Institute for Advanced Development Studies

Setting an example? Spillover effects of Peruvian Magnet Schools

By:

Mariel Bedoya, Bruno Gonzaga,
Alejandro Herrera & Karen Espinoza

Development Research Working Paper Series
No. 01/2019

March 2019

SETTING AN EXAMPLE?
SPILLOVER EFFECTS OF PERUVIAN MAGNET SCHOOLS *

Mariel Bedoya † Bruno Gonzaga ‡ Alejandro Herrera §
Karen Espinoza ¶

March 22, 2019

Abstract

In this paper we use a Multi-Cutoff Fuzzy Regression Discontinuity Design to evaluate spillover effects of students enrolled into Peruvian public magnet schools, *Colegios de Alto Rendimiento* (COAR), on educational outcomes of younger students in their schools of origin. Using administrative data from the Ministry of Education for 2016, we find that having at least one student admitted in a COAR school causes some negative spillover effects on math test scores of students from the following cohort. No evidence of statistically significant results is found for verbal and history test scores, nor for self-reported educational expectations. We discuss potential causes and reasons that may explain our results.

Keywords: magnet school, role model, educational achievement, educational expectations, education spillover.

JEL Classification: I20, I24, I28, O53.

*The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the official position of the institutions they are affiliated to. We thank the valuable feedback and suggestion from Libertad Gonzales, Luigi Pascali, Juan José Díaz, Gianmarco León, Alessandro Tarozzi, Marcos Agurto, Fernando Fernandez and Paul Soto. We also thank comments from the Ministry of Education of Peru, participants in the seminar at the Group for the Analysis for Development (GRADE), and from participants at the 2018 Annual Congress of the Peruvian Association of Economists (APE) and at the Ninth Bolivian Conference on Development Economics (BCDE). All remaining errors are ours.

†Innovations for Poverty Action. 396 Manuel Gonzales la Rosa Street, Magdalena. Lima, Peru. +511 2996846. mbedoya@poverty-action.org.

‡Central Reserve Bank of Peru. 441-445 Santa Rosa Street, Cercado de Lima. Lima, Peru. +511 6132000, ext. 23898. bruno.gonzaga@bcrp.gob.pe.

§Institute for Advanced Development Studies (INESAD). Irpavi. 6775, Altamirano Ave. La Paz, Bolivia. +591 22146069. aherrera@inesad.edu.bo.

¶Barcelona Graduate School of Economics. karen.espinoza@barcelonagse.eu.

1 Introduction

How does a *successful student* affect younger peers' academic performance and expectations in school? Despite evidence suggesting significant effects of role models such as older siblings, movie characters, and faculty members on educational outcomes of students (Nixon and Robinson, 1999; Nguyen, 2008; Qureshi, 2011; Nicoletti and Rabe, 2014; and Nielsen et al., 2015), there are no empirical evidence that have assessed the effect of senior successful students -potentially perceived as role models- on educational outcomes of younger cohorts from the same school. In this paper, we exploit the quasi-random admission process of public magnet schools in Peru, *Colegios de Alto Rendimiento* (COAR, hereinafter), in order to identify the spillover effects of third-grade enrolled students (successful students or role models, henceforth) on the educational achievement and expectations of second-grade high school students in schools of origin.

For this, we compare educational outcomes of second-grade students in high schools which in the previous year had a student who scored barely above the admission cutoff, with educational outcomes in schools which in the previous year had a student who scored barely below the admission cutoff. In the second case it is important to understand that this also implies no students from these schools were admitted into a COAR. The plausibly exogenous variation in admission scores generates exogenous variation in the existence of successful students in schools of origin. We hypothesize that successful students leaving their schools of origin have the potential to be perceived as role models by their younger peers, that is by second-grade students who can become eligible for admission to a COAR school at the end of the year. The admission of a role model might inform and incentivize younger students to increase efforts during the school year, which could in turn impact on their educational achievement and aspirations. Our analysis differs from typical regression discontinuity (RD) settings in the fact that we attempt to assign the running variable at the school level in order to retrieve the causal estimates of having enrolled students in COAR on educational outcomes and expectations of younger students. This identification strategy raises important methodological issues which we address in Section 5. Nonetheless, it must noted that the fact that only up to the top three students of second grade are eligible to apply makes it easier to determine the running variable at the school level without threatening our identification strategy.

COAR schools are public boarding schools introduced in 2015 that aim to provide talented students from the second year of high school an education service of the highest standards during the last three years of secondary education.[1] Conditional on satisfying specific eligibility requirement, the most important of which is being among the top three of the class at the end of the academic year, admission of students into the COAR of their region is determined by the order of merit in which applicants place according to a final score obtained after the selection process. By 2016, the intervention of COAR offered 2,400 vacancies for twenty two operating COAR schools nationwide (there is one COAR school per region) and only the top 22% of applicants were offered admission.

The COAR intervention represents an interesting setting to address this question for several reasons. First, nowadays there are many governments who invest a large amount of public resources in selective schools such as *Colegios de Alto Rendimiento* in Peru,

[1] We do not use admission results of 2015 due to data limitations.

charter and magnet schools in the US, *Liceos Bicentenarios de Excelencia* in Chile, and Grammar Schools in the UK, among others. In the case of Peru, in 2016, the average expenditure per student per month of the COAR intervention was around US$ 2,058. Second, literature surrounding selective schools has focused on first order effects, where the balance is mostly positive on diverse educational outcomes (Cullen, Jacob, and Levitt, 2005; Lillard and Else-Quest, 2006; Bifulco, Cobb, and Bell, 2009 and Esposito, 2010); however, evidence on second order effects such as spillover effects on younger cohorts is scant. Thirdly, and perhaps most importantly, from a policy perspective it is important to understand and measure externalities of public interventions that have the potential to influence motivation and subsequent effort of students and teachers, as it is through the integral understanding of such policies that we will be able to design, scale, and justify interventions that albeit costly may be capable of setting an example and increasing average performance of traditional schools through multiplier effects.

In this first attempt of quantifying potential spillover effects of COAR, our findings suggest there is no evidence of positive effects of enrolled students in the subsequent educational achievement and aspirations of students of second grade in schools of origin. Even more, we find some evidence of negative spillover effects on students in schools of origin. Concretely, using a non-parametric specification, schools with at least one admitted student in 2016 are around 0.45 and 0.53 standard deviations below the mean of the control group in standardized math test scores. Results are statistically significant at the five and ten percent level. Note, however, that we need to interpret this apparent unwanted effect carefully. Although in our case we do not possess yet all necessary data to conclude on the mechanisms that may drive this negative effect, we offer a preliminary hypothesis: we hypothesize this result may be related to indirect composition effects arisen due to the positive signaling of schools with admitted students and selective sorting of top students in control schools in order to increase their probability of being eligible for admission into a COAR at the end of the academic year. Some preliminary supportive evidence of this is that when we perform separate regressions for the bottom and upper quintile of the test score distribution, we find that negative effects are concentrated in the upper quintile, suggesting that selective sorting of top students into control schools may be happening.

The rest of the document is organized as follows, Section 2 presents a brief literature review of first order effects of selective schools as well as recent attempts to identify spillover effects of role models. Section 3 presents a description about the COAR intervention and the admission process; Section 4 presents information about data sources used for the analysis; Section 5 discusses the identification strategy and the validity of its assumptions; Section 6 presents spillover effects of admission and enrollment of potential role models in COAR schools on younger cohorts. Finally, Section 7 discusses possible interpretations of the effect and offers concluding remarks.

