Does Selective High School Improve Student Achievement? Effects of Exam Schools in Beijing

Kathryn Anderson, Xue Gong, Kai Hong, Xi Zhang

Exam Selective School

- Exam selective schools are schools that admit students primarily by exam scores. They can be found in China, United States, United Kingdom, Singapore, Romania, Trinidad and Tobago.
- In developing countries, exam selective schools are tradeoffs between equality and efficiency.
- Exam selective schools are perceived to have positive effects on student achievement, because:
  - Tracking students by ability – peer effect
  - Higher quality of teachers
  - More investment in infrastructure.
  - More opportunities (award, competition, exchange program)
- Parents highly value exam selective school
- Whether exam selective schools indeed promote student achievement remains unclear.

Literature on Exam Schools

- Existing literature shows mixed picture about the effect of exam schools.
  - 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.

Research Design

- Most challenging issue: 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.
- Our paper evaluates the effect of model schools in Beijing on academic performance using the regression discontinuity design (RDD) strategy.
  - Each school has a minimum test score for admission.
  - Students with test scores around the cutoff should have similar underlying ability.
  - Within a neighborhood of cutoff, high school admission is mimic to a random procedure.
- One modification: fuzzy RDD
  - The relation between test score and school attendance is not deterministic.
  - Similar idea as instrumental approach: eligibility serves as instrument

Data

- 3868 students, 11 high schools, Daxing District, Beijing
- Students took the entrance exam to high school in 2005, and the entrance exam to college in 2008.
- Students were admitted to high schools by the Boston mechanism.
- 3570 students took the entrance exam to college. 2277 were in the art track and 1293 were in the science track.
- 2 out of 11 high schools are model schools.
- Variables:
  - Student achievements: SEEH, SEEC
  - Student demographics: gender, age, parental education, parental occupation
  - School characteristics: location, enrollment, student/teacher ratio, teacher certificate and experience

Estimation

- Validity Test:
  - McCrary's density test: no manipulation around cutoffs.
  - Balance of Background Variables: no relevant variables other than the treatment jump at the cutoff.
  - Balance of Self-Choice Characteristics: balanced self-choice characteristics (choice of track and attending entrance exam).
- Three types of effects:
  - Intent-to-treat (ITT) effect: the effect of eligibility
  - Treatment-on-treated (TOT) effect: the effect of attending specific high school
  - Local average treatment effect (LATE): the effect of attending specific high school for compliers who would attend a high school if eligible and would not if not eligible.
- Estimation methods:
  - Parametric estimation: global polynomial estimation
  - Nonparametric estimation: local polynomial (linear) estimation

ITT Estimation

	Art Track		Science Track	
	Parametric Estimation	**Non-Parametric Estimation**	**Parametric Estimation**	**Non-Parametric Estimation**
1	0.015 (0.075) [1558]	0.205*** (0.049) [513]	0.131 (0.207) [683]	0.183 (0.345) [232]
2	-0.039 (0.159) [1730]	-0.118 (0.189) [1005]	-0.176* (0.090) [833]	-0.032 (0.120) [519]
3	0.220 (0.166) [1684]	0.251 (0.209) [1028]	0.148 (0.194) [884]	0.150 (0.170) [444]
4	-0.094 (0.104) [1525]	-0.069 (0.124) [982]	0.102 (0.199) [888]	0.131 (0.199) [498]
5	-0.030 (0.104) [1575]	0.002 (0.120) [993]	0.263 (0.191) [900]	0.198 (0.200) [597]
6	0.000 (0.260) [1441]	-0.008 (0.272) [999]	-0.535** (0.191) [867]	-0.403 (0.239) [491]
7	0.083 (0.180) [1104]	-0.039 (0.134) [697]	0.298* (0.162) [752]	0.126 (0.165) [406]
8	0.084 (0.226) [274]	-0.049 (0.167) [179]	-0.187 (0.203) [303]	-0.222 (0.140) [233]
9	0.209 (0.259) [261]	0.006 (0.196) [166]	-0.078 (0.204) [294]	-0.136 (0.130) [216]
10	0.161 (0.172) [286]	-0.069 (0.177) [164]	0.099 (0.122) [307]	0.051 (0.110) [188]

Why No Effect Is Observed?

