Model Schools



Note: The figure presents the raw SEEH in five point wide bins. The probabilities that the father and mother have college degrees are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC, are dropped.

Figure 5: Graphic Check of Probabilities that Parents Are Farmers, Model Schools

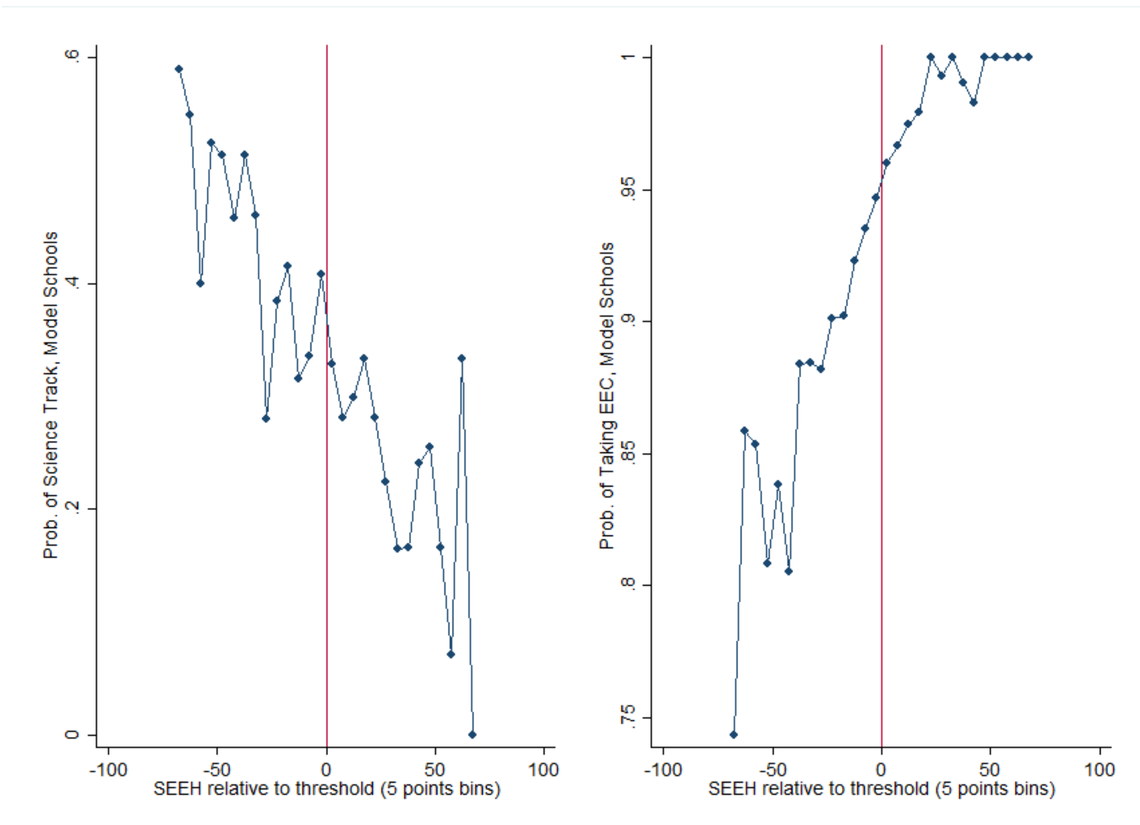


Note: The figure presents raw SEEH in five point wide bins. The probabilities of being a farmer of the father and mother are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC, are dropped.

3.2.3 Validity Test of Self-Choice Characteristics

Students can also strategically sort in terms of self-choice based on other variables. Such manipulation is not always detected by the density test, as it does not necessarily lead to unbalanced density. We examine variables including the probability of choosing the science track and the probability of taking the SEEC. Students may strategically choose the track and decide whether to take the exam according to their expectation of performance. For example, students who are marginally admitted may be more likely to choose the track with their highest expected score on the SEEC than those who are marginally rejected. Similarly, students who are marginally admitted may also be more likely to not take the EEC exam. In both cases the estimation of effect will be upward biased. The graphic analysis in Figure 6 shows no strong evidence of strategic sorting around the cutoff for model schools. Panel B of Appendix Table A5 presents the details of the validity checks of both variables for individual schools. The test results indicate that self-choice is not a threat to the RDD.

Figure 6: Graphic Check of Probability of Choosing the Science Track and Probability of Taking the SEEC, Model Schools



Note: The figure presents the raw SEEH in five point wide bins. Standardized SEEC and the probability of enrolling are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC, are dropped.

4. Effect of Exam Schools on Student Achievement

4.1 ITT Effect of Exam Schools

We first perform the ITT analysis which provides insight into the effect of high school eligibility. The parametric estimation fits Equation (7) by OLS where $f\left(A_{ijt'}\right)$ is a cubic polynomial of SEEH[4]. The non-parametric estimation fits Equation (7) by local linear regression, with the data-driven optimal bandwidth derived by Calanico et al. (2014)[5], a triangle kernel, and a linear function $f\left(A_{ijt'}\right)$. Table 3 show the results for the model schools and all of the individual exam schools[6]. Overall we find little evidence of a positive ITT effect of exam schools on the SEEC, for both tracks along the distribution of the SEEH.

Table 3: ITT Estimation of Effect on SEEC

	Art Track		**Science Track**	
	Parametric Estimation	**Non-Parametric Estimation**	**Parametric Estimation**	**Non-Parametric Estimation**
Model School	-0.039 (0.159) [1730]	-0.118 (0.189) [1005]	-0.176* (0.090) [833]	-0.032 (0.120) [519]
No.1	0.015 (0.075) [1558]	0.205*** (0.049) [513]	0.131 (0.207) [683]	0.183 (0.345) [232]
Xinhua	-0.039 (0.159) [1730]	-0.118 (0.189) [1005]	-0.176* (0.090) [833]	-0.032 (0.120) [519]
No.2	0.220 (0.166) [1684]	0.251 (0.209) [1028]	0.148 (0.194) [884]	0.150 (0.170) [444]
No.3	-0.094 (0.104) [1525]	-0.069 (0.124) [982]	0.102 (0.199) [888]	0.131 (0.199) [498]
No.5	-0.030 (0.104) [1575]	0.002 (0.120) [993]	0.263 (0.191) [900]	0.198 (0.200) [597]
No.8	0.000 (0.260) [1441]	-0.008 (0.272) [999]	-0.535** (0.191) [867]	-0.403 (0.239) [491]
Xingda	0.083 (0.180) [1104]	-0.039 (0.134) [697]	0.298* (0.162) [752]	0.126 (0.165) [406]
Yufa	0.084 (0.226) [274]	-0.049 (0.167) [179]	-0.187 (0.203) [303]	-0.222 (0.140) [233]
Weishanzhuang	0.209 (0.259)	0.006 (0.196)	-0.078 (0.204)	-0.136 (0.130)

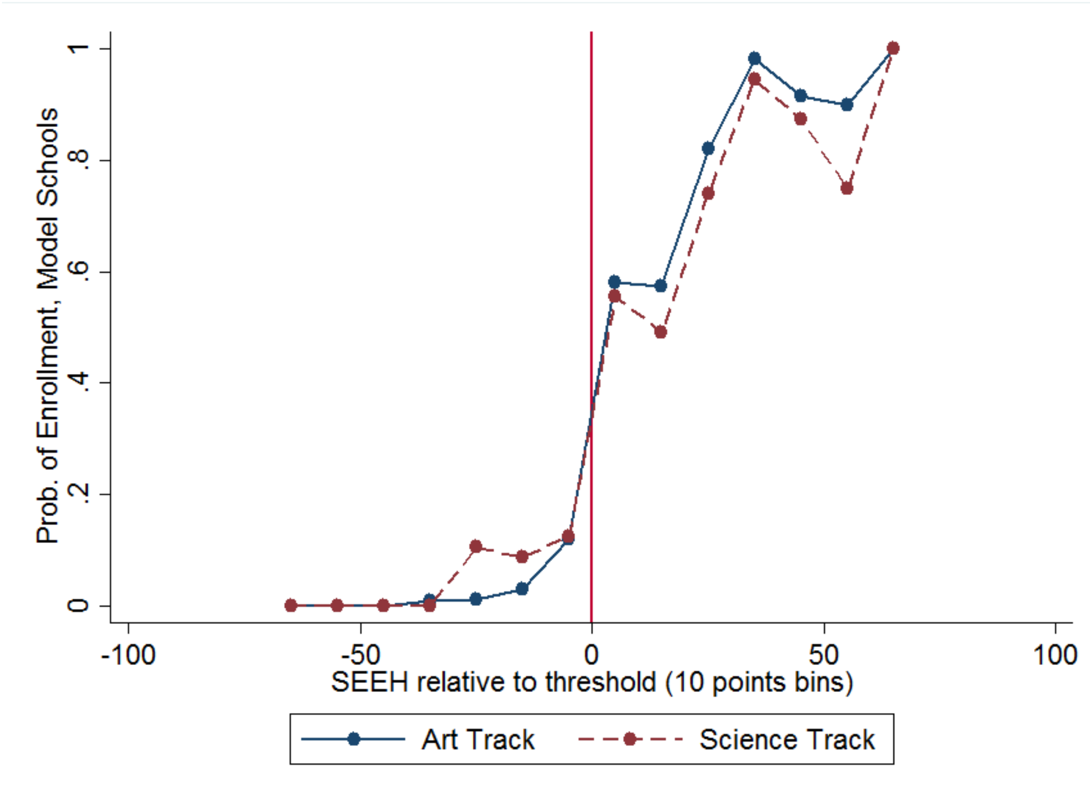
[4] We explore different polynomials in the section on robustness checks.
[5] We also conduct robustness checks with respect to other optimal bandwidths such as the MSE-optimal bandwidth (Imbems and Kalyanaraman, 2012) and the cross-validation optimal bandwidth (Ludwig and Miller, 2007). According to Calanico et al. (2014) those bandwidths tend to be too large. The bandwidth selection is summarized in Appendix Table A6.
[6] We do not estimate the effect of Jiugong School which has the lowest cutoff.

	[261]	[166]	[294]	[216]
Caiyu	0.161 (0.172) [286]	-0.069 (0.177) [164]	0.099 (0.122) [307]	0.051 (0.110) [188]

Note: Parametric estimations controls for a cubic function of the standardized SEEH, and the students covered are within one standard deviation of the cutoffs. Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel, which is the boundary optimal (Imbens and Lemieux, 2008). The standard errors which are robust and clustered on school are shown in parentheses. The numbers of observations are in brackets.

The result indicates that eligibility for a seat in an exam school does not necessarily lead to improvement in student achievement. The result even holds for those more selective model schools, with one exception -- art students in No.1 school. The admission system provides students incentive to manipulate by proposing a lower ranked school, especially for the most selective ones. One possible reason for the lack of effect is that the eligibility does not lead to a higher probability of enrollment; students who are marginally eligible for a highly ranked school propose a lower rank school for safety. To explore this possibility, we estimate the ITT effect of eligibility on the probability of enrollment using the same estimation strategy. Figure 7 shows how the probability of enrollment in model schools changes with the SEEH around the cutoff. Details of the estimation can be found in Appendix Table A8.

Figure 7: Probability of Enrollment, Model Schools



Note: The figure presents the raw SEEH in five point wide bins. Standardized SEEC and the probability of enrolling are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC, are dropped.

Students are more likely to attend some schools when they are eligible with SEEH higher or equal to the cutoff, but in many other cases eligibility does not significantly increase the possibility of enrollment. One possible reason is that students and their parents do not have sufficient information, and they have no idea about the true ranks of schools or the students' relative performance among peers. Given that those students had learned with the same peers for at least three years, we view this possibility quite unlikely, especially for the most selective schools that had well established reputations for decades.

Under the Boston mechanism, parents in Beijing are overcautious as they play "safe" strategies by top ranking their lower ranked school too often (He, 2012). This is a potential explanation for the continuity in the probably of enrollment at the cutoff. In this study, if a student who was eligible for a school attended a school with a lower cutoff, such students must list the school she was eligible for at the lower rank. Appendix Table A9 summarizes the patterns of over caution. It is found that there is a substantial number of overly cautious students, and the percentages of overly cautious students over all eligible students are the highest in those two model schools. Moreover, students with relatively lower achievement, who are more likely to marginally pass the cutoff, are more likely to be overly cautious, compared with students whose expected scores are much higher than the cutoff. In Appendix Table A8 we show that on average the SEEH of non-overly cautious student is significantly higher than the SEEH of the overly cautious students for every school. The weak jump of treatment status at the cutoff is likely due to manipulation with respect to the admission mechanism; for safety students with relatively lower SEEH are more likely to top rank schools with lower cutoffs. Details about the manipulation in the Boston mechanism are discussed in Appendix B.

Since the probability of manipulation may decrease with SEEH, eligibility may increase the slope of enrollment, which can be interpreted as the marginal effect of SEEH on the probability of enrollment. To explore this possibility we adopt the RKD described by Equation (10), where $f(A_{ijt'})$ is a cubic polynomial of SEEH in the parametric estimation and a linear polynomial in the non-parametric estimation. Table 4 shows the results from the non-parametric estimation. The overcautious manipulation is found for the most selective school (No. 1). The patterns for Xinhua and probably the No. 2 school show little evidence of overcautious manipulation; the probability of enrollment jumps at the cutoff, then decreases with SEEH as students are more likely to top rank a higher ranked school (No. 1 for Xinhua and No. 1 and Xinhua for No. 2) when they expect higher SEEH. For the other schools, RDD or RKD approach seems to be problematic with insufficient discontinuity at the cutoff[7]. Although it is still possible for us to find kinks with a higher order polynomial, the results for these schools should be interpreted with caution and for comparison purposes only.

Table 4: ITT Estimation of Effect on Enrollment, RKD Specifications

	Art Track				**Science Track**			
	Parametric Estimation		**Non-Parametric Estimation**		**Parametric Estimation**		**Non-Parametric Estimation**	
	Jump (*b*)	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**
Model School	1.363** (0.448) [2061]	-0.341 (0.235) [2061]	0.418 (0.377) [1584]	-0.148 (0.438) [1584]	2.380*** (0.561) [1049]	-6.879** (2.776) [1049]	0.814* (0.353) [541]	-0.943* (0.486) [541]
No.1	-12.16*** (0.328) [1918]	24.44*** (1.415) [1918]	-2.331*** (0.579) [948]	2.522*** (0.563) [948]	-13.63*** (2.664) [890]	27.37*** (6.233) [890]	-2.763*** (0.092) [375]	2.883*** (0.109) [375]
Xinhua	-0.366 (0.558) [2061]	3.421 (2.786) [2061]	0.676* (0.363) [1584]	-0.547 (0.385) [1584]	0.524* (0.245) [1049]	0.322 (1.993) [1049]	0.892* (0.391) [541]	-1.029* (0.461) [541]
No.2	0.466 (0.478) [2077]	-0.014 (2.285) [2077]	0.503 (0.291) [1623]	-0.232 (0.232) [1623]	0.460 (0.555) [1079]	0.044 (0.267) [1079]	0.539* (0.280) [684]	-0.254 (0.477) [684]
No.3	0.243 (0.237) [1991]	0.387 (0.528) [1991]	0.167 (0.163) [1431]	-0.090 (0.213) [1431]	0.215 (0.202) [1129]	0.156 (0.502) [1129]	0.159 (0.150) [656]	-0.131 (0.337) [656]
No.5	0.289 (0.186) [2031]	-1.462 (1.158) [2031]	0.312 (0.221) [1347]	-0.793 (0.621) [1347]	0.336 (0.296) [1126]	-0.454 (0.573) [1126]	0.292 (0.252) [693]	-0.485 (0.457) [693]
No.8	0.098 (0.111) [1897]	-0.230 (0.293) [1897]	0.104 (0.114) [1085]	-0.379 (0.401) [1085]	0.243 (0.251) [1114]	0.591 (0.708) [1114]	0.081 (0.084) [766]	-0.135 (0.195) [766]
Xingda	1.902 (1.735) [1644]	2.864 (3.314) [1644]	0.605 (0.368) [995]	0.786 (0.494) [995]	0.159 (0.215) [1025]	-1.025 (0.788) [1025]	0.401 (0.231) [752]	0.548 (0.337) [752]
Yufa	1.927 (5.122) [507]	-0.655 (5.896) [507]	1.032 (1.101) [188]	0.537 (0.569) [188]	5.464 (6.925) [472]	5.743 (7.690) [472]	0.661 (0.612) [379]	0.364 (0.327) [379]
Weishan zhuang	-18.26 (23.59) [479]	-24.62 (30.89) [479]	0.804 (0.911) [209]	0.245 (0.424) [209]	5.802 (13.19) [458]	9.945 (15.08) [458]	0.141 (0.415) [341]	-0.015 (0.282) [341]
Caiyu	53.63 (40.66)	73.04 (53.01)	-2.394** (0.905)	-1.026* (0.458)	-18.39** (8.046)	-24.87** (9.312)	-1.916*** (0.482)	-0.988** (0.325)

[7] When there is no sufficient discontinuity at the cutoff, we have the problem of weak instruments, which distorts the confidence interval. For both model schools and the aggregate model school, the F statistics of the ITT estimation of the effect on enrollment with the RKD specification are greater than 10, which according to Staiger and Stock (1997) is a threshold to safely assume that the analysis does not have a weak instrument. Therefore our main results with respect to model schools seem not to suffer from the weak instrument problem.