2 Literature Review

2.1 Magnet Schools in Perspective

The impact of magnet schools on academic attitudes and educational performance has been broadly documented in the economic literature. Several studies have analyzed the first order effects of magnet schools on their enrolled students' educational performance compared to those who attend regular public schools, and in general, the balance is positive. Experimental studies find mixed results when analyzing educational achievement, some studies show that magnet school students achieve higher results in GPA, math, and language test scores (Betts et al., 2006; Goldrig and Phillips, 2008 and Bifulco, Cobb, and Bell, 2009); while other studies find positive but not significant differences in educational achievement between the magnet schools students and the regular ones (Ballou, Goldring, and Liu, 2006; Ballou, 2007; and Esposito, 2010). Additional positive and significant results related to the academic behavior are found in this literature, such as the increase in credit accumulation, reduction of dropout ratios and reduction of absenteeism (Crain et al., 1992; Kemple and Snipes, 2000; and Cullen, et al. 2005). There is also evidence of positive effects on academic attitudes and aspirations, a greater sense of community at school and more peer support (Lillard and Else-Quest, 2006; and Cobb et al. 2009).

Similarly, a handful of papers have used a quasi-experimental approach to answer the same question and they also find some positive results on educational outcomes. Mixed results on GPA and other academic tests emerge from this literature (Adcock and Phillips, 2000; Goldschmidt and Martinez, 2004; Yang et al., 2005; Dohrmann et al., 2007; Judson, 2014; and Betts et al., 2015); however, effects on academic behavior show a positive balance in variables such as advance course-taking (Rice et al., 2015), and graduation ratios (Silver, Saunders, and Zarate, 2008; and Siegel-Hawley and Frankenberg, 2011). Furhtermore, some of these effects are stronger in specific groups, especially if the students face any limitations in their learning process (below average readers, sub-urban schools, etc.).

To our knowledge, there are no studies that separately identify the first order effect and the potential spillover effects of magnet schools; however, the studies that measure spillover effects of charter schools in the US and UK are related to our research (Bifulco and Ladd, 2006; Sass, 2006; Booker et al., 2008; and Clark, 2009)[2]. These authors exploit the variation in location and year of creation of charter schools to analyze the first and second order effects over test scores of traditional public schools' students; however, impact on student achievement remains unclear. Jinnai (2011) makes a broad analysis separating the spillover effect based on the creation of new charter schools in neighborhoods of North Carolina. The author finds negative spillover effects on students from grades that do not directly compete with the grades offered by charter schools, he suggests that these findings are due to the introduction of market competition in the local monopoly of public schools. Thus, the traditional public schools reallocate their resources to increase the quality of the grades that directly compete with the charter school to maintain the level of enrollment, at the expense of subtracting resources from the grades that face less competition.

Although studies about spillover effects of charter schools can be compared with our case of analysis, the divergence falls on the *theory of change* they present to explain their findings. In our case, the creation of COAR schools could not be considered as an introduction of direct competition to the public schools services, mainly because the supply of COAR services is scarce (only one COAR per region) and the regular public schools have financial limitations to imitate the COAR services. Nevertheless, we believe that the competitive environment can modify the strategy of traditional public schools, if having a student admitted in a COAR generates a positive quality signal for the school of

[2]Magnet and charter schools have many features in common, however, charter schools have less coverage than the first ones and there is variation regarding the quality of services they offer.

origin and it has the potential to attract students from other public schools. However, it could still be the case that control schools also modify their behavior due to the pressure of the competition, and in order to maintain the enrollment level they attempt to improve the quality of their services; then the net effect is unclear.

2.2 About Role Models

The economic literature regarding the effects of role models on educational outcomes is recent and scarce; also the channels of transmission of this effect have not been clearly established. For instance, research about sibling effects on educational outcomes is related to this argument, proposing that the educational achievement or behavior of older siblings may have an impact on younger; and siblings' achievement (Qureshi, 2011; and Nielsen et al., 2015). Nicoletti and Rabe (2014) find that only high-achieving older siblings positively affect the academic performance of the younger ones; no effects have been found in the opposite direction. The authors point out that this effect can be explained by the frequent interaction of the siblings at home generating an imitation process, where the older sibling becomes the main reference for the younger siblings and can influence their educational performance and aspirations. Finally, authors argue that transmission of information between siblings is fundamental to generate a positive effect: older siblings transmit relevant knowledge about the cost and benefit of exerting effort, and share specific information about the school and teachers. Recently, a study of Barron, Basurto and Cuadra (2018) examining peer effects of enrolled students into a high-achiever's school of Peru finds similar results; admitted students increased their siblings' GPA by 0.33 standard deviations and their math grades by 0.22 standard deviations.

The effect of role models on educational outcomes is also studied in areas such as investment in human capital and gender studies. Nguyen (2008) compare the effects of providing information about the returns in education versus the introduction of a role model. The author finds that providing information about returns improves test scores, and the effect is larger for those students whose initial expected returns were below than the provided information. However, the introduction of a role model only has positive effects on test scores when the role model presents a similar socioeconomic profile to that of students in the school. This is consistent with Ray (2006) who suggest that role models might be influential only when they come from a similar background or ethnic group.

Regarding gender studies, Nixon and Robinson (1999) find that *female role models*, are important for girls' performance in high school. Educational attainment of female students is positively and significantly correlated with the percentage of female faculty, but the magnitude of the correlation is small. Bettinger and Long (2005) find similar effects, female faculty members in charge of the initial courses in college do increase student interest in a subject and the likelihood that a female student will take additional credit hours or major in particular subjects. Both male and female faculty staff may be role models for young women, however, in these studies female role models are more relevant for young women in the sense that they directly represent what they can achieve in the future.

We highlight two ideas from these papers. First, direct and clear information about experiences of role models to the treated group is needed to observe effects. Second, the influence of the role model depends on how easily students imagine themselves achieving what their role model has achieved. This connection is usually based on observable characteristics that both sides have in common, and the shared experiences become a realistic expectation of what they can achieve by behaving similarly.

3 About the COAR Intervention

Colegios de Alto Rendimiento (COAR) is an educational program implemented in 2015 that seeks to provide high-performing students from public secondary schools of Peru a high-quality educational service that strengthens their academic, soft skills, artistic and sporty competences during the last three years of high school. COAR schools function as boarding schools. By 2016, the program offered 2,400 vacancies in 22 regions of Peru (out of 24); each region had its own COAR school offering 100 vacancies, except for Lima (the capital city) which offered 300 vacancies. [3]. Before the implementation of the COAR program there were two operating private schools for high-achieving students in two regions of Peru: in Lima (the capital city) and in Ayacucho. Both schools became part of the program later in 2015.

Students admitted in COAR schools are offered a scholarship that covers tuition and living expenses. These include, but are not limited to, meals, housing, and school supplies such as pencils, notebooks, books and even a personal laptop. The scholarship does not cover relocation costs. In addition, students have the option to earn a diploma of completion of the International Baccalaureate program and have access to several university scholarships after graduation. In 2016, the expenditure per student in a COAR was around US$2,058 (this includes administrative costs as well). Since its implementation, the intervention has not assigned a specific budget for diffusion; and thus one of the most effective channels for diffusion in the first years of implementation has been through current COAR students.

