Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya*

Adrienne M. Lucas and Isaac M. Mbiti

Draft: October 2012

DO NOT CITE WITHOUT PERMISSION

Abstract

The most desirable Kenyan secondary schools are elite government schools that admit the best students from across the country. We exploit the random variation generated by the centralized school admissions process in a regression discontinuity design to obtain causal estimates of the effects of attending one of these elite public schools on student progression and test scores in secondary school. Despite their reputations, we find little evidence of positive impacts on learning outcomes for students who attended these schools, suggesting that their sterling reputations reflect the selection of students rather than their ability to generate value-added test-score gains.

Keywords: Education, Kenya, returns to secondary school

JEL Codes: H52, I21, O12, O15

*Lucas: Department of Economics, University of Delaware, alucas@udel.edu. Mbiti: Department of Economics, Southern Methodist University, lmbiti@smu.edu. We are especially grateful to the Kenya National Examination Council, the Kenyan Ministry of Education, and the Teachers Service Commission for providing us with data. The contents of this paper do not reflect the opinions or policies of any of these agencies. The authors bear sole responsibility for the content of this paper. For useful comments and suggestions we thank Kehinde Ajayi, Josh Angrist, David Autor, Abhijit Banerjee, Kristin Butcher, Elizabeth Cascio, Damon Clark, Esther Duflo, Eric Edmonds, Scott Imberman, Michael Kremer, Kevin Lang, Daniel Millimet, Patrick McEwan, Muthoni Ngatia, Moses Ngware, Owen Ozier, Christian Pop-Eleches, Tavneet Suri, Shobhana Sosale, Miguel Urquiola, and numerous seminar and conference participants. Lucas acknowledges support from a Wellesley College faculty research grant. Mbiti undertook some of this research while a MLK Jr. Visiting Assistant Professor at MIT. Mbiti acknowledges support from a National Academy of Education/Spencer foundation post-doctoral fellowship, a SMU Research Grant, and the Brown University Population Studies and Training Center. We dedicate this paper to the memory of David Mbiti who was instrumental in our data collection and provided crucial support and insights.

1 Introduction

Elite, prestigious government secondary schools are common in education systems throughout the world. These schools are very selective, admit a relatively small number of high achieving students each year, cost significantly more than other public school options, and are highly coveted by both students and their parents in part because their graduates achieve the highest test scores on nation-wide assessments. Their alumni are luminaries in business, politics, and civil service and have a disproportionate influence on the economic progress of their countries as a result of their careers in the upper echelons of government and business.[1] For example in Kenya, the setting for this study, fewer than two percent of secondary school students attend these elite schools, and they charge over US$1000 per year in comprehensive fees, well above GDP per capita in Kenya. Competition for the limited places in these schools is intense, with parents expending considerable resources on private primary school fees and out of school tutoring in an effort to increase their child's primary school exit exam score and therefore the likelihood of admission into an elite school.[2] While elite secondary schools are perceived to be academically superior, whether their reputations reflect their selective admission of the best students or their production of better value-added achievement outcomes is unclear.

In this paper, we employ a regression discontinuity design to estimate the impact of attending one of the elite secondary schools (known as national schools) in Kenya on student graduation and achievement at the end of secondary school. The key conceptual difficulty in assessing the impact of attending an elite school on student outcomes is the endogenous selection of students into schools. To address this difficulty, we rely on Kenya's secondary school admissions rule that creates a discontinuity in national school matriculation. Specifically, students are admitted to a single government secondary school based on their scores on the nationally stan-

[1] Four Ghanaian presidents and two other African Presidents, including Robert Mugabe, attended elite public secondary schools in Ghana. Alumni of similar schools in Kenya include President Mwai Kibaki, two former vice-presidents, and numerous senior political leaders and CEOs.

[2] For example, approximately 41 percent of primary school aged children in Kenya attended extra tutoring sessions in 2010 (Uwezo 2010).

dardized primary school exit exam, district-specific quotas, and school preferences that students express prior to taking the exit exam. National schools admit the top achieving students from each district in the country, while lower ranked students are admitted to less renowned government schools. Students with the lowest scores are denied admission to secondary school (Ozier 2011). The precise admissions algorithm results in students from the same district with very similar scores potentially being admitted to schools of different levels of prestige.

Given the admissions rule, our regression discontinuity design compares students who were barely admitted to a national school to those who barely missed admission. At the conclusion of secondary school, these students' achievement scores on the national secondary school exit exam and probability of graduation should differ only by the effect of the quality of their secondary school.[3] An unusual feature of our design is that within each national school the admissions rule generates a separate discontinuity for each district. Our empirical approach combines these multiple discontinuities across districts and schools. Therefore our estimates do not rely on the lowest scoring students within a school for identification since the marginal admitted student from each district is not necessarily the lowest scoring student in the school.

Prior studies have found surprisingly mixed evidence on the impacts of attending selective schools on students' educational outcomes in both developing and developed country settings. Much of this recent literature relies on evidence from regression discontinuities, where students who are just above the threshold for attending a high quality school are compared to their counterparts who are just below the admissions threshold. Using this estimation strategy studies such as Jackson (2010), Park, Shi, Hsieh, and An (2010), and Pop-Eleches and Urquiola (2011) found that students who attended higher quality schools had improved test scores. In contrast, a number of other studies such as Abdulkadiroglu, Angrist, and Pathak (2011), Bui, Craig, and Imberman (2011), Clark (2010), de Hoop (2010), Fryer and Dobbie (2011), and Sekhri and Rubinstein (2010) did not find any evidence of such schooling improving learning

[3] Quality includes all attributes that vary between national and lower tiered schools (e.g. infrastructure, peers, teachers, other resources, etc). We discuss these attributes further in Section 2.

outcomes.

Due to the specific admissions algorithm and the context in Kenya, our study builds on the prior literature in a number of important ways. First, most of the prior studies exploited a single qualification threshold for each school, effectively comparing the lowest scoring students in elite schools to the highest scoring students in less desirable schools. If teachers target the level of instruction to the median student, then these low scoring students who just qualify for admittance may not fully benefit from the more desirable school (Duflo, Dupas, and Kremer 2011). In the current study, as a result of multiple discontinuities within each national school, the marginal students in our regression discontinuity estimates are located throughout the within-school baseline test score distribution. Therefore, we can conduct meaningful inference about the effect of attending an elite school on students who are not the lowest ability students within a national school. Second, the quality of the alternative school available to those who are not admitted to a national school is province specific, enabling us to explore the heterogeneous impacts of national school attendance by the quality of the alternative schooling option. If students who are not admitted to the most elite schools have good quality alternatives, then this could reduce the potential impact of attending the elite schools (see Deming et al. 2011 for an example in Charlotte-Mecklenberg). Third, our data contain the universe of students who graduated from primary school in Kenya in 2004. All students who graduate from primary school automatically apply to secondary school. Therefore, our sample is not selected based on who chose to apply to the elite schools. Finally, by focusing on the best students in Kenya, our analysis provides some insights into the production of future leaders who are likely to shape the economic development of the country.

Our results show that being admitted to a national school was not associated with an improvement in the probability of timely progression through secondary school. We also find little evidence of positive impacts of national school graduation on test scores. We do not find a statistically significant effect on the composite test score despite the superior peer quality and resources found in national schools. Our estimates are sufficiently precise that we can rule out moderately sized effects of approximately one

tenth of a standard deviation or larger in the composite test score, even though peers at national schools scored about one half of a standard deviation higher on the baseline test relative to peers at alternative schools. For students from the North Eastern province, the poorest province, the change in peer quality for those who attended national schools was an even more dramatic one and a half standard deviations. Yet even with this substantial jump in peer quality, we do not find a statistically significant effect of national school graduation on achievement and can rule out effect sizes larger than 0.27 standard deviations in the composite test score. We also do not find heterogeneous effects based on the within-school baseline test score distribution, gender, or socioeconomic status. Additionally, we find that students with many high quality alternative schooling options might have been better served by attending non-national schools. Finally, we find that national school students were exposed to a more diverse peer group, which is not surprising given the national school admission process.

We do, however, find a robust causal association between national school graduation and a higher Swahili subject test score. We posit two possible explanations. First, as national schools are the most ethnically diverse schools in the country, students' use of Swahili as a lingua franca in these schools could have increased their proficiency. Students in non-national schools with more localized catchment areas could more easily converse in local, ethnic languages. Second, in accordance with their original mandates to increase national unity, some national schools spend more time on Swahili, thereby emphasizing the importance of Swahili as a national language.

Overall, we argue that the estimated lack of effect is meaningful due to our ability to rule out moderately sized effects, the lack of significant results even for students with the lowest quality alternative options, and the presence of multiple discontinuities within each school such that we are not relying on the student with the lowest test score in each national school for identification. Our findings suggest that the recent $30 million government expansion of the national school system will do little to boost learning levels, yet a common language and exposure to diverse peers could

be important justifications for a national school system in a country like Kenya with a high degree of ethnic polarization.

