Does School Quality Improve Student Performance? New Evidence from Ghana

Kehinde F. Ajayi*
Boston University

January 31, 2014

Abstract

This study examines whether school assignment affects student performance in Ghana. I track a cohort of 160,000 students who applied to 650 secondary schools across the country under a merit-based admission system. Using a selection on observables approach and regression discontinuity design, I find that students admitted to more selective schools are more likely to stay in the same school and to complete on time, but demonstrate only marginal improvements in overall completion rates and exam performance. I also find substantial heterogeneity in effects, suggesting that both school quality and match quality determine student outcomes in this context.

*Email: kajayi@bu.edu. I am grateful to the Computerised School Selection and Placement System Secretariat, Ghana Education Service, Ghana Ministry of Education, SISCO Ghana, and the West Africa Examinations Council for providing data and background information. I have benefited from numerous discussions throughout this project and especially thank David Card, Varanya Chaubey, David Deming, Joshua Goodman, David Levine, Justin McCrary, Edward Miguel, Amanda Pallais, Sarath Sanga and seminar participants at Boston University, the Center for the Study of the Economies of Africa, the Spring 2012 Greenline Labor Meeting, and DITE 2013 workshop for helpful comments. This research was supported by funding from the Spencer Foundation, the Institute for Business and Economic Research, the Center for Effective Global Action, and the Center for African Studies at UC Berkeley. Marric Buessing provided excellent research assistance.

1 Introduction

Parents, policymakers, and researchers have long grappled with the question of whether attending a high performing school improves a student's future outcomes. A number of recent studies advance the existing literature by estimating the effects of gaining admission to a selective school, using administrative data from school choice systems with merit-based admission procedures.[1] Researchers generally interpret these estimated admission effects as capturing the effects of improvements in school quality; however, this common interpretation neglects the potential importance of another factor – the *match* between a school's attributes and a student's individual characteristics.

This paper examines both the school quality and match quality effects of admission to selective secondary schools in Ghana. Using administrative data from Ghana's centralized school choice system, I track the cohort of 160,000 students who applied to 650 secondary schools across the country in the first year the system was introduced, and link students' application and school assignment information to their academic performance at the end of secondary school. In contrast to the emphasis in earlier work, I focus on the fact that admission to selective schools not only offers students access to better educational inputs and higher achieving peers but also enables them to attend a school that better matches their preferences.

The secondary school application context in Ghana provides an ideal opportunity to tackle a key empirical challenge: finding an exogenous source of variation in school assignment in order to separate observed school attendance from unobserved factors such as student motivation, family resources, and intrinsic academic ability. Ghana's centralized secondary school application system allocates students based on their ranking of a list of preferred choices and their performance on a standardized exam. I can therefore observe student exam scores and preferences, and exploit exogenous variation in school assignment resulting from the merit-based admission rule. Additionally, there is a substantial amount of noncompliance with school assignments because students are able to switch into different schools through unofficial channels. This laxity in the Ghanaian school choice system allows me to observe student switches as a proxy for the quality of initial student-school matches.

I use two research designs to estimate the effects of admission to selective schools, which

[1] The results of these studies have been mixed. Researchers have found positive effects of admission to selective schools in Trinidad and Tobago, Malawi, and Romania (Jackson (2010), de Hoop (2010), and Pop-Eleches and Urquiola (2013), respectively). However, Lucas and Mbiti (forthcoming) find weak effects of admission to elite public schools in Kenya and de Janvry, Dustan, and Sadoulet (2013) find significant negative effects in Mexico. Studies on elite high schools in the United Kingdom and the United States have also found limited impacts on student performance (Clark (2010), Abdulkadiroğlu, Angrist, and Pathak (forthcoming) and Dobbie and Fryer (forthcoming)).

allows me to confirm the validity of these alternative methodologies. First, I use a selection on observables approach, based on the premise that information on student application choices (their revealed preferences) provides a sufficient means to control for unobserved factors that could lead to selection bias in estimates of the relationship between school quality and student outcomes. Second, I use a regression discontinuity (RD) design that draws on the fact that students on either side of a cutoff for admission to a specific school have similar academic ability and preferences but could nonetheless get assigned to schools of differing quality. Thus, I compare the academic outcomes of students on opposite sides of an admission cutoff as a measure of the effects of gaining admission to a given school.

These two research designs yield reassuringly similar results despite the fact that they rely on different identifying assumptions, each with their own limitations. The selection on observables approach generates estimates for a more general population, but relies on the strong assumption that student preferences are a sufficient control for endogenous sorting into schools. The RD design relies on more credibly exogenous variation in school assignment but only identifies effects for students near a threshold for admission to a given school, so does not readily allow researchers to generalize about effects for students in the rest of the population. Both designs are commonly used in studies on the effects of school quality – for example, Dale and Krueger (2002) and Altonji, Elder, and Taber (2005) use selection on observables approaches to estimate the effects of attending selective colleges and Catholic schools in the United States; and Lee and Lemieux (2010) include several school quality studies in their review on the use of RD designs in the field of economics. However, researchers typically use a single approach for a given study, which limits their ability to draw conclusions about external validity (a notable exception is Jackson (2010) who uses multiple approaches in his analysis of secondary school admission effects in Trinidad and Tobago).

Using both approaches, I find that students admitted to selective schools generally receive better matches, as demonstrated by their increased likelihood of staying in their officially assigned school. There is a large amount of noncompliance with school assignment in Ghana and under 60 percent of students who complete secondary school do so at the school they were initially assigned to attend.[2] Nonetheless, admission to a selective school significantly reduces the likelihood that a student switches schools. Students are 40 percent more likely to stay in the same school until the end of secondary for every one standard deviation increase

[2]The extent of compliance with school assignments varies substantially across school choice systems. For example, Pop-Eleches and Urquiola (2013) report that the incidence of switching is rare in Romania, while Jackson (2010) finds that under 60 percent of students in Trinidad and Tobago take their secondary school certification exam in the school to which they were initially assigned, similar to the case in Ghana. The majority of noncompliant students in Ghana shift to schools that are equally as or less selective than their initial assignment (schools to which they would initially have had access).

in the average test scores of their peers. Additionally, they are 10 percent more likely to complete the exit exam at the end of secondary school in the normative time of three years, and 5 percent more likely to complete secondary school at all. However, I find only small effects on overall scores and the number of subjects passed on the secondary school certification exam.

I also find substantial variation in the effects of school assignment, which reinforces the potential importance of match quality. I begin by looking at heterogeneity on the student side and find that certain students benefit more from admission to selective schools than others – female students experience larger increases in the number of core subjects passed, while students with higher baseline exam scores, students from high-performing elementary schools, and students who applied to selective secondary schools experience larger improvements in exam scores but smaller effects on pass rates and retention.

On the school side, I look at heterogeneity based on school attributes and evaluate what types of schools excel in delivering specific outcomes by estimating the correlation between school characteristics and the residuals from a series of value-added regressions. I find that the characteristics that influence student retention differ from those that influence exam performance. Schools with boarding facilities, a higher share of female teachers, and better teacher qualifications produce higher-than predicted gains in completion rates but not in exam performance; female only schools disproportionately increase exam performance but do not increase completion rates; and public schools have high value-added effects on exam performance as well as on the likelihood that students comply with their initial school assignment. Altogether, these results suggest that multiple dimensions of student and school characteristics play a role in determining student performance.[3]

These results have two main policy implications: First, if the effects of school quality are heterogeneous and low on average, then providing generic information on school quality is not likely to enable students or parents to make individually-optimal choices. Several studies have examined demand-side responses to the provision of school quality information, with mixed results: Hastings and Weinstein (2008) and Andrabi, Das, and Khwaja (2013) find positive effects in Chicago and Pakistan respectively; however, Banerjee, Banerji, Duflo, Glennerster, and Khemani (2010) and Mizala and Urquiola (2013) find no effects of information provision in India and Chile. Demand-side responses are important to the extent that they incentivize

[3]These findings echo results from Duflo, Dupas, and Kremer (2011), who conduct a randomized evaluation of an ability-based tracking program in Kenya. They find direct benefits to being in a class with high-performing peers (a peer quality effect) but also find positive impacts of being tracked into a class with peers of similar ability (a match quality effect). Cullen, Jacob, and Levitt (2005) and Hastings, Kane, and Staiger (2008) similarly examine the impacts of match quality in their studies on the effects of school choice in Chicago and North Carolina.

schools to improve their quality (Hoxby (2003)). My results suggest that rather than simply providing information on schools with the best academic performance or even those with the highest value added, stakeholders might generate larger responses by accounting for differences in match quality and by tailoring information to a given student type. For instance, parents may be more likely to respond to information about the highest performing school for girls versus boys, or for high versus low achieving students; the best performing public or single sex school; or the best school for increasing completion rates versus the best for improving test scores.

Second, heterogeneity in school quality effects suggests that offering students opportunities to opt out of their initial assignments or facilitating post-enrollment switches could alleviate the risks of receiving poor matches in settings with limited information. Schools that may appear to be a good fit for a given student based on limited information ex ante could turn out to be a bad match ex post. In a system with some degree of fluidity in school enrollment, students with bad matches may be able to respond by transferring to a different school in order to complete their studies. In the case of an extremely rigid system where there is little room for transfers, students may be more likely to drop out of school altogether if faced with an unfavorable assignment.[4]

Finally, this study also relates to literature on the non-cognitive effects of school quality. Several studies exploiting random and quasi-experimental variation in school assignment document significant impacts on non-cognitive indicators such as dropout, engagement in criminal activity, and labor market earnings, even in cases where authors find only marginal improvements in test scores (see for example Card and Krueger (1992), Gould, Lavy, and Paserman (2004), Cullen, Jacob, and Levitt (2006), Lavy (2010), Deming (2011), and Chetty, Friedman, Hilger, Saez, Schanzenbach, and Yagan (2011)). If securing a better match enables students to switch schools less frequently and increases students' satisfaction with their education, then this could be another channel through which access to higher quality schools improves students' future outcomes. Ultimately, these results suggest that shifting the current school quality debate towards a greater emphasis on the importance of match quality could provide valuable insights into understanding how schools affect student outcomes.

The remainder of the paper is structured as follows: Section 2 presents an institutional background on Ghana's secondary education system. Section 3 describes the data used in this study. Section 4 outlines the main empirical strategy, based on a selection on observables approach. Section 5 provides an alternative set of estimates using a discontinuity design.

[4]Indeed, de Janvry, Dustan, and Sadoulet (2013) find that admission to elite high schools increased dropout rates for marginally admitted students in Mexico when the prospects of transferring after assignment to an elite school were low.

Section 6 discusses students' school switching behavior. Section 7 explores the link between school quality and observable school characteristics, and Section 8 concludes.

2 Institutional Background

Ghana's centralized secondary school admission process provides a systematic source of variation in school assignment. Each year, students completing junior high school (JHS) compete for admission to secondary or senior high school (SHS), where they proceed to their tenth year in Ghana's 6-3-3 sequence of primary, junior high, and secondary school.[5] Ghana adopted a computerized school selection and placement system (CSSPS) in 2005 to increase the transparency and efficiency of this secondary school transition process. In the first year of the new system, students submitted a ranked list of up to three secondary school choices, with each choice consisting of a school and a program track offered in that school.[6] Students then took the Basic Education Certification Exam (BECE) at the end of JHS. Finally, students who performed well enough on the BECE to qualify for admission to secondary school were assigned to a program, with priority based on their BECE performance.

The CSSPS assigns qualified students to secondary school programs through a deferred acceptance algorithm that proceeds in the following way:

- *Round 1*: Each student applies to the first program on her submitted list of ranked choices. Each school has a pre-specified number of vacancies available in each program track offered and tentatively accepts applicants one at a time in order of their aggregate BECE score (a raw score out of 600). It rejects remaining applicants once all of its vacancies are tentatively filled.

- *Round k*: Each student who has been rejected in the previous round ($k-1$) applies to the next program on her submitted list. Each program then compares the set of applicants it has already tentatively accepted to the set of new applicants and again tentatively fills its vacancies by considering all applicants one at a time in order of their BECE performance and rejects remaining applicants once all of its vacancies have been tentatively filled.[7]

[5]The first nine years of schooling constitute Ghana's free and compulsory basic education and the final three years are considered secondary education. I therefore use the terms secondary school and senior high school interchangeably throughout the remainder of this paper.

[6]Available tracks included Agriculture, Business, General Arts, General Science, Home Economics, Technical Studies, Visual Arts, or a technical or vocational program. Students could pick from over 2,300 program options from any of the 650 secondary schools in the country.

[7]The deferred acceptance feature of this assignment mechanism means that student scores ultimately determine admission priorities. For example, a student who lists a program as her second choice could

- *Final round*: The algorithm terminates when no vacancies remain in any of the programs selected by students who have been rejected in the previous round (i.e., when these students have been rejected from all three of the choices on their list). At this point, each student who has been tentatively admitted to a program is placed in that program as the student's final assignment.

Students who qualified for admission to secondary school but were rejected from all three of their chosen programs were given an opportunity to select one of the 110 districts and one of the 10 regions as a preferred choice and were again assigned in order of merit to schools with remaining vacancies in a given district or region. At this point in the placement process, there was no guarantee about the characteristics or quality of the schools to which students would be admitted, but the available schools tended to have fewer resources and to consist of lower-performing students since they were undersubscribed in the initial assignment rounds.

In sum, Ghana's computerized school selection and placement system generated discontinuous changes in student assignments once a program reached its capacity. Two students who had the same list of ranked choices could be assigned to schools of very different quality simply by virtue of having a one point difference in their BECE scores.