The admission process of students into COAR starts around November each year and results are published at the end of February. For the purposes of our study, we will describe eligibility requirements and the admission process implemented in 2016. In order to be eligible to apply to a COAR school in 2016, students had to fulfill the following requirements: *(i)* be a Peruvian citizen or resident, *(ii)* be 15 years old until march 31, *(iii)* hold a parental or legal tutor authorization, *(iv)* have completed the first two years of high school in a public school, *(v)* have an annual GPA of 15 over 20 or more in the second year of high school, *(vi)* have occupied one of the top three places of the class in the second year or one of the five first places in a national contest organized by the Ministry of Education of Peru. [4]

Once a student was eligible to apply to COAR, the selection process of applicants consisted of three phases. The first phase consisted on a cognitive ability test, which measures competences related to the curricular requirements of the Ministry of Education. The second and third phases involved an individual interview and a soft skills test. After the three phases had concluded, an index was constructed as a weighted average of the

[3] These schools offer 60 vacancies for students studying in public schools from the same region. Lima offers 180 vacancies. In 2016 There are two regions and one constitutional province that do not have a COAR school in the region: Ancash, Tumbes and Callao. Each of the latter are offered fixed vacancies in other COARs conditional on applicants meeting the specific eligibility requirements. Ancash has 40 fixed vacancies, and Tumbes and Callao have 10 fixed vacancies each.

[4] Only 0.55% of enrolled students in 2016 were admitted due to the latter case.

scores obtained in each phase, where the cognitive ability test weighted fifty percent, the interview weighted thirty percent and the soft skills test weighted twenty percent. This index became the final score which determined admission of students into a COAR.

In the beginning of the selection process, students were asked to choose two COAR schools to which they would be willing to attend if they successfully passed the selection process mentioned above. Students from schools in regions where there was an existing COAR had to necessary choose the COAR from the same region of their current school (or school of origin from hereafter) as first option. On the other side, students from schools where there was no COAR operating could choose as first option any COAR in the country.

Once the final score was computed, the admission and assignment process of students into COAR schools followed a specific procedure. First, final scores were sorted descending by region and the first 60 students in the ranking of each region with a COAR were automatically assigned to the COAR of that region. Second, students from regions with no COAR were assigned; the top 40 students of the Ancash region were assigned to their first choice, while the top 10 of Tumbes and Callao were assigned to their first choice as well. Third, final scores of the rest of applicants were sorted descending at the national level and the last 900 vacancies were offered to the first 900 students among this new sorted list. Students selected in this phase were offered a vacancy to their first or second choice of COAR upon availability. Fourth, if among the latter 900 admitted students their first and second choices had no remaining spots available, students were offered a vacancy in other COARs with availability. Finally, rejections from the previous four steps are counted and students following in the ranking were offered a vacancy to fill available spots in each COAR.

4 Data

We examine the impact of having at least one enrolled in COAR in 2016 (one year after the program started) on educational achievement and expectations of students from the next cohort from schools of origin. Since regions of Lima and Ayacucho had selective schools prior to COARs which served as local experiences for the design of the COAR intervention, we exclude both regions from the analysis. We also exclude regions with no operating COAR. Administrative data on the 2016 application process was obtained from the Ministry of Education of Peru. This data includes information about all the steps of the COAR admission process in 2016: scores for the cognitive ability, interview and soft skills tests, final score computed, admission results and final enrollment, final order of merit according to the steps of the assignment process described in Section 3, and some specific information of applicants such as their school of origin, their first and second choices of COAR, among others. In 2016, there were 10,454 applicants from 4,527 schools out of which 2,351 (22.5%) from 1,604 different schools finally enrolled in a COAR. In addition, despite the fact that 72.23% of schools with applicants had more than 1 student who applied to a COAR school, only 32% of schools with admitted students sent more than one student to a COAR.

We measure educational achievement and self-reported expectations using the National Student Assessment (ECE) test scores of 2016. The ECE is administered by the Ministry of Education to students of second year of secondary school and measures competence in verbal, mathematics and history reasoning according to the National Curriculum[5]. The evaluation is conducted annually since 2015 for all private and public high schools of Peru in November. Jointly with the ECE, students are asked to fill in a questionnaire with information about the schooling level of their parents, household and individual characteristics such as mother tongue, expectations about their highest educational attainment and additional questions about the students' opinion of teachers as well as other features about the school environment. The timing of the evaluation assures that we will be able to observe any change in test scores that might be related to having older students admitted or enrolled in a COAR in the beginning of the same year. Finally, additional covariates at the school level are obtained from the National School Census, a public database collected by the Ministry of Education annually which contains information about school characteristics related to infrastructure, equipment, teachers, enrollment, among others.

5 Empirical Strategy

The admission of students into COAR schools of each region is a deterministic and discontinuous function of the merit of order in which a student places according to the final score obtained during the admission process described in Section 3. It must be noted however, that surpassing the threshold only increases the probability a student will end up going to a COAR. This jump in the probability of enrollment suggests the use of a Fuzzy Multi-Cutoff Regression Discontinuity (RD) Design to evaluate spillover effects of COAR on students from schools of origin (Calonico et al., 2016).

In order to analyze spillover effects on younger students from schools of origin, it is crucial to define treatment status and the running variable at the school level. Redefining the running variable diverges from typical RD applications and thus raises important methodological concerns about the underlying identifying assumptions which will be discussed below. To examine the effect of interest, we define treatment status as *having at least one student* enrolled in a COAR, and the running variable as *the maximum score of applicant students*. Setting the running variable as the maximum score among applicants from a specific school ensures that the running variable can be interpreted as the *minimum effort* the school would have had to make in order to have *at least one* admitted (and potentially enrolled) student[6]. Formally, the pooled estimand in this context is given by:

$$\tau_{FRD} = \frac{\lim_{\varepsilon\to 0^+} E\left[Y_{rsi}|X_{sr} - C_r = \varepsilon\right] - \lim_{\varepsilon\to 0^+} E\left[Y_{rsi}|X_{sr} - C_r = -\varepsilon\right]}{\lim_{\varepsilon\to 0^+} E\left[D_{sr}|X_{sr} - C_r = \varepsilon\right] - \lim_{\varepsilon\to 0^+} E\left[D_{sr}|X_{sr} - C_r = -\varepsilon\right]} \tag{E1}$$

where Y_{rsi} is the outcome of educational achievement for student i of second grade of secondary school s in region r; X_{rs} is the running variable, which as described above varies across schools and is determined as the maximum grade obtained by third grade students who applied to COAR[7]. Finally, C_r is the regional cutoff, determined by the minimum score at the school level of admitted third grade students in the region r; and

[5]For more details, see: http://www.minedu.gob.pe/curriculo.

[6]For instance, in a school where the three top students applied and were admitted into a COAR, the maximum score among the three is imputed as the running variable at the school level. Regardless of the application of the other two students, the application of the student with the maximum grade alone determines the treatment status of the school.

[7]We expect spillover effects on second graders to occur between March (month in which regular classes start) and November (month in which the National Student Assessment is taken each year).

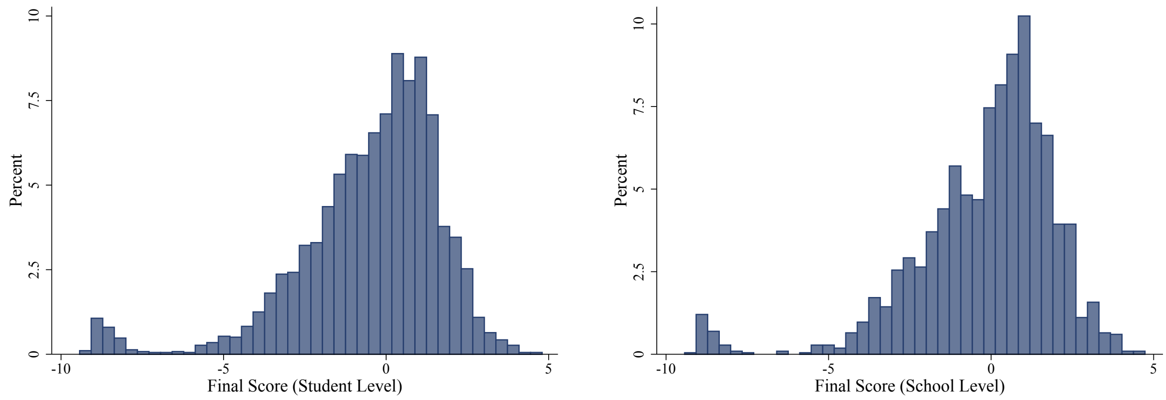
D_{sr} is the potential treatment status for a given school in the region r. As mentioned above, treatment status is determined as having at least one student being admitted into a COAR in 2016.