2 Secondary Education in Kenya

2.1 Background

Kenya follows an 8-4-4 system of education, where primary school consists of eight years and secondary school and university are each four years. Both primary and secondary school conclude with nationwide standardized exams that are centrally graded. The exam scores act as proof of completion for those entering the labor force and are used to qualify for additional education. Upon completion of primary school pupils take the Kenya Certificate of Primary Education (KCPE) exam. The KCPE comprises 5 compulsory subjects and is graded from 0 to 500 points. At the conclusion of secondary school students take the Kenya Certificate of Secondary Education (KCSE) exam. For the KCSE students take seven to nine subject exams out of the 30 possible examination subjects. English, Swahili and mathematics are compulsory subjects as are at least two sciences, one humanity, and one practical subject.[4] The maximum score on the KCSE is twelve points. Student may take up to nine subject exams with the KCSE score computed as the average of seven scores: the compulsory subjects, the top two science scores, the top humanities score, and the top practical score.

In 2004, almost 655 thousand students graduated from the approximately 21 thousand government and private primary schools that administered the KCPE. Four years later in 2008, 35 percent of this cohort graduated from one of 5158 secondary schools and took the KCSE. Across all secondary grades the 2004 gross enrollment rate was 48 percent with a 40 percent net enrollment rate (World Bank 2004).

[4]Sciences are biological science, biology, chemistry, physics, and physical science. Humanities are Christian religious education, geography, Hindu religious education, history, and Islamic religious education. Practical subjects are accounting, Arabic, art and design, aviation, construction, computer studies, commerce, drawing design, economics, electronics, French, German, home science, metalwork, music, and woodwork.

Each Kenyan government secondary school belongs to one of three tiers: national, provincial, or district. The national schools are the most elite government schools and the most prestigious secondary schools in the country. They are also among the oldest schools, often modeled on British public schools but with government funding.[5] In 2008, these eighteen single sex boarding schools admitted approximately 3000 of the top primary school candidates from across the nation with places reserved for students from each district.[6] The almost 1000 provincial schools, the second tier, admitted the top remaining students from within a province. The approximately 3000 district schools, the bottom tier, drew students from the district who could not gain admission into national or provincial schools. Almost 80,000 students graduated from provincial schools and over 115,000 students graduated from district schools in 2008. For students who took the KCPE and wanted to continue their education but opt out of government schools, approximately 850 private schools admitted students from KCPE primary schools and followed the same curriculum and utilized many of the same teaching materials as government secondary schools. The quality of private schools varied. Although some private schools were selective, on average they were most similar to district schools based on incoming student KCPE scores and KCSE scores at graduation. In 2008 only 12 percent of secondary school graduates graduated from private schools.[7]

National schools have better physical and human capital resources on a number of margins. Relative to other schools, they have better facilities, offer a larger variety of courses, are staffed by teachers with more education and experience, and provide a higher quality peer group. They have on average 1.5 times the landholdings per

[5] For example Lenana School and Nairobi School were founded by the colonial administration as Duke of York School and Prince of Wales School. They originally admitted only white students, but were integrated following independence.

[6] After gaining independence from the United Kingdom in 1963, the Kenya Commission on Education promoted the use of secondary schools as vehicles to promote national unity, resulting in the three tiered system with admissions based on both merit and region (Gould, 1973).

[7] This does not include students that graduated from international private schools. Students in the international private school system do not take the KCPE, are not eligible for admissions to government secondary schools, do not take the KCSE, and are not in our data. Most of these schools follow the Internationl Baccalaurate (IB) or International General Certificate of Secondary Education (IGCSE) curriculum.

student relative to other government schools, allowing for additional recreational and classroom space. They generally have well equipped computer labs, electricity, and modern buildings and toilets, while provincial schools and district schools are far less well equipped often lacking electricity and indoor plumbing (NCKEF 2004). For example the Nairobi School, a national boys school in Nairobi, has a campus of over 200 acres that includes a swimming pool, tennis and basketball courts, and woodwork and metalwork facilities (Nairobi School 2012). Another national school, Mangu High, owns a small aircraft for use by its aviation students. Additionally, national schools offer a wider variety of subjects. While the full national curriculum contains 30 subjects, most schools offer fewer than twelve subjects due to the high cost of providing appropriate facilities and instructors. In 2008, the national schools offered an average of sixteen subjects, the average provincial school offered about twelve subjects, and district and private schools each offered on average about eleven subjects. Almost all national schools offered computer studies, French, and German while few provincial or district schools offered these classes. Two national schools were the only government schools that offered the aviation KCSE course. National schools have similar pupil-teacher ratios compared to other government secondary schools, but their teachers have more training and experience (Ministry of Education 2012). In national schools, 80 percent of teachers had degrees beyond secondary school, compared with 68 percent in other government secondary schools. National school teachers were almost twice as likely to hold advanced degrees and had on average one additional year of teaching experience relative to their provincial school counterparts. Finally, the KCPE score of incoming national school students in 2004 was two standard deviations higher than the average student matriculating to a non-national school.

To provide these resources, national schools charge higher comprehensive fees. In 2012, national schools charged 90,000 Ksh to 130,000 Ksh (US$1071 to US$1547), while provincial schools charged 20,000 Ksh to 50,000 Ksh (US$238 to US$595).[8] Na-

[8]Personal correspondence from school administrators. Conversion to US$ based on the January 2012 to August 2012 average exchange rate of 84.05 Ksh to US$1.

tional schools receive approximately the same amount per pupil from the government as the other types of schools (20,000 Ksh in 2006), but their total spending per pupil is greater because of the higher student fees (Onsomu et al. 2006).

2.2 Selection into Secondary Schools

Admissions to secondary schools are conducted centrally by the Ministry of Education using a computerized system. Each national school has a set of district quotas, the number of students to be admitted from each district, determined by the Ministry of Education. When students register for the KCPE they list two preferred national schools, two preferred provincial schools, and two preferred district schools. These preferences, KCPE score, and district quotas determine to which secondary school a student is admitted. A student is admitted to at most one secondary school.

The secondary school admissions algorithm ranks students within each district and gender by descending KCPE score. The admissions process occurs separately for boys and girls as national schools are sex-specific. The highest ranked student in a district (by gender) is placed in his or her first choice national school. Each subsequent student in the district is placed in his first choice national school, if the school still has an opening. If a student's first choice school's district-specific quota is already full, the placement algorithm considers the second choice school. The student will be placed in the second choice national school if the district-specific quota has not been filled. If both preferred national schools are already full, then the student will be admitted to a preferred provincial school following the same algorithm, even if other, non-preferred, national schools have openings for students from the same district. Therefore under this admissions mechanism two students with the same stated preferences and KCPE scores only separated by one point (out of 500) could be admitted to schools of different tiers. Students are notified of their school placement prior to the start of the school year in January. An unofficial second round occurs after the initial formal placements. Some students who are unhappy with their placement directly apply to an alternative school and are admitted at the discretion of the school principal, provided that there are available places due to an admitted student not matriculating

at the start of the school year. Therefore, while adherence to the admissions rule is quite high, it is not strictly binding.[9]

3 Empirical Strategy

If students were placed randomly into schools, then we could estimate the treatment effect of attending a national school as follows:

$$Y_i = \alpha + \beta T_i + \varepsilon_i \tag{1}$$

Where Y is the outcome variable (i.e. secondary school test score or the probability of timely graduation) for student i, T is a binary variable that indicates if the student was subject to the treatment, attending a national school for 4 years, and ε is the idiosyncratic error. In Kenya, students are not randomly allocated to schools. Students with higher ability and unobserved characteristics, likely correlated with secondary school outcomes, that cause a principal to admit them in the second round attend national schools. Therefore it is likely that T and ε would be correlated, and thus the OLS estimates of Equation (1) would produce biased estimates of $\widehat{\beta}$, the treatment effect.

In Kenya, the Ministry of Education admissions algorithm provides exogenous variation in which students are admitted to national schools. We exploit this variation for identification through a regression discontinuity (RD) framework. Intuitively, we compare the students barely admitted to a national school to those who barely missed admission. First, we implement the Ministry of Education admissions algorithm with the actual district quotas and student KCPE scores and stated preferences. This generates the list of students admitted (and not admitted) to each national school. These lists create a number of discontinuities where students whose scores could be as close as one point different were admitted to schools of different tiers. Second, we solve for the admissions threshold for each district-school pair. The KCPE score of

[9] Admission to provincial and district schools occurs in a similar manner. Lack of documentation and poor adherence to the rule based assignment prevent us from implementing our empirical strategy for these lower tiers.

the last student admitted from a district is the effective score cutoff, c_{sj}, for school s for district j, i.e. the minimum score that a student from district j needed to exceed in order to be admitted to school s. For students admitted to a national school, c_{sj} is the district-specific cutoff of the school to which they were admitted. For students who did not gain admission to any national school, c_{sj} is the district-specific cutoff of the school in their preference set to which their KCPE score was closest.[10] Since cutoffs vary by school and the applicant's district we follow Pop-Eleches and Urquiola (2011) and define the running variable, r_i, as the distance (in points) between student i's KCPE score, $KCPE_i$, and the binding cutoff, c_{sj}:

$$r_i = KCPE_i - c_{sj} \tag{2}$$

A student was admitted to a national school if $KCPE_i \geq c_{sj}$, while those students whose scores were below the cutoff were not admitted to national schools. While binding for most students, adherence to the rule-based admission was imperfect for two reasons. First, a student admitted to a national school might not have matriculated or could dropped out prior to completion. Second, places in national schools that are not claimed by those initially admitted were allocated at the discretion of the principal to direct applicants during the unofficial second round. These second round students had test scores below the admissions threshold. Due to this imperfect compliance around the cutoff, we employ a "fuzzy" regression discontinuity design that we estimate with a two stage least squares model (Angrist and Lavy 1999; Hahn, Todd, and Van der Klaauw 2001; Lee and Lemieux 2010). We define treatment as graduating from a national school, and use national school (rule based initial) admission assignment as an instrument for graduating from a national school as follows[11]:

[10] An alternative strategy is to treat the unit of observation as a student-choice instead of a student, allowing two observations per student. The empirical results using this alternative methodolgy are substantively similar.