Despite this systematic initial assignment process, student placements are weakly enforced. Students received their placement outcomes in September and were supposed to enroll at the school where they were admitted in October. Not all students complied with their official assignments, however. Schools could unofficially admit students at their discretion if they had any unfilled vacancies resulting from admitted students failing to enroll or from under subscription in the official placement process. Regardless of the school they ultimately attend, students sit the same Secondary School Certification Exam (SSCE) at the end of secondary school. Both the BECE and SSCE are centrally administered by the West African Examinations Council (WAEC), so student performance is comparable across schools.

3 Data

To study the effect of school assignment on academic performance, I draw on a linked dataset that allows me to track the secondary school performance of students who completed the BECE and qualified for admission to secondary school in 2005. My analysis incorporates two main sources of data: i) CSSPS administrative data on student characteristics, application choices, BECE scores, and admission outcomes, for all applicants to secondary school; and

displace a student who has a lower score but listed that same option as her first choice and was tentatively accepted to that program in an earlier round.

ii) WAEC examination results which report student performance on the Secondary School Certification Exam (SSCE) at the end of secondary school. The data cover the universe of CSSPS applicants to secondary schools in Ghana in 2005 and the universe of students who took the Secondary School Certification Exam in 2008 and 2009.

Senior high school is three years long, so students who entered SHS in 2005 should sit the SSCE in 2008 if they completed SHS in normative time. The dataset incorporates SSCE information on students from the 2008 and 2009 cohorts to allow students who entered secondary school in 2005 to complete in the normative time of three years or to delay secondary school completion by one additional year. Students do not have a unique identifier so individuals are linked across the CSSPS and WAEC samples using their name, sex, and date of birth. The final dataset links BECE students to their SSCE results for 72 percent of the 2005 cohort of applicants. Aggregate data on the total number of BECE and SSCE candidates suggest that 80 percent of secondary school admits complete the SSCE. Thus, my linked sample covers approximately 90 percent of the target population.

Table 1 presents a set of summary statistics. Altogether, 160,377 students were eligible for admission to secondary school in 2005. I drop students who had taken the BECE in earlier years and deferred their application to secondary school because I do not observe their BECE scores or the ranked list of choices they submitted. I use the remaining sample of 159,607 students for my analysis. Out of this set of students, 69.4 percent were admitted to one of their three ranked choices through the deferred acceptance algorithm. Another 19.8 percent of students were assigned to an undersubscribed school in their preferred district, and the remaining 10.8 percent of students were admitted to an undersubscribed school in their preferred region. BECE scores range from 162 to 477 out of a maximum of 600. I standardize them to have mean 0 and standard deviation 1.

One limitation of the data is that I do not observe BECE scores for the 10.8 percent of students who did not gain admission to one of their three chosen programs or their preferred district but were instead admitted to a program in their preferred region. Due to a data overwrite during the placement process, I simply observe a code indicating that these students were assigned to a program in their region of choice. I do, however, observe secondary school exam scores for all students who complete the SSCE, including those with missing BECE scores. I therefore conduct a series of bounding exercises to ensure that my main results are robust to alternative assumptions about the distribution of missing exam scores.

My analysis focuses on admission to a selective school as the main treatment of interest. School selectivity captures several related factors: i) peer quality, because selective schools admit higher-performing students; ii) perceived school quality, to the extent that student demand reflects subjective beliefs about school quality; and iii) match quality, because students admitted to a higher-ranked choice generally gain admission to a more selective school (as illustrated in Figure 1).[8] Most existing studies have focused on peer quality as the main mechanism through which admission to a selective school affects students outcomes (Abdulkadiroğlu, Angrist, and Pathak (forthcoming), particularly emphasize this channel). I argue that match quality also has important implications.

As a first step in examining the relationship between school selectivity and student performance, I estimate a set of naïve linear regressions of the following form:

$$Y_{ijs} = \alpha Q_s + \beta Q_j + \gamma BECE_i + \boldsymbol{X}_{ij}^{'}\boldsymbol{\delta} + \epsilon_{ijs}$$

where Y_{ijs} is a student's performance in secondary school, Q_s is the mean BECE score of students admitted to the secondary school to which student i was assigned, Q_j is the mean BECE score in a student's junior high school, $BECE_i$ is a student's BECE score at the end of junior high school, and X_{ij} is a vector of student characteristics (age, gender, and an indicator for attending a public JHS). I cluster standard errors at the assigned secondary school level.

The main coefficient of interest, α, has an intent-to-treat interpretation because it captures the correlation between the selectivity of a student's *assigned* secondary school and that student's future outcomes. This addresses potential bias from selection into compliance with school assignments. Nonetheless, this estimate could still be biased by endogeneity in applications to selective schools to begin with. The direction of this bias is not obvious – on one hand, there could be positive selection if students who applied to selective schools have higher intrinsic motivation and individual resources that would lead to high academic performance; on the other hand, there could be negative selection if students who applied to selective schools tended to come from privileged family backgrounds but had lower innate academic ability and were less willing to exert effort studying.

Keeping this potential bias in mind, the results reported in Table 2 suggest that students who gain admission to more selective secondary schools experience little improvement in their academic performance. For every one standard deviation increase in the mean BECE score of assigned secondary school peers, a student is 2.1 percentage points more likely to take the SSCE at the end of secondary school (from a mean of 72.3 percent), 4.4 percentage points more likely to take the SSCE within the normative time of three years (from a mean of 54.5 percent), and 14 percentage points more likely to take the SSCE in her assigned school (from a mean of 42.1 percent). With respect to exam performance, students who take the

[8]Figure 1a illustrates that students rank more selective schools higher up on their list of three choices. Figure 1b indicates that students admitted to higher-ranked choices are exposed to peers of higher ability.

SSCE score 1.141 points higher on the core subjects (on a mean of 13.9 points out of 40), but there is no significant increase in the number of core SSCE subjects passed (the mean is 2.865 subjects out of 4). Altogether, school selectivity appears to have the biggest effect on the likelihood that a student complies with her school placement and graduates from her initially assigned school, but has only marginal effects on overall attainment or exam performance.

Several pre-secondary school characteristics are also correlated with secondary school performance. Interestingly, students from high-performing junior high schools are less likely to take the SSCE and have lower performance on the exam. This suggests that the benefits of attending a high-performing school at the junior high level may be primarily through preparing students to take the standardized entrance exam, and not in fact through increasing students' underlying academic ability. This finding is also consistent with a story of negative selection into selective schools, because students from high-performing junior high schools are more likely to apply to selective secondary schools, even conditional on their individual BECE scores (Ajayi (2013)). Additionally, there are significant gender differences in secondary school performance – boys are less likely to complete secondary school but those who do complete have higher performance on the SSCE than their female peers.

4 Selection on Observables

Although the preceding analysis provides suggestive evidence on the link between school selectivity and student performance, there is still some potential selection bias because students who apply to selective schools may systematically differ from students who do not. The first approach I use to address this potential bias is a selection on observables strategy (similar to that used by Dale and Krueger (2002) and Dale and Krueger (2011)). The basic premise is that students' secondary school applications reveal their innate academic potential. If we assume that students with similar potential apply to a similarly selective set of schools, then two students who submit an identical ranked list of choices will be likely to have similar unobserved characteristics.

4.1 Econometric Model

To implement this approach, I estimate the effect of school selectivity on students outcomes after controlling for student preferences and BECE scores. I use the following specification to evaluate a given education outcome Y_{ibps} for student i with BECE score b and preference

p who was assigned to secondary school s:

$$Y_{ibps} = \alpha Q_s + \beta Q_j + \gamma BECE_i + \boldsymbol{X}'_{ij}\boldsymbol{\delta} + \lambda_p + \nu_{ibps}$$

where Q_s again indicates the mean BECE score of students in the secondary school to which student i was assigned, Q_j is the mean BECE score in a student's junior high school, $BECE_i$ is a student's BECE score, X_{ij} is a vector of student characteristics, and λ_p is a control for student preferences indicated by their ranked list of selected choices. The identifying assumption underlying this approach is that school selectivity is independently assigned after controlling for observed student preferences.

I define student preferences in three successively restrictive ways. First, I estimate a *self-revelation*[9] model that controls for the median selectivity of schools in a students' ranked list of choices. This allows me to retain a sample of 142,366 students with information on the selectivity of schools in their choice list. Second, I estimate a *matched school list* model in which I match students based on their ranked list of schools and restrict the sample to students who submitted a ranked list of school choices that was identical to the list submitted by at least one other student. This yields a sample of 97,402 students, with 18,419 preference listings. I then define λ_p as a fixed effect for each of the preference combinations arising from students' ranked lists of selected schools. Third, I estimate a *matched program list* model in which I match students based on their ranked list of school and program choices. The resulting sample consists of 33,149 observations with 12,068 distinct preferences. This third approach is arguably the most convincing because students are directly matched on their exact lists of ranked choices, but this comes at the cost of excluding a large part of the sample because I exclude all students with unique preference listings.

Each of these three models identifies school assignment effects using a distinct source of variation. To clarify the identification strategies, I briefly outline an example for each design below. In the first model, I compare two students with the same BECE score who applied to a set of three schools with the same average selectivity, but where one student gained admission to the first choice on her list and the other student gained admission to her third choice, for example. This could be the case if one student applied to three schools of average selectivity and the other applied to two highly selective, and one non-selective school. We may still wonder why these students applied to different specific schools, however. In the matched school list model, I compare two students with the same BECE score who applied to the same ranked list of three schools. The variation in assignment here could result from students selecting different programs at each school. For example, the BECE cutoff for

[9] Dale and Krueger (2002) use this term to convey the notion that students' application choices reveal their unobserved academic potential.

admission to the General Science program at the most selective school in 2005 was 404 but was 376 for General Arts. Thus, even if both students had listed this school as their first choice, they could get assigned to different schools if they had listed different programs as their first choice. In the matched program list model, I compare two students who submitted the same combination of program and school choices. This design most closely approximates the RD design discussed in more detail later, and identification primarily comes from students on the margin of admission to a given program. For example, the CSSPS splits ties based on students' scores in the key subjects required for a particular program, so two students on a threshold with an aggregate score of 400 may have different scores on Maths and Integrated Science and so could still be assigned to different schools.

4.2 Main Results

Table 3 reports the main results. Columns (1), (3), and (5) present the baseline regressions for each of the three samples: self-revelation, matched school choices, and matched school and program choices. The even columns present the selection on observables estimates for the three samples: Column (2) controls for the median selectivity of selected schools (self-revelation model); Column (4) includes fixed effects for each ranked list of school choices (matched schools model); and Column (6) includes fixed effects for each ranked list of school and program choices (matched programs model).

Overall, I find that the largest effects of admission to selective schools are on the likelihood of staying in the same school and there are also small but statistically significant effects on the likelihood of timely completion and of taking the SSCE at all. The coefficients in the selection on observables regressions indicate that a one standard deviation increase in the mean BECE scores of secondary school peers in a student's assigned school is associated with an 2.0 to 4.3 percentage point (5 percent) increase in the likelihood of taking the SSCE at all (Panel A) and 4.7 to 7.4 percentage point (10 percent) increase in the likelihood of timely completion (Panel B). Meanwhile, results in Panel C indicate that the same change in admission outcomes implies a 15.9 to 20.2 percentage point increase in the likelihood of staying in the same SHS. Given that 42 to 48 percent of students in these samples take the SSCE in their assigned SHS, these coefficients translate to approximately a 40 percent increase in school retention.

Evidence of effects on SSCE performance is mixed, depending on assumptions about selection into taking the exam. As I mentioned in the data section, I do not have a unique identifier for students so I link BECE candidates from the 2005 cohort to SSCE candidates in 2008 and 2009 using their names and birthdates. I am able to deterministically link 72 percent of students in the sample. I therefore do not observe outcomes for the remaining set of students. I deal with this missing data on outcomes in three main ways. First, I drop students with unobserved SSCE scores from the analysis altogether and calculate the effects using only the linked sample. Second, I assume that students' SSCE scores are missing at random and I assign the median SSCE score to students who I am unable to link. Third, I assume that unobserved students drop out of school and do not complete the SSCE in which case, I assign a score of 0 to students with missing scores and include them in my analysis. These three approaches allow me to construct a set of bounds on the estimated treatment effect.

Panel A in Table 4 focuses on students who complete the SSCE and shows an increase of 0.68 to 1.04 points in SSCE performance on a mean of 14 points out of a maximum of 40. Panels B and C impute missing scores for students who are not observed taking the SSCE, based on two assumptions (assigning the median SSCE score to missing students, or assigning them a 0). Both imputations lead to a larger positive and significant increase in SSCE scores for students assigned to a more selective school, with coefficients ranging from 0.83 to 1.74. In Panels D to F, I repeat the same analysis but with the number of core SSCE subjects passed as my main outcome. The estimated coefficients range from -0.01 to 0.16 on a mean of 2.07 to 2.95 (depending on the sample). Altogether, this analysis suggests that admission to a more selective school may have had a small but significant positive effect on exam performance.

4.3 Heterogeneity

In addition to looking at average outcomes, I also examine whether there is any significant heterogeneity in the estimated effects by looking at variation along four dimensions: sex, individual BECE performance, junior high school performance, and preferences for school quality. Several studies on the effects of school assignment have found differences by sex. In general, girls tend to be more likely to benefit from improvements in schooling environments. In line with this, I find some evidence that girls are more likely to stay in the same school when admitted to a selective school (Column 3 in Table A.1) and more likely to have an increase in the number of SSCE subjects passed (Columns 7 to 9); however, there are no significant differences in effects on overall completion rates (Columns 1 and 2) or on SSCE scores (Columns 4 to 6).

Next, I look for differences based on students' individual BECE performance (Table A.2). I find that there are smaller effects on school retention and the number of subjects passed for higher achieving students, but larger effects on SSCE scores. These results suggest that there are some differential effects of school selectivity along the intensive and extensive margins. Given that high achieving students are not likely to be on the margin of completing secondary school or passing the core subjects, it is not surprising that they experience the largest gains in their exam scores. In contrast, lower performing students who might be in danger of not passing core exams or taking the SSCE at all experience larger gains along these extensive margins.