5.1 Validity of Identifying Assumptions

The RD approach used in this paper requires three identifying assumptions. First, the density of the running variable should be continuous around the threshold. Any evidence of manipulation around the cutoff would invalidate the quasi-random nature of the COAR admission process. Given that we re-define the running variable at the school level, in our setting it is important to verify the continuity of the distribution of final scores around the cutoff both at the student level (applicants) and at the school level.

Manipulation of the final score at the student level is unlikely to happen due to the fact that the cutoff value which determined admission was defined only after all final scores were computed and students ranked accordingly. Nonetheless, we replicated the assignment process described in Section 3 and confirm that students who achieved high enough scores were automatically admitted into the program. The distribution of final scores of applicants can be found in Panel A of Figure 1.

FIGURE 1: Density function of the COAR scores

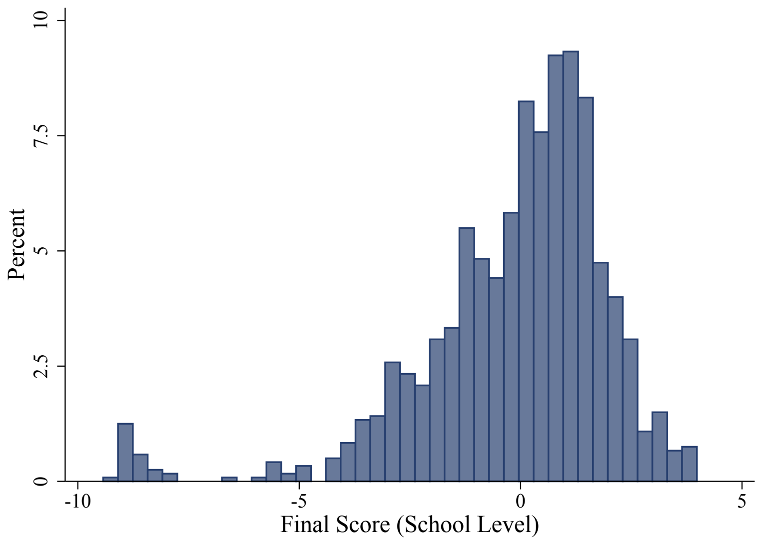


(A) (B)

A simple visual inspection of Figure 1 evidences continuity of final scores of applicants around the cutoff; however, this does not appear to be the case at the school level. Panel B of Figure 1 shows some agglomeration of schools above the threshold. We perform the formal density test proposed by McCrary (2008) to confirm our suspicions. The null hypothesis of the test is the non existence of discontinuity at the cutoff point and involves an estimation of the density near the threshold. Such density estimation is based upon a non-parametric local polynomial density estimator, first introduced by Cheng, Fan, and Marron (1997). As expected, the null hypothesis cannot be rejected for the density function of final scores of applicants (p-value $= 0.4178$), but it is rejected for the density function constructed at the school level (p-value $= 0.0406$).

This latter finding is not necessary a sign of manipulation of scores done by schools. One of the implications of redefining the running variable at the school level is that we may be biasing the density of the running variable at the threshold if applicants just below the cutoff come from schools with at least one admitted student. This implies that schools more prone to bias the density are those with two or more applicants. In order to address this concern we test in which regions agglomeration of schools just above the threshold is more severe. We run the McCrary test by region for the COAR score at the school-level, and find that regions with higher bunching are likely to be those with more schools and thus with more applicants[8]. To be conservative, we opt to keep the regions in which the test yielded a p-value equal or higher than 0.30. Such regions are Cajamarca, Pasco, Loreto, Piura, Huancavelica, Moquegua, Junin, San Martin, Amazonas, Ucayali and Lambayeque; we will refer to this sample as the *sub-sample* in the remaining of the document. For the sub-sample (whose density function can be found in Figure 2), the McCrary test cannot reject the null hypothesis of continuity (p-value = 0.1643).

FIGURE 2: Sub-sample density function of the COAR scores



[8]Such regions are Cusco, La Libertad, Puno, Arequipa, Ica, Apurimac, and Huanuco.

Next, following the standard procedure, we assess the plausibility of the *unconfoundedness* assumption, which requires the error term to be continuous at the threshold in the sub-sample chosen. The usual way of testing this passes through checking the non-evidence of discontinuity in observable characteristics. It is crucial to verify this assumption since it implies that schools below the threshold are valid counterfactuals for schools above. In our case, we test the null hypothesis of discontinuity in the following observable characteristics of 2014 [9]: area of location (rural/urban), percentage of students with a mother tongue different than Spanish, number of students in the school, number of teachers and number of classrooms. We estimate the reduced model using a linear and a quadratic polynomial fit, along with a non-parametric specification. Results are presented in Table 1. We additionally include three variables from 2015: schools with at least one student admitted into a COAR the previous year and standardized math and language test scores.[10] Although we do not necessarily expect variables of 2015 to be balanced across treatment and control schools since these variables belong to the first year of treatment (and thus scores of 2016 could have been affected due to results of 2015), we include these variables because it can potentially help us explain results of found in our outcomes of interest.

As expected, we observe no statistically significant differences at the threshold in pre-treatment covariates of 2014. Nonetheless, schools above the cutoff seem more likely to be rural, have more students with a native language different from Spanish, have less students overall, and have slightly less teachers and classrooms. It is interesting to note that we do not observe statistically significant differences among schools slightly above and below the threshold in the likelihood of having at least one student enrolled in a COAR in 2015, nor in standardized math and language test scores of 2015. This latter finding allows us to interpret main results of 2016 without considering potential accumulated effects arisen from the first year of intervention. Altogether, we consider findings as supporting evidence of the validity of the *unconfoundedness* assumption.

Finally, we verify the relevance of our instrument. In other words, we verify that the probability of participating in the program discontinuously increases once the cutoff score is surpassed. As anticipated, in Figure 3 we observe a jump in the probability of having at least one enrolled student in a COAR when the running variable is above the threshold. Concretely, the probability of having at least one enrolled student increases in approximately 50 percentage points once the threshold is surpassed. First stage estimates are also presented in Table 2.

6 Results

6.1 Results for the sub-sample

Table 2 presents estimates of the effect of having at least one student admitted (*reduced form* estimates) and enrolled (*fuzzy regression discontinuity* estimates) in a COAR school[11]. We include reduced form results since we consider that the effect of having a

[9]We use characteristics of 2014 since in 2015 the program had already started.

[10]Unlike 2016, in 2015 a history test was not administered to students.