	[529]	[529]	[216]	[216]	[482]	*[482]*	[325]	[325]

Note: b and c refer to the same parameters in Equation (10). Parametric estimations control for a cubic function of the standardized SEEH, and the students covered are within 1.5 standard deviation of the cutoffs. Italic numbers indicate that there are coefficients of higher order of kink which are significant at the 90% level at least. Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimation controls for a quadratic function of standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local polynomial estimation with a triangle kernel. The standard error which is robust and clustered on school is shown in parentheses. The number of observations is in brackets.

For those less selective schools, there is another potential explanation for the weak jump at the cutoff. Hastings et al. (2005) show that there is trade-off between school test scores and proximity. Parents who are elastic to school quality, which is approximated by test scores, are willing to travel a long distance for an even modest gain in school scores while parents who are inelastic to school quality are more likely to choose a local school. For those less selective schools, if parents do not value the higher cutoff of a school far away as a significant change in school quality, they may just top rank a local school with a lower cutoff.

The last four schools – Yufa, Weishanzhuang and Caiyu – are located in separate satellite towns surrounding the center part of Daxing District[8]. They have very closed cutoffs, which indicates that they are comparable in school quality, but they are quite far away from each other. The distances between three schools are 19km between Yufa and Weishanzhuang, 19km between Weishanzhuang and Caiyu, and 31km between Yufa and Caiyu[9]. All of these schools are more than 10km away from the center part of Daxing District. Parents in those towns may just choose the local school if their children are not eligible for those much higher ranked school in the center of the district. The same situation may also apply to children who live in the center part of the district; although all of the remaining seven schools are not far away from each other (all of them allocated in a rectangular area of 3km × 4km), some parents may not bother sending their children

[8] The names of the towns are consistent with the names of the schools. They are Yufa town, Weishanzhuang Town and Caiyu town.
[9] Those distances are measured as the crow flies approximately in Google Maps. Usually the traveling distances are longer.

to a non-neighbor school for a small increase in school quality. Appendix Table A10 shows the pattern of enrolled students who could attend a school with a higher cutoff. It is clear that only the most selective schools are attractive for students from satellite towns. This is also the case for students from lower ranked schools in the center area of the district. A model describes the tradeoff between school quality and commuting distance can be found in Appendix C.

All of the schools enrolled students with a SEEH less than the cutoff, especially the most selective schools. Appendix Table 11 summarizes the enrolled non-compliers by school. There is a substantial amount of enrolled non-compliers in all schools, although in some lower ranked schools the ratio is smaller. Taking into account the overcautious manipulation and the trade-off between distance and perceived school quality, we estimate the effects of groups of schools for which we have significant change in the treatment status at the cutoff. Table 5 shows the effect on the treatment status for those school groups: model schools (No. 1 and Xinhua), center area schools (No. 1, Xinhua, No. 2, No. 3, No. 5, No. 8 and Xingda) which can be divided into two subgroups as center area selective schools (No. 1, Xinhua and No. 2) and center area less selective schools (No. 3, No. 5, No. 8 and Xingda). In the rest of the paper we focus on those three groups of schools but still report in a separate section of the paper the effect of individual schools which are robust to the presence of weak instruments[10].

Table 5: ITT Estimation of Effect on Enrollment by Groups, RKD Specifications

	Art Track				**Science Track**			
	Parametric Estimation		**Non-Parametric Estimation**		**Parametric Estimation**		**Non-Parametric Estimation**	
	Jump (*b*)	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**	**Jump (*b*)**	**1st Kink (*c*)**
Specification 1								

[10] Under fuzzy RD settings the estimated effects with weak jump in the treatment status tend to over-reject the null hypothesis of no effect (Marmer et al., 2014). We estimate the effect which is robust to weak instruments by the method proposed in various literatures, including the Anderson-Rubin test (Anderson and Rubin, 1949), the conditional likelihood-ratio (CLR) test (Moreira, 2003) and the robust CLR test (Finlay and Magnusson, 2009).

Model School	1.363** (0.448) [2061]	*-0.341* *(0.235)* *[2061]*	0.418 (0.377) [1584]	-0.148 (0.438) [1584]	2.380*** (0.561) [1049]	*-6.879*** *(2.776)* *[1049]*	0.814* (0.353) [541]	-0.943* (0.486) [541]
Specification 2								
Center Area School	1.331 (2.108) [1644]	1.821 (4.567) [1644]	0.793** (0.328) [995]	1.268** (0.419) [995]	-0.991 (1.842) [1025]	-3.208 (3.866) [1025]	0.429 (0.389) [752]	0.732 (0.514) [752]
Specification 3								
Center Area Selective School	0.195 (0.473) [2077]	1.952 (1.477) [2077]	0.488* (0.260) [1623]	0.558* (0.299) [1623]	0.165 (0.439) [1079]	2.369 (1.705) [1079]	0.416 (0.260) [684]	0.969* (0.498) [684]
Center Area Less Selective School	0.902 (2.095) [1644]	*0.135* *(4.578)* *[1644]*	0.713* (0.333) [995]	1.210** (0.396) [995]	-1.739 (1.850) [1025]	-5.359 (3.885) [1025]	0.266 (0.374) [752]	0.530 (0.459) [752]

Note: Parametric estimation controls for a cubic function of the standardized SEEH, and the students covered are within 1.5 standard deviation of the cutoffs. Italic numbers indicate that there are coefficients of higher order of kink which are significant at the 90% level at least. Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimation controls for a quadratic function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local polynomial estimation with a triangle kernel. The standard error which is robust and clustered on school is shown in parentheses. The number of observations is in brackets.

4.2 TOT (LATE) Effect of Exam Schools on SEEC

For schools where a discontinuity is found, we fit Equation (8) to obtain the LATE effect of exam schools, which provides the effect of enrollment in each exam school on the SEEC[11]. The parameters and estimation strategy are the same as in Equation (7). In addition, to employ the possible kink at the cutoff for identification especially for schools where there is no evidence of a discontinuity, we also fit Equation (10), where $f(A_{ijt'})$ is a cubic polynomial and linear function of the SEEH in parametric and non-parametric estimations respectively. Tables 6 shows the results from both the RKD and RDD specifications.

Table 6: LATE Estimation of the Effects of School Groups on SEEC

	Art Track		**Science Track**	
	Parametric Estimation	**Non-Parametric Estimation**	**Parametric Estimation**	**Non-Parametric Estimation**
RKD				
Model School	-0.415	-0.448	0.261	0.013

[11] Before we estimate the model we re-perform the validity test in Sections 4.2.2 and 4.2.3 using the RKD specification. The results are summarized in Appendix Tables A4 and A6. Overall we do not find strong evidence that RKD is invalid.

	(0.625) [2061]	(0.697) [1584]	(0.403) [1049]	(0.429) [541]
Center Area School	0.405*** (0.149) [1644]	0.351 (0.260) [995]	-0.899** (0.451) [1025]	-0.714 (0.532) [752]
Center Area Selective School	0.190 (0.145) [2077]	0.094 (0.146) [1623]	0.369* (0.220) [1079]	-0.087 (0.212) [684]
Center Area Less Selective School	0.150 (0.141) [1644]	0.333 (0.251) [995]	-0.608** (0.278) [1025]	-0.820 (0.586) [752]
RDD				
Model School	-0.108 (0.468) [1730]	-0.407 (0.804) [1005]	-0.393 (0.264) [833]	-0.079 (0.271) [519]
Center Area School	-0.238 (0.521) [1104]	0.140 (0.427) [697]	-1.005 (0.661) [752]	-0.387 (0.530) [406]
Center Area Selective School	0.379 (0.269) [1684]	0.518 (0.404) [1028]	0.243 (0.311) [884]	0.363 (0.350) [444]
Center Area Less Selective School	-0.141 (0.281) [1104]	0.132 (0.401) [697]	-0.660* (0.347) [752]	-0.482 (0.698) [406]

Note: Parametric estimations controls for a cubic function of the standardized SEEH, and the students covered are within 1.5 standard deviations of cutoffs for the RKD and 1 standard deviation of the cutoffs for the RDD. Non-parametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The standard error which is robust and clustered on school is shown in parentheses. The number of observations is in brackets. The number in bold indicates that the result is robust to weak instruments.

We find little evidence of a positive effect of school group on exam performance from both of the RKD and RDD specifications. For the most selective schools with higher cutoffs which are perceived to improve student achievement the most, at best we find no effect on college exam score.

4.3 *Effect of Exam Schools on Other Outcomes*

We also examine the effect of exam schools on other outcomes, including the probability of taking the SEEC, college admission and the probability of attending higher and lower ranked exam schools. The probability of taking the SEEC captures the attrition status. College admission outcomes are correlated with the SEEC, and they also reflect the relative rank of students in a larger area, as the college admission procedure is centralized at the Beijing City level. By examining the effect on the probability of attending higher and lower ranked exam schools we are able to find evidence of strategic behaviors of students. Table 7 shows the results of non-parametric estimation from the RKD specification, and the results of non-parametric estimation from the RDD specifications are shown in Appendix Table 13.

Table 7: Estimation of the Effects of School Groups on Various Outcomes

		Art Track			**Science Track**		
	Taking SEEC	**College Qualification**			**College Qualification**		
		Key 4-year	**4-year**	**3-year**	**Key 4-year**	**4-year**	**3-year**
RKD							
Model School	-0.015 (0.130) [2600]	-0.100 (0.235) [1587]	-0.157 (0.341) [1587]	-0.297 (0.194) [1587]	0.065 (0.286) [541]	0.061 (0.203) [541]	-0.028 (0.247) [541]
Center Area School	-0.308 (0.417) [2665]	-0.051 (0.068) [998]	0.157 (0.093) [998]	0.246 (0.209) [998]	0.087 (0.093) [753]	0.324 (0.209) [753]	-0.648 (0.445) [753]
Center Area Selective School	0.022 (0.086) [2379]	0.095 (0.061) [1626]	0.227*** (0.061) [1626]	-0.206** (0.087) [1626]	0.088 (0.080) [684]	0.027 (0.159) [684]	-0.088 (0.057) [684]
Center Area Less Selective School	0.310 (0.756) [2665]	-0.048 (0.060) [998]	0.149 (0.096) [998]	0.233 (0.189) [998]	0.099 (0.108) [753]	0.371 (0.254) [753]	-0.744 (0.505) [753]
RDD							
Model School	-0.007 (0.045) [2007]	-0.269 (0.290) [1008]	-0.443 (0.583) [1008]	-0.032 (0.158) [1008]	-0.025 (0.169) [519]	-0.023 (0.127) [519]	0.020 (0.178) [519]
Center Area School	-0.432** (0.207) [1503]	-0.059 (0.066) [698]	0.123 (0.106) [698]	0.079 (0.261) [698]	0.106 (0.102) [407]	0.266 (0.183) [407]	-0.238 (0.279) [407]
Center Area Selective School	0.080*** (0.029) [1676]	0.152 (0.190) [1031]	0.278 (0.170) [1031]	0.113* (0.061) [1031]	0.210 (0.270) [444]	0.154 (0.188) [444]	0.047 (0.166) [444]
Center Area Less Selective School	-0.431* (0.225) [1503]	-0.055 (0.057) [698]	0.118 (0.101) [698]	0.074 (0.272) [698]	0.132 (0.151) [407]	0.332 (0.262) [407]	-0.297 (0.378) [407]

Note: Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The standard error which is robust and clustered on school is shown in parentheses. The number of observations is in brackets.

Table 7 shows the effect on other outcomes, including attrition measured as the probability of taking the EEC and qualifications to key universities, 4-year universities and 3-year colleges. We find some significant effects, especially for those center area selective schools; students attending those schools are more likely to take the EEC exam, and art students from those schools have an increased probability of qualifying for a 4-year university, accompanied by a decreased probability of qualifying for a 3-year college. The latter implies that not all of students who were not eligible for a 4-year university can benefit from the center area selective schools. While some of them become qualifying for a 4-year university, some others even lose the qualification for a 3-year college. Given that a student who qualifies for a 4-year university must also qualify for a 3-year college, the net negative effect on the probability of qualifying for a 3-year college only is even larger.

4.4 Weak Instrument and Effect of Individual School

When we estimate the effect of individual schools, we have weak instrument problems in many cases; there is no significant change in attendance at the cutoff. At the presence of weak instrument variable, we are not able to get an unbiased estimator which is fully robust in the fuzzy RDD framework. The two stage least squares (2SLS) estimator is biased with a non-normal t-statistic. Other estimators such as the limited information maximum likelihood (LIML) estimator and the Fuller-k estimator are less biased, but they are still not fully robust[12]. We can still test the hypothesis of zero effect and propose the corresponding confidence set. The tests we use are the Anderson-Rubin (AR) test when there is only one instrument and the conditional likelihood-ratio (CLR) test when there are multiple instruments. Both tests are unbiased in the presence of weak instrument. A brief introduction of those two tests and other weak-instrument-relevant issues can be found in Appendix D. Table 8 reports the non-parametric estimation of the effects of individual schools on exam scores.

Table 8: LATE Estimation of the Effects of Individual Schools on SEEC

	Art Track	**Science Track**

[12] LIML and Fuller-k estimators have higher tolerance to weak instrument, especially when the number of instrument variables is large. Their superiority is not remarkable when there is only one instrument.

	Effect 2SLS	Wald Confidence Interval	AR Test	Effect 2SLS	Wald Confidence Interval	AR Test
	RKD					
No.1	0.369 [948]	[-0.801, 1.539]	0.479	-0.402 [375]	[-9.251, 8.446]	0.929
Xinhua	-0.466 [1584]	[-1.900, 0.968]	0.440	0.012 [541]	[-0.752, 0.775]	0.976
No.2	0.296 [1623]	[-0.549, 1.142]	0.523	-0.234 [684]	[-1.620, 1.153]	0.696
No.3	-0.122 [1431]	[-1.255, 1.011]	0.830	-3.023 [656]	[-11.04, 4.997]	0.000 [-83.23, -0.934] ∪ [0.995, 77.18]***
No.5	0.053 [1347]	[-0.266, 0.372]	0.764	-1.661 [693]	[-5.049, 1.727]	0.012 [-54.72, -0.067] ∪ [2.486, 51.40]***
No.8	0.141 [1085]	[-0.837, 1.118]	0.813	-3.226 [766]	[-14.90, 8.444]	0.003 [-119.9, -0.186] ∪ [3.088, 113.5]***
Xingda	0.970 [995]	[-0.960, 2.900]	0.152	-1.467 [752]	[-3.785, 0.851]	0.013 [-37.76, -0.085] ∪ [1.806, 34.83]***
Yufa	1.203 [188]	[-7.689, 10.10]	0.796	2.596 [379]	[-0.522, 5.715]	0.000 [-46.24, -1.808] ∪ [0.933, 51.43]***
Weishanzhuang	0.060 [209]	[-1.399, 1.518]	0.934	-1.301 [341]	[-4.322, 1.720]	0.006 [-31.51, -0.635] ∪ [0.091, 28.91]***
Caiyu	-0.243 [216]	[-1.046, 0.559]	0.514	-1.072 [325]	[-7.426, 5.282]	0.338
	RDD					
No.1	-3.899 [513]	[-8.388, 0.590]	**0.000 [-41.57, -2.012]***	-3.198 [232]	[-11.98, 5.579]	0.560
Xinhua	-0.337 [1005]	[-1.544, 0.871]	0.510	-0.071 [519]	[-0.551, 0.409]	0.772
No.2	0.513 [1028]	[-0.509, 1.536]	0.206	0.334 [444]	[-0.411, 1.080]	0.353
No.3	-0.549 [982]	[-1.821, 0.722]	0.559	0.960 [498]	[-2.831, 4.751]	0.488
No.5	0.006 [993]	[-0.536, 0.547]	0.984	0.584 [597]	[-0.405, 1.572]	0.299
No.8	-0.027 [999]	[-1.895, 1.841]	0.977	-2.008 [491]	[-6.549, 2.533]	0.076 [-47.41, -0.097] ∪ [4.453, 43.40]*
Xingda	0.266 [697]	[-1.522, 2.054]	0.758	-0.701 [406]	[-2.476, 1.075]	0.421
Yufa	2.486 [179]	[-14.71, 19.68]	0.751	2.468 [233]	[-2.237, 7.173]	0.091 [-44.58, -5.735] ∪ [0.111, 49.52]*
Weishanzhuang	0.019 [166]	[-1.130, 1.168]	0.974	13.46 [216]	[-333.5, 360.5]	0.264
Caiyu	0.201 [164]	[-0.646, 1.048]	0.673	0.255 [188]	[-0.965, 1.475]	0.619

Note: Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The number of observations is in brackets. The AR test is robust to heterogeneity and clustering on school. We first report the p-value of the AR test and then report the significant confidence interval if the null hypothesis of zero effect is rejected.

For the art track, most of the results are insignificant. Nevertheless, we have weak evidence showing that the most selective school – No. 1 School – actually decreases the exam performance for art students by at least two standard deviation. For students in the science track, although some less selective schools have a significant effect on the SEEC, we are not able to determine the direction of the effect for any of them. We also analyze the robust effects of individual schools on other outcomes. The results are shown in Appendix Tables 12 and 13. We do not have many significant and informative results. The most noteworthy one is that the No. 1 school increases the probability of being qualified for all of the three levels of colleges.

5. Linking Exam Schools with Student Achievement

In this section we analyze the relation between exam schools and student achievement, in hopes of finding out the possible reasons why exam schools do not have strong positive effects on achievement. We explore two channels: peer quality and schooling quality. We focus on the most selective model schools; the results for the other exam schools are shown in the Appendix.