[11] An alternative specification of treatment could be the initial matriculation to a national school, but data limitations preclude this option. We prefer our specification for a number of reasons. First, a student who matriculated but droped out is not fully treated by a national school. Second, almost all students who graduated from a national school attended that school for all four years. Third, transferring to a national school (after the second round of admissions) is extrememly rare and never occurs after the start of the third year. Nevertheless, a few students who we observe graduating from a provincial school might have initially matriculated to a national school, and are therefore

$$T_i = \delta 1\{r_i \geq 0\} + f(r_i) + \mathbf{X}_i'\mathbf{\Gamma} + \nu_i \tag{3}$$

where $1\{r_i \geq 0\}$ is an indicator function that takes a value of 1 for students admitted to national schools, i.e. $KCPE_i \geq c_{sj}$, $f(r_i)$ is a smooth function of the running variable allowed to vary on either side of the discontinuity, $\mathbf{X}$ is a vector of control variables that includes dummy variables for sex, public primary school, district, school choices, and school choice by district interactions, while v_i is an idiosyncratic error term assumed to be independent across districts but allowed to be correlated between students within the same district. We include school choice by district interactions to control for the "contest" in which a student participates, since students only competed for national school places against students from their own district with the same preferred national school. In our baseline specification $f(r_i)$ is a piecewise linear function that we allow to vary discontinuously at $r_i = 0$, the effective cutoff. In robustness checks we include a third degree polynomial that varies at $r_i = 0$ as an alternative $f(r_i)$. A student's KCPE score is absorbed in our specifications as it is a linear combination of the score distance to the cutoff, the district fixed effects, and contest controls.

The second stage is then

$$Y_i = \alpha + \beta T_i + f(r_i) + \mathbf{X}_i'\boldsymbol{\theta} + \varepsilon_i \tag{4}$$

where we instrument for the treatment, T_i, with Equation 3. Other notation is as in previous equations. As with all regression discontinuities, the identification assumption is that as the discontinuity threshold is approached from above or below the individuals are essentially identical prior to treatment. Thus we would expect that in the absence of differential treatment, these students would have similar outcomes at the conclusion of secondary school.

We use this strategy to estimate student specific outcomes of timely completion and test scores as well as test for discontinuities in school characteristics, replacing Y_i with the appropriate outcome. For the tests of discontinuities in school characteristics

partially treated, although this is also rare.

we use the number of subjects available, peer KCPE scores, the coefficient of variation of peer KCPE scores, and two measures of school diversity as dependent variables in different specifications.

A number of features of our application of the RD design are particularly noteworthy. First, all students who took the KCPE effectively applied to all types of secondary schools, including national schools, in contrast to other school systems that require an additional application for selective schools. Second, we are not relying exclusively on the lowest scoring students in a national school for identification. A common critique of other papers in education that use a similar regression discontinuity design is that the teachers in an elite school are probably not teaching to the student who was barely admitted to the school, but instead to a student whose score was closer to the median. In most settings, students who are closer to the median of the school are far away from the admissions threshold. Because of the multiple district-specific effective thresholds within each school, the marginal student admitted from one district could have had a score that was substantially higher than the marginal student from another district. For example, for Alliance Boys High School, the oldest boys' national school in Kenya, the effective KCPE cutoff score was 459 (out of 500) for students from Mbeere district in the Eastern province and 346 for students from Ijara district in the North Eastern province. Therefore, the last boy admitted from Mbeere had a score well above the minimum score in the school and even above the school median of 443 and is in our RD sample. The range of district-specific cutoff scores within a national school averaged 142 points. Multiple marginal students at different places within a single school's score distribution allow us to test for differential effects by relative percentile. Third, the quality of the non-national school option for each student varies by home province. Girls in Nairobi province, for example, had higher quality provincial school options than girls from the North Eastern province. The first feature enhances our study's external validity, while the other two provide avenues to test for heterogeneous effects and strengthen our findings.

4 Data

In this study, we use administrative test scores and district quotas for students who graduated from primary school in 2004. Our test score data are the KCPE administrative records from the Kenya National Examination Council (KNEC). The KCPE data contain secondary school preferences, KCPE scores, gender, district, and primary school type for the universe of students who took the KCPE in 2004. We match the 2004 KCPE data to the administrative examination records of all students who took the KCSE in 2008.[12] The KCSE records contain each student's composite and subject KCSE scores and school in which the student was enrolled at the time of the exam. Table 1 contains select summary statistics by school type. Not surprisingly, the average KCPE and KCSE scores of national school graduates were the highest.

We combine the KCPE score and student preferences data with the Ministry of Education's district quotas for each national school. From these data, we generate the rule-based secondary school admission for each student of the 2004 KCPE cohort. The effective cutoff for each district-school pair is the lowest KCPE score of an admitted student from a particular district. District-specific effective cutoffs for national schools ranged from 234 to 467 points with a mean of 419.

For most of our analysis we restrict our sample to those students whose KCPE scores are within one half of a standard deviation, 34 points, of the binding national school cutoff, $|r_i| \leq 34$. Figure 1 illustrates the KCPE and KCSE scores for all students and those within the RD sample of plus and minus 34 points of the threshold. The KCPE exam is scored on an integer scale from 0 to 500 (Frames A and B). The KCSE is scored on an integer scale from 1 to 12 (Frames C and D). As expected, students who are within the RD sample had higher scores on both exams when comparing Frame A to B and Frame C to D. One potential concern is that test scores reflect incomplete information due too many scores at the maximum or minimum. No students in our RD sample earned a 0 or a perfect 500 on the KCPE exam (Frame B). Five percent of our sample earned a perfect 12 on the KCSE and 0.02 percent earned

[12] Of the students who took the KCSE in 2008 we are able to match 97% to their KCPE scores.

the minimum score of 1 (Frame D). Given the few students with perfect scores and that they are evenly split between national and provincial schools, we do not expect top coding in test scores to significantly affect our findings.

5 Results

In Section 5.1 we show the strong first stage relationship between being admitted to a national school and graduating from a national school. We then use the regression discontinuity design outlined in Section 3 to document the differences in school characteristics across the admissions threshold. Next, we estimate the effect of being admitted to a national school on timely progression and the (instrumented) effect of graduating from a national school on student achievement. Finally, we test for heterogenous effects by student ability, socioeconomic status, and gender. In all of the initial specifications we limit our sample to those within a 34 point window (one half of a standard deviation) on either side of the national school threshold. In Section 6 we provide evidence that our findings are robust to alternative sample designations.

5.1 Graduation from a National School

Figure 2 illustrates the first stage relationship between national school graduation and students' KCPE scores relative to the binding effective national school cutoff, our running variable. Students with scores above the threshold were admitted to national schools in the first round and those with scores below the threshold were not. Each circle represents the portion of students the same score distance from the national school cutoff who graduated from a national school, e.g. 75 percent of students with scores 10 points above the cutoff graduated from a national school. The data are plotted using one point bins, the smallest possible bin size given the integer nature of the KCPE score. The solid lines are fitted values from a bivariate linear specification, estimated separately on either side of the threshold.

The figure shows that the admissions rule had substantial, but imperfect, adherence. Students who scored exactly at the threshold (i.e. $r_i = 0$) were substantially

more likely to graduate from a national school than students whose scores were one point below the threshold (i.e. $r_i = -1$). The vertical distance between the solid lines at the threshold approximates δ in Equation 3 without the additional $\mathbf{X}_i$ covariates.

Table 2 contains estimates of Equation 3, confirming the statistical significance of the discontinuity seen in Figure 2. Being admitted to a national school had a positive and significant effect of about 50 percentage points on the probability that a student graduated from a national school. Therefore, the admissions rule was followed with strong, but imperfect, adherence, and we use a student's rule-based admission as an instrument in our implementation of the fuzzy regression discontinuity below.