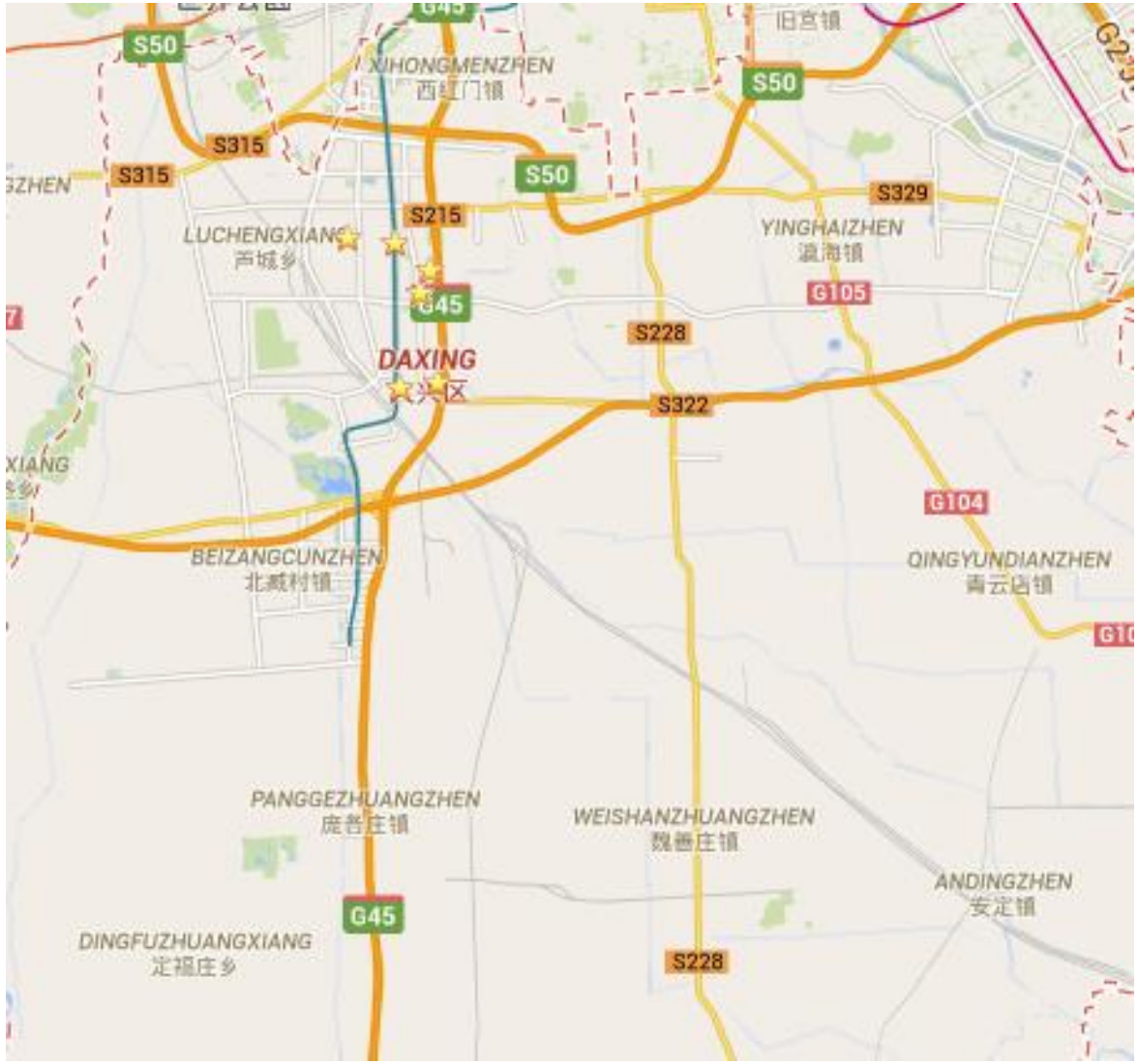
- Conjecture 1: Manipulation in the Boston Mechanism (Pathak and Sonmez 2008, 2013; Abdulkadiroglu and Sonmez, 2003; He, 2012)
  - Students tend not to truthfully report preference, for safety.
  - Students round the cutoff are more likely to manipulate.
- Conjecture 2: Trade-off between Commuting Cost and Benefit (Hastings et al. 2005)
  - Students from satellite towns are not willing to attend schools which are far away just for slight improvement in perceived quality.
- Both of Conjecture 1 and 2 lead to insignificant jump of the probability of enrollment at the cutoff.
- Conjecture 3: There is no real improvement of school quality at the cutoff.
- Conjecture 4: All of the measures of school quality, which show jump, do not have effect on student achievement.
- Conjecture 3 and 4 are about the effect of specific educational inputs.