5.1 Peer Quality in Exam Schools

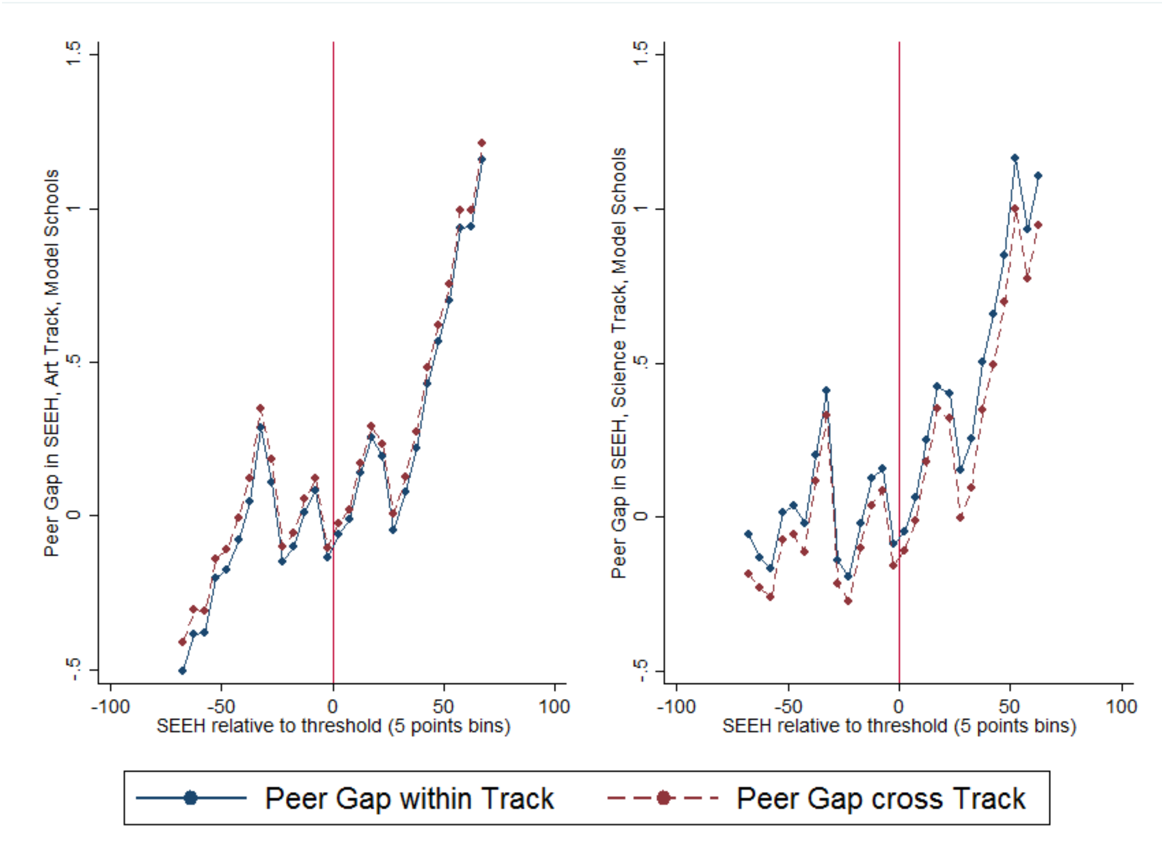
Peer quality is likely to jump when students transit from middle school to high school under the admission rule described in Section 3.2, especially in the most selective schools. According to the Big-Fish-Little-Pond effect (BFLPE) by Marsh, Chessor, Craven and Roche (1995), such a jump in peer achievement is perceived to have a negative effect on the student's own achievement. The marginal students in the model schools are high-achieving students in general; they are usually at the top percentage in their middle school classes. Since they are accepted at the margin, compared with their classmates in high school, they actually are the weaker students in the model school and are the least likely to rise to the top. Such a dramatic drop in relative rankings at both the classroom and school levels can harm the academic self-perception of students in these programs and affect their performances on exams.

We use the following peer gap to denote the relative ranking of a student:

$$P_{ij} = A_{ijt'} - \bar{A}_{ijt'}$$

where $\bar{A}_{ijt'}$ is the average SEEH of the same track at school j . Since students in different tracks are on the same campus, to take into account potential interactions between tracks, we also define a cross peer gap by replacing $\bar{A}_{ijt'}$ with the average SEEH over both tracks at school j. Figure 8 illustrates how peer gaps P_{ij} change with SEEH. There seems to be a certain change in peer gaps in both tracks. Panel A of Appendix Table A14 shows numerical ITT estimations for all schools. Compared with the probability of enrollment, the peer gap shows more significant discontinuity at the cutoffs of individual schools, and the patterns are consistent with the finding on the probability of enrollment. Students in No. 1, No. 3 and No. 5 schools experienced increased peer gap while students in other schools (except for No. 8) experienced decreased peer gap.

Figure 8: Peer Gap in SEEH, Model Schools

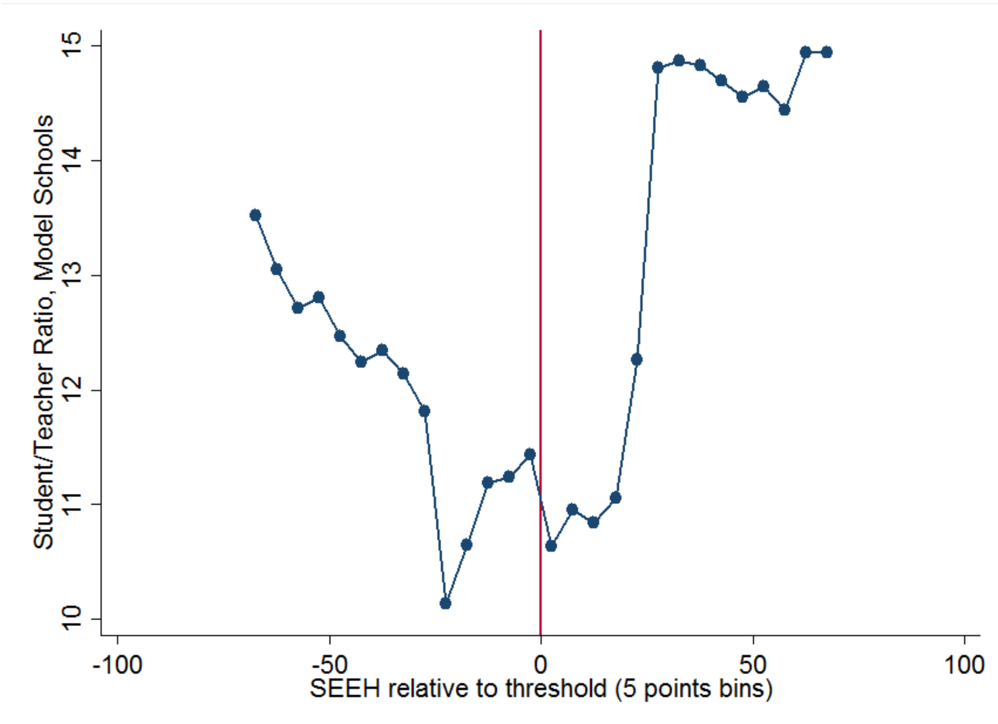


Note: The figure presents the raw SEEH in five point wide bins. Standardized SEEC scores and the probability of enrolling are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC are dropped.

5.2 Schooling Quality of Exam Schools

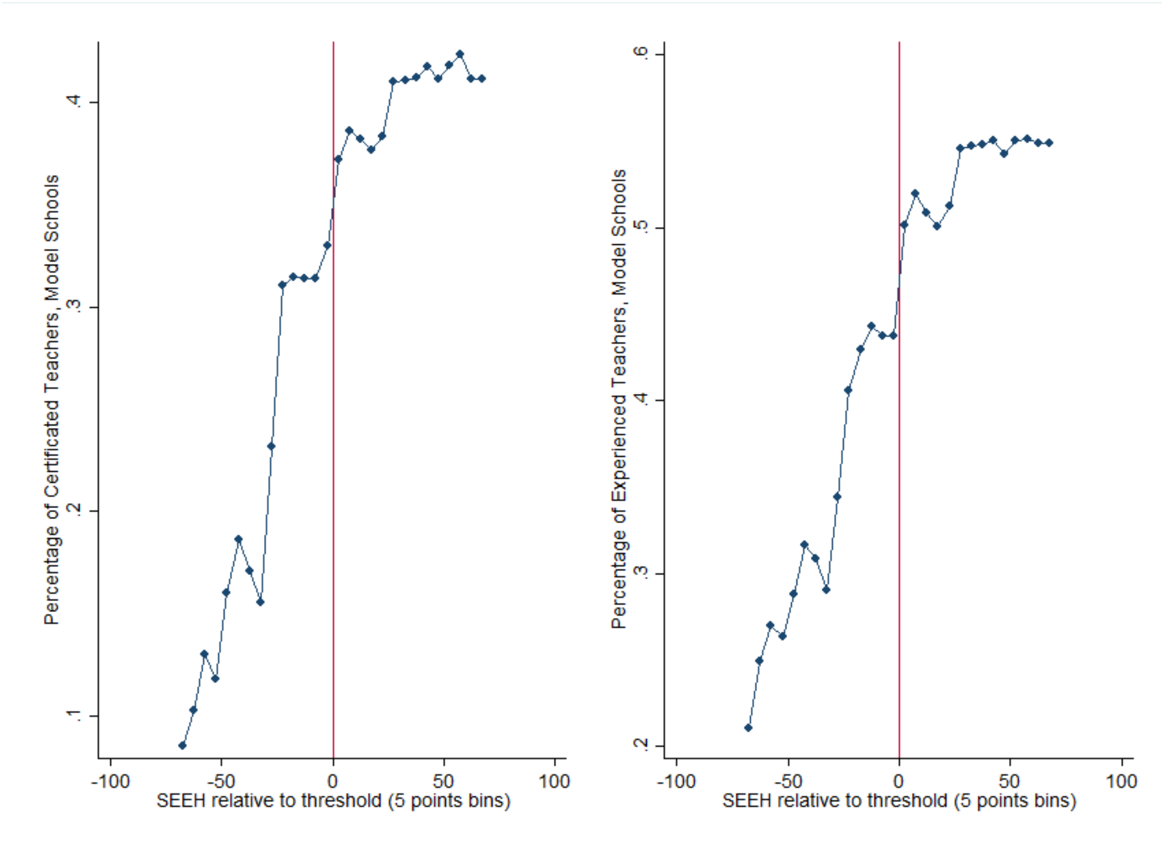
The most selective exam schools are perceived to have better facilities, more experienced teachers and more advanced curricula. We examine three indicators of schooling quality: student/teacher ratio, percentage of teachers with an advanced certificate and the percentage of teachers older than 35. The changes in those variables at the cutoff of model schools are illustrated in Figure 9, and the details of the numerical estimation for individual schools are in Panel B of Appendix Table 14. Overall we find significant change in the student/teacher ratio at the cutoffs of the most selective school and some less selective schools in the urban area, but we only find pieces of evidence of the change in the other two measures of teacher quality in less selective schools. The change in school quality measures are not as remarkable as the change in peer quality.

Figure 9: Student/Teacher Ratio in Model Schools



Note: The figure presents the raw SEEH in five point wide bins. Standardized SEEC scores and the probability of enrolling are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC are dropped.

Figure 10: Percentage of Certificated and Experienced Teachers, Model Schools



Note: The figure presents the raw SEEH in five point wide bins. Standardized SEEC scores and the probability of enrolling are presented within 70 points from the cutoff. Each dot represents the average outcome of the corresponding bin. Students with missing school identity, SEEH and SEEC are dropped.

5.3 Effect of Peer Quality and School Quality on Achievement

The treatment of model schools and some other exam schools includes higher peer achievement, better school quality and other unobserved determinants of achievement. The estimated effect of those schools on achievement is likely to be a mixed result, which can be decomposed into the contributions from each determinant. The effect of attending model schools can be derived from the treatment effect (TOT or LATE) as follows.

$$\gamma = \frac{\partial y_{ij}}{\partial H_{ij}} = \sum_{k=1}^{K} \frac{\partial y_{ij}}{\partial U_{ij}^{k}} \frac{\partial U_{ij}^{k}}{\partial H_{ij}},$$

where $y_{ij} = y_{ij}(U_{ij}^1, U_{ij}^2, \dots, U_{jt}^K)$ is the achievement equation, U_{ij}^k is the k-th educational input in school j (or model schools) for student i, and in total there are K types of educational inputs.

Peer achievement and school quality are educational inputs that are directly affected by eligibility for model schools. There could be other inputs that are also affected[13]. Pooling those unobserved inputs together, we have the following decomposition:

$$\gamma = \frac{\partial y_{ij}}{\partial H_{ij}} = \frac{\partial y_{ij}}{\partial P_{ij}}\frac{\partial P_{ij}}{\partial H_{ij}} + \sum_{k=1}^{3}\frac{\partial y_{ij}}{\partial Q_{ij}^k}\frac{\partial Q_{ij}^k}{\partial H_{ij}} + \frac{\partial y_{ij}}{\partial U_{ij}}\frac{\partial U_{ij}}{\partial H_{ij}},$$

where P_{ij} is the peer gap, and, in school district j for student i, Q_{ij}^k denotes the student/teacher ratio, the percentage of teachers with an advanced certificate and the percentage of teachers older than 35, with $k = 1{,}2$ and 3 respectively. U_{ij} denotes all of the other unobserved inputs that are affected by eligibility for model schools. If we assume that those unobserved inputs can only be indirectly affected by eligibility for model schools through the observed inputs,

$$\frac{\partial U_{ij}}{\partial H_{ij}} = \frac{\partial U_{ij}}{\partial P_{ij}}\frac{\partial P_{ij}}{\partial H_{ij}} + \sum_{k=1}^{3}\frac{\partial U_{ij}}{\partial Q_{ij}^k}\frac{\partial Q_{ij}^k}{\partial H_{ij}},$$

then the estimated effect is the sum of direct effect of observed inputs and indirect effects through unobserved inputs. We can decompose the effect of total expenditure in a similar way into the effect of construction capital and other expenditures[14].

[13] Those unobserved factors may include student self-perception and parent expectation. One of the main goals of those more selective schools is to place as many students as possible in top universities. It is very likely that students at different positions in the achievement distribution are treated differently, e.g., teachers may pay more attention to students in the top percentage of the class. There is research from which we can infer the goals of high schools. Deng and Treiman (1997) provides a historical view of China's education system. Ding and Lehrer (2007) introduce the current educational secondary education system in Jiangsu Province, which is very similar to the one in Beijing.

[14] For the unobserved inputs which also have direct effects on achievement, e.g., facilities and curriculum, the assumption does not hold, and we get omitted inputs bias. Nevertheless, it is very likely that such bias is positive, as the more selective schools tend to have better unobserved inputs. In this case our estimated effects of the observed inputs on achievement serve as upper bounds.

To get the effect of observed inputs we estimate Equation 10, using a set of indicators of the first five schools eligibility as instrument variables. We estimate models with single observed inputs and the model with four observed inputs. The results are presented in Table 9.

Table 9: LATE of Peer Quality and School Quality on SEEC

	Art Track					**Science Track**				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
	RDD									
Peer Gap, Own	-0.34 [0.59]	-	-	-	-0.38 (0.43)	-0.36 [0.33]	-	-	-	-1.03 (0.64)
Student/Teacher Ratio	-	-0.05 [0.80]	-	-	-0.03 (0.08)	-	0.00 [0.06]	-	-	-0.06 (0.13)
Percentage of Teachers with Advanced Certificate	-	-	1.05 [0.62]	-	4.57 (8.14)	-	-	-0.21 [0.08]	-	-11.9 (9.70)
Percentage of Teachers Older than 35	-	-	-	0.34 [0.68]	-4.33 (6.24)	-	-	-	0.31 [0.10]	6.99 (7.98)
No. of Observations	2272	2272	2272	2272	2272	1290	1290	1290	1290	1290
	RKD									
Peer Gap, Own	-0.85 [0.98]	-	-	-	1.88 (14.7)	0.01 [0.62]	-	-	-	-1.11 (1.19)
Student/Teacher Ratio	-	-0.07 [0.05]	-	-	-0.67 (3.04)	-	0.02 [0.74]	-	-	-0.17 (0.16)
Percentage of Teachers with Advanced Certificate	-	-	2.39 [0.54]	-	-0.59 (87.4)	-	-	-0.79 [0.67]	-	-16.0 (18.8)
Percentage of Teachers Older than 35	-	-	-	1.19 [0.00]	-22.2 (104)	-	-	-	-0.81 [0.76]	8.37 (11.9)
No. of Observations	2272	2272	2272	2272	2272	1290	1290	1290	1290	1290

Note: Parametric estimations control for a cubic function of the standardized SEEH. The standard error which is robust and clustered on school is shown in parentheses. The p-value of the CLR test is shown in brackets. We use indicators of eligibility for the No.1, Xinhua, No.2, No.3, and No.5 schools as instruments.

In columns (1) to (4) and (6) to (9) we estimate the models with one observed input at a time, and in Columns (5) and (10) we estimate the models with multiple observed inputs. We are not able to find any significant peer effects. This is consistent with the findings by Abdulkadiroglu et. al. (2014). In most cases the results show little effect of the three measures of school quality on outcomes. Some effects are significantly different from zero, but we are not able to determine the direction of those effects because of the weak instruments. We also conduct the analysis with respect to a single input for each individual school. The results are summarized in Appendix Table A15. We find mixed peer effects; we find no effect in those selective schools, positive effects in the less selective schools in the center area but negative effects in the schools in rural towns. The mixed peer effects may be due to different omitted variables or different strategies of schools that treat students differently according to their relative position in the distribution of achievement. For school quality measures, overall we do not find consistently strong evidence of positive effects of any of the quality indicators on the SEEC.