[11]We first estimate the model described above for the total sample as reported in Table 8 in Appendix. Nevertheless, in this section we present and explain the estimated results for the sub-sample as defined

Table 1: Estimates of discontinuity in covariates per school

	Coefficients (standard errors)			
Dependent variables	[1]	[2]	[3]	Mean non-admitted
Rural schools(2014)	0.018	0.075	0.035	0.282
	(0.079)	(0.128)	(0.088)	
Mother tongue(2014)	2.381	7.685	4.648	5.983
	(4.370)	(6.999)	(4.887)	
No. of students(2014)	-41.155	-35.841	-36.579	224.9
	(45.681)	(69.395)	(50.105)	
No. of teachers(2014)	-1.191	0.002	-0.608	16.22
	(2.465)	(3.608)	(2.667)	
No. of classrooms(2014)	-0.165	-0.031	-0.096	1.978
	(0.297)	(0.428)	(0.318)	
Treated schools(2015)	-0.013	-0.046	-0.019	0.247
	(0.085)	(0.140)	(0.097)	
Math test scores(2015)	-0.041	-0.036	-0.043	0.000
	(0.116)	(0.169)	(0.125)	
Language test scores (2015)	-0.002	0.012	0.006	0.000
	(0.108)	(0.170)	(0.122)	
No. Schools	482	482	482	
Specification	Linear	Quadratic	Kernel	

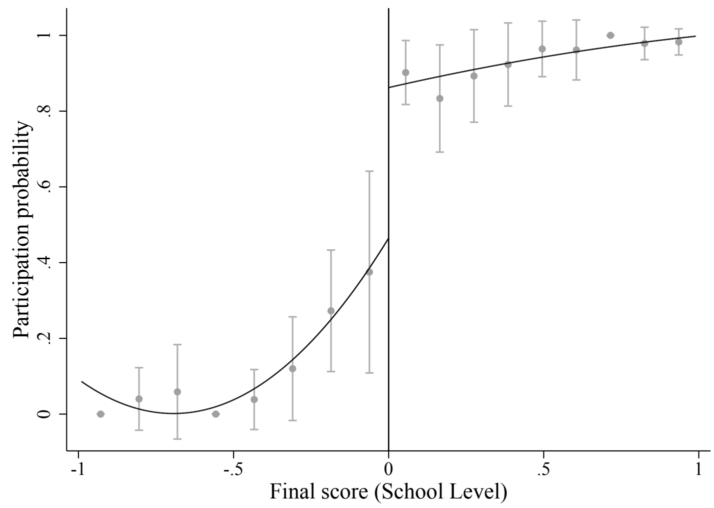
Note: *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. Robust standard errors reported in parentheses. Each coefficient comes from a different regression to test for discontinuity in pre-treatment covariates at the school level. Columns 1 to 3 include a polynomial of degree 1, 2 and a Kernel specification. Two variables from first year of the program (2015) are included: percentage of treated schools in 2015 and math test scores. Regressions from the last row are estimated at the student level with clustered errors at the school level.

successful student in schools of origin could arise just because of having an admitted student in the school, not solely due to having admitted students who enrolled in the COAR school.

Panel A of the table shows our results from the first stage, which measures the discontinuity in the probability of having at least one enrolled student if at least one student is admitted. Panels B, C and D show instead the effect of having at least one student admitted or enrolled in a COAR on language, math and history test scores for the following cohort, respectively. Additionally, we explore if results are robust to different parametric specifications of the model and to the inclusion of baseline characteristics of 2014 as control variables. Control variables included are rural location, number of students, number of teachers, classrooms per school, and a dummy variable equal to one if the school had students enrolled in a COAR school in 2015. Columns 1 to 3 include, in order, a polynomial of degree one, two and a Kernel specification on the running variable

and justified in the previous section.

FIGURE 3: Program eligibility



Note: the graph above shows the proportion of schools who sent at least 1 student to a COAR school. The running variable is the final score of the admission process, normalized to zero. A quadratic polynomial on each side of the cutoff point is fit to the data.

(these polynomials are interacted with the eligibility indicator). The inclusion of a non-parametric approach is justified by the advantage of not relying on a specific functional form assumption, thus avoiding inaccurate model specification.

As can be seen from Panel A, estimates indicate that the probability that at least one student in the school of origin enrolls in a COAR increases in a range of 48 to 53 percentage points when the eligibility criteria is met. These results are significant at the ten and five percent level and also are robust across different specifications. Results for the language and history test scores show negative and non statistically significant results. For instance, Panel B shows that estimates from the reduced form model range from -0.16 to -0.07 standard deviations below the mean of the control group, while the fuzzy RD estimates indicate that the existence of previous cohort students enrolled in COAR decreases the performance in language test scores of the following cohort in the school of origin in a range of 0.13 to 0.41 standard deviations. In Panel D, reduced form estimates range from -0.13 to -0.07 standard deviations below the mean for history tests whereas the Fuzzy RD estimates range from 0.06 to 0.01 standard deviations.

In contrast to previous results, math test scores do show some statistically significant results of having at least one student enrolled in a COAR. Specifically, reduced form estimates presented in Panel C range from -0.38 to -0.15 standard deviations and are significant at the five percent level for the quadratic and Kernel specifications, while Fuzzy RD estimates range from -0.28 to -0.96 and are statistically significant at the ten percent level only for the non-parametric model. We interpret this findings as suggestive evidence of negative spillover effects of having at least one senior student enrolled in a COAR on math test scores of younger cohorts. This result is consistent with previous empirical evidence that suggests math test scores are more sensitive to educational interventions. Finally, in Panel E we present results for the academic expectations indicator. This indicator is a dummy equal to one if students expect to attain university and even post

TABLE 2: Second stage estimates for the sub-sample (Bandwith = 1)

	Coefficient (standard errors)					
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: First stage						
Z (score $\geq$ 0)	0.526***	0.398**	0.484***	0.526***	0.394**	0.480***
	(0.098)	(0.168)	(0.117)	(0.096)	(0.162)	(0.112)
Panel B: Language						
Reduced Form	-0.106	-0.163	-0.129	-0.068	-0.124	-0.088
	(0.114)	(0.205)	(0.138)	(0.097)	(0.173)	(0.115)
Fuzzy RD	-0.202	-0.411	-0.266	-0.129	-0.315	-0.183
	(0.227)	(0.579)	(0.307)	(0.190)	(0.474)	(0.249)
Panel C: Mathematics						
Reduced Form	-0.185	-0.384**	-0.258**	-0.148	-0.336**	-0.215**
	(0.117)	(0.184)	(0.131)	(0.102)	(0.156)	(0.109)
Fuzzy RD	-0.351	-0.965	-0.533*	-0.282	-0.851	-0.448*
	(0.239)	(0.681)	(0.323)	(0.205)	(0.573)	(0.262)
Panel D: History						
Reduced Form	-0.109	-0.134	-0.119	-0.071	-0.087	-0.076
	(0.091)	(0.161)	(0.109)	(0.078)	(0.140)	(0.092)
Fuzzy RD	-0.207	-0.337	-0.245	-0.135	-0.221	-0.158
	(0.183)	(0.459)	(0.245)	(0.153)	(0.383)	(0.201)
Panel E: Expectations						
Reduced Form	0.007	0.018	0.009	0.017	0.026	0.017
	(0.029)	(0.049)	(0.033)	(0.025)	(0.041)	(0.028)
Fuzzy RD	0.013	0.046	0.019	0.032	0.065	0.036
	(0.054)	(0.120)	(0.068)	(0.047)	(0.104)	(0.059)
Observations	27,026	27,026	27,026	27,026	27,026	27,026
No. of schools	482	482	482	482	482	482
Polynomial	Linear	Quadratic	Kernel	Linear	Quadratic	Kernel
Other controls	No	No	No	Yes	Yes	Yes

Note: The table reports estimates of the spillover effects of admission into a COAR (Reduced Form) and enrollment in COAR (Fuzzy RD) on three educational outcomes: Language Test Scores (Panel B), Mathematics Test Scores (Panel C), History Test Scores (Panel D) and self-reported educational expectations (Panel E). Eligibility for admission is an indicator equal to one if there was at least one student admitted into a COAR in the previous cohort such school; the running variable is defined as the maximum score observed among applicants of the previous cohort students. Columns 1–3 include, in order, a polynomial of degree 1, 2 and Kernel on the running variable (these polynomials are interacted with the eligibility indicator) following Calonico, et al. (2014a) methodology. Columns 4–6 additionally control for pretreatment characteristics of schools for the year 2014. Standard errors clustered at the school level in parenthesis. *, **, and *** denote significance at 10%, 5%, and 1% levels.

graduate education in the future. We do not find statistically significant results for this indicator, although coefficients are positive across specifications.