I also find that effects vary based on the quality of students' previous academic environments. The direction of expected heterogeneity is ambiguous. On one hand, we may expect that students who are coming from less-privileged backgrounds and who attended lower-performing junior high schools would be more likely to benefit from admission to selective schools. On the other hand, students from lower-performing JHSs might be underprepared and may struggle to keep up in a more competitive environment. My estimates provide support for both hypotheses. Table A.3 reports the results. As with individual BECE scores, I find smaller effects on school retention and pass rates but larger effects on exam scores for students who came from junior high schools with above median performance on the BECE.

Lastly, I look for differences in effects based on whether students had a preference for attending an elite school or selective school. I define elite schools as the 22 schools that were established under the British colonial administration before Ghana gained independence in 1957 and are classified in Category A under a categorization scheme created by Ghana Education Service. These are the oldest schools in the country and they are more selective than the average school. I split the sample based on whether students applied to at least one elite senior high school. Students who applied to elite schools applied to schools that had a median incoming BECE score of 330, while students who did not apply to elite schools chose schools with a median incoming BECE score of 285. Once again, I find that students who applied to elite schools had smaller increases in the likelihood of taking the SSCE on time or at all and in pass rates, but larger improvements in exam performance (Table A.4). A similar pattern emerges if I split the sample based on whether students applied to a portfolio of schools that had above median selectivity (Table A.5).

Altogether, this selection on observables analysis yields larger estimated effects of school selectivity on retention and generally yields smaller estimated effects on SSCE performance. Moreover, the estimated retention effects are larger for female students than for males, while effects on exam scores are larger for high performing students and those who attended and applied to higher performing schools.

5 Discontinuity Design

To further examine the effects of school assignment, I use a similar discontinuity design to what is commonly used in existing literature on this topic. The identifying assumption underlying this analysis is that two students with similar initial test scores who applied to the same school would have had similar outcomes later in their academic careers if they had both been on the same side of an admission cutoff. If instead, one student gains admission to their selected choice while the other student is narrowly rejected and assigned to a less selective alternative school then we can attribute differences in their future outcomes to the differences in their school assignments.

Although this identification strategy is cleaner and more compelling, it comes with the cost of larger data requirements and a reduced sample size. Most importantly, a researcher must know the admission cutoffs for each program and be able to observe where individual students lie in the distribution of exam scores relative to these cutoffs in order to restrict the sample to the region around the cutoffs.

For the purposes of the subsequent analysis, I limit the linked CSSPS sample to focus on students who applied to a program where the capacity constraint was binding (i.e., where the number of applications exceeded the number of available spaces). This allows me to examine cases in which there was a discontinuous change in admission outcomes for students with similar BECE scores on either side of the admission cutoff. These selection criteria result in a sample of 201,610 observations of students applying to a total of 576 programs (students are included multiple times if they did not gain admission to their first choice school and ended up applying to lower-ranked choices). The sample shrinks to 43,784 observations when I focus on students with BECE scores within 20 points of an admission cutoff. Table B.1 presents summary statistics on my RD sample.

5.1 Econometric Model

I use a two stage least squares estimator to evaluate the effect of school assignment on students' academic performance. The instrumental variables specification for this analysis is:

$$\begin{aligned} Q_{is} &= \gamma 1\{\theta_i \geq \underline{\theta}_p\} + a(\theta_i) + \eta_i && (1) \\ Y_{is} &= \delta \mathrm{E}\left(Q_{is} \mid \theta_i\right) + a(\theta_i) + \mu_i && (2) \end{aligned}$$

where $1\{\theta_i \geq \underline{\theta}_p\}$ is an indicator for whether student i's score exceeds the cutoff for admission to program p in school s and $a(\theta_i)$ is a control for a linear function of exam scores. In the first stage, I estimate the effect of the admission rule on students' school assignment, Q_{is}. In particular, I examine how much the selectivity of a student's assigned secondary school increases as a result of scoring above the cutoff for admission. The second stage of my analysis estimates the effect of this initial assignment on students' future academic outcomes Y_i. Since the data are pooled across several programs, I include cutoff fixed effects and estimate robust standard errors clustered at the program level.

5.2 Data Considerations

Altogether, the administrative data pose two challenges for using a discontinuity design to study the effect of school placement on academic performance: 1) unobserved admission cutoffs; and 2) missing BECE scores for students who receive a regional placement (10.8 percent of the sample). I discuss these issues and empirical strategies for addressing them in the remainder of this section.

In order to pool the data for applicants across the individual programs, I redefine students' BECE scores relative to the admission cutoff for the program they are applying to such that the marginal admitted student has a score of 0. Identifying the appropriate cutoff to use for each program is somewhat challenging because I do not explicitly observe the cutoffs in my data. I deal with this in two ways – my preferred approach is to replicate the placement procedure described in the CSSPS handbook and retrieve the cutoffs that would have emerged if the procedure had been implemented strictly. I am able to replicate admission outcomes for 90% of students. The error likely arises from the splitting of ties (I split ties arbitrarily, while the actual implementation split ties using subject-specific scores which I do not observe), and from schools which do not report vacancies. I exclude these schools from my analysis, however the CSSPS administrators may have dealt with missing vacancies in a different manner.

As a robustness check, I also perform my analysis using the lowest observed score of students admitted to each program as the cutoff. This approach assumes that the deferred acceptance assignment algorithm was implemented accurately such that the observed cutoffs are the actual cutoffs that would obtain under strict assignment. The challenge of using a discontinuity design in a context where cutoffs are not directly observed as well as the issue of noncompliance have been addressed in other studies (see Jackson (2010) for example). My empirical procedures are in line with existing literature on this topic.

A second data-related factor to note is that I am missing baseline (BECE) scores for about 10 percent of students in the 2005 cohort who did not gain admission to one of their chosen schools and were instead assigned to a school in their preferred region with remaining spaces.

I do not observe BECE scores for these students but instead observe a code indicating that they received a regional placement. I am, however, still able to observe outcome variables for these students in cases where they complete secondary school and take the SSCE. I therefore use a bounding exercise to check the sensitivity of my analysis to alternative assumptions about the missing data.

5.3 Main Results

As a first test of the validity of the RD identification assumptions, Figure 2 displays the distribution of student BECE scores around the pooled cutoffs for admission to oversubscribed programs. The figure illustrates that there is a discontinuous decrease in the density of students below the admission cutoffs. This change in density suggests manipulation of the running variable which is students' relative BECE scores or nonrandom attrition from the sample (McCrary (2008)). Indeed, as mentioned earlier, 10 percent of students received a regional placement. Figure 3 provdes additional evidence of nonrandom attrition because there is a discontinuous change in the performance of junior high schools attended for students who are on either side of the admission cutoffs. This suggests that the assumption of exogenous assignment is not valid since students across the thresholds do not appear to have similar observable characteristics.

Table B.2 formally estimates the extent to which the assumption of exogenous assignment is violated, by regressing student characteristics on an indicator for scoring above the admission cutoff. This analysis confirms that there is a significant increase in the performance of junior high schools attended, although I do not find any significant differences in age, the proportion of male students or likelihood attending a public junior high school for students on either side of the admission cutoffs. I first attempt to address this issue by controlling for observable student characteristics as well as by excluding students who are within 5 points of the cutoff based on the notion that selection of students may be more severe around the admission cutoff and that my simulation of the assignment system may have lead to some error in correctly identifying the actual locations of admission cutoffs.

Figure 4 illustrates the discontinuous change in admission chances for students on opposite sides of the threshold for admission to an oversubscribed program. In particular, students experience a 74 percentage point increase in their admission chances if their test scores are above the admission threshold. This jump rises to an 88 percentage point boost in admission chances if I exclude the students within 5 points of the cutoff. This increase in admission chances translates into improved school quality - students above the threshold are assigned to secondary schools with peers who have average BECE scores that are 0.65

standard deviations higher than peers for students below the threshold. Moreover, their assigned schools have a 10 percentage point higher pass rate on the SSCE in 2008 (Table B.3).

In addition to these significant differences in exposure to school quality, there is some evidence that assignment to a more selective school has lasting effects on students' academic performance. Figure 5 displays the differences in students' academic attainment and performance on the SSCE. Specifically, I find that students who score just below the admission cutoffs appear to be less likely to comply with their school assignment and more likely to transfer into a different school by the time they complete secondary school. These new schools are generally not closer to a student's junior high school but are better performing than their initially assigned schools (Figure 6). However, among the students who do complete senior high school, there are only marginal differences in academic performance.

Tables 5 and 6 present regression estimates. A one standard deviation increase in BECE scores of assigned peers increases the likelihood of taking the SSCE by 4.7 percentage points, the likelihood of taking the SSCE in three years by 9.6 percentage points, and the likelihood of taking the SSCE in a students' assigned school by 22.7 percentage points. All of these estimated coefficients are again larger than estimates from a naïve regression. Estimates of the effects on exam performance are not significantly different from the comparable ordinary least squares estimates. A one standard deviation increase in BECE scores of assigned peers increases SSCE scores by 0.732 to 1.086 points (on a mean of 12.499 to 15.953) and increases the number of core subjects passed by 0.037 to 0.158 (on a mean of 2.408 to 3.074).

5.4 Bounding and Robustness Checks

To bound the effects of sample selection on the baseline estimates, I perform a series of robustness checks and examine whether unobserved selection in the likelihood of receiving a regional placement (and thus having no BECE score reported in the data) can explain the initial results. To begin, I assume that BECE scores follow a normal distribution. In particular, I simulate a normal distribution using the density of students to the right hand side of each admission cutoff and at the low end of the left tail as moment conditions. This effectively assumes that there is no selective attrition to the right hand of the cutoff and at the very low end of the left tail, so all of the selection occurs in the region within 75 points below the admission cutoff. I then fit a normal distribution which satisfies these conditions and calculate the density of missing students given the disparity between the expected and observed density of students in the middle-range of the distribution (Figure 7 illustrates the result of this exercise). Finally, I assign BECE scores to students who are missing scores based on the distributions that would be observed under three scenarios: 1) students' outcomes are unrelated to their initial BECE scores (i.e., no selection); 2) students with the best secondary school outcomes had the highest BECE scores (lower bound); and 3) students with the best secondary school outcomes had the lowest BECE scores (upper bound). I then run the same regression analysis presented earlier under each of these scenarios.

Tables B.4 and B.5 report the results of this analysis. The effects of school selectivity on staying in the same secondary school remain positive and significant under all specifications, although the effects on timely progression and on taking the SSCE at all are not robust to alternative assumptions about selection into regional placements. Similarly, the effects on SSCE scores are sensitive to alternative assumptions about the distributions of missing baseline BECE scores. Estimates for effects on SSCE scores range from -2.021 to 4.856 percentage points, and for SSCE passes from -0.231 to 0.795.

As an additional robustness check on my baseline estimates in Tables 5 and 6, I also estimate the same OLS and IV regressions using the observed minimum scores of students admitted to each school as the admission cutoff, instead of imputing the cutoffs using the official admission rule. The estimated coefficients are qualitatively similar (results available on request). I also perform the same analysis using different approaches to link BECE candidates in 2005 with the SSCE candidates in 2008 and 2009 (namely, using a probabilistic matching technique to identify potential matches). The results are robust to these alternative approaches as well.

6 Compliance with School Assignment

A key result that emerges from the preceding analysis is that admission to a selective school in Ghana primarily influences the likelihood of attending a particular school, rather than affecting whether or not students ultimately complete secondary school at all. Across all specifications, the effects on taking the SSCE in a student's initially assigned school are at least four times as large as the effects on overall SSCE completion and are also much larger than effects on SSCE scores. This finding calls for a deeper understanding of what types of students tend to change schools and where they end up moving to. Only 58 percent of students who take the SSCE at the end of secondary school do so in the school to which they were initially admitted. This means that 42 percent of students who complete secondary school do not comply with their initial secondary school assignment. In this section, I examine the factors that determine the likelihood of changing schools and document the characteristics of new schools compared to initial assignments.

Figure 6 illustrates the basic patterns in characteristics of students who move and the schools to which they move, within the discontinuity design framework. Several observations stand out. First, students who are assigned to less-selective and lower-performing schools on either side of the cutoffs are more likely to move (Panels A and B). Second, students assigned to schools outside their JHS district are more likely to move (Panels C and D). Third, there are opposite movements on either side of the cutoffs in terms of school quality – students above the thresholds tend to move to less selective and lower-performing schools, while students below the thresholds tend to move to equally as selective, but slightly higher-performing schools. Moreover, students tend to move to schools that are further away than their initial assignments, and this is particularly the case for students assigned to more selective schools. The overall effect of these noncompliance decisions is that the gap in quality of schools actually attended by students on either side of the thresholds decreases, and on average, students tend to complete secondary school at a school that is further away than their initially assigned school. (See transition matrices in Appendix C for more details.)

At a larger level, this analysis speaks to the role of re-optimization behavior in potentially explaining cross-study differences in estimates of the effects of school assignment. It suggests that differences in the ease of noncompliance with initial school assignments in a given school system may help to explain differences in estimated effects of school assignment on student performance across different contexts.

7 Correlates of School Quality

Up to this point, my analysis has focused on changes in student performance resulting from assignment to schools of differing selectivity levels. The merit-based nature of Ghana's application system makes school selectivity a natural indicator of school quality, however, this measure has limited meaning in and of itself. To deepen our understanding of the mechanisms through which school assignment might affect student retention and exam performance, I examine the link between school performance and observable school characteristics (including facilities, pupil-teacher ratios, and teacher qualifications). Here, I draw on data from the annual school censuses conducted by Ghana Education Service for their Education Management Information System (EMIS). The EMIS data provide information on a range of school attributes (Table D.1 in the appendix lists the summary statistics).