Although for some schools such as the most selective ones and those in the rural towns, the peer achievement and other school quality measures change remarkably, different inputs which include unobserved ones seem to work in different directions and cancel out each other. As a result, we do not find strong evidence of significant effects of individual school inputs on student achievement.

6. Sensitivity Analysis

We check whether our main results are robust to controlling for different polynomials of the SEEH, optimal bandwidths calculated with different criteria, different weighting variables and the identification strategy using the kink only. Table 10 shows the results for the art track, and the Appendix Table A16 shows the results for the science track.

Table 10: Robustness Checks, School Groups, Art Track

			Polynomial		**Optimal Bandwidths**		
	AR Test (1)	**Uniform Weight (2)**	**Quadratic (3)**	**Cubic (4)**	**IK (5)**	**CV (6)**	**Distance to the nearest other cutoff (7)**
	RDD						
Model School	0.510	-0.246 (0.652) [1005]	-0.358 (0.538) [1005]	-0.453 (0.466) [1005]	-0.229 (0.532) [1998]	-0.214 (0.572) [1902]	0.783 (0.240) [346]
Center Area School	0.758	0.063 (0.875) [697]	0.013 (0.984) [697]	0.102 (0.940) [697]	-0.232 (0.466) [2268]	-0.229 (0.479) [2267]	11.88 (0.636) [293]

Center Area Selective School	0.206	0.329 (0.330) [1028]	0.497 (0.186) [1028]	0.720 (0.134) [1028]	0.276 (0.205) [1869]	0.270 (0.202) [1887]	0.365 (0.366) [352]
Center Area Less Selective School	0.758	0.054 (0.875) [697]	0.012 (0.984) [697]	0.144 (0.940) [697]	-0.116 (0.466) [2268]	-0.113 (0.479) [2267]	2.241 (0.636) [293]
	RKD						
Model School	0.440	-0.337 (0.534) [1584]	-0.247 (0.570) [1584]	-0.526 (0.409) [1584]	-0.001 (0.998) [2241]	-0.127 (0.678) [2213]	0.850 (0.257) [346]
Center Area School	0.152	0.411 (0.290) [995]	0.083 (0.897) [995]	0.274 (0.545) [995]	-2.339 (0.122) [1949]	-7.121 (0.115) [1828]	3.907 (0.824) [293]
Center Area Selective School	0.523	-0.225 (0.456) [1623]	0.271** (0.043) [1623]	0.414* (0.059) [1623]	0.212 (0.447) [2246]	0.160 (0.573) [2234]	0.448 (0.271) [352]
Center Area Less Selective School	0.152	0.367 (0.290) [995]	0.073 (0.897) [995]	0.280 (0.545) [995]	-0.281 (0.122) [1949]	-0.329 (0.115) [1828]	0.809 (0.824) [293]

Note: The p-values of the AR test are shown in parentheses. The number of observations is in brackets. For the purpose of comparison, in Columns (2), (3) and (4) we use the same bandwidth as the main model, which is not necessarily the optimal bandwidth under the current settings. We also calculate the new optimal bandwidths and perform the analyses with them. We get quite similar results, which are available upon request.

Column (1) shows the p-value of the AR test for school groups. The results confirm the insignificant effects in Table 6. In Column (2) we estimate the model using a uniform weight variable, which assigns equal weight to each student regardless of the distance between his/her SEEH and the corresponding cutoff. Columns (3) and (4) show the results when we estimate the model by controlling for quadratic and cubic functions of SEEH rather than a linear polynomial[15]. We also re-estimate the model with different bandwidths –the IK optimal bandwidth by Imbens and Kalyanaraman (2012) and the CV optimal bandwidth used in Ludwig and Miller (2007) in Columns (5) and (6) respectively. Finally in Column (7) we restrict the bandwidth to the distance to the nearest neighborhood cutoff. This excludes the potential effect of other cutoffs. All of those robustness checks provide us similar results which are consistent with our main ones. The robustness checks for the science track also show that our main results

[15] Ganong and Jager (2014) find that the cubic specification is more desirable than the linear and quadratic specifications in a RKD study with curvature in the global function of SEEH.

are not sensitive to the choice of weighting variable, polynomial or bandwidth. We perform the same set of analyses for individual schools. The results are consistent with the main findings and are available upon request.

7. Heterogeneous Effect

We explore the possibility that our effects can be a function of gender and parental education of the student. It is commonly believed in China that boys are relatively good at science and girls are relatively good at art. In our sample girls account for a larger portion in the art track than the science track. It is possible that the effects of exam schools are different for boys and girls, and the gender difference in effects largely depends on the track.

Effects of model schools may also vary among students from different family backgrounds. We focus on two types of families: parents with college degree and parents who are farmers. In developing countries students from farm families are disadvantaged in studying and are perceived to have lower achievement than other students. However, it is not necessarily the case that the students from farm families benefit less from attending model schools.

Table 11: Heterogeneous Effects, Art Track

	Gender		**Parental Education**		**Parental Occupation**	
	Girl	**Boy**	**College Degree**	**No College Degree**	**Agricultural**	**Non-Agricultural**
	RDD					
Model School	-0.520 (0.457) [541]	-0.326 (0.625) [464]	7.312 (0.446) [204]	-0.180 (0.643) [800]	0.096 (0.898) [424]	-2.226 (0.016) [496]
Center Area School	-0.129 (0.839) [362]	0.489 (0.222) [335]	-0.301 (0.758) [181]	0.007 (0.993) [490]	0.231 (0.705) [299]	-93.87 (0.479) [135]
Center Area Selective School	0.462 (0.364) [550]	0.601 (0.135) [478]	1.327 (0.151) [274]	0.192 (0.421) [753]	0.750 (0.466) [404]	0.510 (0.168) [463]
Center Area Less Selective School	-0.122 (0.839) [362]	0.462 (0.222) [335]	-0.204 (0.758) [181]	0.007 (0.993) [490]	0.231 (0.705) [299]	-1.124 (0.479) [135]
	RKD					
Model School	-0.574 (0.435)	-0.374 (0.435)	-1.200 (0.538)	-0.319 (0.221)	-0.151 (0.883)	-0.885 (0.244)

	[811]	[773]	[404]	[1169]	[690]	[675]
Center Area School	0.182 (0.568) [528]	0.570** (0.004) [467]	-0.000 (1.000) [258]	0.297 (0.320) [706]	0.332 (0.145) [427]	4.117 (0.222) [237]
Center Area Selective School	0.151 (0.557) [843]	-0.007 (0.978) [780]	0.419*** (0.001) [388]	-0.019 (0.912) [1220]	-0.092 (0.687) [710]	0.575*** (0.010) [666]
Center Area Less Selective School	0.169 (0.568) [528]	0.553*** (0.004) [467]	-0.000 (1.000) [258]	0.298 (0.320) [706]	0.335 (0.145) [427]	-2.121 (0.222) [237]

Note: The standard errors which are robust and clustered on school are shown in parentheses. The number of observations is in brackets. The bandwidth used is one standard deviation of the SEEH. Bold and italic font indicates that the AR test and the corresponding confidence interval are consistent with the results.

Table 12: Heterogeneous Effects, Science Track

	Gender		**Parental Education**		**Parental Occupation**	
	Girl	**Boy**	**College Degree**	**No College Degree**	**Agricultural**	**Non-Agricultural**
	RDD					
Model School	-1.247* (0.023) [155]	0.316 (0.247) [364]	6.375 (0.502) [149]	-0.169 (0.196) [369]	-0.298 (0.504) [180]	-0.234 (0.343) [281]
Center Area School	-0.646 (0.562) [135]	-0.199 (0.723) [271]	-2.190 (0.168) [88]	0.041 (0.951) [282]	-0.439 (0.552) [184]	-10.69 (0.230) [90]
Center Area Selective School	0.862 (0.176) [134]	0.179 (0.615) [310]	0.683 (0.551) [133]	0.238 (0.524) [310]	24.10 (0.002) [145]	0.050 (0.839) [231]
Center Area Less Selective School	-0.742 (0.562) [135]	-0.252 (0.723) [271]	-1.635 (0.168) [88]	0.056 (0.951) [282]	-0.533 (0.552) [184]	9.686 (0.230) [90]
	RKD					
Model School	-1.309* (0.095) [163]	0.467 (0.269) [378]	-9.485 (0.445) [156]	-0.145 (0.566) [384]	-0.681 (0.439) [187]	0.018 (0.963) [290]
Center Area School	-1.104 (0.458) [252]	-0.621** (0.054) [500]	-9.667 (0.157) [176]	-0.241 (0.520) [513]	-0.392 (0.180) [313]	5.868 (0.051) [219]
Center Area Selective School	-0.232 (0.340) [220]	0.022 (0.933) [464]	-0.088 (0.827) [214]	-0.075 (0.789) [468]	0.019 (0.957) [234]	-0.062 (0.825) [331]
Center Area Less Selective School	-1.089 (0.458) [252]	-0.731* (0.054) [500]	-2.331 (0.156) [176]	-0.317 (0.520) [513]	-0.440 (0.180) [313]	16.90 (0.051) [219]

Note: The standard errors which are robust and clustered on school are shown in parentheses. The number of observations is in brackets. The bandwidth used is one standard deviation of the SEEH. Bold and italic font indicates that the AR test and the corresponding confidence interval are consistent with the results.

Overall the difference in effects is insufficient for us to draw any strong conclusions that are different from the main results; attending more selective schools does not improve performance on the college entrance exam. Nevertheless, the results show heterogeneous patterns in some specifications. Boys in the art track are more likely to benefit from attending less selective schools. Students in the art track who are from families that are relatively more advanced in terms of social class, indicated by parental college degree and non-agricultural occupation, are likely to benefit from attending the most selective exam school. For the science track while there no positive effects are found, we find gender heterogeneity in returns to attending exam schools in many specifications, but those evidences are insufficient for us to draw a general conclusion about which gender get more from attending exam schools. Those findings provide us potential directions in which the exam schools can work as expected.

8. Conclusion

While selective exam schools that admit students solely by exam scores are established in many countries across the world in the hope of improving student outcomes, there is less consensus among researchers about the effect of selective exam schools on student achievement. In the United Kingdom (Clarke, 2010), Romania (Pop-Eleches and Urquila, 2013) and Trinidad and Tobago (Jackson, 2010) exam schools are shown to have positive effect on student achievement while in the United States (Abdulkadiroglu et al., 2014; Dobbie and Fryer, 2014) and Kenya (Lucas and Mbiti, 2014) there is no evidence of positive effect of exam school.

Using regression discontinuity (kink) design, our paper examined the effect of exam school in China, including model schools, which are of higher quality and more selective than other schools. However, we find modest to no effect of these schools on test score performance. What is even worse is that we find significant negative effects among several subgroups. Selective exam schools have higher-quality peers, but there are also other omitted inputs. The mixture of various inputs lead to insignificant effects of exam schools. The non-positive effects are more prevalent for girls. The evidence of heterogeneous effects by parental education and occupation is mixed. These findings cast doubt on the policy of labeling schools by quality. The quality of schools is more important in improving student test scores than the model school label which in the extreme may only serve as a signal for positive ability matching between students and schools[16].

Our study does not prove that selective exam schools does not benefit students at all, because of several noteworthy points. First, the primary outcome we examine in this paper is the college exam score, which is only one of many important short-run outcomes. While we find some positive effects on other exam-relevant outcomes, such as the probabilities of taking an exam and qualifying for college admission, exam schools can help students in other ways that are not evaluated on exams. For example, students in more selective schools are more likely to make high-quality friends and can benefit from the higher level of networks they enter. Second, the RDD (RKD) can only provide us the estimated effect of model schools on a subsample of students - compliers at the cutoff. These are the students whose SEEH is at the cutoff and will surely attend model schools if they are eligible. They will never attend if their SEEH is below the cutoff. Using the RDD (RKD) framework, we can say nothing precise about the effect on other students. It is possible that other students in the class can benefit from exam schools; Duflo et al. (2011) find that the teachers in elite schools may pay more attention to median students than to students who are marginally admitted. Finally, because of the data limitations we are not able to identify the effect of other unmeasured school quality indicators, which may have significant effects on student achievement. We are not able to examine the effect of potential change in student attributes including psychological or behavioral change during the transition to selective high schools. Those unmeasured inputs may work through different channels. These other effects

[16] Even the signal may work only in the short-run. In the long-run parents and students can recognize the true quality by other indicators, such as the performance of students on exams in previous years.

are important for the overall evaluation of exam school policy. Nevertheless, our results are disappointing for the parents who wait outside the building holding the entrance exam to high school. There is no conclusive evidence of that those selective exam schools work as expected to improve the score of entrance exam to college. For those parents such score is probably the most important dimension of achievement related to high school as the score directly determines the ability to get into a prestigious college.

Appendix

A. Exam, Admission and Track Rules

A1 Entrance Exam to High Schools

Students in middle schools participate in an entrance exam if they want to be admitted to a high school. The entrance exam to high school is held once per year, in late June. The score serves as the only criterion for enrollment in high schools with two exceptions. One is that students with excellent awards, such as "Jin Fan" and "Yin Fan" awards, are able to gain admission without taking the exams[17]. They are not included in our data because we drop observations whose SEEH is missing. The other is that students who satisfy one of several personal conditions, such as being a minority race or the child of a martyr, can obtain additional scores. They are included as non-compliers. We do not know which students fall into these two groups. The exam consists of six sub-exams covering Chinese, Mathematics, English, Physics, Chemistry and Physical Education. The maximum scores are 120, 120, 120, 100, 80 and 30 respectively, or a maximum total SEEH of 570.

A2 Admission to High School and Track Choice

The application and admission procedures are outlined in the main text. Here we list some complementary notes. There is also a rule to rank students with the same score. First, some pre-announced priorities are considered, such as being the child of a serviceman or diplomat. If some remaining students still have the same score, the students are ranked by scores in Mathematics, Chinese and English sequentially. Finally, students are ranked by lottery number, with the lowest numbers selected first.

[17] Those awards are for distinguished achievement in competitions of techniques, arts or sports. They are awarded to students in elementary, middle and high schools. In our study they can be ignored because few students can earn them; in 2005 14 middle school students were rewarded and only one of them came from Daxing district.

Students who are admitted by a school are not able to reject the offer, unless they decide to leave the system of public schools. For example, if a student lists school A as the top and is admitted by school A, he cannot regret and choose to attend school B which is ranked lower by him, even if his score is higher than the cutoff of school B. The reason is that, when admitted by school A, the student is moved out of the pool, and none of the other schools is able to admit him. When a student does not attend the schools for which he is also qualified, the only reason is that he does not list those schools before the one he attends in his reported preference ranking.

In middle school there is no track difference; students in middle schools take the same courses. The subjects covered in the entrance exam to high school are also the same for all students. In high school students still take the same courses until the end of the first year, when students indicate their track preferences. Before the second year, based on their preferences, students are re-allocated to science or art classes, where the courses and exam materials can differ.

A3 Entrance Exam of Colleges

The entrance exam for colleges in Beijing is also held each June. The subjects covered in the exam are different for the two high school tracks. The exam for the art track includes Chinese, Mathematics for Art, Foreign Language and Integrated Art, while the exam for the science track includes Chinese, Mathematics for Science, Foreign Language and Integrated Science. Integrated Art combines History, Politics and Geography while Integrated Science combines Physics, Chemistry and Biology. The full scores for Chinese, Mathematics for Art or Science and Foreign Language are all 150. Full score for each integrated subject is 300. Thus the full score for the exam for both tracks is 750.

Students satisfying certain personal conditions, such as minority race, can obtain additional scores in the admission phase. However this does not affect the SEEC; the score serves as the only criterion for enrollment in colleges in most cases. There is also a complicated admission process for enrollment in colleges. However, this process has little to do with the SEEC and other outcomes related to the SEEC and goes beyond the scope of our study.

A4 Model and Regular Schools

There were 68 model schools in Beijing in 2008. This kind of high school originated from the key schools policy started in the 1950s when the government allocated more of its limited resources to certain selected schools in hopes of improving education quality and student achievement. Model schools replaced key schools in the 1990s. Unlike key schools, model schools focus on multiple outcomes rather than exam scores. Because the admission to college largely depends on exam scores, the pedagogy of model schools is still restricted to the scope of the test. All schools other than model schools are classified as regular schools. The differences between the two types of schools can be summarized as follows. First, model schools receive more government financial support. They can also collect more money from external funding because of their excellent reputations. With more funding, model schools can support better learning conditions such as newer facilities, higher-quality teachers, more exchange opportunities and more advanced instructional equipment. Model schools appeal to most middle school students, if not all of them, and many excellent students list them as the top choice.

No.1 High School was founded in 1956 and appointed as the only key school of Daxing District in 1978. In 2002 it was selected into the first batch of model schools of Beijing City. No.1 High School is equipped with modern educational facilities, such as a standard stadium and a library with hundreds of thousands of books. More than 20 percent of the teachers have at least a master's degree. It is not surprising that No.1 High School has great outcomes; on average 95% of its graduates are admitted to college, and 30% of them go to key universities.