5.2 Differences in School Characteristics

Before we present the effect of national schools on achievement, we explore the differences in school characteristics across the admissions threshold using the fuzzy regression discontinuity design. As outlined in Section 2.1, national and non-national schools differ on a number of margins. Due to data constraints, we can only empirically test for discontinuities in a limited number of these characteristics. To implement this process we use various measures of school characteristics as outcomes in the two stage least squares specification defined by Equations 3 and 4. Table 3 contains the results. In column 1 we use the number of KCSE subjects offered by a school as a proxy for school resources. All schools must offer at least seven subjects: mathematics, English, Swahili, two science subjects, one humanities subject, and one practical subject. The complete secondary curriculum contains 30 subjects, and the inclusion of subjects beyond the minimum is at each school's discretion. While an admittedly coarse measure of school resources, offering more subjects requires additional specialized teachers, classrooms, and perhaps special equipment such as aircraft.[13] Table 3 column 1 shows that students who graduated from national schools had about 2.6 more subjects available to them, reflecting schools with better resources.

[13] Since the marginal cost of an additional subject might be cheaper for larger schools and national schools are larger than some provincial schools, we include the total size of the graduating cohort as an additional control variable. The finding is robust to the removal of this control.

We consider the mean and coefficient of variation of peer KCPE scores, peer characteristics that other studies (e.g. Duflo, Dupas, and Kremer 2011) have found to have important impacts on a student's own achievement. We test for a discontinuity in peer quality (as measured by standardized peer KCPE scores) and find that students in national schools were exposed to higher quality peers (column 2). The IV estimates indicate that the increase in peer quality was about one half of a standard deviation in national schools. However, we do not find evidence that the coefficient of variation of peer scores varied discontinuously around the admissions threshold (column 3). Overall, the first three columns in Table 3 reinforce the conventional belief in Kenya that national schools provide a higher quality educational experience than other government schools.

National schools also differed from other government schools on a number of additional margins that do not have clear achievement implications. By design, national schools should be more ethnically diverse than non-national schools because they draw students from the whole nation, rather than from only a particular province or district, and ethnicity and district are closely related in Kenya. In columns 4 and 5 we explicitly test for a discontinuity in within-school diversity. Due to data limitations we base our measure of diversity on students' home districts, a reasonable proxy for differences in ethnicity and native language in Kenya. In column 4 we use an Herfindahl-Herschman index (HHI) of diversity of graduates' home districts. A school in which all graduates came from the same home district would have a value of 1, while a more diverse school would have an HHI closer to 0. Each student is assigned his or her school-specific HHI measure. For the students in our sample the average value of the HHI is 0.3. Students who graduated from national schools were exposed to more diverse peers (lower HHI) as can be seen with the negative and statistically significant value of -0.307 in column 4. As an additional measure of the diversity of students, we calculate the portion of graduates from a secondary school that were not from the student's home district. Consistent with the HHI findings, students who graduated from national schools have a 27.5 percentage point increase in the portion of students not from their home district (column 5). Overall these results indicate

that national schools were succeeding in their mandate to bring together students from across the country. Previous studies have shown that social interactions in diverse schools can promote better inter-ethnic or inter-racial relations (Boisjoly et al. 2006), however the effect of increased ethnic diversity on achievement is unknown.

All of these characteristics as well as the superior facilities and more educated teachers with more experience outlined in Section 2.1 combined in the production of student outcomes at the end of secondary school.

5.3 Timely Progression

A student who progresses through secondary school on pace would complete it in four years. Figure 3 illustrates the relationship between the probability of graduating from secondary school four years after taking the KCPE and a student's score relative to the binding national school score threshold. The circles are the portion of students at each score distance that graduated from secondary school in exactly four years.[14] The solid lines are the fitted values from linear specifications, estimated separately on either side of the threshold. The visual inspection of these lines suggests no differential graduation probability on either side of the discontinuity. The estimate of Equation 3 with graduation from any secondary school as the dependent variable appears in Table 4 and confirms the lack of a statistically significant break and small point estimate (0.002). Therefore, students admitted to national schools were not differentially likely to graduate on time. Additionally, these estimates indicate no differential selection of students by school type into taking the KCSE exam.

5.4 Achievement

At the conclusion of secondary school, all students take the KCSE exam. A student's KCSE score is the average of seven subject exams: three core subjects (math, Swahili, and English), the top two science scores, the top humanities score, and the top practical score. This score is used to determine admission to colleges and universities and

[14] The increase in noise at higher values of the score distance reflects a smaller sample of students with such high scores.

some employers set minimum KCSE achievement requirements. Section 5.2 showed that national schools had more resources and higher ability and more diverse peers. The achievement effects we estimate are the overall effect of these differences along with any other unmeasured quality differences between the two school types.

Figure 4 illustrates the reduced form relationship between being admitted to a national school and the standardized KCSE score at the conclusion of secondary school. Despite the substantial differences between national schools and other government schools, we observe very small differences in KCSE performance across the national school admissions threshold.[15] Table 5 contains a number of estimations to directly test the statistical significance of any effect. Columns 1 and 2 contain a naïve OLS specification of Equation 1, controlling for KCPE score and district, sex, public primary school, and school choice dummy variables. When estimated over the full sample (column 1), graduating from a national school was associated with an increase in the KCSE score of 30 percent of a standard deviation. Once the sample is limited to similar students, the point estimate decreases substantially to less than 5 percent of a standard deviation (column 2). Column 3 contains the reduced form estimation of the effect of being admitted to a national school on a student's KCSE score. Being admitted to a national school had no statistically significant effect on the KCSE score.[16] Column 4 uses the instrumental variables approach specified in Equations 3 and 4 to examine the effect of national school graduation on KCSE scores. The results again show that national schools did not have a statistically significant effect on KCSE scores. Moreover, the standard errors are small enough that we can rule out moderate effect sizes of 0.12 standard deviations or larger. Therefore, despite superior resources and peers, attending a national school did not result in a higher KCSE score.

In columns 5-8 we estimate the effect of national school graduation on the average

[15] As discussed in the previous section, we found that the probability of graduation from secondary school did not differ around the national school threshold. This alleviates concerns about differential selection into taking the KCSE exam around the threshold.

[16] From Figure 4, one might expect a statistically signficant effect when those students who score at the cutoff are removed from the sample. The effect estimated over the re-defined sample is similarly statistically insignificant.

of the required subjects exams (English, math, and Swahili) and each required subject score separately. We do not find a statistically significant effect of national schools on the required subject average or English score. National schools appear to have negatively affected mathematics scores, but this result is only marginally significant in this specification and is not robust to the alternative specifications presented in Section 6. National schools did have a large, statistically significant, positive effect of 0.257 standard deviations on Swahili scores (column 8). We are not able to empirically test the source of this difference, but a number of hypotheses are consistent with this findings. First, our findings in Section 5.2 showed that national schools were more diverse than other schools. Given the diversity of mother tongues in national schools, Swahili, one of two official languages in Kenya and often referred to as the national language, could be the lingua franca for daily communication, while in provincial schools a local language could be used.[17] Second, a greater emphasis could be placed on Swahili in national schools because of their original mandate to create national unity. Indeed, according to the official teaching timetables some national schools devote more time to Swahili instruction than provincial schools do. Third, high quality Swahili teachers could be relatively more scarce than teachers of other subjects, resulting in the a larger ability difference between Swahili teachers and other subject teachers in national schools versus provincial schools. Due to data limitations, we cannot empirically distinguish among these hypotheses.[18]

Even though we do not find an effect of national schools on composite KCSE scores, benefits could accrue heterogeneously by baseline ability, the quality of the non-national school alternative, gender, or socioeconomic status. We test each of these below.

As discussed in Section 3, the admissions algorithm creates multiple regression discontinuities within the same school, resulting in substantial KCPE score variation

[17] English is the other official language of Kenya, but is less commonly used than Swahili in day to day communications.

[18] We are able to reject the hypothesis that students in national schools took fewer subjects and therefore had more time to devote to required subjects. Based on the instrumental variables technique, we find that students in national schools took 0.059 more subjects tests, but this is not statistically significant with a standard error of 0.052.

within a single school. We can, therefore, test for heterogeneous effects by students' relative KCPE scores and place in a school's KCPE distribution. We first examine the extent to which national schools differentially affect students by their prior academic preparation.[19] In column 1 of Table 6 the coefficient on the interaction between national school graduation and student's KCPE score is negative but not statistically significant. Therefore, the impact of national schools did not vary by a student's KCPE score. Second, in column 2 we test for potential heterogeneity in the effect of national schools by students' within-school relative score positions. The multiple discontinuities within a single national school allow us to test if students benefit differentially based on their initial knowledge relative to their within-school peers. We determine a student's within-school percentile based on incoming KCPE scores.[20] Our IV estimates in column 2 show that there are no differential effects of national school graduation if a student is in the same percentile in both national and non-national schools, i.e. the coefficient on the interaction term is small and statistically insignificant.[21]

Another potential source of differential effects was the substantial variation in the quality of the non-national school option for students who were not admitted to a national school. The schools preferred after national schools are provincial schools. As the name suggests, these schools admit students from a single province. The heterogeneous quality of these schools creates an additional source of variation between students from difference provinces at the national school admissions threshold. For example, for girls from Nairobi province the top provincial school, Precious Blood,

[19] Formally, we use admission to a national school times KCPE score as an additional exogenous regressor in our IV estimation in which we treat graduation from a national school times KCPE score as an additional endogenous regressor. As discussed above, the KCPE score is absorbed by the set of control variables.