To estimate the explanatory power of individual school characteristics, I first calculate a set of school-level average residuals from the self-revelation regressions in Column 2 of Tables 3 and 4, and take these as a measure of school value-added. I then regress these school-level residuals on a vector of school characteristics. Table 7 reports the results of this exercise. Altogether, it appears that retention depends on teacher characteristics and school facilities, while exam performance depends on peers. In particular, I find that being assigned to a school with boarding facilities, a higher share of qualified teachers, and a higher share of female teachers are all correlated with higher value-added in terms of the likelihood of taking the SSCE (at all, on time, or in the initially assigned school). Public schools are better at getting students to stay in their assigned schools and at improving student scores on core subjects. Single sex schools tend to generate higher value added on SSCE scores; however, female-only schools are associated with increases in the number of core subjects passed while male-only schools are associated with decreases.

This analysis complements a set of studies in the literature on school quality (including Black and Smith (2006)), that emphasize the importance of accurately measuring school quality, particularly by using multiple proxies. In a closely related paper, Lai, Sadoulet, and de Janvry (2011) find that teacher qualifications explain most of the predictive power of school fixed effects in their study on the effects of school quality on student performance in Beijing. Overall, the fact that different school attributes deliver different academic outcomes reiterates the importance of thinking about match quality instead of viewing school quality as a homogeneous good.

8 Conclusions

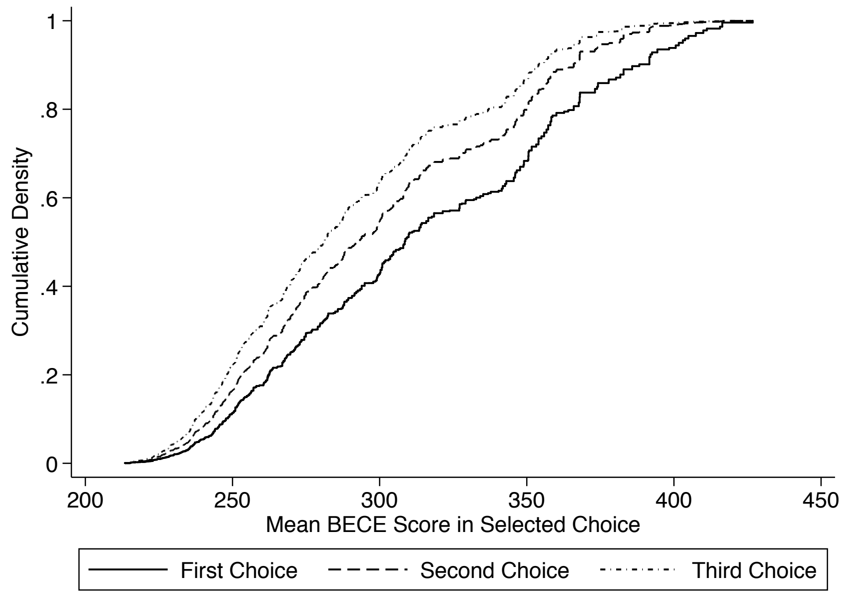
This paper reexamines the effects of school assignment on students' academic performance. I build on a growing number of studies that use selection on observables approaches and RD designs to address concerns about endogenous selection into schools. Existing studies in this literature typically focus on changes in school quality as the main treatment of interest. A central contribution of this paper is to demonstrate that changes in match quality offer an additional mechanism for school assignment to affect student outcomes.

My empirical analysis on secondary school students in Ghana provides two pieces of evidence in support of this alternative interpretation. First, students admitted to more selective schools generally receive better matches but do not experience large gains in academic performance. They are less likely to switch schools and more likely to complete secondary school on time, but there are only marginal effects on overall completion rates and performance on the secondary school certification exam. Second, not all students who gain admission to selective schools benefit from this opportunity to attend a high quality school. There are significant differences in effects by student gender, ability, preferences, and academic background, with some subgroups experiencing negative admission effects and with certain types of schools being more effective and generating specific outcomes than others. The large effects on school retention along with smaller effects on exam performance and substantial heterogeneity collectively highlight the importance of considering match quality as a channel through which admission to selective schools impacts student outcomes.

One limitation of this study is its focus on academic performance as a primary outcome. A striking finding from an array of quasi-experimental studies is the robust relationship between school assignment and improvements in non-cognitive outcomes such as career aspirations and propensity for crime, even in the absence of effects on academic performance. Although I am unable to test this hypothesis directly, my findings point to the possibility that improvements in match quality (and associated reductions in the likelihood of disliking or switching schools) ultimately lead to improvements in non-academic outcomes.

Finally, this study has direct implications for policies aimed at facilitating school choice by providing information on school quality. A fundamental objective of information provision is to incentivize schools to improve their quality by empowering consumers to choose the best performers. However, typical interventions provide information on average school performance and do not differentiate between alternative outcomes (such as test scores versus dropout rates) or particular subgroups of students. Given the likelihood that there is substantial heterogeneity in school assignment effects as well as in preferences for school attributes, parents and students may be rational not to respond to generic information. In contrast, providing information on school performance along multiple dimensions and by student type could enable individuals to choose the best school that matches their needs.

(a) Peer Quality by Rank of Selected Choices



(b) Peer Quality by Rank of Assigned Schools

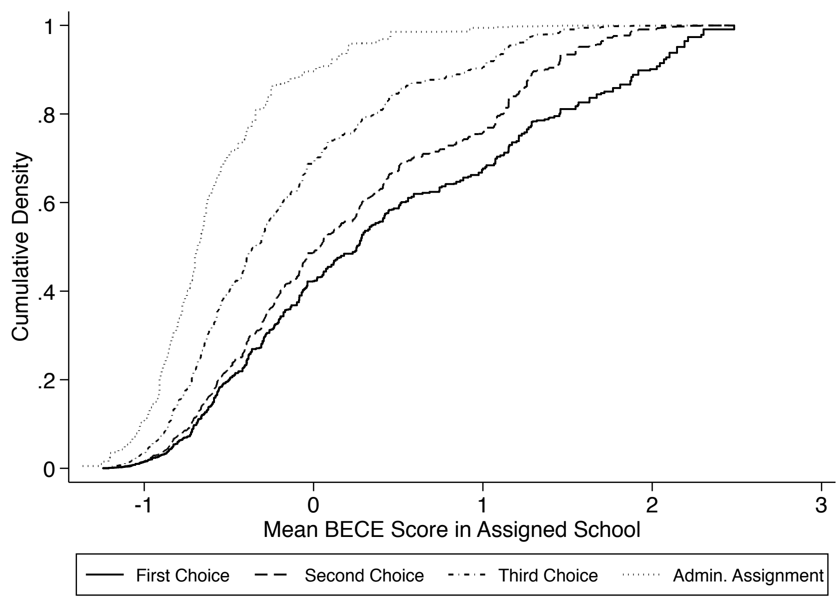


Figure 1: Distribution of Peer Quality by Rank of Selected and Assigned Choices

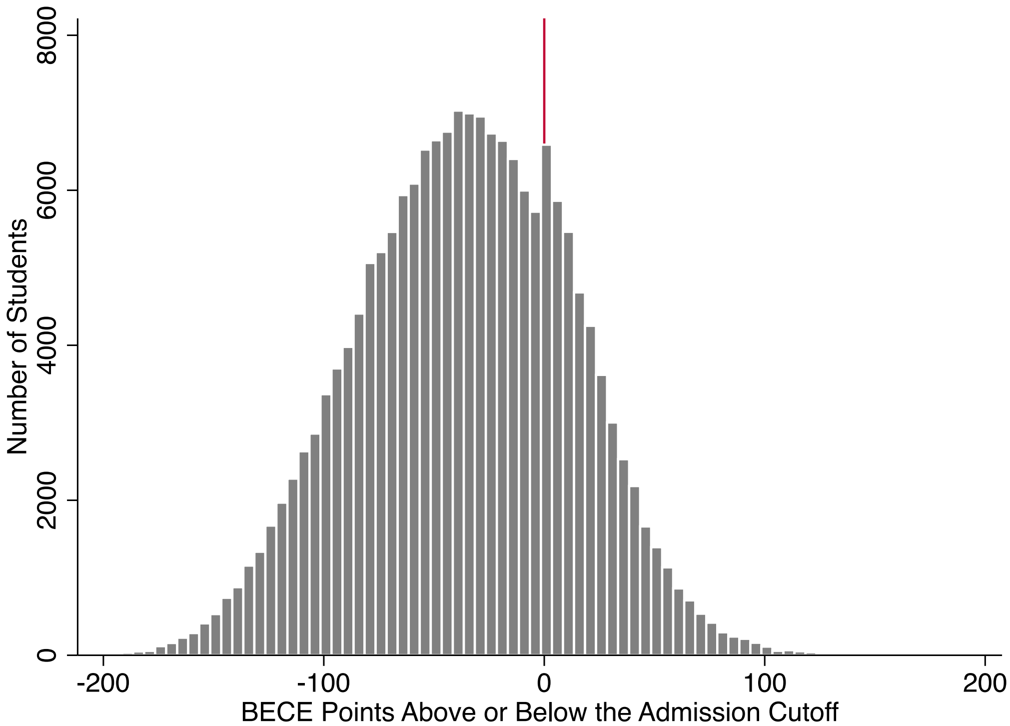
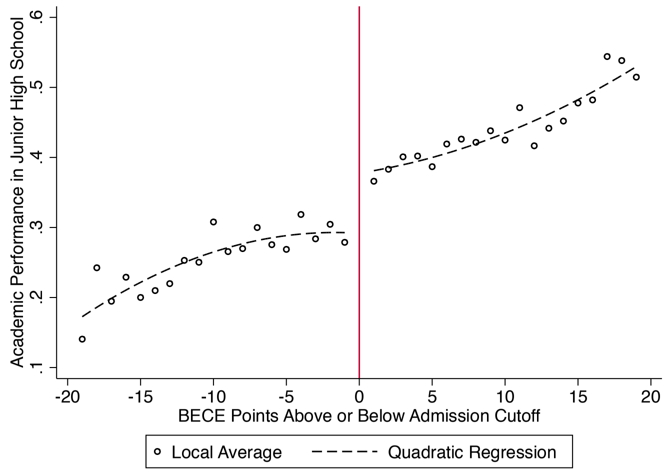
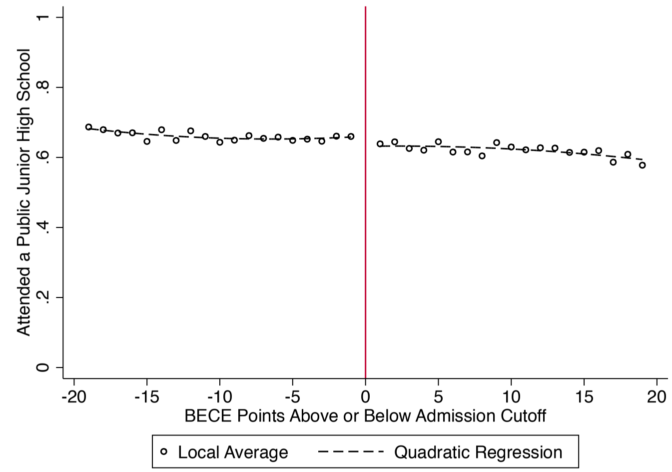


Figure 2: Density of Students around Admission Cutoffs

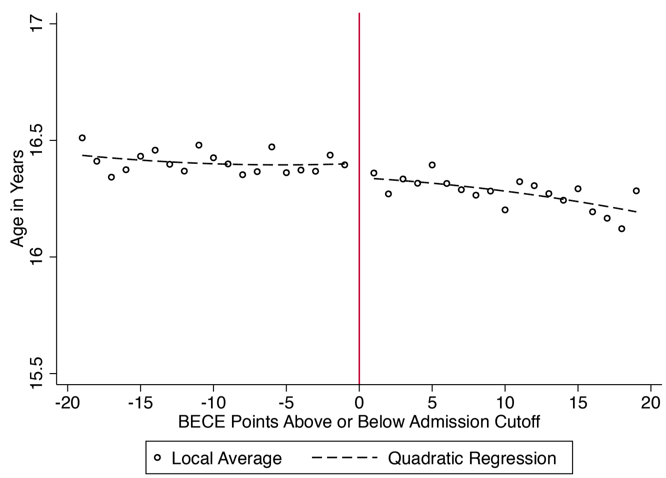
(a) Mean BECE of JHS Peers



(b) Attended a Public JHS



(c) Age in Years



(d) Male

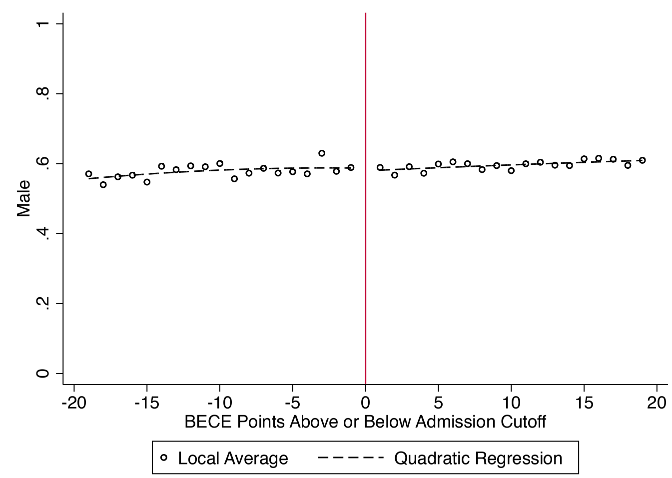
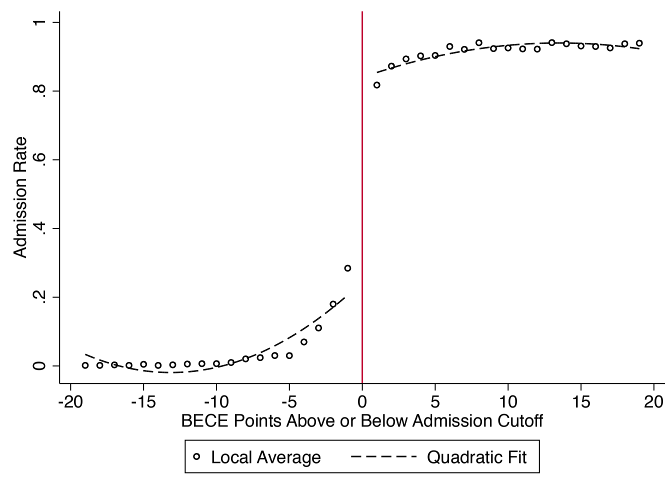
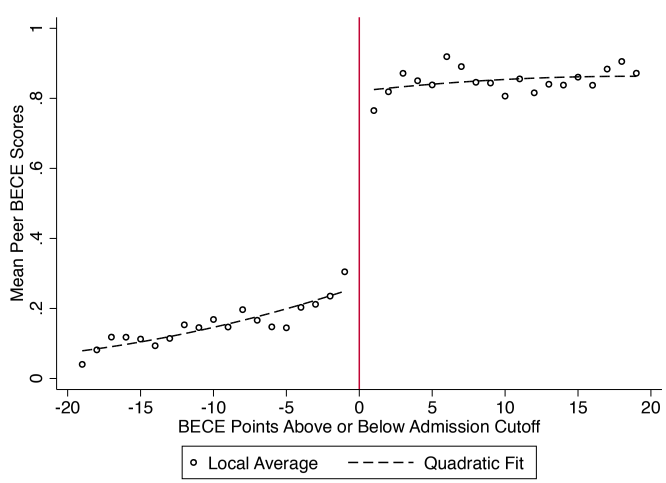