Manipulating the Boston Mechanism

- Eligible Non-Compliers by Attending Schools

	1	2	3	4	5	6	7	8	9	10
1	--	566	603	607	607	608	609	609	609	609
2	**94**	--	437	463	448	464	464	464	464	464
3	**97**	**327**	--	508	505	519	541	548	548	548
4	**3**	**18**	**47**	--	**149**	168	178	200	200	200
5	**5**	**27**	**57**	252	--	255	280	289	289	289
6	**29**	**75**	**99**	**181**	**150**	--	261	289	289	289
7	**1**	**10**	**46**	**315**	**214**	**453**	--	512	512	512
8	**0**	**0**	**0**	**2**	**2**	**3**	**67**	--	170	**170**
11	**0**	**0**	**0**	**1**	**0**	**1**	**66**	**240**	**240**	**239**
9	**0**	**1**	**1**	**5**	**4**	**13**	**64**	**253**	--	**253**
10	**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)

Trade-off Between Commuting Cost and Benefit



Solutions

- Regression Kink Design (RKD)
  - It shares similar ideas and technical issues with RDD.
  - In RDD, what is supposed to change at the cutoff is the probability of attending a specific high school.
  - In RKD, what is supposed to change at the cutoff is the marginal effect of test score on the probability of attending a specific high school.
- Cluster Schools by Groups
  - Model School (1,2)
  - Center Area School (1-7)
  - Center Area Selective School (1-3)
  - Center Area Less Selective School (4-7)
- Weak Instrument Problem
  - In fuzzy RDD framework, weak discontinuity at cutoff is equivalent to weak instrument.
  - We can have robust confidence interval, but no robust estimator.

LATE Effect of School Groups on SEEC

	Art Track		Science Track	
	RKD	**RDD**	**RKD**	**RDD**
Model School	-0.448 (0.697) [1584]	-0.407 (0.804) [1005]	0.013 (0.429) [541]	-0.079 (0.271) [519]
Center Area School	0.351 (0.260) [995]	0.140 (0.427) [697]	-0.714 (0.532) [752]	-0.387 (0.530) [406]
Center Area Selective School	0.094 (0.146) [1623]	0.518 (0.404) [1028]	-0.087 (0.212) [684]	0.363 (0.350) [444]
Center Area Less Selective School	0.333 (0.251) [995]	0.132 (0.401) [697]	-0.820 (0.586) [752]	-0.482 (0.698) [406]

LATE Effect of Individual Schools on SEEC

	Art Track			Science Track		
	Effect 2SLS	**Wald Confidence Interval**	**AR Test**	**Effect 2SLS**	**Wald Confidence Interval**	**AR Test**
1	-3.899 [513]	[-8.388, 0.590]	**0.000 [-41.57, -2.012]***	-3.198 [232]	[-11.98, 5.579]	0.560
2	-0.337 [1005]	[-1.544, 0.871]	0.510	-0.071 [519]	[-0.551, 0.409]	0.772
3	0.513 [1028]	[-0.509, 1.536]	0.206	0.334 [444]	[-0.411, 1.080]	0.353
4	-0.549 [982]	[-1.821, 0.722]	0.559	0.960 [498]	[-2.831, 4.751]	0.488
5	0.006 [993]	[-0.536, 0.547]	0.984	0.584 [597]	[-0.405, 1.572]	0.299
6	-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]***
7	0.266 [697]	[-1.522, 2.054]	0.758	-0.701 [406]	[-2.476, 1.075]	0.421
8	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]***
9	0.019 [166]	[-1.130, 1.168]	0.974	13.46 [216]	[-333.5, 360.5]	0.264
10	0.201 [164]	[-0.646, 1.048]	0.673	0.255 [188]	[-0.965, 1.475]	0.619

Linking Exam Schools with Student Achievement

- A variation of Hanushek's educational production function (Hanushek, 1979)

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

  - Suppose on average there is no change in background characteristics B, then the two remaining elements, peer effects and school inputs, should capture the effect of attending a specific school. (But to identify their causal effects we need additional assumptions).

- Linking exam schools with peer effect and three observed school inputs: student/teacher ratio, teacher certificate and experience.

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

- If we assume that the unobserved inputs can only be indirectly affected by eligibility through the observed inputs

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

LATE of Peer Quality and School Quality on SEEC

	Art Track					Science Track				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
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

Conclusion Remarks

- We find modest to no effect of exam schools in China, including those which are of higher quality and more selective.
  - What is even worse is that we find significant negative effects among several subgroups.
- Various mechanical settings may explain such finding: manipulation in Boston mechanism and trade-off between commuting cost and benefit. Technically both of them lead to the weak instrument problem.
  - However, after taking care of them, we can still hardly find any remarkable positive effect.
- Selective exam schools have higher-quality peers, but there are also other omitted inputs.
  - The mixture of various inputs lead to insignificant effects.
- Our study does not prove that selective exam schools do not benefit students at all.
  - Long-term benefits, weak external validity, unmeasured school quality indicators
- There is no conclusive evidence of that those selective exam schools work as expected to improve test score.