[20] The score percentile is based on the school from which a student graduated. We prefer this measure of percentile instead of one calculated based on the school to which a student was admitted. First, this better reflects the actual peers with which a student interacted. Second, the admissions rule for provincial school is *de jure* the same as the one for national schools, but the level of adherence is substantially lower. Therefore, we are not confident that we can correctly locate the relevant admitted peer group for each non-national school student.

[21] Formally, we use admission to a national school times within-school score percentile as an additional exogenous regressor in our IV estimation in which we treat graduation from a national school times within-school score percentile as an additional endogenous regressor. We include within school score percentile as an additional regressor.

is highly regarded and had a higher average incoming KCPE score than some of the girls' national schools. In contrast, for girls from North Eastern province the top provincial school had an average incoming KCPE score one standard deviation below the average of the least selective national school. Since students with the highest scores who were not admitted to a national school select from the entire range of provincial schools within their home province, we use the average incoming KCPE score for the 2008 provincial school graduates as our measure of provincial school quality, a quasi revealed preference ranking. We rank all schools separately by gender using this average score and consider the top 10 percent of provincial schools for each gender in the country "highly selective."[22] The portion of total provincial seats in a province that were "highly selective" is our proxy measure for the overall quality of non-national school options available to each student. This measure varies from 0 for both genders in North Eastern province to 0.54 for girls in Nairobi. In column 3 we use our IV methodology to test for a differential effect by this provincial quality.[23] The interaction term on our measure of provincial quality and graduation from national school is negative and statistically significant. Therefore, consistent with prior studies, we find that students with greater access to "highly selective" provincial school seats benefited less from attending a national school.[24] These estimates suggest that students from Nairobi and Central province, where more than 27.5% of spaces in provincial schools are "highly selective," may have been better served by provincial rather than national schools.

Finally, we examine whether the effects differ by gender or by our proxy measure of socioeconomic status. We do not find any differential effects by gender (column 4) or by public primary school graduation (a proxy for socioeconomic status, column

[22] Even in mixed-gender provincial schools, quotas and admissions are separate for each gender. Therefore, we rank by gender.

[23] Formally, we use admission to a national school times provincal school quality as an additional exogenous regressor in our IV estimation in which we treat graduation from a national school times provincal school quality as an additional endogenous regressor. We cannot include provincial school quality as a separate regressor as it is absorbed by district and district times school choice fixed effects.

[24] We cannot empirically reject that this measure of provincial quality is correlated with other provincial attributes that might cause students to perform better or worse in a national or provincial school environment.

5).[25]

6 Robustness Checks

One potential concern with any regression discontinuity is manipulation of the assignment around the threshold. We think this is unlikely in our setting for a number of reasons. First, the effective threshold for each national school is district, school, and year specific, is not established until all exams have been graded, and depends on all of the exam scores and preferences within a district. Second, all exams are graded centrally and the graders do not know the students. Third, manipulation by graders would require knowledge about student preferences, and this information does not appear on the exams. Finally, once KCPE exams are scored, the admissions decisions are computer generated based on these scores, student preferences, and district quotas. On the other hand, manipulation can occur in the second round when students directly approach school principals for admission. For this reason, we rely on the first round rule-based admission, not the second round admissions decisions, when we construct our instrument for graduating from a national school. Figure 5 provides additional evidence about the validity of our regression discontinuity design. The first frame provides the number of students who scored in each one-point score bin. As expected there is no break as the cutoff is approached from above or below. The number of students who scored exactly at the threshold is substantially larger than those who score one point below and above. This spike is due to the design of the cutoff. By assigning the lowest student score in each district-school pair to serve as the threshold, all district-school pairs had a student exactly at the threshold, but the next higher (or lower) student score in a district could be multiple points away from the cutoff, thus the number of students in each point above and below the threshold was smaller than at the threshold.

Unfortunately, we only have limited demographic data on students to test for ad-

[25] Lucas and Mbiti (2012) found that the likelihood of private primary school attendance is increasing in parental education.

ditional discontinuities at the threshold. Frame B displays the percent of students who graduated from a public primary school, an approximation of socioeconomic status since richer students are more likely to attend private primary schools. Frame C plots the average student age. Both measures become noisier for the scores that are more than 25 points above the threshold because of the relatively smaller sample size for those scores. Neither demographic measure appear to be visually discontinuous at the admissions threshold. Table 7 columns 1 and 2 empirically test for the statistical significance of any discontinuity and find neither to be statistically significant. Our empirical strategy passes the typical concerns about the internal validity of a regression discontinuity.

Unobserved student behaviour could vary discontinuously at the threshold. For example, students admitted to a national school might rely on the name cache of their school and not study as hard as those students in provincial schools who felt like they need to overcome the provincial school stigma (MacLeod and Urquiola 2009). This would bias the results towards not finding an achievement effect for national schools. Unfortunately we cannot empirically measure student effort, but in Kenya the KCSE score is used for admission to universities and to qualify for some jobs, without regard to secondary school name, partially mitigating this possibility as a valid strategy for students who wish to continue their studies or seek employment after secondary school.

Table 8 provides a number of specification checks. Each specification is a modification of the baseline estimates that we repeat in column 1 from earlier tables to aid comparison. In column 2 we limit the controls to the piecewise linear function of the running variable, KCPE score, and female and district dummy variables. In column 3 we replace the piecewise linear function of the running variable with a third degree polynomial that we allow to vary discontinuously at the threshold. In column 4 we re-estimate the results using the Imbens-Kalyanaraman optimal bandwidth for our data (Imbens and Kalyanaraman 2009). Finally, we limit the estimation window to only include observations within 17 points of the admissions threshold in column 5. Overall, our results are robust to these specification checks with the exception of the

mathematics subject score in Panel E. The previously statistically significant math result is not robust to specifications with a third degree polynomial (column 3) or narrower sample (column 5). In fact, the point estimate becomes positive, but statistically insignificant, in the specification with the third degree polynomial (column 3).

7 Conclusions

This paper exploits the centralized system of admitting students to government secondary schools in Kenya to estimate the effect of elite schools on student academic achievement. Using a regression discontinuity design we find that students who were admitted to national schools (the most elite schools in Kenya) were equally likely to graduate on time from secondary school as their peers who were not admitted. Further, we find no statistically significant difference between student composite scores on the secondary school exit exam for students in national versus other schools. The lack of an effect is similar across genders, socioeconomic status, academic preparation, and within-school score percentile. We do find that the national school graduates had higher Swahili scores, consistent with its use as a lingua franca in national schools and more time being devoted to Swahili in some national schools. Our results also suggest that students from provinces with higher quality provincial schools could have benefited from provincial instead of national school graduation.

Even though we find no average impact of national schools on achievement, parents, students, and the government highly value and expend substantial resources on these elite schools. One possibility is that national schools may deliver other benefits that we are not able to measure. For instance, national school graduation could act as a signal, alter occupational choice and earnings, provide access to enhanced networking opportunities. Alternatively, stakeholders could be effectively running the OLS regression in which they see that students who graduate from national schools have the among the highest scores in the country at the conclusion of secondary school, not taking into account that these students also had among the highest test scores at the

start of secondary school. Future research will seek to disentangle these alternatives and provide additional insight into national schools, an institution often credited with developing the past and future leaders of Kenya.

References

Abdulkadiroglu, A., J. D. Angrist, and P. A. Pathak (2011): "The Elite Illusion: Achievement Effects at Boston and New York Schools," NBER Working Paper 17264.

Angrist, J., and V. Lavy (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114(2), 533–575.

Boisjoly, J., G. J. Duncan, M. Kremer, D. Levy, and J. Eccles (2006): "Empathy or Antipathy? The Impact of Diversity," *The American Economic Review*, 96(5), 1890–1905.

Bui, S. A., S. G. Craig, and S. A. Imberman (2011): "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Achievement," NBER Working Paper 1709.

Clark, D. (2010): "Selective Schools and Academic Achievement," *The B.E. Journal of Economic Analysis and Policy*, 10(1).

de Hoop, J. (2010): "Keeping Children in School: Why School Quality Matters in Sub-Saharan Africa," Tinbergen Institute Mimeo.

Deming, D. J., J. S. Hastings, T. Kane, and D. O. Staiger (2011): "School Choice, School Quality and Postsecondary Attainment," NBER Working Paper 17438.

Dobie, W., and R. Fryer (2011): "Exam High Schools and Academic Achievement: Evidence from New York City," NBER Working Paper 17286.

Duflo, E., P. Dupas, and M. Kremer (2011): "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *The American Economic Review*, 101(5), 1739–1774.