Figure 3: Distribution of Covariates around Admission Cutoffs

(a) First Stage: Admission Probability



(b) First Stage: Mean BECE Score of SHS Peers



(c) First Stage: SSCE Performance of SHS (2008)

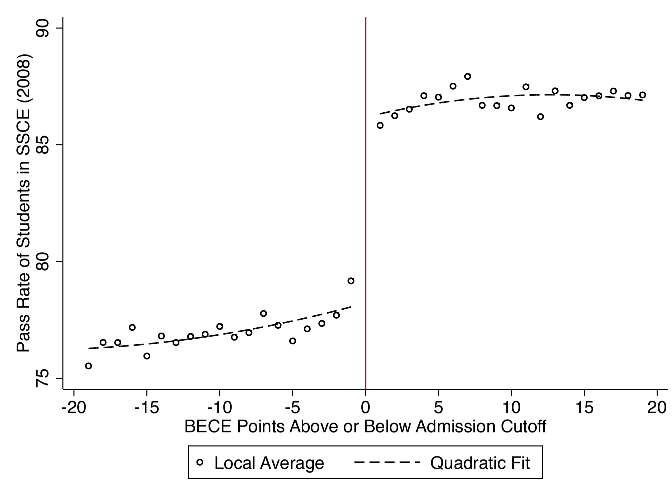
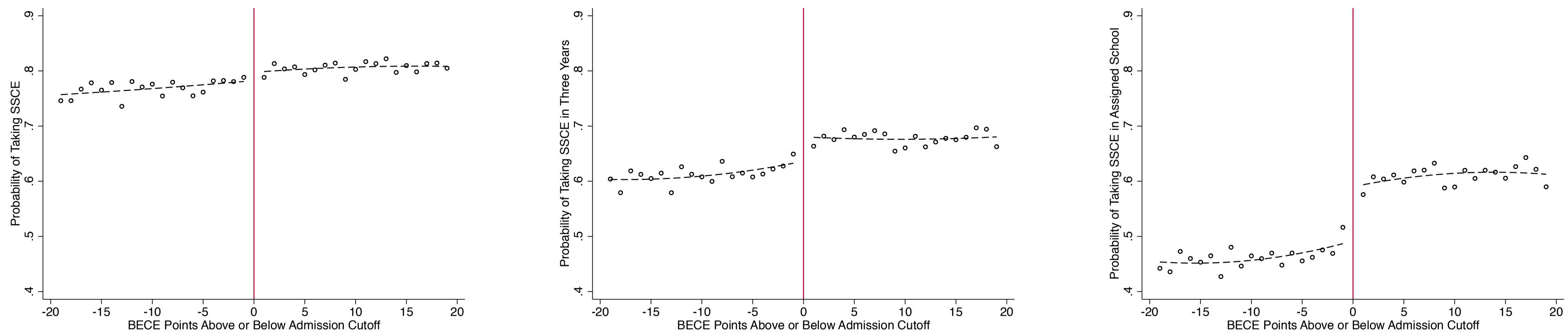


Figure 4: First Stage Effects

(a) School Retention



(b) SSCE Performance

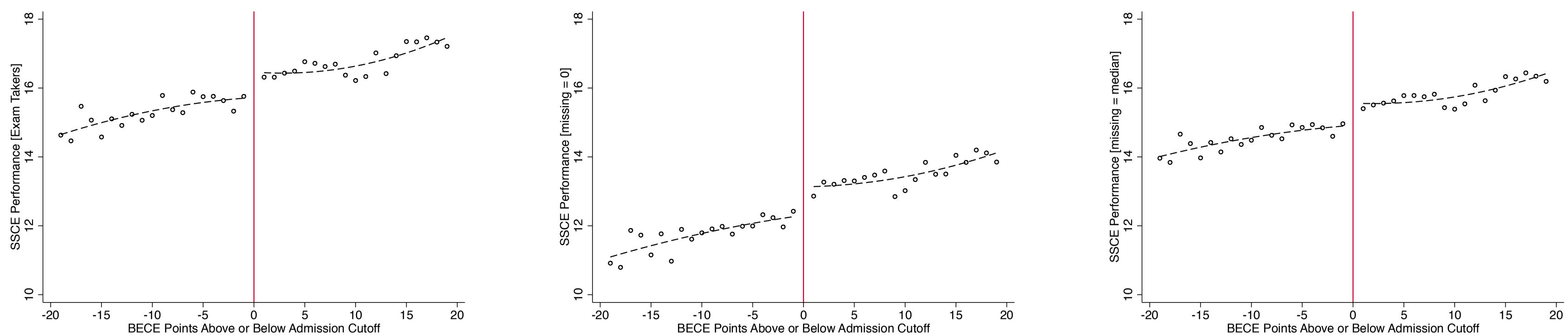
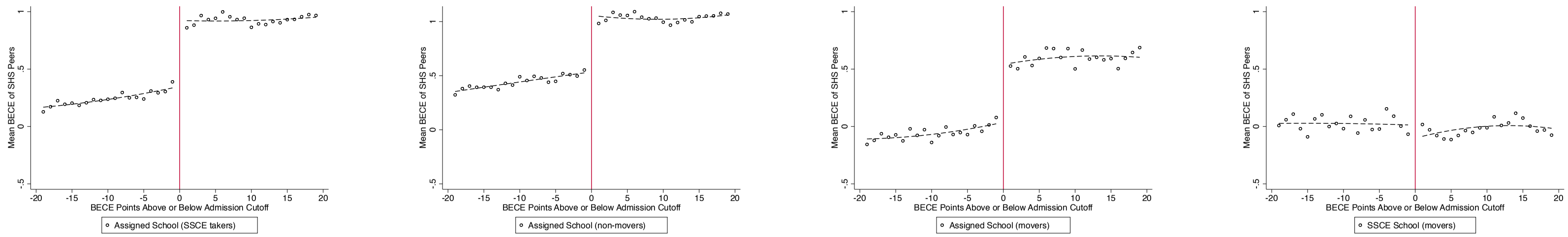
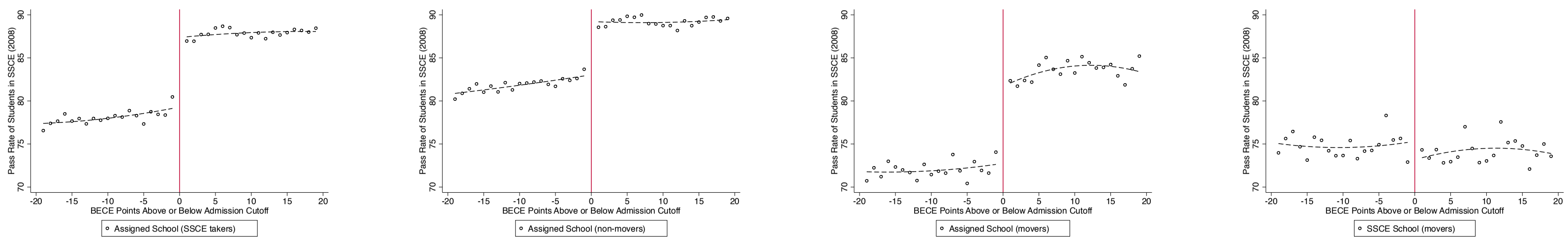


Figure 5: Reduced Form Outcomes

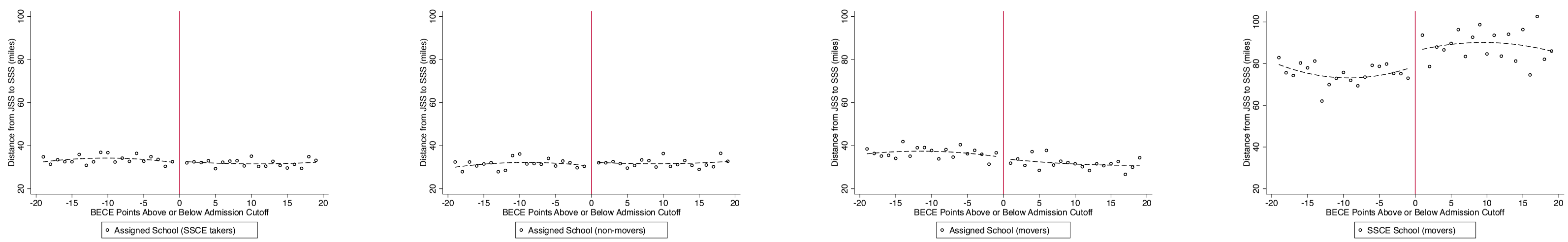
(a) Peer Quality



(b) SSCE Pass Rate



(c) School Distance



(d) Staying in JHS District

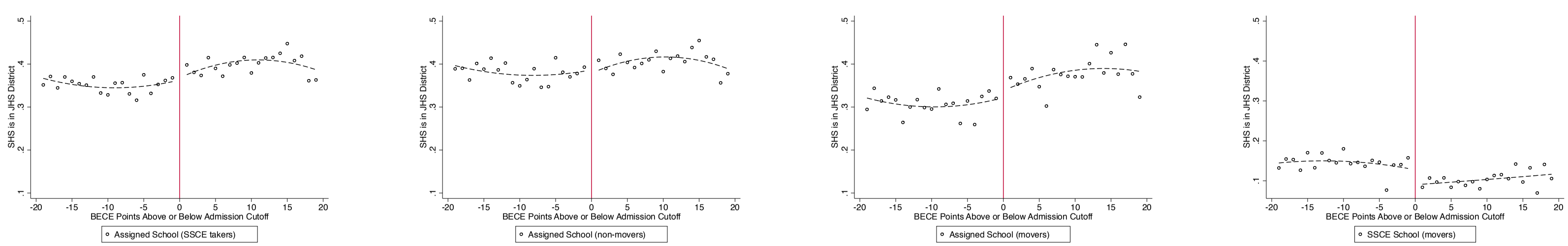


Figure 6: Evidence of Students Moving to Better Schools

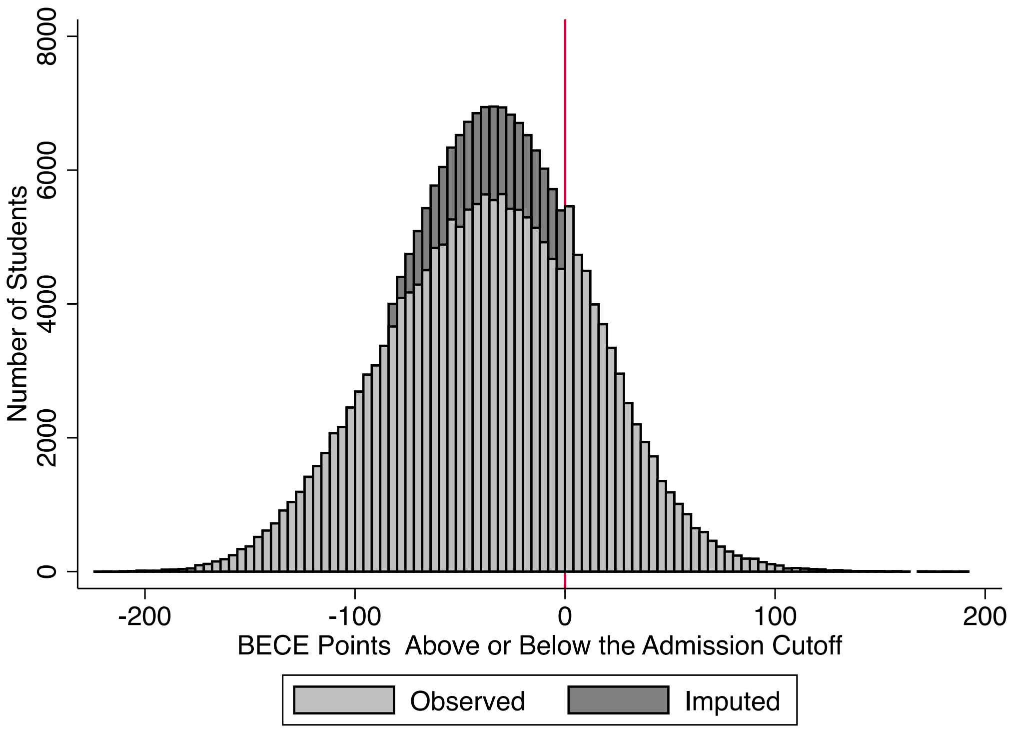


Figure 7: Bounding Selection Effects

Table 1: Summary Statistics

	All	Take the SSCE	Do not take the SSCE
	(1)	(2)	(3)
Student Characteristics			
Age	16.618	16.445	17.081
Male	0.583	0.571	0.616
JHS Public	0.749	0.727	0.809
Standardized BECE score	0.000	0.129	-0.337
Mean BECE of JHS peers	0.000	0.071	-0.187
Mean BECE of SHS peers	0.000	0.101	-0.263
Admission Outcomes			
First choice program	0.274	0.285	0.243
Second choice program	0.187	0.196	0.162
Third choice program	0.233	0.227	0.250
District of choice	0.198	0.179	0.251
Region of choice	0.108	0.113	0.093
Secondary School Performance			
Take SSCE	0.728	1.000	
Take SSCE in three years	0.550	0.756	
Take SSCE in assigned school	0.421	0.579	
SSCE score	9.958	13.687	
SSCE core passes	2.068	2.843	
SSCE total passes	4.114	5.655	
N	159607	116118	43489