Xinghua High School was a normal school for teachers after its establishment in 1949 and changed to a high school in 1997. It was selected as a model school of Beijing City (the second model school in Daxing District) in 2005. Although it operated as a high school for only 10 years, Xinghua High School also did well on the outcomes of interest. In recent years 90% of the students at Xinghua High School were able to go to college. No.2 High School was founded in 1972 and became the affiliated high school of Beijing Normal University in 2006. Historically No.2 High School was not one of the top high schools in the district. Recently the school rose rapidly in student achievement. In 2000 only 29% of students were admitted to college, but in 2009 the proportion increased to 80%. Even though it is still a regular high school, the gap between it and other model schools is narrowing

B. Manipulation in the Boston Mechanism

It is shown that in the Boston Mechanism, parents/students are likely to be sophisticated, which means they tend to choose not report their preference truthfully. In this appendix we will illustrate how parents/students manipulate in a gaming framework. Suppose we have the following settings about the game.

1. There are I students and M high schools.
2. Each school has capacity of c_m seats and in total there are $N \geq \sum_{m=1}^{M} c_m$ students.
3. Each student has a strategy space $\theta_i = \{r_{ik}\}$, where r_{ik} is a ranking of schools.
4. Each school rank students by exam score in its priority list $q_m = \{r_m\}$.
5. School does not have payoffs.
6. Each student has a payoff $u_i = u(r_i, r_{-i}, s_i, s_{-i})$.

Suppose there are two types of students: sophisticated and sincere. For sophisticated students, the strategy space is a set of permutations of all possible ranks, thus $\theta_i^{so} = \{r_{ik}\}_{k=1}^{K}$. For sincere students, the strategy space only includes one element which is the true preference, thus $\theta_i^{si} = \{r_{ik}\}_{k=t}$. Therefore, at the equilibrium, sincere students report the true preference as they always do so, but it is not sure whether sophisticated students will also report their true preferences. By distinguishing sophisticated and sincere students, we only allow sophisticated students to strategically respond to others' choices.

Without loss of generality, we impose the following assumptions[18].

1. The quality of school is well known for everyone. Thus school quality is common knowledge, especially for those elite schools.

[18] The existence of manipulation does not depend on those assumptions. We impose them only for the purpose of illustration.

2. Schools are distinguishable. Therefore, we exclude the possible of tied schools in the student preference.
3. There are exactly $N = \sum_{m=1}^{M} c_m$ students, and N is large enough, thus each student can be affected by the choice of others.

With those assumptions, the following simple example of a static game provides us some insight about the manipulation in the Boston mechanism.

Example: Suppose there are three schools, a, b and c, each with only one seat and there students i_1, i_2 and i_3. The exam scores are A_1, A_2 and A_3, and $A_3 > A_2 > A_3$. Therefore, each school has the same priority list $q_m = \{i_3, i_2, i_1\}$. We have school quality s_1, s_2 and s_3, and $s_1 > s_2 > s_3$. Suppose school quality is the only factor that determines the preference of student, then the student utilities representing preferences $u_i = (u_{i1}, u_{i2}, u_{i3})$ are as follows:

	a	**b**	**c**
$\boldsymbol{u_{i1}}$	3	2	1
$\boldsymbol{u_{i2}}$	3	2	1
$\boldsymbol{u_{i3}}$	3	2	1

Thus three students have the same preference over schools and their payoffs of being assigned to each school are also the same.

Student 1 and 2 are sophisticated and student 3 is sincere[19]. Hence, the strategy space of student 1 and 2 are $\theta_1 = \theta_2 = \{abc, acb, bac, bca, cab, cba\}$ and the strategy space of student 3 is $\theta_3 = \{abc\}$. Then we have the following Boston game:

		Student 2					
		abc	**acb**	**bac**	**bca**	**cab**	**cba**
Student 1	**abc**	(1,2,3)	(2,1,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)
	acb	(1,2,3)	(2,1,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)

[19] This assumption indicates that if student 3 knows that his score is the highest, he/she does not need to manipulate at all. The assumption is reasonable because in practice the top school has a certain amount of seats. Students can predict their relative performance by many channels such as their previous performance, feedback/suggestion from teachers, and their own feelings. Therefore, at least for some best students, even though they are still not able to perfectly predict their ranks, they are quite sure that they can get in the top school. Thus there is no need for them to manipulate their reported preferences.

	bac	(2,1,3)	(2,1,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)
	bca	(2,1,3)	(2,1,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)
	cab	(1,2,3)	(1,2,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)
	cba	(1,2,3)	(1,2,3)	**(1,2,3)**	**(1,2,3)**	(2,1,3)	(2,1,3)

We do not list student 3 in the table as his/her strategies and payoffs are always the same ($\theta_3 = \{abc\}$ and $u_3 = 3$).

There are 12 Nash equilibria (in boldface) with the same equilibrium payoff profile $(1,2,3)$. We have the following useful remarks about the equilibria:

1. Student 1 does not care which strategy to choose, as all of them lead to the same payoff. Hence, as long as other students stay in the equilibria, he/she can freely choose strategies, there is no difference. This is the undesirable freedom for the student with the worst performance.
2. Student 2 does not choose the truth-telling strategy. He/she reports school b as the first choice. This is the manipulation of the majority of students whose exam scores are not so high or low that they can ignore others' choices.
3. Student 3 does not care others' strategy, no matter whether they stay in the equilibria or deviate. This is the desirable freedom for the student with the best performance.

Student 2 does not report the true preference because he/she does not want take the risk that he/she may fail to get a place in the second choice which would be available if he/she listed that school as the first choice. For example, if he/she reports the true preference, student 1 will strategically list school b as the top preference and crowd him/her out to school c. Therefore, for safety student 2 list school b as the top preference at the equilibria.

C. Choice of Distance and Quality

We propose a simple model to illustrate how students make choice between school quality and commuting distance in the Boston mechanism. Suppose we have the following settings:

1. There are M schools and each has quality Q_m, which can be observed[20], and capacity k_m.
2. For student i, the cost of attending school m is c_{im}. In this paper we focus on commuting cost and assume it is proportional to commuting distance.
3. For student i, the probability of being offered a seat from school m with quality Q_m is $p(Q_m, c_m, g_i, r_i, g_{-i}, r_{-i})$, where g_i is the exam score of student i, r_i is the reported preference of student i, g_{-i} and r_{-i} are the exam score and reported preference of students other than i. We focus on Q_m and c_{im}, thus leave all of the other relevant variables to a vector A. Moreover, we assume that $p'(Q) < 0$, $p'(g) > 0$, $p''(Q) > 0$.
4. If rejected by the top preferred school m, student i is offered a seat by other school j with the probability of $q^m(Q_j, c_j, g_i, r_i, g_{-i}, r_{-i})$. Moreover, we have $q'(Q) < 0$ and $q'(g) > 0$.

The optimization problem of student i can be written as follows:

$$\max_{Q} p(Q, A)(Q - c) + (1 - p(Q, g_i, A)) \sum_{j=-Q} q(Q^j, A^j)(Q^j - c^j)$$

where we also have the following assumptions:

1. Student i is risk neutral. The utility in each state is $u_i^m = Q_m - c_{im}$.
2. The total amount of schools M is large enough such that student i can make decisions regarding a continuous quality Q approximately.

[20] In practice people usually observe a number of proxies of school quality rather than school quality itself. However, people can always form a subjective expectation of school quality and optimize based on it.

We can loosen those assumptions, but this does not bring us more insights at the cost of more complication. Solving the problem, we have the following condition characterizing the optimal choice:

$$p(Q,A) = p'(Q,A)(\sum_{j=-Q} q(Q^j,A^j)(Q^j - c^j) - (Q - c))$$

Regarding the optimal choice the optimal condition above has the following implications:

1. $\sum_{j=-Q} q(Q^j,A^j)(Q^j - c^j) < Q^* - c^*$. The optimal top ranked school always leads to higher payoff than the expected payoff from other schools if the student is rejected by the top ranked school.
2. For a school Q to be a potential first choice, it must satisfies $\sum_{j=-Q} q(Q^j,A^j)(Q^j - c^j) < Q^* - c^*$. Therefore, given the characteristics of other schools, there exists a cutoff $\tilde{t}_m = \sum_{j=-Q} q(Q^j,A^j)(Q^j - c^j)$ for each school. When the utility of getting in school m, which is $u_m = Q_m - c_{im}$, is smaller than the cutoff $\tilde{t}_m$, there will be no student ranking m as the top choice.
3. If c_{im} increases, to have students still choose school m as the top choice, we also need the quality Q_m to increase. Therefore, given other conditions fixed, students are less likely to choose a school which is far away from home for a moderate increase in school quality.

In the Boston mechanism, introducing distance (thus cost) have ambiguous impacts on the manipulation behavior. To illustrate this, we sequentially add three sets of distance into the game discussed in section A2. The first set is $c_1 = \{c_{1a} = 1, c_{1b} = 2, c_{1c} = 0\}$, $c_2 = \{c_{2a} = 1, c_{2b} = 0, c_{2c} = 1\}$ and $c_3 = \{c_{3a} = 1, c_{3b} = 1, c_{3c} = 0\}$. We can see that student 3 will still truthfully report abc. Incorporating the distance (cost) into payoffs, we have the following modified game:

	Student 2					
	abc	**acb**	**bac**	**bca**	**cab**	**cba**

Student 1	**abc**	**(1,2,2)**	(0,0,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)
	acb	**(1,2,2)**	(0,0,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)
	bac	(0,0,2)	(0,0,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)
	bca	(0,0,2)	(0,0,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)
	cab	(0,0,2)	(0,0,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)
	cba	(1,2,2)	(1,2,2)	**(1,2,2)**	**(1,2,2)**	(0,0,2)	(0,0,2)

In addition to the original 12 Nash equilibria, now we have two more equilibria, in which student 2 tell his/her true preference. Moreover, in one equilibrium Student 1 will also tell the true preference. Thus we have one equilibrium where all of the students report their true preferences. (Note now the true preference of student 1 becomes acb). In the second set of distances we set $c_1 = c_2 = c_3 = \{1,1,2\}$. Then it is obvious that the equilibria remain the same and we still do not have an equilibrium where all of the students tell truth. The introducing of distance may also change the equilibria where at least one student tells the truth. Suppose $c_1 = \{c_{1a} = 0, c_{1b} = 0, c_{1c} = 0\}$, $c_2 = \{c_{2a} = 0, c_{2b} = 2, c_{2c} = 0\}$ and $c_3 = \{c_{3a} = 0, c_{3b} = 1, c_{3c} = 0\}$. Then the game becomes:

		Student 2					
		abc	**acb**	**bac**	**bca**	**cab**	**cba**
Student 1	**abc**	(1,0,3)	**(2,1,3)**	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**
	acb	(1,0,3)	**(2,1,3)**	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**
	bac	**(2,1,3)**	**(2,1,3)**	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**
	bca	**(2,1,3)**	**(2,1,3)**	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**
	cab	(1,0,3)	(1,0,3)	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**
	cba	(1,0,3)	(1,0,3)	(1,0,3)	(1,0,3)	**(2,1,3)**	**(2,1,3)**

Now in total we have 18 equilibria, among which there are 6 equilibria having at least one student telling the true preference. (Note: now the true preference of student 2 becomes acb).

D. Weak Instrument, Partial Robust Estimator, and Robust Test and Inference

Suppose we have the following regression:

$$y = bx + e$$

where y, x and e are $n \times 1$ vectors, and $cov(x, e) \neq 0$. Therefore, OLS estimator $\hat{b}_{ols}$ is biased. If we use instrument variable (IV) approach to solve the endogeneity problem, we have the following equations:

$$\begin{cases} x = cz + u \\ y = bx + e \end{cases}$$

where z is a $n \times 1$ vector and it serves as instrument variable. Therefore we have the reduced form as follows:

$$\begin{cases} x = cz + u \\ y = czb + bu + e = czb + v \end{cases}$$

where $v = bu + e$, thus $(v_i, u_i) \sim N(0, (\sigma_1^2, \rho\sigma_1\sigma_2, \rho\sigma_1\sigma_2, \sigma_2^2))$. The OLS estimator and 2SLS estimator can be derived as follows:

$$\hat{b}_{ols} = \frac{x'y}{x'x} = \frac{(c'z' + u')(czb + v)}{(c'z' + u')(c'z' + u')}$$

$$\hat{b}_{2sls} = \frac{x'z(z'z)^{-1}z'y}{x'z(z'z)^{-1}z'x} = \frac{(c'z' + u')P_z(czb + v)}{(c'z' + u')P_z(c'z' + u')}$$

where $P_z = z(z'z)^{-1}z'$.

It is shown in literatures that the 2SLS estimator is not necessarily less biased than the OLS estimator, although the 2SLS is consistent as long as $cov(x, z) \neq 0$. We have the following asymptotic biases of OLS and 2SLS:

$$plim\hat{b}_{ols} = b + corr(x, e)\frac{\sigma_2}{\sigma_x}$$

$$plim\hat{b}_{2sls} = b + \frac{corr(z, u)}{corr(z, x)}\frac{\sigma_2}{\sigma_x}$$

where $corr$ indicates correlation and σ_x is standard deviation of x. We can see that if the endogenous variable and the instrument variable are only weakly correlated, it is possible that we have larger asymptotic bias from 2SLS. In an extreme case, if $corr(z, x) = 0$, which means the endogenous variable and the instrument variable are not correlated at all, the asymptotic bias of 2SLS is the same as the one of OLS. Moreover, in this extreme case, $\hat{b}_{2sls}$ is even not consistent, as it does not converge to a deterministic limit. If we apply the standard t-test with t-statistic in the presence of weak instrument, we will be likely to over-reject the null hypothesis.

So far as we know, there is no estimator which are fully robust to weak instrument[21]. However, there are some partial robust estimators which are less biased than 2SLS estimator. Typical examples are k-class estimators including limited-information-maximum-likelihood (LIML) estimator and Fuller-k estimator. A k-class estimator can be written as follows:

$$\hat{b}_{k-class} = \frac{x'(I_n - kQ_z)y}{x'(I_n - kQ_z)x}$$

where I_n is the identity matrix and $Q_z = I_n - P_z$. Different k-class estimators have different values of k[22]. Those partial robust estimators have larger critical value for the F-statistic weak instrument test than the 2SLS estimator, especially when we have a large amount of instruments; for a specific value of F-statistic, we may conclude that the 2SLS estimator suffer from the problem of weak instrument by a k-class estimator does not. In the most cases of our paper we only have one instrument, and it is shown that with only one instrument the k-class estimators do

[21] By fully robust we mean that an estimator is consistent and under the null hypothesis the t-test has standard normal distribution, no matter whether the instrument is weak or not. We do not have such estimator because in the extreme case with zero correlation, the b parameter can not be identified, thus it can not be consistently estimate either.

[22] For example, k is the smallest root of the equation $\det(U'U - JU'Q_zU) = 0$ for LIML, where $U = [y, x]$ and J is the number of instruments. For Fuller-k estimator, k is the one for LIML minus a adjust term which is related with the sample size and the number of instruments.

not superior to 2SLS estimator (Stock and Yogo, 2005). Therefore, we do not explore those alternative estimators.

Despite we do not have fully robust estimators, we have fully robust test and associated confidence interval for the null hypothesis like $b = b_0$. In our paper we explore two of the fully robust tests: Anderson-Rubin (AR) test and conditional likelihood ratio (CLR) test. AR test rewrite the null hypothesis of $b = b_0$ to $d = 0$, where d is absorbed from the following equation:

$$y - b_0 x = zd + e$$

Under the new null hypothesis the AR statistic is

$$AR = \frac{e' P_z e}{(e'e - e' P_z e)/(n - J - 1)}$$

which converges in distribution to a chi-square distribution. We reject the null hypothesis at a certain confidence level, thus can not distinguish b and b_0, iff AR is greater or equal to the critical value. The confidence interval at the same confidence level can be calculated accordingly. Details about AR test can be found in Anderson and Rubin (1949).

AR test has strong power when we only have one instrument, but its power decreases dramatically with the number of instruments. In our study we have multiple instrument variables only when we want to explore the roles of multiple mediating educational inputs, so we intensively use AR test and confidence interval. When we have multiple instrument variables, we use CLR test. CLR test starts from the likelihood function (the joint distribution of y and x) of the 2SLS model we introduced above and derive the LR statistic. The critical values at a certain confidence level conditional on model parameters can be obtained by simulation. Then the corresponding confidence interval can also be calculated. Details about CLR test can be found in Moreira (2003).

Table A1: Descriptive Statistics by School, Art Track

	Model School		Regular School								
	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Jiugong**	**Weishanzhuang**	**Caiyu**
Exam Score											
SEEC	0.971 (0.740)	0.358 (0.655)	0.256 (0.703)	-0.283 (0.798)	-0.095 (0.590)	0.240 (0.655)	-0.394 (0.804)	-1.098 (0.584)	-1.149 (0.539)	-1.243 (0.679)	-1.360 (0.799)
SEEH	1.271 (0.399)	0.643 (0.231)	0.492 (0.453)	0.032 (0.367)	0.108 (0.259)	0.212 (0.610)	-0.186 (0.394)	-0.825 (0.406)	-1.038 (0.596)	-1.174 (0.833)	-1.429 (0.814)
Student Characteristics											
Male	0.525	0.474	0.478	0.390	0.453	0.521	0.545	0.424	0.517	0.4	0.347
Age	18.464 (0.613)	18.703 (0.699)	18.771 (0.753)	18.650 (0.713)	18.698 (0.718)	18.660 (0.681)	18.848 (0.749)	18.815 (0.710)	18.907 (0.679)	18.692 (0.703)	18.625 (0.638)
Parental Backgrounds											
College Father	0.312	0.006	0.124	0.106	0.050	0.021	0.569	0	0.076	0.046	0
College Mother	0.362	0.009	0.131	0.065	0.057	0.021	0.405	0	0.059	0.054	0
Farmer Father	0.369	0.696	0.121	0.541	0.182	0.535	0.664	0.957	0.722	--	1
Farmer Mother	0.364	0.698	0.115	0.598	0.226	0.576	0.752	0.956	0.768	--	1
No. of Obs	459	323	314	123	159	144	343	92	118	130	72

Note: Sample means for the 2008 cohort. No. of Obs is the number of observations with at least one non-missing value for the variable listed (or the number of observations with non-missing values on the SEEH which is the forcing variable in the RDD).