Hahn, J., P. Todd, and W. V. der Klaauw (2001): "Identification and Estimation of Treatment Effects with Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

Imbens, G., and K. Kalyanaraman (2009): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *forthcoming in Review of Economic Studies*.

Jackson, C. K. (2010): "Do Students Benefit from Attending Better Schools? Evidence From Rule-Based Student Assignments in Trinidad and Tobago," *Economic Journal*, 20(549), 1399–1429.

Lee, D. S., and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

MacLeod, W. B., and M. Urquiola (2009): "Anti-Lemons: School Reputation and Educational Quality," NBER Working Paper 1512.

Onsomu, E., D. Muthaka, M. Ngware, and G. Kosimbei (2006): "Financing of Secondary Education in Kenya: Costs and Options," KIPPRA Discussion Paper 55.

Ozier, O. (2011): "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis," World Bank Mimeo.

Park, A., X. Shi, C. tai Hsieh, and X. An (2010): "Does School Quality Matter?: Evidence from a Natural Experiment in Rural China," Preliminary Draft.

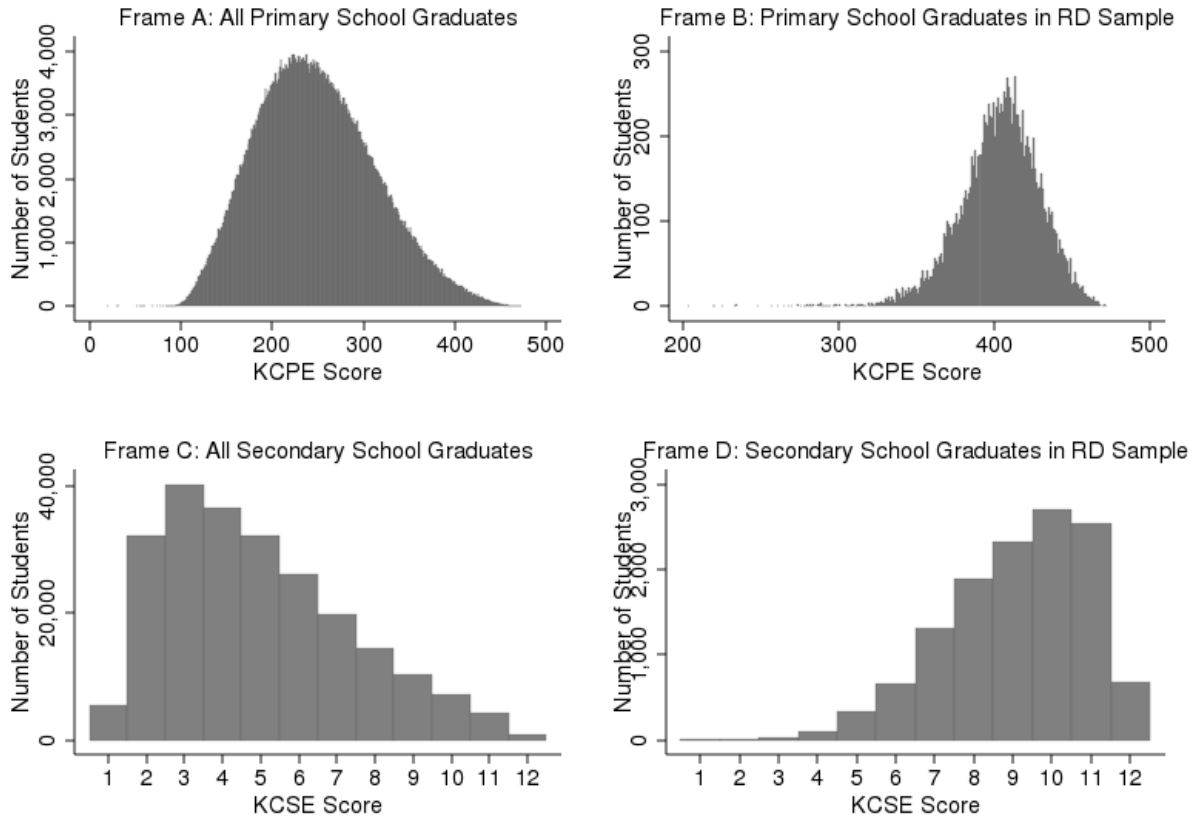
Pop-Eleches, C., and M. Urquiola (2011): "Going to a Better School: Effects and Behavioral Responses," NBER Working Paper 16886.

Sekhri, S., and Y. Rubinstein (2010): "Do Public Colleges in Developing Countries Provide Better Education than Private ones? Evidence from General Education Sector in India," Mimeo, University of Virginia.

Uwezo (2010): "Are Our Children Learning? Annual Learning Assessment Kenya 2010," Uwezo report.

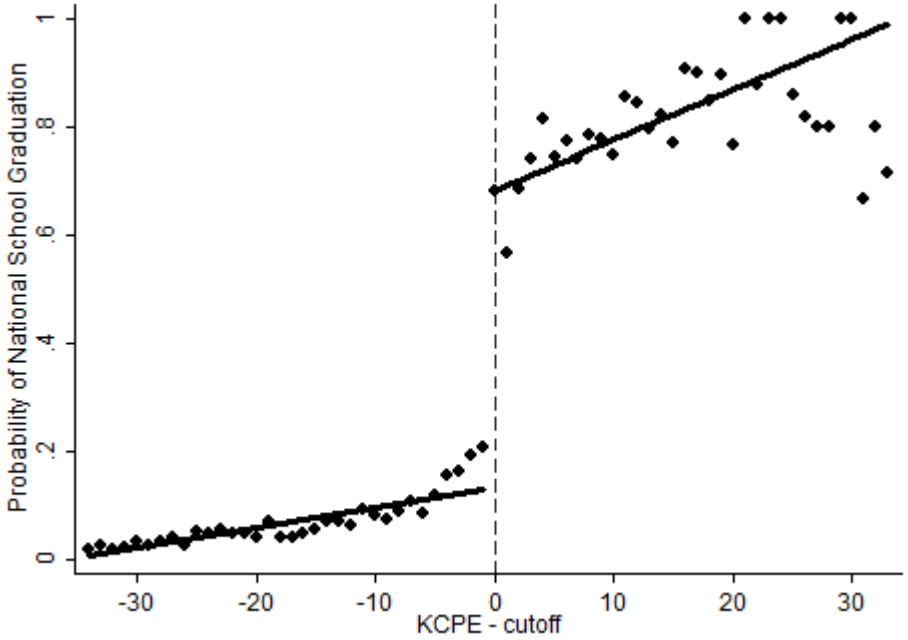
World Bank (2004): "Strengthening the Foundation of Education and Training in Kenya: Opportunities and Challenges in Primary and General Secondary Education," Report No. 28064-KE.

Figure 1: KCPE and KCSE Distributions



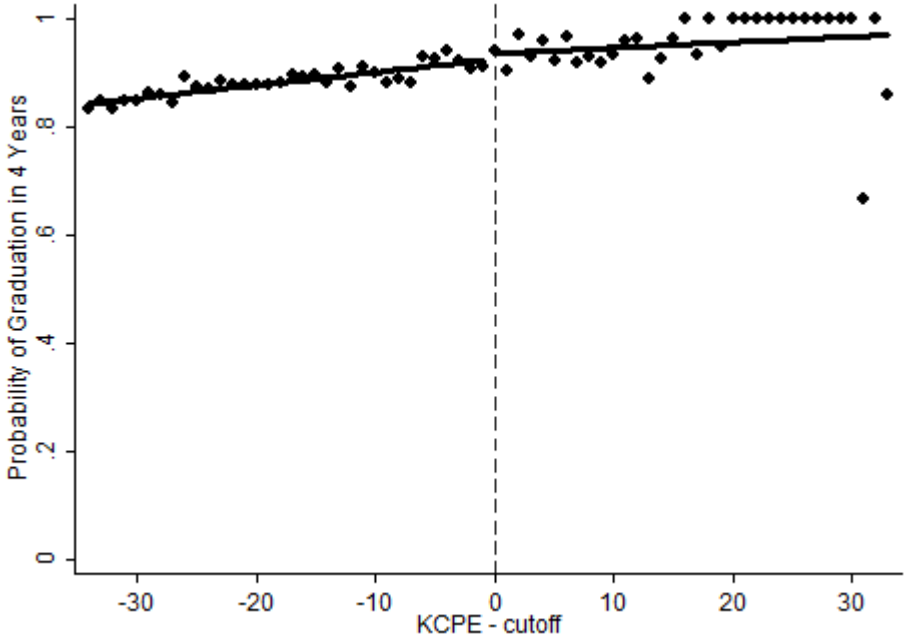
Notes: Panels B and D: sample limited to students within the +/- 34 point window of a national school cutoff.