Table 2: Descriptive Analysis: School Retention and Exam Performance

	School Retention			Exam Performance	
	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE score	SSCE passes
	(1)	(2)	(3)	(4)	(5)
Mean BECE of SHS peers	0.021	0.044	0.140	1.141	0.000
	(0.004)***	(0.006)***	(0.008)***	(0.134)***	(0.012)
Mean BECE of JHS peers	-0.031	-0.049	-0.084	-2.425	-0.249
	(0.003)***	(0.004)***	(0.005)***	(0.080)***	(0.011)***
Individual BECE	0.082	0.130	0.108	4.395	0.464
	(0.003)***	(0.004)***	(0.004)***	(0.119)***	(0.011)***
Male	-0.026	-0.063	-0.049	1.018	0.150
	(0.004)***	(0.005)***	(0.006)***	(0.123)***	(0.011)***
Age	-0.028	-0.012	-0.008	-0.618	-0.074
	(0.001)***	(0.001)***	(0.002)***	(0.020)***	(0.004)***
JHS Public	-0.006	-0.011	0.010	0.066	-0.013
	(0.003)	(0.005)**	(0.005)*	(0.067)	(0.010)
R^2	0.057	0.088	0.130	0.402	0.125
N	142424	142424	142424	102980	102980
Mean Outcome	0.723	0.545	0.421	13.900	2.865

Notes: Table displays results from a set of OLS regressions with the following dependent variables: (1) an indicator for taking the SSCE exam; (2) an indicator for taking the SSCE in three years; (3) an indicator for taking the SSCE in the initially assigned school; (4) aggregate performance on the core SSCE subjects (English, Maths, Social Science and Integrated Science), letter grades for each subject have been converted to a scale of 1 to 10 and aggregated for a total score out of 40; and (5) the sum of indicators for passing each of the four core SSCE subjects. Robust standard errors are clustered at the senior high school level and reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table 3: Selection on Observables Estimates of Effects on School Retention

	Self-revelation		Matching on school choices		Matching on program choices	
	Baseline	Control for selectivity of choices	Baseline	School list fixed effects	Baseline	Program list fixed effects
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Take SSCE						
Mean BECE of SHS peers	0.021	0.020	0.022	0.028	0.030	0.043
	(0.004)***	(0.004)***	(0.005)***	(0.005)***	(0.007)***	(0.008)***
R^2	0.057	0.057	0.057	0.259	0.060	0.432
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	0.723	0.723	0.736	0.736	0.747	0.747
Panel B. Take SSCE on Time						
Mean BECE of SHS peers	0.044	0.047	0.044	0.060	0.054	0.074
	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.008)***	(0.010)***
R^2	0.088	0.088	0.087	0.294	0.092	0.462
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	0.545	0.545	0.564	0.564	0.584	0.584
Panel C. Take SSCE in Assigned School						
Mean BECE of SHS peers	0.140	0.159	0.146	0.184	0.163	0.202
	(0.008)***	(0.008)***	(0.009)***	(0.008)***	(0.011)***	(0.011)***
R^2	0.130	0.142	0.132	0.359	0.136	0.512
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	0.421	0.421	0.444	0.444	0.471	0.471

Notes: Baseline regressions include controls for student BECE scores, age, gender, average BECE score in JHS, and an indicator for attending a public JHS. Selection on observables regressions additionally include: i) controls for the average selectivity of schools selected (column 2) or ii) fixed effects for student preferences, given by their ranked list of school choices (column 4) and ranked list of program choices (column 6). School selectivity is measured by the mean BECE score of peers in a student's initially assigned school. Robust standard errors are clustered at the level of the assigned SHS and reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table 4: Selection on Observables Estimates of Effects on Exam Performance

	Self-revelation		Matching on school choices		Matching on program choices	
	Baseline	Control for selectivity of choices	Baseline	School list fixed effects	Baseline	Program list fixed effects
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. SSCE Score						
Mean BECE of SHS peers	1.143	1.040	1.037	0.676	1.122	0.931
	(0.134)***	(0.136)***	(0.141)***	(0.115)***	(0.173)***	(0.197)***
R^2	0.402	0.403	0.420	0.593	0.445	0.730
N	102957	102957	71659	71659	24763	24763
Mean Dep. Variable	13.900	13.900	14.198	14.198	14.815	14.815
Panel B. SSCE Score [missing=median]						
Mean BECE of SHS peers	1.177	1.106	1.084	0.829	1.259	1.216
	(0.122)***	(0.123)***	(0.127)***	(0.097)***	(0.155)***	(0.148)***
R^2	0.323	0.324	0.341	0.492	0.369	0.627
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	13.374	13.374	13.617	13.617	14.103	14.103
Panel C. SSCE Score [missing=0]						
Mean BECE of SHS peers	1.434	1.348	1.343	1.168	1.620	1.737
	(0.118)***	(0.117)***	(0.128)***	(0.106)***	(0.170)***	(0.182)***
R^2	0.288	0.289	0.300	0.450	0.321	0.586
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	10.053	10.053	10.445	10.445	11.067	11.067
Panel D. SSCE Passes						
Mean BECE of SHS peers	0.000	-0.009	0.001	-0.008	-0.003	-0.000
	(0.012)	(0.012)	(0.014)	(0.014)	(0.019)	(0.026)
R^2	0.125	0.126	0.128	0.376	0.128	0.562
N	102957	102957	71659	71659	24763	24763
Mean Dep. Variable	2.865	2.865	2.893	2.893	2.937	2.937
Panel E. SSCE Passes [missing=median]						
Mean BECE of SHS peers	0.022	0.016	0.020	0.014	0.022	0.029
	(0.009)**	(0.009)*	(0.010)*	(0.010)	(0.014)	(0.017)*
R^2	0.082	0.083	0.087	0.280	0.090	0.445
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	2.903	2.903	2.921	2.921	2.953	2.953
Panel F. SSCE Passes [missing=0]						
Mean BECE of SHS peers	0.086	0.076	0.085	0.098	0.112	0.159
	(0.016)***	(0.016)***	(0.018)***	(0.017)***	(0.027)***	(0.031)***
R^2	0.131	0.131	0.132	0.317	0.136	0.475
N	142366	142366	97402	97402	33149	33149
Mean Dep. Variable	2.072	2.072	2.128	2.128	2.194	2.194

Notes: same as above.

Table 5: Discontinuity Design Estimates of Effects on School Retention

	OLS	RF	RF	IV	IV
	Within $(-20, 20)$ of cutoff	Within $(-20, 20)$ of cutoff	Excluding $(-5, 5)$	Within $(-20, 20)$ of cutoff	Excluding $(-5, 5)$
Panel A. Take SSCE					
Mean BECE of SHS peers	0.038			0.032	0.047
	(0.005)***			(0.015)**	(0.021)**
1{BECE ≥ cutoff}		0.018	0.031		
		(0.008)**	(0.014)**		
R^2	0.067	0.066	0.070	0.067	0.071
N	43784	43784	34392	43784	34392
Mean Dep. Variable	0.786	0.786	0.783	0.786	0.783
Panel B. Take SSCE on Time					
Mean BECE of SHS peers	0.065			0.086	0.096
	(0.006)***			(0.018)***	(0.024)***
1{BECE ≥ cutoff}		0.048	0.062		
		(0.010)***	(0.015)***		
R^2	0.104	0.102	0.107	0.104	0.108
N	43784	43784	34392	43784	34392
Mean Dep. Variable	0.643	0.643	0.640	0.643	0.640
Panel C. Take SSCE in Assigned School					
Mean BECE of SHS peers	0.179			0.218	0.227
	(0.006)***			(0.019)***	(0.024)***
1{BECE ≥ cutoff}		0.121	0.148		
		(0.010)***	(0.016)***		
R^2	0.162	0.144	0.153	0.162	0.167
N	43784	43784	34392	43784	34392
Mean Dep. Variable	0.531	0.531	0.528	0.531	0.528

Notes: Regressions include cutoff fixed-effects as well as controls for BECE score, BECE score × 1{BECE ≥ cutoff}, gender, and average BECE score in student's JHS. Robust standard errors clustered at the cutoff level are reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table 6: Discontinuity Design Estimates of Effects on Exam Performance

	OLS	RF	RF	IV	IV
	Within $(-20, 20)$ of cutoff	Within $(-20, 20)$ of cutoff	Excluding $(-5, 5)$	Within $(-20, 20)$ of cutoff	Excluding $(-5, 5)$
Panel A. SSCE Score					
Mean BECE of SHS peers	0.604			0.886	0.732
	(0.083)***			(0.235)***	(0.312)**
1{BECE $\geq$ cutoff}		0.502	0.485		
		(0.133)***	(0.206)**		
R^2	0.466	0.465	0.465	0.465	0.466
N	34398	34398	26945	34398	26945
Mean Dep. Variable	15.968	15.968	15.953	15.968	15.953
Panel B. SSCE Score [missing=0]					
Mean BECE of SHS peers	1.078			1.116	1.086
	(0.100)***			(0.291)***	(0.415)***
1{BECE $\geq$ cutoff}		0.620	0.706		
		(0.161)***	(0.269)***		
R^2	0.305	0.303	0.304	0.305	0.306
N	43784	43784	34392	43784	34392
Mean Dep. Variable	12.545	12.545	12.499	12.545	12.499
Panel C. SSCE Passes					
Mean BECE of SHS peers	0.040			0.057	0.037
	(0.013)***			(0.041)	(0.056)
1{BECE $\geq$ cutoff}		0.032	0.024		
		(0.023)	(0.037)		
R^2	0.128	0.128	0.134	0.128	0.134
N	34398	34398	26945	34398	26945
Mean Dep. Variable	3.071	3.071	3.074	3.071	3.074
Panel D. SSCE Passes [missing=0]					
Mean BECE of SHS peers	0.150			0.135	0.158
	(0.018)***			(0.055)**	(0.081)*
1{BECE $\geq$ cutoff}		0.075	0.102		
		(0.031)**	(0.052)*		
R^2	0.120	0.119	0.124	0.120	0.125
N	43784	43784	34392	43784	34392
Mean Dep. Variable	2.413	2.413	2.408	2.413	2.408

Notes: same as above.

Table 7: Correlates of School Quality

Dependent variable:	School Retention			Exam Performance	
School-level average residuals from selection on observables regressions	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE Score	SSCE Passes
	(1)	(2)	(3)	(4)	(5)
Profile					
Urban locality	0.006	0.005	0.012	-0.096	0.025
	(0.005)	(0.007)	(0.010)	(0.132)	(0.021)
Public	-0.011	0.014	0.113	0.747	-0.009
	(0.013)	(0.018)	(0.034)***	(0.401)*	(0.054)
Boys only	-0.023	-0.057	-0.083	1.418	-0.185
	(0.023)	(0.032)*	(0.051)	(0.624)**	(0.057)***
Girls only	-0.008	-0.056	-0.015	2.709	0.148
	(0.016)	(0.048)	(0.023)	(0.549)***	(0.047)***
Number of female students	-0.000	-0.000	0.000	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)
Number of male students	0.000	0.000	0.000	-0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)***	(0.000)
Infrastructure					
Boarding only	0.020	0.027	0.029	-0.321	-0.014
	(0.015)	(0.018)	(0.025)	(0.379)	(0.057)
Some boarding	0.010	0.014	0.019	0.032	0.014
	(0.005)**	(0.006)**	(0.009)**	(0.126)	(0.022)
Drinking water available	0.010	0.011	0.020	-0.120	-0.014
	(0.007)	(0.010)	(0.014)	(0.167)	(0.025)
Electricity functional	0.018	0.022	0.023	0.111	0.039
	(0.013)	(0.015)	(0.018)	(0.243)	(0.041)
Girls toilet seats per female student	0.013	-0.254	-0.416	-2.827	0.466
	(0.207)	(0.287)	(0.168)**	(2.320)	(0.282)*
Boys toilet seats per male student	0.096	0.389	0.345	1.354	-0.569
	(0.176)	(0.277)	(0.351)	(2.492)	(0.350)
Pupils per classroom	0.000	0.000	0.000	0.003	-0.000
	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)
Temporary classrooms	0.012	0.014	0.008	-0.579	-0.009
	(0.008)	(0.011)	(0.019)	(0.216)***	(0.037)
Classrooms needing major repair	0.015	0.013	-0.005	0.185	0.010
	(0.009)	(0.012)	(0.021)	(0.261)	(0.039)
Classrooms needing minor repair	0.012	0.000	-0.004	-0.256	0.015
	(0.009)	(0.012)	(0.017)	(0.233)	(0.035)