Table A2: Descriptive Statistics by School, Science Track

	Model School		Regular School								
	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Jiugong**	**Weishanzhuang**	**Caiyu**
Exam Score											
SEEC	1.101 (0.651)	0.494 (0.835)	0.566 (0.639)	-0.224 (0.797)	0.350 (0.517)	0.259 (0.712)	-0.344 (0.892)	-0.560 (0.591)	-0.230 (0.462)	-0.966 (0.744)	-1.135 (0.719)
SEEH	1.060 (0.515)	0.558 (0.326)	0.432 (0.431)	-0.129 (0.490)	-0.013 (0.413)	-0.069 (0.842)	-0.270 (0.588)	-0.952 (0.389)	-1.021 (0.415)	-1.439 (0.942)	-1.838 (0.949)
Student Characteristics											
Male	0.724	0.586	0.721	0.768	0.768	0.647	0.579	0.690	0.705	0.660	0.669
Age	18.434 (0.599)	18.737 (0.706)	18.730 (0.650)	18.913 (0.742)	18.725 (0.662)	18.741 (0.639)	18.719 (0.744)	18.690 (0.785)	18.852 (0.736)	18.691 (0.672)	18.821 (0.684)
Parental Backgrounds											
College Father	0.524	0	0.335	0.174	0.145	0.082	0.478	0	0.011	0.053	0
College Mother	0.364	0.008	0.288	0.130	0.101	0.047	0.362	0	0.023	0.053	0
Farmer Father	0.153	0.667	0.038	0.319	0.176	0.353	0.703	0.958	0.679	--	1
Farmer Mother	0.343	0.667	0.036	0.377	0.203	0.424	0.758	0.986	0.759	--	1
No. of Obs	145	133	215	69	69	85	178	71	88	94	145

Note: Sample means for the 2008 cohort. No. of Obs is the number of observations with at least one non-missing value for the variable listed (or the number of observations with non-missing values on the SEEH which is the forcing variable in the RDD).

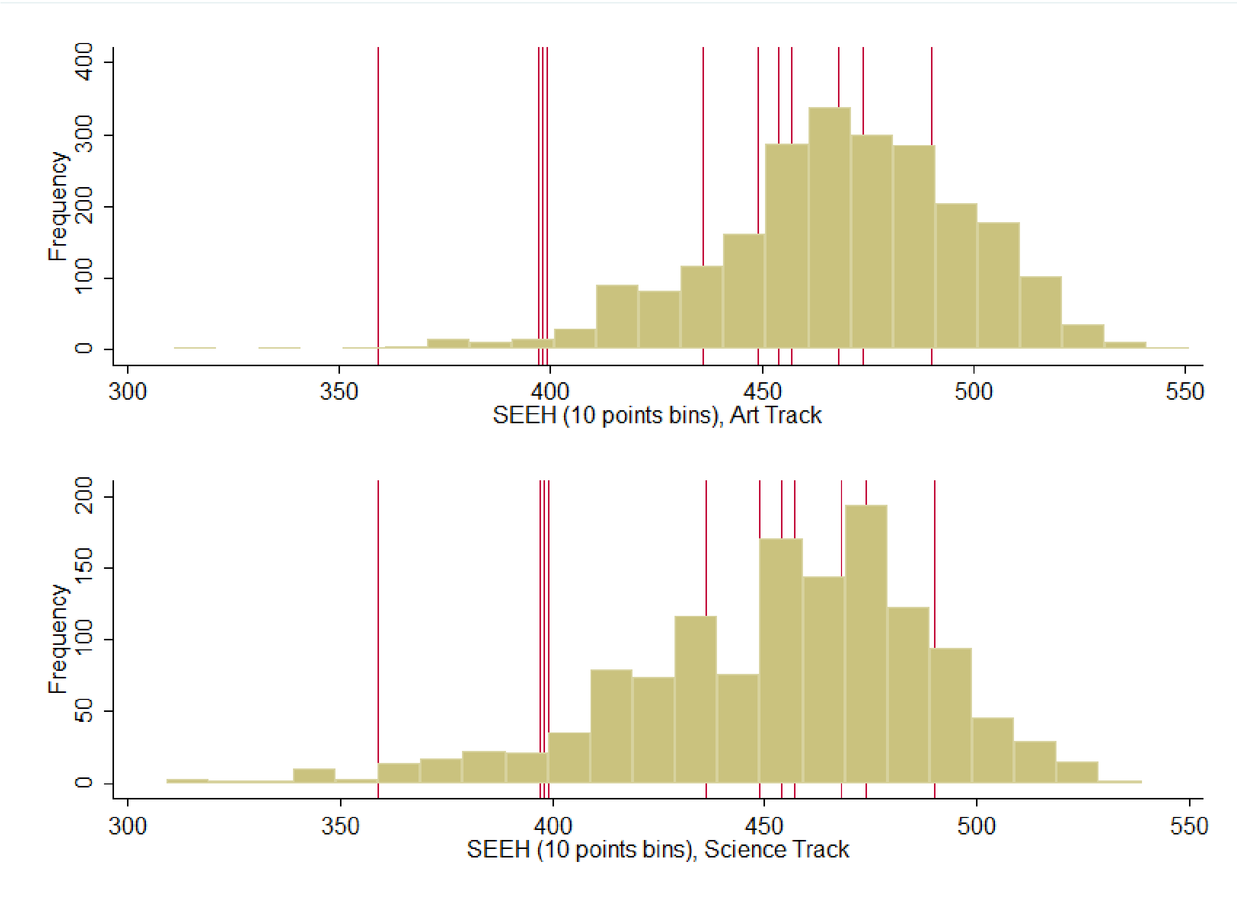
Table A3: Descriptive Statistics of Schools, by School

	Model School		Regular School								
	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Jiugong**	**Weishanzhuang**	**Caiyu**
Urban	1	1	1	1	1	1	1	0	0	0	0
Number of Enrollment	2286	1502	1877	817	1023	950	1704	804	823	966	801
Number of Teacher	153	156	171	63	85	83	182	71	68	63	67
Student/Teacher	14.941	9.628	10.977	12.968	12.035	11.446	9.363	11.324	12.103	15.333	11.955
Percentage of Teachers with Advanced Certificate	0.412	0.429	0.287	0.317	0.329	0.494	0.258	0.155	0.118	0.063	0.075
Percentage of Teachers Younger than 35	0.451	0.410	0.626	0.508	0.518	0.434	0.648	0.789	0.618	0.841	0.806
Minimum Score of SEEH for Admission in 2005	490	474	468	454	457	449	436	398	359	397	399

Table A4: Density Test, All Schools

	RDD		RKD	
	Art Track	**Science Track**	**Art Track**	**Science Track**
No.1 School	-0.095 (0.088)	-0.076 (0.151)	-0.002 (0.003)	-0.002 (0.002)
Xinhua School	0.106 (0.085)	-0.067 (0.108)	-0.001 (0.005)	-0.004 (0.008)
No.2 School	0.063 (0.085)	0.433*** (0.106)	-0.011** (0.005)	0.003 (0.009)
No.3 School	0.373*** (0.094)	0.616*** (0.113)	-0.009* (0.005)	-0.016* (0.008)
No.5 School	0.222** (0.088)	0.183 (0.112)	-0.009 (0.006)	-0.006 (0.009)
No.8 School	0.766*** (0.106)	0.773*** (0.123)	-0.003 (0.005)	-0.003 (0.005)
Xingda School	-0.338** (0.161)	-0.108 (0.157)	0.004** (0.002)	0.002 (0.002)
Yufa School	-0.133 (0.377)	0.376 (0.247)	0.000 (0.000)	0.000 (0.000)
Weishanzhuang School	-0.609 (0.433)	0.115 (0.259)	0.000 (0.000)	0.000 (0.000)
Caiyu School	0.276 (0.331)	0.519** (0.241)	0.000 (0.000)	-0.000 (0.000)

Figure A2: Density of Students by SEEH: All Specific Schools



Note: Sample includes all students in the corresponding track in Daxing District of Beijing in 2008. SEEH is censored at 300 from the left. For students with missing school identity, SEEH and SEEC are dropped.

Table A5: Balance Check of Background Variables and Self-Choice Variables, RDD

	Model School		Regular School							
	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Weishanzhuang**	**Caiyu**
Panel A: Background Variables										
Art Track										
Male	0.012	0.082	-0.015	-0.066	-0.124**	0.144**	0.036	0.135	0.049	-0.262**
	(0.172)	(0.078)	(0.052)	(0.042)	(0.048)	(0.059)	(0.060)	(0.192)	(0.178)	(0.090)
Age	-0.169	-0.129	0.029	-0.028	0.001	-0.045	0.023	0.726***	0.642***	0.754***
	(0.092)	(0.086)	(0.110)	(0.073)	(0.097)	(0.115)	(0.130)	(0.133)	(0.062)	(0.058)
College Father	0.078	-0.087	0.061	-0.175**	-0.159	0.325	-0.075	0.018	0.024	-0.166
	(0.056)	(0.059)	(0.072)	(0.076)	(0.096)	(0.231)	(0.062)	(0.105)	(0.104)	(0.142)
College Mother	-0.046*	-0.065	0.073	-0.090	-0.123	0.155	0.005	0.130	0.140	-0.434
	(0.021)	(0.052)	(0.061)	(0.051)	(0.083)	(0.156)	(0.067)	(0.234)	(0.243)	(0.258)
Farmer Father	0.062	0.253***	-0.243	-0.094	-0.136	-0.045	0.077	-0.162	-0.117	0.339
	(0.054)	(0.069)	(0.134)	(0.125)	(0.159)	(0.220)	(0.137)	(0.207)	(0.200)	(0.209)
Farmer Mother	0.035	0.231***	-0.218	-0.127	-0.110	0.035	0.041	0.166	0.189	0.378**
	(0.041)	(0.065)	(0.151)	(0.116)	(0.147)	(0.236)	(0.154)	(0.244)	(0.261)	(0.160)
Science Track										
Male	-0.026	-0.092	0.028	-0.026	0.014	-0.110	0.146	-0.080	-0.162***	0.049
	(0.099)	(0.104)	(0.042)	(0.053)	(0.089)	(0.089)	(0.108)	(0.065)	(0.031)	(0.107)
Age	-0.219	-0.039	0.083	0.230*	0.050	-0.032	0.224**	0.219	0.251	0.121
	(0.416)	(0.096)	(0.115)	(0.119)	(0.090)	(0.097)	(0.097)	(0.233)	(0.228)	(0.293)
College Father	-0.025	-0.137	0.117	-0.047	-0.137	0.133	-0.029	0.002	0.011	0.025
	(0.118)	(0.111)	(0.116)	(0.094)	(0.116)	(0.224)	(0.102)	(0.060)	(0.059)	(0.050)
College Mother	-0.043	-0.115	0.068	-0.038	-0.083	0.148	0.010	-0.115	-0.140	-0.107
	(0.071)	(0.074)	(0.088)	(0.060)	(0.090)	(0.128)	(0.083)	(0.145)	(0.149)	(0.151)
Farmer Father	0.173	0.292**	-0.270*	-0.094	-0.079	-0.032	0.114	0.153	0.178	0.106
	(0.136)	(0.094)	(0.132)	(0.110)	(0.124)	(0.239)	(0.205)	(0.092)	(0.105)	(0.071)
Farmer Mother	0.277**	0.251*	-0.176	-0.053	-0.065	-0.018	0.041	0.229	0.353**	0.272
	(0.078)	(0.109)	(0.160)	(0.103)	(0.126)	(0.242)	(0.225)	(0.144)	(0.124)	(0.152)
Panel B: Self-Choice Variables										
Prob. of Science Track	-0.041	-0.060	0.193***	-0.010	-0.061	0.051	0.041	0.078	0.068	0.082
	(0.063)	(0.061)	(0.042)	(0.079)	(0.060)	(0.076)	(0.057)	(0.245)	(0.242)	(0.148)
Prob. of Taking EEC	0.004	-0.002	0.039	-0.033	-0.036	-0.036	0.101*	-0.035	-0.004	-0.088
	(0.019)	(0.017)	(0.026)	(0.049)	(0.057)	(0.113)	(0.055)	(0.115)	(0.118)	(0.116)

Note: Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The optimal bandwidth is the CCT optimal bandwidth. The standard error which is robust and clustered on school is shown in parentheses.

Table A6: Balance Check of Background Variables and Self-Choice Variables, RKD

	Model School		Regular School							
	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Weishanzhuang**	**Caiyu**
Panel A: Background Variables										
Art Track										
Male	-0.045 (0.090)	0.262*** (0.076)	0.260*** (0.068)	-0.105 (0.124)	0.103 (0.135)	-0.356*** (0.053)	-0.060 (0.090)	-0.069 (0.099)	-0.035 (0.085)	0.056 (0.044)
Age	-0.080 (0.098)	-0.201* (0.095)	-0.131 (0.233)	-0.003 (0.096)	-0.082 (0.205)	0.109 (0.136)	-0.005 (0.124)	-0.282** (0.094)	-0.200*** (0.054)	-0.254*** (0.074)
College Father	0.085 (0.048)	0.028 (0.094)	0.093 (0.283)	-0.559 (0.426)	-0.588 (0.561)	-0.646 (0.443)	0.122 (0.103)	0.010 (0.054)	0.018 (0.051)	0.081 (0.074)
College Mother	0.079** (0.031)	-0.025 (0.062)	-0.077 (0.229)	-0.366 (0.303)	-0.322 (0.384)	-0.328 (0.292)	0.004 (0.085)	-0.060 (0.108)	-0.057 (0.100)	0.144 (0.101)
Farmer Father	0.021 (0.080)	0.470** (0.177)	0.213 (0.423)	0.294 (0.487)	0.501 (0.608)	0.035 (0.327)	-0.140 (0.125)	0.080 (0.100)	0.050 (0.090)	-0.106 (0.076)
Farmer Mother	-0.095*** (0.023)	0.511** (0.181)	0.317 (0.428)	0.254 (0.474)	0.440 (0.578)	-0.118 (0.348)	-0.124 (0.144)	-0.053 (0.096)	-0.051 (0.092)	-0.138* (0.071)
Science Track										
Male	0.001 (0.083)	-0.242 (0.181)	-0.018 (0.245)	0.268 (0.154)	0.298* (0.164)	0.192 (0.119)	-0.131 (0.087)	0.035 (0.049)	0.058 (0.038)	0.007 (0.076)
Age	-0.144 (0.245)	-0.021 (0.223)	-0.511** (0.211)	-0.328 (0.244)	-0.350 (0.234)	0.112 (0.094)	-0.081 (0.128)	-0.065 (0.062)	-0.076 (0.066)	-0.057 (0.092)
College Father	0.002 (0.129)	-0.265 (0.241)	0.114 (0.422)	-0.127 (0.466)	0.003 (0.556)	-0.237 (0.309)	-0.013 (0.120)	0.036 (0.031)	0.032 (0.030)	0.022 (0.030)
College Mother	-0.032 (0.092)	-0.187 (0.155)	-0.021 (0.398)	-0.021 (0.198)	0.021 (0.336)	-0.159 (0.186)	-0.025 (0.098)	0.044 (0.046)	0.048 (0.048)	0.043 (0.051)
Farmer Father	0.070 (0.085)	0.617** (0.202)	0.055 (0.470)	-0.075 (0.434)	0.034 (0.514)	0.009 (0.334)	-0.023 (0.166)	-0.078 (0.062)	-0.080 (0.057)	-0.049 (0.056)
Farmer Mother	0.067 (0.039)	0.580* (0.252)	0.185 (0.448)	-0.083 (0.491)	0.082 (0.550)	-0.042 (0.330)	-0.028 (0.192)	-0.088 (0.074)	-0.110 (0.064)	-0.082 (0.062)
Panel B: Self-Choice Variables										
Prob. of Science Track	-0.017 (0.066)	-0.117 (0.076)	0.158 (0.114)	0.084 (0.085)	0.142 (0.139)	0.072 (0.071)	-0.007 (0.044)	0.003 (0.060)	0.006 (0.057)	-0.011 (0.040)
Prob. of Taking EEC	0.003 (0.021)	-0.007 (0.065)	0.029 (0.123)	0.004 (0.106)	0.041 (0.132)	0.013 (0.146)	-0.041 (0.048)	0.033 (0.040)	0.027 (0.041)	0.050 (0.039)

Note: Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The optimal bandwidth is the CCT optimal bandwidth. The standard error which is robust and clustered on school is shown in parentheses.