Figure 2: Probability of National School Graduation



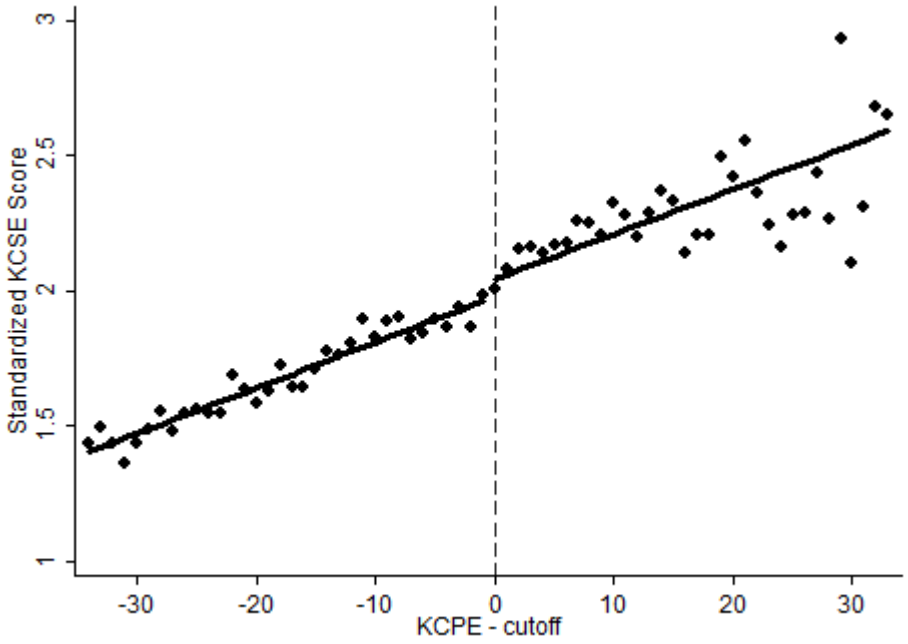
Note: "KCPE-cutoff" is the KCPE score minus the effective national school cutoff score. See text for details on the calculation of the cutoff. Each point is the mean of the probability of graduation from a national school within non-overlapping 1 point bins. The solid lines are fitted values from a linear specification, separately estimated on each side of the cutoff.

Figure 3: Timely Progression through Secondary School



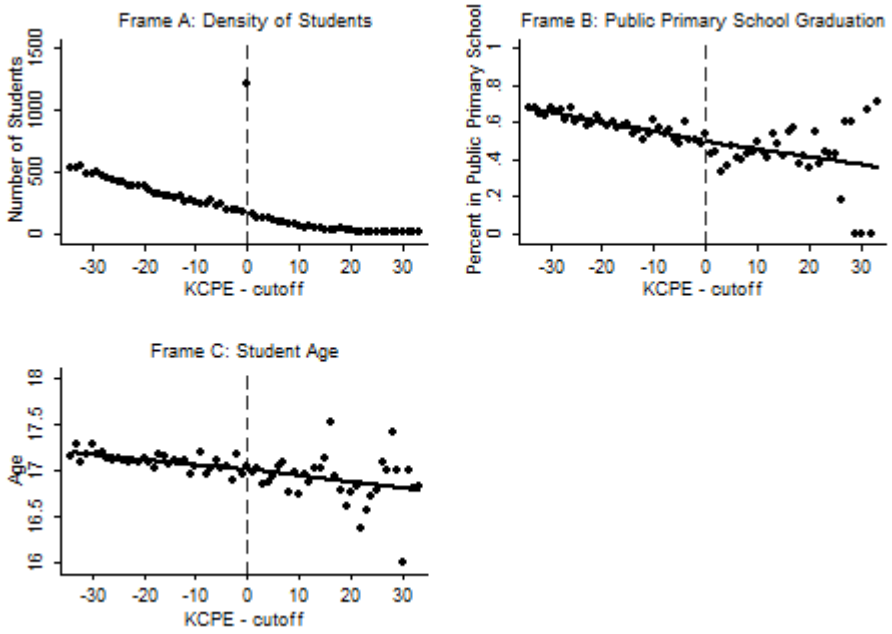
Note: "KCPE-cutoff" is the KCPE score minus the effective national school cutoff score. See text for details on the calculation of the cutoff. Each point is the mean of the probability of graduation from any secondary school within four years of taking the KCPE within non-overlapping 1 point bins. The solid lines are fitted values from a linear specification, separately estimated on each side of the cutoff.

Figure 4: Student Achievement



Note: "KCPE-cutoff" is the KCPE score minus the effective national school cutoff score. See text for details on the calculation of the cutoff. Each point is the mean of the standardized KCSE score within non-overlapping 1 point bins. The solid lines are fitted values from a linear specification, separately estimated on each side of the cutoff.

Figure 5: Validity of the Regression Discontinuity



Note: "KCPE-cutoff" is the KCPE score minus the effective national school cutoff score. See text for details on the calculation of the cutoff. Frames B and C: Each point is the mean of y-axis variable within non-overlapping 1 point bins.

Table 1: Summary Statistics

	Primary School Graduates (2004)	Secondary School Graduates (2008)						
		All	National School	Provincial School	District School	Private School	Disability School	Unknown Type
Number of Students	651,647	229,503	3,100	79,394	115,435	28,578	309	2,687
Number of Schools		5,158	18	943	3,190	859	6	142
Average KCPE Score (out of 500)	246.0 (67.8)	288.4 (60.6)	414.4 (33.2)	322.5 (49.8)	266.0 (50.5)	273.7 (66.9)	331.7 (85.0)	251.1 (58.9)
Average KCSE Score (out of 12)		4.92 (2.4)	9.58 (1.79)	6.23 (2.3)	4.05 (1.9)	4.38 (2.6)	6.97 (3.1)	3.19 (1.9)
Average Number of Subject Tests Offered		11.0 (1.3)	16.4 (2.7)	12.2 (1.3)	10.8 (1.0)	10.9 (1.4)	12.5 (1.8)	10.1 (1.4)
Percent Male	51.7	52.6	56.0	53.7	53.0	48.2	49.5	48.6
Percent Graduating from Public Primary Schools	92.6	87.1	49.2	81.6	93.7	80.4	57.6	89.2

Notes: Standard deviations appear in parenthesis. Source: Calculations based on Kenya National Examination Council data.

Table 2: First Stage Relationship: Probability of National School Graduation

Dependent Variable:	Graduate from National School
Admitted to a National School	0.497***
	(0.020)
Window	+/- 34
Observations	14,157
Rsquared	0.54

Notes: Sample limited to students within indicated window of national school cutoff. Linear probability model. Controls: Piecewise linear function of KCPE score minus cutoff; district, male, national school preferences, and public primary school dummy variables; and interactions between school preference and district dummy variables. Standard errors clustered at the district level appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Differences in School Characteristics at the National School Admissions Threshold

Dependent Variable:	School Resources	School Peers			
	Number of Subjects Offered	Peer KCPE Scores	KCPE Coefficient of Variation	Ethnic HHI	Portion from Other Districts
	(1)	(2)	(3)	(4)	(5)
Graduated from a National School	2.585***	0.450***	0.515	-0.307***	0.275***
	(0.156)	(0.059)	(0.325)	(0.032)	(0.044)
Window	+/- 34	+/- 34	+/- 34	+/- 34	+/- 34
Observations	12,467	12,466	12,466	12,467	12,467
Rsquared	0.48	0.52	0.07	0.59	0.46

Notes: Standard errors clustered at the district level appear in parentheses. All columns are instrumental variables regressions with admission to a national school an instrument for graduation from a national school. Sample limited to students who graduated from secondary school and are within stated window of a binding national school threshold. Controls: Piecewise linear function of KCPE score minus cutoff; district, male, national school preferences, and public primary school dummy variables; and interactions between school preference and district dummy variables. Column 1: the number of distinct subject exams completed by all peers. Column 2: standardized peer KCPE scores. Column 3: coefficient of variation of peer KCPE scores. Column 4: HHI calculated (max = 1) based on home districts of students who graduated from the same secondary school. Column 5: portion (measured 0-1) of students from districts other than student's home district. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Effect of National Schools on Timely Progression

Dependent Variable:	Graduate from Any Secondary School
Admitted to a National School	0.002
	(0.009)
Window	+/- 34
Observations	14,157
Rsquared	0.10

Notes: Linear probability model. Sample limited to students within indicated window of national school cutoff. Controls: Piecewise linear function of KCPE score minus cutoff; district, male, national school preferences, and public primary school dummy variables; and interactions between school preference and district dummy variables. Standard errors clustered at the district level appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Effect of National Schools on Achievement

Dependent Variable:	Standardized KCSE Score				Standardized Subject Scores			
	OLS		Reduced Form	IV Estimation	Required Subjects	English	Math	Swahili
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduate From National School	0.294***	0.046*		0.015	0.032	-0.008	-0.116*	0.257***
	(0.0261)	(0.0270)		(0.053)	(0.051)	(0.045)	(0.063)	(0.058)
Admitted to a National School			0.008					
			(0.028)					
Window	all pupils	+/- 34	+/- 34	+/- 34	+/- 34	+/- 34	+/- 34	+/- 34
Observations	211,937	12,467	12,467	12,467	12,467	12,467	12,467	12,467
Rsquared	0.64	0.36	0.36	0.36	0.40	0.42	0.28	0.35