Continued on next page

	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE Score	SSCE Passes
Management					
School management committee	0.003	0.004	0.017	-0.272	-0.048
	(0.008)	(0.012)	(0.016)	(0.215)	(0.030)
SMC meetings per year	0.001	0.000	-0.003	-0.015	0.023
	(0.004)	(0.005)	(0.008)	(0.105)	(0.014)
Circuit supervisor visits per year	-0.000	0.004	0.005	0.015	0.007
	(0.003)	(0.004)	(0.005)	(0.069)	(0.011)
In-service trainings in last year	0.001	-0.001	-0.004	-0.006	0.007
	(0.003)	(0.004)	(0.005)	(0.081)	(0.012)
In-service training on HIV/AIDS	0.001	0.002	-0.000	0.109	0.051
	(0.006)	(0.008)	(0.012)	(0.164)	(0.025)**
HIV/AIDS issues taught	-0.014	-0.017	-0.027	0.030	-0.018
	(0.006)**	(0.008)**	(0.012)**	(0.159)	(0.028)
Teacher Characteristics					
Male	-0.081	-0.098	-0.065	1.101	-0.052
	(0.025)***	(0.033)***	(0.056)	(0.643)*	(0.104)
Median age	0.001	0.001	0.001	-0.017	0.001
	(0.001)	(0.001)	(0.001)	(0.019)	(0.003)
Ranked superintendent	0.031	0.008	-0.006	-0.451	-0.038
	(0.017)*	(0.022)	(0.034)	(0.471)	(0.073)
Fulltime	-0.044	0.019	0.134	0.481	0.008
	(0.030)	(0.036)	(0.054)**	(0.750)	(0.107)
Primary function teaching	0.002	-0.022	0.078	0.899	0.228
	(0.077)	(0.096)	(0.154)	(1.614)	(0.280)
University graduate	0.034	0.036	0.028	-0.371	-0.002
	(0.011)***	(0.014)**	(0.024)	(0.277)	(0.040)
Passed teaching qualification	0.027	0.033	0.034	0.177	0.074
	(0.010)***	(0.014)**	(0.021)	(0.295)	(0.047)
Pupil teacher ratio	0.000	-0.000	0.000	0.007	0.001
	(0.000)	(0.000)	(0.000)	(0.007)	(0.001)
R^2	0.224	0.232	0.285	0.204	0.057
N	447	447	447	447	447
Mean Dep. Variable	0.006	0.008	0.010	0.096	-0.011

Notes: Table displays results from a set of OLS regressions with the school-level average residuals from selection on observables estimates as the dependent variable (based on self revelation specification in column 2 of Tables 3 and 4). *p<0.1, **p<0.05, ***p<0.01.

References

ABDULKADIROĞLU, A., J. D. ANGRIST, AND P. A. PATHAK (forthcoming): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*.

AJAYI, K. F. (2013): "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," *Working Paper*.

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113(1), 151–184.

ANDRABI, T., J. DAS, AND A. I. KHWAJA (2013): "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," Working paper.

BANERJEE, A. V., R. BANERJI, E. DUFLO, R. GLENNERSTER, AND S. KHEMANI (2010): "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," *American Economic Journal: Economic Policy*, 2(1), 1–30.

BLACK, D. A., AND J. A. SMITH (2006): "Estimating the Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics*, 24(3), 701–728.

CARD, D., AND A. KRUEGER (1992): "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100(1), 1–40.

CHETTY, R., J. FRIEDMAN, N. HILGER, E. SAEZ, D. SCHANZENBACH, AND D. YAGAN (2011): "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 4, 1593–1660.

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

CULLEN, J., B. JACOB, AND S. LEVITT (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.

——— (2006): "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 74(5), 1191–1230.

Dale, S., and A. B. Krueger (2011): "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data," Working Paper 17159, National Bureau of Economic Research.

Dale, S. B., and A. B. Krueger (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *The Quarterly Journal of Economics*, 117(4), 1491–1527.

de Hoop, J. (2010): "Selective Secondary Education and School Participation in Sub-Saharan Africa: Evidence from Malawi," Tinbergen Institute Discussion Papers 10-041/2, Tinbergen Institute.

de Janvry, A., A. Dustan, and E. Sadoulet (2013): "Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City," *Working Paper*.

Deming, D. J. (2011): "Better Schools, Less Crime?," *The Quarterly Journal of Economics*.

Dobbie, W., and R. G. Fryer (forthcoming): "The Impact of Attending a School with High-Achieving Peers: Evidence from New York City Exam Schools," *American Economic Journal: Applied Economics*.

Duflo, E., P. Dupas, and M. Kremer (2011): "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, 101(5), 1739–74.

Gould, E. D., V. Lavy, and M. D. Paserman (2004): "Immigrating to Opportunity: Estimating The Effect of School Quality Using a Natural Experiment on Ethiopians in Israel," *The Quarterly Journal of Economics*, 119(2), 489–526.

Hastings, J., T. Kane, and D. Staiger (2008): "Heterogeneous Preferences and the Efficacy of Public School Choice," *National Bureau of Economic Research Working Paper*, (No. 12145 and No. 11805 combined).

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

Hoxby, C. M. (2003): "School Choice and School Productivity: Could School Choice Be a Tide that Lifts All Boats?," in *The Economics of School Choice*, ed. by C. M. Hoxby, NBER Chapters, pp. 287–342. University of Chicago Press.

Jackson, C. K. (2010): "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," *Economic Journal*, 120(549), 1399–1429.

Lai, F., E. Sadoulet, and A. de Janvry (2011): "The Contributions of School Quality and Teacher Qualifications to Student Performance Evidence from a Natural Experiment in Beijing Middle Schools," *Journal of Human Resources*, 46(1), 123–153.

Lavy, V. (2010): "Effects of Free Choice Among Public Schools," *The Review of Economic Studies*, 77(3), 1164–1191.

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

Lucas, A. M., and I. M. Mbiti (forthcoming): "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2), 698 – 714.

Mizala, A., and M. Urquiola (2013): "School Markets: The Impact of Information Approximating Schools' Effectiveness," *Journal of Development Economics*, 103, 313–335.

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

A Appendix: Heterogeneity in School Quality Effects

This section of the appendix examines heterogeneity in selection on observables estimates of the effects of school quality on student performance. In particular, I examine heterogeneity based on: sex, BECE performance, JHS quality, applying to an elite school, and average selectivity of schools applied to.

Table A.1: Selection on Observables Estimates - Heterogeneity by Sex

	School Retention			Exam Performance					
	Take SSCE (1)	Take SSCE on Time (2)	Take SSCE in Assigned School (3)	SSCE score (4)	SSCE score [missing= median] (5)	SSCE score [missing=0] (6)	SSCE passes (7)	SSCE pass [missing= median] (8)	SSCE pass [missing=0] (9)
Panel A. Matching on school and program choice									
Mean BECE of SHS peers	0.033	0.065	0.212	1.049	1.271	1.663	0.056	0.071	0.169
	(0.010)***	(0.012)***	(0.013)***	(0.237)***	(0.181)***	(0.209)***	(0.031)*	(0.021)***	(0.038)***
Male	-0.034	-0.065	-0.067	1.068	0.878	0.466	0.202	0.170	0.067
	(0.011)***	(0.012)***	(0.012)***	(0.171)***	(0.121)***	(0.173)***	(0.036)***	(0.026)***	(0.037)*
Mean BECE × Male	0.018	0.015	-0.017	-0.198	-0.089	0.121	-0.095	-0.068	-0.016
	(0.010)*	(0.011)	(0.011)	(0.230)	(0.189)	(0.210)	(0.027)***	(0.020)***	(0.034)
R^2	0.432	0.462	0.512	0.730	0.627	0.586	0.562	0.445	0.475
N	33149	33149	33149	24763	33149	33149	24763	33149	33149
Mean Dep. Variable	0.747	0.584	0.471	14.815	14.103	11.067	2.937	2.953	2.194
Panel B. Matching on school choice									
Mean BECE of SHS peers	0.028	0.064	0.201	0.657	0.757	1.088	0.053	0.052	0.134
	(0.005)***	(0.007)***	(0.009)***	(0.137)***	(0.119)***	(0.124)***	(0.016)***	(0.012)***	(0.019)***
Male	-0.021	-0.052	-0.043	1.074	0.783	0.536	0.207	0.147	0.085
	(0.005)***	(0.005)***	(0.006)***	(0.082)***	(0.065)***	(0.080)***	(0.016)***	(0.011)***	(0.016)***
Mean BECE × Male	0.001	-0.007	-0.029	0.033	0.122	0.137	-0.106	-0.065	-0.062
	(0.005)	(0.006)	(0.007)***	(0.145)	(0.126)	(0.121)	(0.014)***	(0.011)***	(0.018)***
R^2	0.259	0.294	0.359	0.593	0.492	0.450	0.376	0.281	0.317
N	97402	97402	97402	71659	97402	97402	71659	97402	97402
Mean Dep. Variable	0.736	0.564	0.444	14.198	13.617	10.445	2.893	2.921	2.128
Panel C. Self-revelation									
Mean BECE of SHS peers	0.019	0.048	0.169	1.231	1.211	1.439	0.062	0.064	0.121
	(0.006)***	(0.008)***	(0.010)***	(0.200)***	(0.185)***	(0.154)***	(0.013)***	(0.010)***	(0.018)***
Male	-0.026	-0.064	-0.056	1.085	0.745	0.434	0.165	0.117	0.039
	(0.004)***	(0.005)***	(0.006)***	(0.100)***	(0.099)***	(0.096)***	(0.011)***	(0.008)***	(0.013)***
Mean BECE × Male	0.002	-0.002	-0.016	-0.321	-0.176	-0.150	-0.119	-0.080	-0.073
	(0.006)	(0.007)	(0.008)**	(0.249)	(0.235)	(0.199)	(0.010)***	(0.009)***	(0.016)***
R^2	0.057	0.088	0.142	0.404	0.324	0.289	0.128	0.084	0.131
N	142366	142366	142366	102957	142366	142366	102957	142366	142366
Mean Dep. Variable	0.723	0.545	0.421	13.900	13.374	10.053	2.865	2.903	2.072

Notes: Regressions include controls for student BECE scores, age, gender, average BECE score in JHS, and an indicator for attending a public JHS. Panels A and B also include fixed effects for student preferences and Panel C controls for the average selectivity of schools selected. Robust standard errors are clustered at the level of the assigned SHS and reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table A.2: Selection on Observables Estimates - Heterogeneity by BECE Performance

	School Retention			Exam Performance					
	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE score	SSCE score [missing= median]	SSCE score [missing=0]	SSCE passes	SSCE pass [missing= median]	SSCE pass [missing=0]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Matching on school and program choice									
Mean BECE of SHS peers	0.177	0.255	0.384	-6.976	-6.733	-4.611	0.209	0.022	0.553
	(0.027)***	(0.030)***	(0.034)***	(0.537)***	(0.456)***	(0.602)***	(0.089)**	(0.060)	(0.100)***
Individual BECE	0.001	0.002	0.001	0.078	0.057	0.069	0.007	0.005	0.008
	(0.000)***	(0.000)***	(0.000)***	(0.003)***	(0.002)***	(0.003)***	(0.001)***	(0.000)***	(0.001)***
Mean BECE × Indiv. BECE	-0.000	-0.001	-0.001	0.023	0.024	0.019	-0.001	0.000	-0.001
	(0.000)***	(0.000)***	(0.000)***	(0.002)***	(0.001)***	(0.002)***	(0.000)**	(0.000)	(0.000)***
R^2	0.432	0.463	0.514	0.740	0.643	0.591	0.562	0.445	0.476
N	33149	33149	33149	24763	33149	33149	24763	33149	33149
Mean Dep. Variable	0.747	0.584	0.471	14.815	14.103	11.067	2.937	2.953	2.194
Panel B. Matching on school choice									
Mean BECE of SHS peers	0.189	0.234	0.331	-7.263	-7.033	-4.771	0.288	0.043	0.609
	(0.015)***	(0.019)***	(0.028)***	(0.361)***	(0.292)***	(0.327)***	(0.055)***	(0.039)	(0.056)***
Individual BECE	0.001	0.002	0.002	0.080	0.061	0.076	0.008	0.005	0.009
	(0.000)***	(0.000)***	(0.000)***	(0.002)***	(0.001)***	(0.002)***	(0.000)***	(0.000)***	(0.000)***
Mean BECE × Indiv. BECE	-0.000	-0.001	-0.000	0.024	0.024	0.018	-0.001	-0.000	-0.002
	(0.000)***	(0.000)***	(0.000)***	(0.001)***	(0.001)***	(0.001)***	(0.000)***	(0.000)	(0.000)***
R^2	0.261	0.296	0.360	0.610	0.515	0.456	0.377	0.280	0.318
N	97402	97402	97402	71659	97402	97402	71659	97402	97402
Mean Dep. Variable	0.736	0.564	0.444	14.198	13.617	10.445	2.893	2.921	2.128
Panel C. Self-revelation									
Mean BECE of SHS peers	0.163	0.178	0.236	-7.473	-7.259	-5.301	0.225	-0.025	0.464
	(0.017)***	(0.025)***	(0.037)***	(0.442)***	(0.349)***	(0.385)***	(0.065)***	(0.047)	(0.064)***
Individual BECE	0.001	0.002	0.002	0.078	0.059	0.075	0.008	0.005	0.009
	(0.000)***	(0.000)***	(0.000)***	(0.001)***	(0.001)***	(0.001)***	(0.000)***	(0.000)***	(0.000)***
Mean BECE × Indiv. BECE	-0.000	-0.000	-0.000	0.025	0.025	0.020	-0.001	0.000	-0.001
	(0.000)***	(0.000)***	(0.000)**	(0.001)***	(0.001)***	(0.001)***	(0.000)***	(0.000)	(0.000)***
R^2	0.060	0.090	0.142	0.436	0.366	0.302	0.127	0.083	0.132
N	142366	142366	142366	102957	142366	142366	102957	142366	142366
Mean Dep. Variable	0.723	0.545	0.421	13.900	13.374	10.053	2.865	2.903	2.072

Notes: same as above.