Table A7: Bandwidth Selection

	Art Track				Science Track			
	RDD		RKD		RDD		RKD	
	CCT	IK	CCT	IK	CCT	IK	CCT	IK
Model School	0.484	1.364	0.866	2.497	0.506	1.133	0.527	1.130
No.1 School	0.293	0.977	0.578	1.054	0.325	0.762	0.545	1.597
Xinhua School	0.484	1.364	0.866	2.497	0.506	1.133	0.527	1.130
No.2 School	0.490	1.182	0.939	2.490	0.389	1.485	0.687	1.796
No.3 School	0.585	2.200	0.933	1.562	0.512	1.276	0.676	1.802
No.5 School	0.574	2.953	0.817	1.554	0.594	1.337	0.728	1.736
No.8 School	0.689	3.161	0.754	1.872	0.574	1.183	0.864	2.846
Xingda School	0.701	3.267	0.926	1.909	0.573	2.475	1.016	1.611
Yufa School	0.733	2.296	0.789	2.595	0.820	2.385	1.172	3.294
Weishanzhuang School	0.725	2.270	0.853	2.817	0.806	2.333	1.139	3.344
Caiyu School	0.647	2.244	0.840	2.554	0.695	2.224	1.027	3.310

Table A8: ITT Estimation of Enrollment

	Art Track		Science Track	
	Parametric Estimation	**Non-Parametric Estimation**	**Parametric Estimation**	**Non-Parametric Estimation**
Model School	0.365 (0.218)	0.290* (0.156)	0.449 (0.276)	0.414* (0.186)
No.1 School	0.083 (0.098)	-0.053 (0.037)	0.042 (0.089)	-0.057 (0.063)
Xinhua School	0.438* (0.202)	0.351* (0.166)	0.508 (0.290)	0.456* (0.234)
No.2 School	0.447 (0.275)	0.489 (0.312)	0.467* (0.254)	0.448 (0.263)
No.3 School	0.193 (0.205)	0.126 (0.139)	0.217 (0.219)	0.137 (0.148)
No.5 School	0.440 (0.343)	0.413 (0.306)	0.358 (0.320)	0.339 (0.301)
No.8 School	0.260 (0.274)	0.276 (0.284)	0.219 (0.226)	0.201 (0.209)
Xingda School	-0.211 (0.158)	-0.148* (0.071)	-0.234 (0.157)	-0.180 (0.123)
Yufa School	0.138 (0.123)	-0.020 (0.031)	-0.110 (0.109)	-0.090 (0.085)
Weishanzhuang School	0.239 (0.151)	0.309 (0.217)	-0.064 (0.164)	-0.010 (0.142)
Caiyu School	-0.525*** (0.115)	-0.341 (0.209)	0.136 (0.147)	0.201* (0.107)

Note: Parametric estimations control for a cubic function of the standardized SEEH, and the sample covers students within one standard deviation of the cutoffs. Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The standard error which is robust and clustered on school is shown in parentheses. The number of observations is the same as in Table 3.

Table A9: Eligible Non-Compliers by Attending Schools

	No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Weishanzhuang**	**Caiyu**
No.1	--	566	603	607	607	608	609	609	609	609
Xinhua	**94**	--	437	463	448	464	464	464	464	464
No.2	**97**	**327**	--	508	505	519	541	548	548	548
No.3	**3**	**18**	**47**	--	**149**	168	178	200	200	200
No.5	**5**	**27**	**57**	252	--	255	280	289	289	289
No.8	**29**	**75**	**99**	**181**	**150**	--	261	289	289	289
Xingda	**1**	**10**	**46**	**315**	**214**	**453**	--	512	512	512
Yufa	**0**	**0**	**0**	**2**	**2**	**3**	**67**	--	170	**170**
Jiugong	**0**	**0**	**0**	**1**	**0**	**1**	**66**	**240**	**240**	**239**
Weishanzhuang	**0**	**1**	**1**	**5**	**4**	**13**	**64**	**253**	--	**253**
Caiyu	**0**	**1**	**1**	**3**	**1**	**6**	**39**	130	132	--
Total Eligible Students	732	1455	1785	2503	2332	2740	3045	3704	3706	3702
Percentage of Overcautious Students	31.30%	31.50%	14.10%	20.30%	22.30%	17.40%	7.80%	13.30%	6.50%	17.90%
Average SEEH of Non-Overcautious Students	1.38 (0.27) [503]	1.03 (0.45) [902]	0.89 (0.50) [1113]	0.74 (0.56) [1352]	0.75 (0.56) [1333]	0.59 (0.63) [1575]	0.54 (0.66) [1657]	0.37 (0.81) [1828]	0.37 (0.82) [1830]	0.37 (0.81) [1827]
Average SEEH of Overcautious Students	1.01 (0.21) [229]	0.76 (0.28) [553]	0.67 (0.31) [672]	0.38 (0.43) [1151]	0.46 (0.40) [999]	0.37 (0.43) [1165]	0.22 (0.52) [1388]	-0.13 (0.76) [1876]	-0.13 (0.76) [1876]	-0.13 (0.76) [1875]
Difference (t test)	0.37*** (0.02)	0.27*** (0.02)	0.22*** (0.02)	0.35*** (0.02)	0.28*** (0.02)	0.22*** (0.02)	0.31*** (0.02)	0.51*** (0.03)	0.50*** (0.03)	0.51*** (0.03)

Note: Numbers in bold indicate eligible non-compliers who attend a school with a lower cutoff.

Table A10: Enrolled Complier by Eligible Schools

	No. 1	Xinhua	No. 2	No. 3	No. 5	No. 8	Xingda	Yufa	Weishanzhuang	Caiyu
No.1	503	**94**	**97**	**3**	**5**	**25**	**1**	**0**	**0**	**0**
Xinhua	503	430	**327**	**18**	**27**	**75**	**10**	**0**	**1**	**1**
No.2	503	430	494	**47**	**57**	**99**	**46**	**0**	**1**	**1**
No.3	503	430	494	166	251	**181**	**315**	**2**	**5**	**3**
No.5	503	430	494	**149**	251	**150**	**214**	**2**	**4**	**2**
No.8	503	430	494	166	251	250	**453**	**3**	**13**	**6**
Xingda	503	430	494	166	251	250	476	**67**	**64**	**39**
Yufa	503	430	494	166	251	250	476	170	**253**	129
Jiugong	503	430	494	166	251	250	476	170	253	129
Weishanzhuang	503	430	494	166	251	250	476	170	253	129
Caiyu	503	430	494	166	251	250	476	**170**	**253**	129
Total Enrolled Compliers	503	430	494	166	251	250	476	170	253	129
Percentage of Overcautious Students	0%	21.9%	66.2%	89.8%	22.7%	72.4%	95.2%	100%	100%	30.2%
Percentage of Overcautious Students who are eligible two Schools higher	0%	0%	19.6%	28.3%	10.8%	60%	66.2%	39.4%	100%	4.7%
Percentage of Overcautious Students who are eligible three Schools higher	0%	0%	0%	10.8%	2.0%	39.6%	45.0%	1.8%	25.3%	2.3%

Note: Numbers in bold indicate enrolled students who could attend a school with a higher cutoff.

Table A11: Enrolled Non-Complier

	No. 1	Xinhua	No. 2	No. 3	No. 5	No. 8	Xingda	Yufa	Weishanzhuang	Caiyu
Number of enrolled Non-complier	106	464	549	201	289	291	521	177	282	234
Total enrollment	609	34	55	35	38	41	45	7	29	105
Percentage of enrolled non-complier	17.4%	7.3%	10.0%	17.4%	13.1%	14.1%	8.6%	4.0%	10.3%	44.9%
Average SEEH of enrolled non-complier	0.46	-0.03	-0.43	-0.76	-0.70	-0.96	-1.42	-2.33	-3.04	-2.57

Table A12: Estimation of the Effects of Individual Schools on Various Outcomes, RDD

		Art Track, College Qualification			Science Track, College Qualification		
	Taking SEEC	**Key 4-year**	**4-year**	**3-year**	**Key 4-year**	**4-year**	**3-year**
No.1	-0.070 (0.816) [946]	-0.088 (0.929) [513]	**-2.272** **(0.056)** **[-35.35, -0.349]*** **[513]**	**-0.678** **(0.047)** **[-11.19, -0.109]*** **[513]**	-2.593 (0.427) [232]	-0.383 (0.820) [232]	-0.563 (0.239) [232]
Xinhua	-0.006 (0.877) [2007]	-0.223 (0.163) [1008]	-0.367 (0.224) [1008]	-0.027 (0.824) [1008]	-0.023 (0.880) [519]	-0.020 (0.863) [519]	0.019 (0.910) [519]
No.2	0.081 (0.122) [1676]	0.151 (0.401) [1031]	0.273 (0.110) [1031]	0.112 (0.253) [1031]	0.193 (0.465) [444]	0.141 (0.341) [444]	0.044 (0.795) [444]
No.3	-0.199 (0.480) [2102]	-0.251 (0.306) [985]	-0.328 (0.246) [985]	0.417 (0.350) [985]	-0.296 (0.032) [-6.359, -0.040] ∪ [0.154, 5.768]** [498]	0.226 (0.742) [498]	0.618 (0.343) [498]
No.5	-0.087 (0.510) [1912]	0.009 (0.917) [996]	-0.188 (0.144) [996]	0.045 (0.788) [996]	0.033 0.701 [597]	-0.089 (0.774) [597]	0.370 (0.115) [597]
No.8	-0.114 (0.739) [1958]	-0.361 (0.125) [1001]	-0.107 (0.788) [1001]	0.474 (0.300) [1001]	-0.773 (0.028) [-17.55, -0.067] ∪ [0.673, 16.01]** [492]	-0.243 (0.727) [492]	-0.677 (0.105) [492]
Xingda	**-0.853** **(0.055)** **[-12.27, -0.108]*** **[1503]**	-0.111 (0.318) [698]	0.238 (0.340) [698]	0.150 (0.765) [698]	0.189 (0.179) [407]	0.476 (0.187) [407]	-0.427 (0.349) [407]
Yufa	0.528 (0.748) [542]	29.99 (0.249) [180]	0.163 (0.997) [180]	11.59 (0.891) [180]	--	-0.029 (0.940) [233]	-0.192 (0.851) [233]
Weishanzhuang	-0.030 (0.970) [480]	0.082 (0.240) [167]	-0.007 (0.954) [167]	-0.107 (0.693) [167]	--	-0.552 (0.870) [216]	-13.91 (0.043) [--] [216]

Caiyu	0.361 (0.417) [662]	-0.070 0.452 [165]	0.019 (0.882) [165]	-0.278 (0.207) [165]	--	0.050 (0.794) [188]	**1.079** **(0.005)** **[0.194, 18.73]**** **[188]**

Note: Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The number of observations is in brackets. The AR test is robust to heterogeneity and clustering on school. We first report the p-value of the AR test in parentheses and then report the significant confidence interval if the null hypothesis of zero effect is rejected.

Table A13: Estimation of the Effects of Individual Schools on Various Outcomes, RKD

		Art Track, College Qualification			Science Track, College Qualification		
	Taking SEEC	**Key 4-year**	**4-year**	**3-year**	**Key 4-year**	**4-year**	**3-year**
No.1	0.025 (0.881) [1382]	**0.283 (0.031) [0.029, 0.645]** [949]**	0.348 (0.554) [949]	-0.126 (0.203) [949]	1.602 (0.584) [375]	-2.546 (0.121) [375]	0.185 (0.688) [375]
Xinhua	-0.016 (0.905) [2600]	-0.104 0.650 [1587]	-0.164 (0.596) [1587]	**-0.309 (0.009) [-4.816, -0.065]** [1587]**	0.059 (0.822) [541]	0.055 (0.761) [541]	-0.025 (0.910) [541]
No.2	0.060 (0.803) [2379]	0.300 (0.097) [-4.222, -0.108] ∪ [0.019, 4.822]* [1626]	0.713 (0.000) [-16.35, -0.279] ∪ [0.200, 17.78]*** [1626]	-0.646 (0.009) [-19.28, -0.012] ∪ [0.287, 17.99]*** [1626]	0.237 (0.231) [684]	0.071 (0.865) [684]	-0.235 (0.188) [684]
No.3	-0.013 (0.972) [3225]	-0.305 (0.019) [-6.927, -0.053] ∪ [0.160, 6.317]** [1434]	-0.630 (0.002) [-17.27, -0.196] ∪ [0.137, 16.01]*** [1434]	1.007 (0.005) [-29.54, -0.402] ∪ [0.088, 31.56]*** [1434]	-0.893 (0.002) [-26.11, -0.236] ∪ [0.269, 24.31]*** [657]	-0.727 (0.062) [-10.15, -0.331] ∪ [0.160, 8.695]* [657]	-1.174 (0.001) [-32.13, -0.367] ∪ [0.377, 29.78]*** [657]
No.5	-0.045 (0.743) [2543]	-0.092 (0.088) [-1.106, -0.012] ∪ [0.101, 0.923]* [1350]	-0.241 (0.001) [-5.381, -0.046] ∪ [0.202, 4.899]*** [1350]	0.369 (0.000) [-5.797, -0.187] ∪ [0.159, 6.536]*** [1350]	-0.466 (0.002) [-13.76, -0.067] ∪ [0.519, 12.82]*** [694]	-0.465 (0.135) [694]	-0.612 (0.025) [-20.46, -0.016] ∪ [0.859, 19.23]*** [694]
No.8	-0.033 (0.924) [2837]	-0.219 (0.050) [-5.129, -0.013] ∪ [0.184, 4.691]** [1088]	-0.089 (0.697) [1088]	0.415 (0.304) [1088]	-0.964 (0.000) [-28.16, -0.256] ∪ [0.398, 26.23]*** [767]	-0.724 (0.156) [767]	-1.427 (0.003) [-52.83, -0.088] ∪ [1.354, 49.98]*** [767]
Xingda	0.281 (0.368) [2665]	-0.139 (0.388) [998]	0.430 (0.128) [998]	0.671 (0.121) [998]	0.178 (0.142) [753]	0.664 (0.090) [-6.839, -1.095] ∪ [0.047, 8.166]* [753]	-1.329 (0.007) [-31.25, -0.069] ∪ [1.250, 28.59]*** [753]
Yufa	0.754	-1.417	0.517	0.568	0.092	0.669	0.788

	(0.382) [1097]	(0.369) [189]	(0.809) [189]	(0.911) [189]	(0.230) [381]	(0.037) [-8.717, -0.441] ∪ [0.123, 10.05]** [381]	(0.281) [381]
Weishanzhuang	-0.188 (0.487) [1054]	0.083 (0.388) [210]	-0.033 (0.793) [210]	-0.125 (0.686) [210]	-0.042 (0.343) [343]	-0.348 (0.069) [-6.338, -0.096] ∪ [0.096, 5.642]* [343]	0.025 (0.963) [343]
Caiyu	0.349 (0.182) [1335]	-0.100 (0.297) [217]	-0.019 (0.858) [217]	**-0.289** **(0.063)** **[-0.741, -0.040]*** **[217]**	-0.077 (0.359) [326]	-0.460 (0.273) [326]	0.863 (0.415) [326]

Note: Nonparametric estimation uses the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The number of observations is in brackets. The AR test is robust to heterogeneity and clustering on school. We first report the p-value of the AR test in parentheses and then report the significant confidence interval if the null hypothesis of zero effect is rejected.

Table A14: ITT Estimation of Effect on Peer Quality and School Quality

		Model School		**Regular School**							
		No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Weishan zhuang**	**Caiyu**
Panel A: Peer Quality											
Art Track											
Peer Gap within Track	Jump	1.570***	-0.278	-0.296*	0.110	0.199**	-0.531	-0.706*	-2.424**	-2.256**	-1.714*
		(0.424)	(0.182)	(0.143)	(0.106)	(0.078)	(0.436)	(0.330)	(0.842)	(0.740)	(0.854)
	Kink	-1.655***	0.555	0.348	0.349	0.494	-0.180	-1.138**	-1.213**	-1.132**	-0.871
		(0.427)	(0.377)	(0.495)	(0.324)	(0.522)	(0.547)	(0.425)	(0.474)	(0.417)	(0.459)
Peer Gap cross Tracks	Jump	1.523***	-0.293	-0.306*	0.112	0.207**	-0.568	-0.737*	-3.360**	-3.171**	-2.240*
		(0.409)	(0.181)	(0.149)	(0.111)	(0.073)	(0.460)	(0.337)	(0.967)	(0.878)	(1.018)
	Kink	-1.601***	0.591	0.373	0.366	0.527	-0.226	-1.175**	-1.637**	-1.552**	-1.116*
		(0.412)	(0.387)	(0.511)	(0.293)	(0.507)	(0.568)	(0.425)	(0.532)	(0.480)	(0.541)
Science Track											
Peer Gap within Track	Jump	1.714***	-0.635***	-0.303*	0.585*	0.281	-0.162	-0.366	-1.671**	-1.644**	-2.152**
		(0.239)	(0.143)	(0.162)	(0.271)	(0.164)	(0.440)	(0.456)	(0.580)	(0.595)	(0.665)
	Kink	-1.860***	1.511***	-0.424	0.863	0.214	0.309	-0.728	-0.986**	-0.991**	-1.212**
		(0.303)	(0.245)	(0.391)	(0.554)	(0.602)	(0.516)	(0.638)	(0.322)	(0.325)	(0.379)
Peer Gap cross Tracks	Jump	2.000***	-0.598***	-0.295*	0.570*	0.269	-0.191	-0.386	-1.616**	-1.590**	-2.053**
		(0.248)	(0.151)	(0.154)	(0.269)	(0.166)	(0.437)	(0.468)	(0.582)	(0.593)	(0.663)
	Kink	-2.143***	1.417***	-0.393	0.853	0.248	0.273	-0.753	-0.947**	-0.950**	-1.150**
		(0.291)	(0.250)	(0.376)	(0.572)	(0.627)	(0.512)	(0.648)	(0.328)	(0.328)	(0.380)
Panel B: School Quality											
Art Track											
Student/Teacher Ratio	Jump	-10.80***	-1.765	0.208	1.779***	1.838***	-0.885	-1.657	0.334	0.965	-0.178
		(3.158)	(1.070)	(0.427)	(0.533)	(0.416)	(0.795)	(1.256)	(4.045)	(3.765)	(3.822)
	Kink	11.82***	1.390	-1.165	4.803	3.892	3.025**	-1.959	-0.639	-0.190	-1.073
		(3.190)	(1.515)	(1.978)	(2.771)	(3.163)	(1.184)	(1.686)	(2.034)	(1.894)	(1.770)
Percent of Advanced Certificate	Jump	-0.064	0.055	-0.012	-0.031	-0.050	0.133*	0.128*	0.282	0.254	0.207
		(0.128)	(0.052)	(0.023)	(0.032)	(0.038)	(0.060)	(0.069)	(0.150)	(0.133)	(0.138)
	Kink	0.062	-0.034	-0.041	-0.197	-0.250	-0.017	0.219**	0.152	0.137	0.118
		(0.139)	(0.081)	(0.088)	(0.149)	(0.174)	(0.120)	(0.097)	(0.081)	(0.071)	(0.072)
Percent of Teachers Older than 35	Jump	-0.096	0.072	-0.030	0.012	-0.012	0.132**	0.117	0.397	0.382	0.176
		(0.185)	(0.077)	(0.035)	(0.037)	(0.039)	(0.059)	(0.086)	(0.254)	(0.263)	(0.149)
	Kink	0.090	-0.048	-0.058	-0.101	-0.195	0.082	0.210*	0.194	0.187	0.114
		(0.202)	(0.116)	(0.104)	(0.182)	(0.218)	(0.120)	(0.115)	(0.103)	(0.109)	(0.067)
Science Track											