Notes: Sample limited to students within indicated window of national school cutoff. Controls: district, male, national school preferences, and public primary school dummy variables and interactions between school preference and district dummy variables. Columns (3)-(8): piecewise linear function of KCPE score minus cutoff also included. Columns (4)-(8): Instrumental variables estimation with national school admission as the instrument. Column (5): required subjects are math, English, and Swahili. Standard errors clustered at the district level appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Heterogeneous Effects by Ability, Provincial Quality, and Demographics

	Dependent Variable: Standardized KCSE Score				
	Heterogeneous Effects				
	By Relative Ability		By Provincial Quality	By Gender	By Socioeconomic Status
	(1)	(2)	(3)	(4)	(5)
Graduate From National School	0.388	-0.132	0.120	0.030	0.025
	(0.259)	(0.103)	(0.075)	(0.063)	(0.058)
Graduate From National School X Standardized KCPE Score	-0.143				
	(0.095)				
Graduate From National School X Within School Percentile		0.001			
		(0.131)			
Within School Percentile		-0.448***			
		(0.071)			
Graduate From National School X Availability of High Quality Provincial Seats			-0.436**		
			(0.206)		
Graduate From National School X Female				-0.031	
				(0.053)	
Graduate From National School X Public Primary School					-0.019
					(0.048)
Female	0.271	-0.227	-0.082	-0.064	-0.468
	(0.516)	(0.216)	(0.259)	(0.499)	(0.873)
Public Primary School	0.138***	0.154***	0.138***	0.138***	0.142***
	(0.034)	(0.035)	(0.034)	(0.034)	(0.035)
Window	+/- 34	+/- 34	+/- 34	+/- 34	+/- 34
Observations	12,467	12,467	12,467	12,467	12,467
Rsquared	0.36	0.37	0.36	0.36	0.36

Notes: All columns are instrumental variables regressions with admission to a national school as an instrument for graduating from a national school. Sample limited to students who graduated from secondary school with KCPE scores within stated window of a binding national school cutoff. Controls: piecewise linear function of KCPE score minus cutoff, district, national school preferences, and public primary school dummy variables; and interactions between school preference and district dummy variables. Column (1): admission to national school times standardized KCPE scores used as an additional instrument. Column (2): percentile determined by school from which a student graduated, admission to national school times within-school percentile used as an additional instrument. Column (3): admission to national school times availability of high quality provincial seats used as an additional instrument, high quality provincial seats defined as the portion of provincial seats within student's home province that are in schools that are in the top 10% of provincial schools in the country based on incoming KCPE scores of graduates. Column (4): Admission to national school times female used as an additional instrument. Column (5): Admission to national school times public primary school graduation as an additional instrument, public primary school graduation used as a proxy for lower socioeconomic status. Standard errors clustered at the district level appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Balance in Observables

	Public Primary School	Age
	(1)	(2)
Admitted to a National School	-0.009	-0.062
	(0.015)	(0.041)
Window	+/- 34	+/- 34
Observations	12,467	12,467
Rsquared	0.29	0.07

Notes: Standard errors clustered at the district level appear in parentheses. Sample limited to students within indicated window of national school cutoff. Controls: district dummy variables and a piecewise linear function of KCPE score minus the cutoff. Column 1: Linear probability model. * significant at 10%; ** significant at 5%; ***

Table 8: Additional Specification Checks

	Dependent Variable: Standardized KCSE Score				
	Baseline	Limited Controls	3rd Degree Polynomial	2 Point (IK Ideal) Bandwidth	+/- 17 Point Window
	(1)	(2)	(3)	(4)	(5)
Panel A: Graduate from a National School					
Admitted to a National School	0.497***	0.535***	0.373***	0.494***	0.401***
	(0.020)	(0.019)	(0.023)	(0.021)	(0.024)
Observations	14,157	14,157	14,157	14,157	6,627
Rsquared	0.54	0.49	0.55	0.54	0.57
Panel B: Standardized KCSE Score					
Graduate From National School	0.015	-0.012	0.064	0.005	0.000
	(0.053)	(0.049)	(0.098)	(0.053)	(0.081)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.36	0.28	0.36	0.36	0.38
Panel C: Standardized Required Subject Score					
Graduate From National School	0.032	0.001	0.113	0.022	0.048
	(0.051)	(0.047)	(0.097)	(0.050)	(0.078)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.40	0.33	0.40	0.40	0.43
Panel D: Standardized English Score					
Graduate From National School	-0.008	-0.008	0.018	-0.016	-0.002
	(0.045)	(0.039)	(0.084)	(0.045)	(0.065)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.42	0.36	0.42	0.42	0.44
Panel E: Standardized Math Score					
Graduate From National School	-0.116*	-0.160***	0.052	-0.126**	-0.068
	(0.063)	(0.058)	(0.135)	(0.062)	(0.107)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.28	0.20	0.28	0.28	0.32
Panel F: Standardized Swahili Score					
Graduate From National School	0.257***	0.221***	0.254**	0.247***	0.236***
	(0.058)	(0.054)	(0.110)	(0.058)	(0.085)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.35	0.28	0.35	0.35	0.39
Panel G: Heterogeneous Effects by Standardized KCPE Score					
Graduate From National School	0.388	0.408*	0.430	0.369	0.344
	(0.259)	(0.210)	(0.270)	(0.255)	(0.343)
Graduate From National School X	-0.143	-0.159**	-0.145	-0.139	-0.133
Standardized KCPE Score	(0.095)	(0.076)	(0.095)	(0.094)	(0.128)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.36	0.28	0.36	0.36	0.38

	Dependent Variable: Standardized KCSE Score				
	Baseline	Limited Controls	3rd Degree Polynomial	2 Point (IK Ideal) Bandwidth	+/- 17 Point Window
	(1)	(2)	(3)	(4)	(5)
Panel H: Heterogeneous Effects by Within School Percentile					
Graduate From National School	-0.132	-0.158	-0.118	-0.141	-0.151
	(0.103)	(0.095)	(0.112)	(0.103)	(0.158)
Graduate From National School X Within School Percentile	0.001	0.014	0.027	-0.007	-0.023
	(0.131)	(0.104)	(0.167)	(0.133)	(0.193)
Within School Percentile	-0.448***	-0.439***	-0.440***	-0.453***	-0.482***
	(0.071)	(0.064)	(0.064)	(0.071)	(0.128)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.37	0.29	0.37	0.37	0.39
Panel I: Heterogeneous Effects by Quality of Provincial School					
Graduate From National School	0.120	0.086	0.163	0.107	0.103
	(0.075)	(0.068)	(0.115)	(0.076)	(0.105)
Graduate From National School X Availability of High Quality Provincial Seats	-0.436**	-0.390**	-0.466**	-0.422**	-0.419*
	(0.206)	(0.177)	(0.220)	(0.206)	(0.237)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.36	0.28	0.36	0.36	0.38
Panel J: Heterogeneous Effects by Sex					
Graduate From National School	0.030	0.023	0.084	0.020	0.022
	(0.063)	(0.059)	(0.111)	(0.063)	(0.098)
Graduate From National School X Female	-0.031	-0.072	-0.041	-0.030	-0.046
	(0.053)	(0.045)	(0.054)	(0.053)	(0.066)
Female	-0.064	0.178***	0.702	-0.055	0.421
	(0.499)	(0.030)	(0.675)	(0.498)	(0.410)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.36	0.28	0.356	0.36	0.38
Panel K: Heterogeneous Effects by Socioeconomic Status					
Graduate From National School	0.025	0.017	0.075	0.015	-0.027
	(0.058)	(0.055)	(0.103)	(0.057)	(0.087)
Graduate From National School X Lower SES	-0.019	-0.059	-0.021	-0.019	0.054
	(0.048)	(0.044)	(0.048)	(0.048)	(0.061)
Lower SES	0.142***	0.161***	0.142***	0.142***	0.081*
	(0.035)	(0.034)	(0.035)	(0.035)	(0.046)
Observations	12,467	12,467	12,467	12,467	6,013
Rsquared	0.36	0.28	0.36	0.36	0.38
Window	+/- 34	+/- 34	+/- 34	+/- 34	+/- 17

Notes: Standard errors clustered at the district level appear in parentheses. Panels B-K: Results from instrumental variable specification with admission to a national school as an instrument for graduation from a national school and additional interacted instruments as in Table 6. Column (1): Panel A from Table 2, Panel B from Table 5, Panel G-K from Table 6. Column (2): controls limited to piecewise linear function of running variable, KCPE score, and female and district dummy variables. Column (3): controls for third degree polynomial of the running variable that varies on either side of the national school threshold instead of piecewise linear function. Column (4): running variable transformed into non-overlapping 2 point bandwidth, the IK ideal bandwidth. Column (5): sample limited to window within 17 points of the national school threshold. * significant at 10%; ** significant at 5%; *** significant at 1%.