A concern about previous results is that they may be sensitive to the bandwidth and sub-sample chosen and hence biased. In order to address this potential issue, we run two robustness checks using more conservative bandwidth levels and a more conservative sub-sample. The first exercise uses two different bandwidth levels: $= 0.7$ and $= 0.9$, whereas the second robustness check restricts the chosen sub-sample to regions of Peru that pass the McCrary test with a p-value higher than 0.33 (in contrast to the previous threshold of 0.3). Results from both robustness checks can be found in Tables 3 and 4 of the Appendix[12]. We observe that choosing a narrower neighborhood around the cut-off and a different sub-sample does not alter the sign of coefficients. Although statistical significance remains unchanged using the bandwidth of 0.9, we lose significance in some of the estimated coefficients when we restrict the bandwidth as much as to 0.7.

6.2 Quantile Regressions

Despite negative results found above (which we will discuss in the following section), evidence from the literature presented in Section 2 suggests that significant effects of role models are more likely to be found when students identify themselves with the role model, mainly because of sharing similar characteristic that makes them think they can achieve what role models achieved before. In the case of COAR, it is possible that only high achieving students from the following cohort identify with their older peers given that they are the only ones that can see themselves achieving what their older peers achieved. At this point it is crucial to remember that an eligibility requirement to be able to apply to COAR is to be among the first top three students of the class at the end of the academic year. Under the assumption that the upper distribution of standardized national test scores corresponds to high achieving students, in this section we present estimates of the effect of interest for the upper and lower quantiles of the distribution of test scores. According to what was found by the previous literature, we expect to find that statistically significant effects are concentrated in the upper quantile, where top students are more likely to be located and relate to previous successful students. In Table 5 in Appedix, we present results from this exercise.

Consistent with findings from the previous section, no statistically significant results are found for language and history test scores although results remain negative. However, an interesting pattern emerges for math test scores. As expected, separate regressions for the lower and upper quintile of the distribution of math test scores show statistically significant results are concentrated in the upper quantile. Specifically, younger students from the upper quantile with at least one enrolled student are 0.84 standard deviations below the mean of the control group. Although mechanisms that might drive this effect are yet impossible to address, in the section below we provide some possible interpretations that may drive these negative spillover effects.

[12]We only present estimates of math test scores, in line with significant coefficients found in this section.

We replicate the same exercise of robustness for this section; results can be seen in Table 6 and 7 of the Appendix. Again, these checks confirm the negative sign in the estimated coefficients and the decrease in the statistical significance when the bandwidth is restricted as much as 0.7. However, coefficients are still significant when evaluated at a bandwidth level of 0.9.

7 Discussion & Conclusions

In this paper we attempt to identify spillover effects of the intervention of COAR on educational achievement and academic expectations of students of the second grade from public high schools in Peru. Using a Multi-Cutoff Fuzzy Regression Discontinuity Design on a sample of regions with schools that had no evidence of manipulation of final scores, we find no evidence of positive spillover effects of having at least one student admitted or enrolled in this type of selective school in the beginning of 2016 on standardized test scores and self-reported expectations, measured by the National Student Assessment (ECE) administered in November of 2016.

Considering the possibility of heterogeneous impacts depending on which students are more likely to perceive enrolled students as role models, we perform separate regressions to isolate the effect of interest for the bottom and upper part of the distribution of students' scores in the ECE. An important assumption of this exercise is that students in the upper part of the distribution are more likely to be in the top rank of their class and have incentives to work harder during the year to be able to be admitted into a COAR school. By doing this, we still find no positive spillover effects of sending at least one student to a COAR school on educational achievement and expectations of younger students.

More strikingly - and contrary to our main research hypothesis - overall results exhibit a negative pattern. Specifically, schools with at least one admitted student are on average 0.22 to 0.26 standard deviations below the mean of schools under the threshold in math test scores; similarly, schools with enrolled students in COAR are, on average, 0.45 to 0.53 standard deviations below the control group's mean in math test scores. These results are statistically significant at the five and ten percent level. We find no significant effects for language and history test scores although coefficients across all specifications have a negative sign. The heterogeneity analysis shows that negative and significant effects are concentrated in the upper part of the distribution (fifth quantile), which we assume are identified as the top ranking or high-achieving students in their school and thus more likely to be affected by the existence of *successful students* in their schools.

Although we do not have yet the necessary data to interpret negative and statistically significant results found in math test scores, we offer some preliminary thoughts. First, we hypothesize effects found could be the result of *indirect composition effects*. These hypothesis is in line with Jinnai (2011), who argues that the presence of selective schools introduces competition among regular public schools. In our case, we suspect that the presence of COAR schools could have generated a competitive environment among public schools, where schools with admitted students send a positive signaling of their quality to both parents and students, attracting regular or low-achieving students from other schools in the region who might be driving the mean of the treatment group downwards. The previous dynamic may be coupled with *student sorting* of high-achievers, as top ranking students prefer to leave treatment schools for control schools in order to increase their probability of being eligible for admission into a COAR in a given year.

Alternative mechanisms of transmission that may help to explain our negative results include less peer support for academic achievement. The possibility of entering a COAR could induce a more competitive environment in the classroom, making bright students wary of their grades, and thus reducing the help that they might have given to their peers otherwise. These interpretations, however, will need to be tested with more detailed information about enrollment and mobility across traditional public schools in 2016 and will therefore remain pending in the research agenda.

Finally, it needs to be noted that a low or even negative correlation between performance in the ECE and the final score of the COAR admission process might introduce attenuation bias (or even opposite expected results, as we find in this study). If students do not have enough incentives to perform well in the ECE, then there is no reason to expect positive nor negative spillover effects. This would be consistent with the possibility that some students only increase effort to achieve a high score in the COAR admission process.

References

Adcock, E. P., & Phillips, G. W. (2000). Accountability evaluation of magnet school programs: A value-added model approach. *Education Resources Information Center*.

Ballou, D. (2007). Magnet schools and peers: Effects on mathematics achievement. *Education Resources Information Center Working Paper*.

Ballou, D., Goldring, E., & Liu, K. (2006). Magnet schools and student achievement. *New York: National Center for the Study of Privatization in Education, Teachers College, Columbia University*, *18*, 2007.

Barron, M., Basurto, P., & Cuadra, G. (2018). On the peer effects of star students. *Applied Economics Letters*, *25*(12), 821–825.

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, *95*(2), 152–157.

Betts, J., Kitmitto, S., Levin, J., Bos, J., & Eaton, M. (2015). What happens when schools become magnet schools? a longitudinal study of diversity and achievement. *American Institutes for Research*.

Betts, J. R., Rice, L. A., Zau, A. C., Tang, Y. E., & Koedel, C. R. (2006). Does school choice work? effects on student integration and achievement.