Student/Teacher Ratio	Jump	-12.29*** (1.050)	-1.708 (1.237)	0.796 (0.675)	2.177*** (0.495)	2.090*** (0.394)	0.177 (0.479)	-0.180 (0.896)	-0.145 (1.376)	0.052 (1.544)	-0.249 (1.966)
	Kink	12.61*** (1.048)	2.421 (2.486)	-4.258 (2.448)	6.787** (2.722)	3.096 (2.997)	4.087** (1.644)	-0.444 (1.148)	-0.563 (0.990)	-0.398 (1.063)	-0.635 (1.292)
Percent of Advanced Certificate	Jump	-0.140 (0.147)	0.095 (0.054)	-0.018 (0.023)	-0.101* (0.052)	-0.058 (0.045)	0.056 (0.059)	0.030 (0.094)	0.232 (0.135)	0.222 (0.136)	0.264 (0.147)
	Kink	0.136 (0.153)	-0.108 (0.101)	0.126 (0.099)	-0.261 (0.153)	-0.170 (0.153)	-0.083 (0.077)	0.090 (0.124)	0.142 (0.078)	0.137 (0.078)	0.156 (0.084)
Percent of Teachers Older than 35	Jump	-0.192 (0.213)	0.120 (0.080)	-0.043 (0.036)	-0.067 (0.064)	-0.023 (0.052)	0.053 (0.054)	-0.004 (0.100)	0.198 (0.117)	0.184 (0.117)	0.238* (0.124)
	Kink	0.203 (0.214)	-0.074 (0.157)	0.120 (0.154)	-0.192 (0.211)	-0.155 (0.208)	-0.026 (0.094)	0.046 (0.125)	0.128* (0.066)	0.121* (0.065)	0.144* (0.069)

Note: Sample means for the 2008 cohort. The optimal bandwidth is the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel.

Table A15: LATE Estimation of Effects of Peer Quality and School Quality

		Model School		Regular School							
		No. 1	**Xinhua**	**No. 2**	**No. 3**	**No. 5**	**No. 8**	**Xingda**	**Yufa**	**Weishanzhuang**	**Caiyu**
Panel A: Peer Quality											
Art Track											
Peer Gap within Track	RDD	3.717 (0.000)	-23.21 (0.510)	-0.914 (0.206)	-0.714 (0.559)	0.014 (0.984)	0.016 (0.977)	-0.164 (0.758)	0.920 (0.751)	-0.211 (0.974)	-5.356 (0.673)
	RKD	-1.173 (0.479)	-1.416 (0.440)	15.61 (0.523)	0.181 (0.830)	-0.239 (0.764)	-0.113 (0.813)	-0.383 (0.152)	-4.939 (0.796)	0.404 (0.934)	-24.25 (0.514)
Peer Gap cross Tracks	RDD	3.323 (0.000)	-18.29 (0.510)	-0.849 (0.206)	-0.615 (0.559)	0.012 (0.984)	0.015 (0.977)	-0.169 (0.758)	0.249 (0.751)	-0.041 (0.974)	1.121 (0.673)
	RKD	-1.412 (0.479)	-1.331 (0.440)	-5.951 (0.523)	0.211 (0.830)	-0.289 (0.764)	-0.113 (0.813)	-0.379 (0.152)	0.477 (0.796)	-0.450 (0.934)	-1.010 (0.514)
Science Track											
Peer Gap within Track	RDD	-2.089 (0.560)	-3.555 (0.772)	-0.516 (0.353)	0.293 (0.488)	0.828 (0.299)	1.055 (0.076) [0.062, 18.14]*	0.305 (0.421)	-0.619 (0.091) [-1.577, -0.044]*	-0.345 (0.264)	0.198 (0.619)
	RKD	0.204 (0.929)	0.030 (0.976)	0.157 (0.696)	-7.861 (0.000)	-21.80 (0.012)	1.488 (0.003) [0.395, 16.75]**	0.598 (0.013) [0.167, 8.571]*	-0.678 (0.000) [-7.556, -0.251]***	-0.505 (0.006) [-3.774, -0.060]***	-0.290 (0.338)
Peer Gap cross Tracks	RDD	-2.358 (0.560)	-7.231 (0.772)	-0.500 (0.353)	0.298 (0.488)	0.838 (0.299)	1.010 (0.076) [0.075, 15.55]*	0.310 (0.421)	-0.688 (0.091) [-1.669, -0.039]*	-0.378 (0.264)	0.232 (0.619)
	RKD	0.201 (0.929)	0.032 (0.976)	0.157 (0.696)	-9.839 (0.000)	-81.26 (0.012)	1.463 (0.003) [0.389, 16.01]**	0.575 (0.013) [0.149, 7.906]*	-0.757 (0.000) [-9.722, -0.272]***	-0.559 (0.006) [-7.421, -0.050]***	-0.337 (0.338)
Panel B: School Quality											
Art Track											
Student/Teacher Ratio	RDD	-0.919 (0.000) [-9.435, -0.459]*	0.173 (0.510)	-24.04 (0.206)	-0.089 (0.559)	0.002 (0.984)	0.003 (0.977)	-0.116 (0.758)	-0.031 (0.751)	0.004 (0.974)	-0.033 (0.673)
	RKD	0.065 (0.479)	0.173 (0.440)	-0.116 (0.523)	0.016 (0.830)	-0.039 (0.764)	-0.017 (0.813)	-0.618 (0.152)	-0.032 (0.796)	0.014 (0.934)	0.050 (0.514)

Percent of Advanced Certificate	RDD	18.92 (0.000) [9.086, 215.3]*	-3.053 (0.510)	-13.06 (0.206)	-21.65 (0.559)	-0.172 (0.984)	-0.054 (0.977)	0.679 (0.758)	3.510 (0.751)	-0.434 (0.974)	2.591 (0.673)
	RKD	-9.510 (0.479)	0.048 (0.440)	-1.916 (0.523)	-0.247 (0.830)	0.333 (0.764)	0.334 (0.813)	1.829 (0.152)	2.312 (0.796)	-0.799 (0.934)	-4.085 (0.514)
Percent of Teachers Older than 35	RDD	17.05 (0.000) [7.798, 201.8]**	-1.909 (0.510)	-4.743 (0.206)	-2.388 (0.559)	0.110 (0.984)	-0.080 (0.977)	0.873 (0.758)	-2.236 (0.751)	0.284 (0.974)	1.093 (0.673)
	RKD	-5.314 (0.479)	-3.579 (0.440)	-1.139 (0.523)	-0.326 (0.830)	0.349 (0.764)	0.586 (0.813)	1.784 (0.152)	-6.057 (0.796)	5.291 (0.934)	-2.114 (0.514)
Science Track											
Student/Teacher Ratio	RDD	-0.263 (0.560)	0.048 (0.772)	0.229 (0.353)	0.152 (0.488)	0.118 (0.299)	0.228 (0.076) [0.019, 0.669]*	0.120 (0.421)	-0.720 (0.091)	-0.712 (0.264)	0.130 (0.619)
	RKD	-1.064 (0.929)	-0.009 (0.976)	0.048 (0.696)	0.413 (0.000)	1.197 (0.012)	0.206 (0.003) [0.057, 1.380]***	0.819 (0.013)	-0.271 (0.000)	-0.251 (0.006)	-0.085 (0.338)
Percent of Advanced Certificate	RDD	-189.8 (0.560)	-0.658 (0.772)	-14.61 (0.353)	-2.160 (0.488)	-5.564 (0.299)	-4.044 (0.076) [-72.54, -0.338]*	-1.558 (0.421)	6.562 (0.091) [0.448, 50.47]*	3.859 (0.264)	-1.756 (0.619)
	RKD	2.683 (0.929)	0.107 (0.976)	-1.456 (0.696)	-14.51 (0.000)	-11.21 (0.012)	-4.933 (0.003) [-80.74, -1.135]**	-3.642 (0.013)	4.097 (0.000) [2.290, 37.49]**	3.416 (0.006) [1.578, 29.61]**	1.715 (0.338)
Percent of Teachers Older than 35	RDD	10.72 (0.560)	-0.363 (0.772)	-4.539 (0.353)	-2.876 (0.488)	-429.2 (0.299)	-6.789 (0.076)	-1.890 (0.421)	4.343 (0.091) [0.232, 59.79]*	2.487 (0.264)	-1.085 (0.619)
	RKD	-3.817 (0.929)	0.051 (0.976)	-2.794 (0.696)	-16.26 (0.000)	-9.577 (0.012)	-8.041 (0.003)	-4.355 (0.013)	3.859 (0.000) [2.210, 24.84]**	3.158 (0.006) [1.478, 12.62]**	1.624 (0.338)

Note: The optimal bandwidth is the CCT optimal bandwidth. Non-parametric estimations control for a linear function of the standardized SEEH, weighted by $max(0, 1 - abs((osd - cutoff)/bandwidth))$. It is equivalent to a local linear estimation with a triangle kernel. The p-value of the AR test is shown in parentheses. The significant AR confidence interval is shown in brackets.

Table A16: Robustness Checks, School Groups, Science Track

	AR Test (1)	Uniform Weight (2)	Polynomial		Optimal Bandwidths		
			Quadratic (3)	Cubic (4)	IK (5)	CV (6)	Distance to the nearest other cutoff (7)
	RDD						
Model School	0.772	-0.258 (0.372) [519]	-0.091 (0.695) [519]	0.168 (0.529) [519]	0.052 (0.822) [896]	-0.140 (0.579) [574]	0.621 (0.162) [206]
Center Area School	0.421	-0.718 (0.157) [406]	-0.276 (0.626) [406]	-0.001 (0.999) [406]	-0.582 (0.249) [1258]	-0.451 (0.351) [427]	-0.321 (0.716) [233]
Center Area Selective School	0.353	0.219 (0.591) [444]	0.250 (0.363) [444]	0.561 (0.157) [444]	0.411** (0.047) [1074]	0.473 (0.231) [350]	0.054 (0.930) [194]
Center Area Less Selective School	0.421	-0.858 (0.157) [406]	-0.314 (0.626) [406]	-0.001 (0.999) [406]	-0.309 (0.249) [1258]	-0.556 (0.351) [427]	-0.366 (0.716) [233]
	RKD						
Model School	0.976	-0.535 (0.379) [541]	-0.039 (0.874) [541]	0.258 (0.394) [541]	0.072 (0.864) [896]	-0.322 (0.424) [708]	0.717 (0.148) [206]
Center Area School	0.013*	0.734 (0.582) [752]	-0.716 (0.311) [752]	-0.703 (0.129) [752]	3.547 (0.573) [1079]	3.100 (0.584) [1071]	-0.368 (0.696) [233]
Center Area Selective School	0.696	0.299 (0.425) [684]	0.007 (0.972) [684]	-0.121 (0.609) [684]	0.935 (0.045) [1165]	0.407* (0.099) [884]	0.151 (0.770) [194]
Center Area Less Selective School	0.013*	-0.552 (0.582) [752]	-0.755 (0.311) [752]	-0.783 (0.129) [752]	-0.162 (0.573) [1079]	-0.158 (0.584) [1071]	-0.472 (0.696) [233]

Note: The p-values of the AR test are shown in parentheses. The number of observations is in brackets. For the purpose of comparison, in Columns (2), (3) and (4) we use the same bandwidth as in the main model, which is not necessarily the optimal bandwidth under the current settings. We also calculate the new optimal bandwidths and perform the analyses with them. We get quite similar results, which are available upon request.

References

Abdulkadiroglu, Atila, Joshua D. Angrist, and Parag A. Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools". *Econometrica*, 82(1): 137-196.

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

Anderson, T., and H. Rubin. 1949. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations". *Annals of Mathematical Statistics*, 20(1): 46-63.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2013. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". *Econometrica* (forthcoming).

Card, David, David S., Lee, Zhuang Pei, and Andrea Weber. 2012. "Nonlinear Policy Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design". NBER Working Paper, No. 18564.

Clark, Damon. 2010. "Selective Schools and Academic Achievement". *The B.E. Journal of Economic Analysis & Policy*, 10(1), Article 9.

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

Deng, Zheng, and Donald J. Treiman. 1997. "The Impact of the Cultural Revolution on Trends in Educational Attainment in the People's Republic of China". *American Journal of Sociology*, 103(2): 391-428.

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

Dobbie, Will, and Roland G. Fryer, Jr. 2014. "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools". *American Economic Journal: Applied Economics*, 6(3): 58-75.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya". *American Economic Review*, 101(5): 1739-1774.

Finlay, Keith, and Leandro M. Magnusson. 2009. "Implementing Weak-Instrument Robust Tests for a General Class of Instrumental-Variables Models". Working Paper No. 0901. Department of Economics, Tulane University.

Ganong, Peter, and Simon Jager. 2014. "A Permutation Test and Estimation Alternatives for the Regression Kink Design". IZA DP No. 8282.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design”. *Econometrica*, 69(1): 201-209.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2005. “Parental Preferences and School Competition: Evidence from a Public School Choice Program”. NBER Working Paper, No. 11805.

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

He Yinghua. 2012. “Gaming the Boston School Choice Mechanism in Beijing”. TSE Working Paper, No. 12-345.

Imbens, Guido W., and Karthik Kalyanaraman. 2012. “Optimal Bandwidth Choice for the Regression Discontinuity Estimator”. *Review of Economic Studies*, 79(3): 933-959.

Imbens, Guido W., and Thomas Lemieux. 2008. “Regression Discontinuity Designs: A Guide to Practice”. *Journal of Econometrics*, 142(2): 615-635.

Jackson, C. Kirabo. 2010. “Do Student Benefit from Attending Better Schools? Evidence from Rule-Based Student Assignments in Trinidad and Tobago”. *The Economic Journal*, 120(549): 1399-1429.

Lai, Fang, Elisabeth Sadouolet, and Alain de Janvry. 2009. “The Contributions of School Quality and Teacher Qualifications to Student Performance: Evidence from a Natural Experiment in Beijing Middle Schools”. *Journal of Human Resources*, 46(1): 123-153.

Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics”. *Journal of Economic Literature*, 48(2): 281-355.

Lucas, Adrienne M., and Issac M. Mbiti. 2014. “Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya”. *American Economic Journal: Applied Economics*, 6(3): 234-263.

Ludwig, Jens, and Douglas L. Miller. 2007. “Does Head Start Improve Children’s Life Chances? Evidence from a Regression Discontinuity Design”. *The Quarterly Journal of Economics*, 122(1): 159-208.

Marmer Vadim, Donna Feir, and Thomas Lemieux. 2014. “Weak Identification in Fuzzy Regression Discontinuity Designs”. Working Paper, University of British Columbia.

Marsh, Herbert W., Danuta Chessor, Rhonda Craven, and Lawrence Roche. 1995. “The Effects of Gifted and Talented Programs on Academic Self-Concept: The Big Fish Strikes Again”. *American Educational Research Journal*, 32(2): 285-319.

McCrary, Justin. 2008. "Testing for Manipulation of the Running Variable in the Regression Discontinuity Design". *Journal of Econometrics*, 142(2): 698-714.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models". *Econometrica*, 71(4): 1027-1048.

Pathak Parag A., and Tayfun Sonmez. 2008. "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism". *American Economic Review*, 98(4): 1636-1652.

Pathak Parag A., and Tayfun Sonmez. 2013. "School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation". *American Economic Review*, 103(1): 80-106.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses". *American Economic Review*, 103(4): 1289-1324.

Salam, Reihan. 2014. "Close Stuyvesant High School: Why Super-Elite Public Magnet Schools Aren't Necessary Anymore". Slate. http://www.slate.com/articles/news_and_politics/politics/2014/07/the_case_for_shutting_down_stuyvesant_high_school_the_best_public_school.html. Accessed on August 24, 2014.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments". *Econometrica*, 65(3): 557-586.

Zhang Hongliang. 2014. "The Mirage of Elite Schools: Evidence from Lottery-Based School Admissions in China". Working Paper. Chinese University of Hong Kong.