Bifulco, R., Cobb, C. D., & Bell, C. (2009). Can interdistrict choice boost student achievement? the case of connecticut's interdistrict magnet school program. *Educational Evaluation and Policy Analysis*, *31*(4), 323–345.

Bifulco, R., & Ladd, H. F. (2006). The impacts of charter schools on student achievement: Evidence from north carolina. *Education Finance and Policy*, *1*(1), 50–90.

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D. (2008). The effect of charter schools on traditional public school students in texas: Are children who stay behind left behind? *Journal of Urban Economics*, *64*(1), 123–145.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2016). Regression discontinuity designs using covariates. *Unpublished, University of Michigan*.

Calonico, S., Cattaneo, M. D., & Rocio, T. (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295-2326. doi: 10.3982/ECTA11757

Cheng, M.-Y., Fan, J., & Marron, J. S. (1997). On automatic boundary corrections. *The Annals of Statistics*, 1691–1708.

Clark, D. (2009). The performance and competitive effects of school autonomy. *Journal of Political Economy*, *117*(4), 745–783.

Cobb, C. D., Bifulco, R., & Bell, C. (2009). Evaluation of connecticut's interdistrict magnet schools. *Storrs: The Center for Education Policy Analysis, University of Connecticut*.

Crain, R. L., et al. (1992). The effectiveness of new york city's career magnet schools: An evaluation of ninth grade performance using an experimental design.

Cullen, J. B., Jacob, B. A., & Levitt, S. D. (2005). The impact of school choice on student outcomes: an analysis of the chicago public schools. *Journal of Public Economics*, *89*(5-6), 729–760.

Dohrmann, K. R., Nishida, T. K., Gartner, A., Lipsky, D. K., & Grimm, K. J. (2007). High school outcomes for students in a public montessori program. *Journal of Research in Childhood Education*, *22*(2), 205–217.

Esposito, C. L. (2010). *School type and mathematics achievement: A comparison of magnet and public secondary schools using the educational longitudinal study of 2002* *data set*. University of Connecticut.

Goldring, E. B., & Phillips, K. J. (2008). Parent preferences and parent choices: The public–private decision about school choice. *Journal of Education Policy*, *23*(3), 209–230.

Goldschmidt, P., & Fernandez, J.-F. (2004). The relationship between school quality and the probability of passing standards-based high-stakes performance assessments. cse technical report 644. *Center for Research on Evaluation Standards and Student Testing CRESST*.

Jinnai, Y. (2011). Who benefits from school choice? school competition, student sorting, and spillover effects. *Department of Economics, University of Rochester - Manuscript*.

Judson, E. (2014). Effects of transferring to stem-focused charter and magnet schools on student achievement. *The Journal of Educational Research*, *107*(4), 255–266.

Kemple, J. J., & Snipes, J. C. (2000). Career academies: Impacts on students' engagement and performance in high school.

Lillard, A., & Else-Quest, N. (2006). The early years: evaluating montessori. *Science*, *313*(5795), 1893–1894.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698 - 714. (The regression discontinuity design: Theory and applications) doi: https://doi.org/10.1016/j.jeconom.2007.05.005

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from madagascar. *Unpublished manuscript*, *6*.

Nicoletti, C., & Rabe, B. (2014). Sibling spillover effects in school achievement.

Nielsen, H., Joensen, J., et al. (2015). Peer effects in math and science. In *2015 Meeting Papers*.

Nixon, L. A., & Robinson, M. D. (1999). The educational attainment of young women: Role model effects of female high school faculty. *Demography*, *36*(2), 185–194.

Qureshi, J. A. (2011). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal*.

Ray, D. (2006). Aspirations, poverty, and economic change. *Understanding poverty*, *409421*.

Rice, J., Lash, A., Huang, M., Tran, L., & Peterson, M. (2015). Advanced course completion in magnet and comprehensive high schools: A study in nevada's clark county school district. *Regional Educational Laboratory West*.

Sass, T. R. (2006). Charter schools and student achievement in florida. *Education Finance and Policy*, *1*(1), 91–122.

Siegel-Hawley, G., & Frankenberg, E. (2011). Magnet school student outcomes: What the research says. research brief no. 6. *National Coalition on School Diversity*.

Silver, D., Saunders, M., & Zarate, E. (2008). What factors predict high school graduation in the los angeles unified school district. *Policy Brief*, *14*.

Yang, Y. N., Li, Y. H., Tompkins, L. J., & Modarresi, S. (2005). Using the multiple-matched-sample and statistical controls to examine the effects of magnet school programs. *Education Resources Information Center*.

Appendix

TABLE 3: First and second stage estimates for the subsample for different bandwidths

	Coefficient (standard errors)					
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Bandwidth = 0.7						
First stage						
Z (score ≥ 0)	**0.494*** **	0.276	**0.411*** **	**0.490*** **	0.273	**0.407*** **
	(0.139)	(0.221)	(0.152)	(0.132)	(0.209)	(0.142)
Second stage: Mathematics						
Reduced Form	-0.228	**-0.378***	**-0.294***	-0.195	-0.259	**-0.228***
	(0.149)	(0.223)	(0.164)	(0.125)	(0.181)	(0.131)
Fuzzy RD	-0.461	-1.372	-0.717	-0.398	-0.947	-0.561
	(0.344)	(1.472)	(0.533)	(0.284)	(1.041)	(0.408)
Observations	17,572	17,572	17,572	17,572	17,572	17,572
No. of schools	317	317	317	317	317	317
Panel B: Bandwidth = 0.9						
First stage						
Z (score ≥ 0)	**0.523*** **	**0.364** **	**0.468*** **	**0.520*** **	**0.363** **	**0.463*** **
	(0.103)	(0.181)	(0.125)	(0.101)	(0.174)	(0.119)
Second stage: Mathematics						
Reduced Form	**-0.249** **	-0.314	**-0.271***	**-0.208** **	-0.271	**-0.227** **
	(0.121)	(0.197)	(0.140)	(0.105)	(0.166)	(0.114)
Fuzzy RD	**-0.476***	-0.863	-0.578	**-0.401***	-0.747	**-0.490***
	(0.259)	(0.766)	(0.365)	(0.220)	(0.629)	(0.293)
Observations	23,902	23,902	23,902	23,902	23,902	23,902
No. of schools	433	433	433	433	433	433
Specification	Linear	Quadratic	Kernel	Linear	Quadratic	Kernel
Other controls	No	No	No	Yes	Yes	Yes

Std. errors clustered at the school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 4: First and second stage estimates for the subsample for different bandwidths (p-value of McCrary test ≥0.33)

	Coefficient (standard errors)					
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Bandwidth = 0.7						
First stage						
Z (score ≥ 0)	**0.471*** **	0.260	**0.392** **	**0.474*** **	0.240	**0.379*** **
	(0.141)	(0.226)	(0.154)	(0.132)	(0.214)	(0.141)
Second stage: Mathematics						
Reduced Form	-0.223	-0.374	**-0.288***	-0.171	-0.217	-0.194
	(0.156)	(0.231)	(0.169)	(0.133)	(0.197)	(0.140)
Fuzzy RD	-0.473	-1.440	-0.734	-0.361	-0.902	-0.512
	(0.376)	(1.644)	(0.574)	(0.308)	(1.250)	(0.456)
Observations	15,347	15,347	15,347	15,347	15,347	15,347
No. of schools	292	292	292	292	292	292
Panel B: Bandwidth = 0.9						
First stage						
Z (score ≥ 0)	**0.507*** **	**0.345***	**0.447*** **	**0.511*** **	**0.342***	**0.442*** **
	(0.107)	(0.183)	(0.128)	(0.103)	(0.177)	(0.120)
Second stage: Mathematics						
Reduced Form	**-0.258** **	-0.300	**-0.269***	**-0.210***	-0.236	-0.200
	(0.129)	(0.203)	(0.146)	(0.114)	(0.180)	(0.123)
Fuzzy RD	**-0.510***	-0.870	-0.603	**-0.412***	-0.692	-0.452
	(0.285)	(0.825)	(0.400)	(0.244)	(0.703)	(0.327)
Observations	21,220	21,220	21,220	21,220	21,220	21,220
No. of schools	402	402	402	402	402	402
Specification	Linear	Quadratic	Kernel	Linear	Quadratic	Kernel
Other controls	No	No	No	Yes	Yes	Yes

Std. errors clustered at the school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5: Quantile regression for the sub-sample

	Coefficient (standard errors)			
	[1]		[2]	
	Q1	Q5	Q1	Q5
Panel A: Language				
Reduced Form	-0.181	-0.226	-0.108	-0.167
	(0.227)	(0.200)	(0.199)	(0.163)
Fuzzy RD	-0.371	-0.472	-0.221	-0.352
	(0.487)	(0.453)	(0.418)	(0.358)
Observations	5141	5349	5141	5349
No. of schools	479	482	479	482
Panel B: Mathematics				
Reduced Form	-0.288	-0.456**	-0.218	-0.405**
	(0.213)	(0.192)	(0.185)	(0.160)
Fuzzy RD	-0.589	-0.936**	-0.448	-0.840**
	(0.493)	(0.469)	(0.419)	(0.378)
Observations	5149	5353	5149	5353
No. of schools	480	482	480	482
Panel C: History				
Reduced Form	-0.142	-0.171	-0.067	-0.115
	(0.170)	(0.177)	(0.152)	(0.140)
Fuzzy RD	-0.294	-0.350	-0.139	-0.236
	(0.370)	(0.390)	(0.322)	(0.301)
Observations	5176	5328	5176	5328
No. of schools	479	482	479	482
Specification	Kernel	Kernel	Kernel	Kernel
Other controls	No	No	Yes	Yes

Note: Standard errors clustered at the school level in parentheses.
*, **, and *** denote significance at 10%, 5%, and 1% levels.

TABLE 6: Second stage estimates for the subsample: different bandwidths

	Coefficient (standard errors)			
	[1]		[2]	
	Q1	Q5	Q1	Q5
Panel A: bandwidth = 0.7				
Second stage: Mathematics				
Reduced Form	-0.314	**-0.530****	-0.218	**-0.444****
	(0.271)	(0.230)	(0.235)	(0.180)
Fuzzy RD	-0.753	**-1.284***	-0.528	**-1.087***
	(0.782)	(0.770)	(0.647)	(0.576)
Observations	3,344	3,481	3,344	3,481
No. of schools	315	317	315	317
Panel B: bandwidth = 0.9				
Second stage: Mathematics				
Reduced Form	-0.296	**-0.484****	-0.226	**-0.433*****
	(0.227)	(0.202)	(0.197)	(0.164)
Fuzzy RD	-0.624	**-1.028***	-0.481	**-0.930****
	(0.549)	(0.528)	(0.466)	(0.419)
Observations	4,552	4,734	4,552	4,734
No. of schools	431	433	431	433
Polynomial	Kernel	Kernel	Kernel	Kernel
Other controls	No	No	Yes	Yes

Standard errors clustered at the school level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 7: Second stage estimates for the subsample: different bandwidths (p-value of McCrary Test $\geq$ 0.33)

	Coefficient (standard errors)			
	[1]		[2]	
	Q1	Q5	Q1	Q5
Panel A: bandwidth = 0.7				
Second stage: Mathematics				
Reduced Form	-0.310	**-0.509****	-0.179	**-0.385****
	(0.274)	(0.235)	(0.249)	(0.181)
Fuzzy RD	-0.780	-1.290	-0.468	-1.009
	(0.831)	(0.823)	(0.721)	(0.622)
Observations	2,918	3,030	2,918	3,030
No. of schools	290	292	290	292
Panel B: bandwidth = 0.9				
Second stage: Mathematics				
Reduced Form	-0.297	**-0.472****	-0.198	**-0.378****
	(0.234)	(0.210)	(0.211)	(0.168)
Fuzzy RD	-0.657	**-1.048***	-0.445	**-0.849***
	(0.596)	(0.573)	(0.522)	(0.449)
Observations	4,036	4,192	4,036	4,192
No. of schools	400	402	400	402
Specification	Kernel	Kernel	Kernel	Kernel
Other controls	No	No	Yes	Yes

Standard errors clustered at the school level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Second stage estimates for the total sample (Bandwith = 1)

	Coefficient (standard errors)					
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: First stage						
Z (score $\geq$ 0)	0.539***	0.357***	0.474***	0.543***	0.355***	0.474***
	(0.073)	(0.120)	(0.088)	(0.071)	(0.118)	(0.085)
Panel B: Language						
Reduced Form	-0.108	-0.045	-0.085	-0.040	-0.023	-0.033
	(0.086)	(0.149)	(0.102)	(0.077)	(0.129)	(0.088)
Fuzzy RD	-0.200	-0.127	-0.179	-0.074	-0.064	-0.070
	(0.163)	(0.424)	(0.220)	(0.142)	(0.364)	(0.187)
Panel C: Mathematics						
Reduced Form	-0.135	-0.174	-0.151	-0.070	-0.141	-0.098
	(0.083)	(0.135)	(0.095)	(0.078)	(0.121)	(0.086)
Fuzzy RD	-0.251	-0.486	-0.318	-0.129	-0.397	-0.206
	(0.158)	(0.431)	(0.213)	(0.143)	(0.373)	(0.186)
Panel D: History						
Reduced Form	-0.116	-0.064	-0.095	-0.055	-0.032	-0.041
	(0.071)	(0.121)	(0.083)	(0.064)	(0.106)	(0.073)
Fuzzy RD	-0.215	-0.178	-0.201	-0.102	-0.092	-0.087
	(0.136)	(0.351)	(0.182)	(0.118)	(0.303)	(0.156)
Panel E: Expectations						
Reduced Form	0.003	0.026	0.012	0.016	0.024	0.018
	(0.021)	(0.033)	(0.024)	(0.018)	(0.030)	(0.021)
Fuzzy RD	0.006	0.074	0.025	0.029	0.068	0.038
	(0.038)	(0.094)	(0.050)	(0.033)	(0.085)	(0.044)
Observations	50,697	50,697	50,697	50,697	50,697	50,697
No. of schools	871	871	871	871	871	871
Specification	Linear	Quadratic	Kernel	Linear	Quadratic	Kernel
Other controls	No	No	No	Yes	Yes	Yes

Note: The table reports estimates of the spillover effects of admission into a COAR (Reduced Form) and enrollment in COAR (Fuzzy RD) on three educational outcomes: Language Test Scores (Panel B), Mathematics Test Scores (Panel C), History Test Scores (Panel D) and self-reported educational expectations (Panel E). Eligibility for admission is an indicator equal to one if there was at least one student admitted into a COAR in the previous cohort such school; the running variable is defined as the maximum score observed among applicants of the previous cohort students. Columns 1–3 include, in order, a polynomial of degree 1, 2 and Kernel on the running variable (these polynomials are interacted with the eligibility indicator) following Calonico, et al. (2014a) methodology. Columns 4–6 additionally control for pretreatment characteristics of schools for the year 2014. Standard errors clustered at the school level in parenthesis. *, **, and *** denote significance at 10%, 5%, and 1% levels.