Table A.3: Selection on Observables Estimates - Heterogeneity by JHS Performance

	School Retention			Exam Performance					
	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE score	SSCE score [missing= median]	SSCE score [missing=0]	SSCE passes	SSCE pass [missing= median]	SSCE pass [missing=0]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Matching on school and program choice									
Mean BECE of SHS peers	0.072	0.118	0.232	0.295	0.484	1.352	0.034	0.043	0.260
	(0.016)***	(0.019)***	(0.019)***	(0.283)	(0.199)**	(0.289)***	(0.049)	(0.031)	(0.055)***
High-performing JHS	-0.020	-0.013	-0.002	-1.838	-1.202	-1.445	-0.157	-0.099	-0.160
	(0.012)*	(0.015)	(0.014)	(0.214)***	(0.154)***	(0.202)***	(0.038)***	(0.024)***	(0.046)***
Mean BECE × High-perf. JHS	-0.036	-0.052	-0.038	0.362	0.553	0.115	-0.069	-0.038	-0.148
	(0.015)**	(0.017)***	(0.017)**	(0.280)	(0.203)***	(0.271)	(0.045)	(0.028)	(0.053)***
R^2	0.431	0.462	0.512	0.723	0.622	0.582	0.558	0.442	0.473
N	33149	33149	33149	24763	33149	33149	24763	33149	33149
Mean Dep. Variable	0.747	0.584	0.471	14.815	14.103	11.067	2.937	2.953	2.194
Panel B. Matching on school choice									
Mean BECE of SHS peers	0.068	0.098	0.206	0.167	0.221	1.038	0.059	0.037	0.241
	(0.008)***	(0.010)***	(0.011)***	(0.154)	(0.119)*	(0.148)***	(0.024)**	(0.017)**	(0.028)***
High-performing JHS	-0.022	-0.019	-0.019	-1.960	-1.340	-1.605	-0.201	-0.146	-0.212
	(0.006)***	(0.007)***	(0.007)***	(0.112)***	(0.094)***	(0.105)***	(0.018)***	(0.012)***	(0.023)***
Mean BECE × High-perf. JHS	-0.049	-0.047	-0.030	0.155	0.373	-0.214	-0.111	-0.055	-0.201
	(0.008)***	(0.009)***	(0.011)***	(0.160)	(0.128)***	(0.146)	(0.022)***	(0.015)***	(0.028)***
R^2	0.259	0.294	0.358	0.579	0.483	0.443	0.371	0.277	0.315
N	97402	97402	97402	71659	97402	97402	71659	97402	97402
Mean Dep. Variable	0.736	0.564	0.444	14.198	13.617	10.445	2.893	2.921	2.128
Panel C. Self-revelation									
Mean BECE of SHS peers	0.052	0.071	0.164	0.201	0.106	0.732	0.065	0.027	0.183
	(0.008)***	(0.010)***	(0.013)***	(0.157)	(0.124)	(0.147)***	(0.025)**	(0.018)	(0.028)***
High-performing JHS	-0.028	-0.033	-0.041	-2.031	-1.351	-1.684	-0.241	-0.163	-0.246
	(0.005)***	(0.006)***	(0.007)***	(0.104)***	(0.097)***	(0.101)***	(0.014)***	(0.010)***	(0.018)***
Mean BECE × High-perf. JHS	-0.042	-0.035	-0.016	0.491	0.799	0.298	-0.129	-0.046	-0.171
	(0.007)***	(0.010)***	(0.012)	(0.191)**	(0.166)***	(0.176)*	(0.025)***	(0.018)***	(0.028)***
R^2	0.057	0.087	0.139	0.391	0.318	0.282	0.120	0.079	0.127
N	142366	142366	142366	102957	142366	142366	102957	142366	142366
Mean Dep. Variable	0.723	0.545	0.421	13.900	13.374	10.053	2.865	2.903	2.072

Notes: same as above. High-performing JHSs are those with above median average BECE scores.

Table A.4: Selection on Observables Estimates - Heterogeneity by Application to Elite School

	School Retention			Exam Performance					
	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE score	SSCE score [missing= median]	SSCE score [missing=0]	SSCE passes	SSCE pass [missing= median]	SSCE pass [missing=0]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Matching on school and program choice									
Mean BECE of SHS peers	0.048	0.078	0.201	0.049	0.349	0.929	-0.001	0.024	0.169
	(0.011)***	(0.013)***	(0.014)***	(0.270)	(0.201)*	(0.242)***	(0.036)	(0.024)	(0.042)***
Mean BECE × Applied elite	-0.008	-0.006	0.001	1.332	1.352	1.257	0.001	0.008	-0.015
	(0.011)	(0.013)	(0.014)	(0.252)***	(0.198)***	(0.231)***	(0.034)	(0.023)	(0.040)
R^2	0.432	0.462	0.512	0.731	0.629	0.587	0.562	0.445	0.475
N	33149	33149	33149	24763	33149	33149	24763	33149	33149
Mean Dep. Variable	0.747	0.584	0.471	14.815	14.103	11.067	2.937	2.953	2.194
Panel B. Matching on school choice									
Mean BECE of SHS peers	0.035	0.063	0.180	0.057	0.186	0.610	0.020	0.021	0.127
	(0.006)***	(0.008)***	(0.010)***	(0.157)	(0.125)	(0.126)***	(0.018)	(0.014)	(0.022)***
Mean BECE × Applied elite	-0.012	-0.004	0.008	1.026	1.092	0.948	-0.046	-0.012	-0.048
	(0.006)*	(0.007)	(0.009)	(0.163)***	(0.138)***	(0.137)***	(0.018)**	(0.013)	(0.022)**
R^2	0.259	0.294	0.359	0.594	0.494	0.451	0.376	0.280	0.317
N	97402	97402	97402	71659	97402	97402	71659	97402	97402
Mean Dep. Variable	0.736	0.564	0.444	14.198	13.617	10.445	2.893	2.921	2.128
Panel C. Self-revelation									
Mean BECE of SHS peers	0.030	0.055	0.161	0.226	0.291	0.654	0.021	0.019	0.110
	(0.006)***	(0.008)***	(0.010)***	(0.184)	(0.156)*	(0.150)***	(0.016)	(0.012)	(0.020)***
Applied to elite SHS	0.004	0.018	0.036	0.099	0.214	0.264	0.006	0.006	0.019
	(0.004)	(0.006)***	(0.008)***	(0.120)	(0.106)**	(0.106)**	(0.013)	(0.010)	(0.016)
Mean BECE × Applied elite	-0.016	-0.012	-0.002	1.290	1.315	1.120	-0.047	-0.005	-0.054
	(0.005)***	(0.007)*	(0.009)	(0.197)***	(0.180)***	(0.169)***	(0.016)***	(0.012)	(0.018)***
R^2	0.058	0.088	0.143	0.409	0.332	0.292	0.126	0.083	0.131
N	142366	142366	142366	102957	142366	142366	102957	142366	142366
Mean Dep. Variable	0.723	0.545	0.421	13.900	13.374	10.053	2.865	2.903	2.072

Notes: same as above. Elite schools are the 22 schools that were established under the British colonial administration before Ghana gained independence in 1957 and are classified in Category A by Ghana Education Service.

Table A.5: Selection on Observables Estimates - Heterogeneity by Application to Selective School

	School Retention			Exam Performance					
	Take SSCE	Take SSCE on Time	Take SSCE in Assigned School	SSCE score	SSCE score [missing= median]	SSCE score [missing=0]	SSCE passes	SSCE pass [missing= median]	SSCE pass [missing=0]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Matching on school and program choice									
Mean BECE of SHS peers	0.073	0.092	0.193	-0.867	-0.503	0.369	0.064	0.035	0.253
	(0.015)***	(0.017)***	(0.018)***	(0.324)***	(0.234)**	(0.305)	(0.048)	(0.031)	(0.054)***
Mean BECE × Applied selective	-0.034	-0.020	0.010	2.031	1.994	1.586	-0.073	-0.007	-0.109
	(0.013)**	(0.015)	(0.017)	(0.315)***	(0.236)***	(0.290)***	(0.042)*	(0.028)	(0.049)**
R^2	0.432	0.462	0.512	0.732	0.630	0.587	0.562	0.445	0.475
N	33149	33149	33149	24763	33149	33149	24763	33149	33149
Mean Dep. Variable	0.747	0.584	0.471	14.815	14.103	11.067	2.937	2.953	2.194
Panel B. Matching on school choice									
Mean BECE of SHS peers	0.063	0.077	0.166	-0.843	-0.659	0.098	0.059	0.018	0.207
	(0.009)***	(0.009)***	(0.012)***	(0.179)***	(0.144)***	(0.164)	(0.027)**	(0.019)	(0.030)***
Mean BECE × Applied selective	-0.041	-0.020	0.021	1.762	1.762	1.267	-0.077	-0.005	-0.129
	(0.009)***	(0.010)**	(0.012)*	(0.185)***	(0.152)***	(0.161)***	(0.025)***	(0.018)	(0.030)***
R^2	0.259	0.294	0.359	0.595	0.496	0.451	0.376	0.280	0.317
N	97396	97396	97396	71654	97396	97396	71654	97396	97396
Mean Dep. Variable	0.736	0.564	0.444	14.197	13.617	10.445	2.892	2.921	2.128
Panel C. Self-revelation									
Mean BECE of SHS peers	0.053	0.063	0.140	-0.707	-0.657	-0.023	0.042	-0.002	0.157
	(0.008)***	(0.012)***	(0.015)***	(0.207)***	(0.164)***	(0.187)	(0.027)	(0.020)	(0.031)***
Applied to selective schools	-0.010	0.007	0.009	-0.650	-0.336	-0.459	-0.031	-0.012	-0.043
	(0.006)*	(0.008)	(0.010)	(0.148)***	(0.125)***	(0.127)***	(0.019)	(0.014)	(0.023)*
Mean BECE × Applied selective	-0.039	-0.018	0.023	2.017	2.086	1.613	-0.061	0.020	-0.098
	(0.008)***	(0.012)	(0.015)	(0.235)***	(0.197)***	(0.210)***	(0.027)**	(0.020)	(0.031)***
R^2	0.058	0.088	0.142	0.412	0.335	0.293	0.126	0.083	0.131
N	142366	142366	142366	102957	142366	142366	102957	142366	142366
Mean Dep. Variable	0.723	0.545	0.421	13.900	13.374	10.053	2.865	2.903	2.072

Notes: same as above. Selective schools are those with above median average BECE scores of assigned students.

B Appendix: Regression Discontinuity Analysis

Table B.1: Discontinuity Design Summary Statistics

	All (1)	Discontinuity Sample (2)	Within $(-20, 20)$ (3)	Excluding $(-5, 5)$ (4)
Student Characteristics				
Age	16.618	16.436	16.347	16.345
Male	0.583	0.554	0.586	0.586
JHS Public	0.749	0.692	0.642	0.642
BECE score	0.000	0.102	0.574	0.571
Mean BECE of JHS peers	0.000	0.115	0.366	0.365
Mean BECE of SHS peers	0.000	0.055	0.497	0.480
Admission Outcomes				
First choice program	0.274	0.113	0.224	0.214
Second choice program	0.187	0.183	0.263	0.270
Third choice program	0.233	0.294	0.330	0.328
District of choice	0.198	0.299	0.183	0.187
Region of choice	0.108	0.111	0.000	0.000
Secondary School Performance				
Take SSCE	0.728	0.746	0.786	0.783
Take SSCE in three years	0.550	0.570	0.643	0.640
Take SSCE in assigned school	0.421	0.417	0.531	0.528
SSCE score	9.958	10.482	12.545	12.499
SSCE core passes	2.068	2.160	2.413	2.408
SSCE total passes	4.114	4.289	4.783	4.768
N	159607	201610	43784	34392

Table B.2: Covariate Balance

	JHS Performance (1)	Male (2)	Age (3)	JHS Public (4)
Panel A. Baseline Sample				
1{BECE $\geq$ cutoff}	0.022	-0.002	-0.020	-0.002
	(0.010)**	(0.007)	(0.029)	(0.008)
Exclude $(-5, 5)$ Range	No	No	No	No
R^2	0.633	0.446	0.223	0.221
N	43784	43784	43768	43775
Mean Dep. Variable	0.339	0.586	16.347	0.642
Panel B. Excluding -5 to 5 Range				
1{BECE $\geq$ cutoff}	0.021	0.011	0.013	0.005
	(0.017)	(0.012)	(0.047)	(0.014)
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.634	0.447	0.231	0.222
N	34392	34392	34380	34385
Mean Dep. Variable	0.338	0.586	16.345	0.642

Notes: Regressions include cutoff fixed-effects as well as controls for BECE score, BECE score × 1{BECE $\geq$ cutoff}, gender, and average BECE score in student's JHS. Robust standard errors clustered at the cutoff level are reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table B.3: Discontinuity Design: First Stage

	Admission Chances			SHS School Quality	
				Peer BECE	SSCE Pass Rate
	(1)	(2)	(3)	(4)	(5)
1{BECE $\geq$ cutoff}	0.739	0.739	0.882	0.650	10.171
	(0.014)***	(0.014)***	(0.010)***	(0.019)***	(0.576)***
Mean BECE of JHS peers		0.010	0.009	0.062	0.604
		(0.002)***	(0.002)***	(0.005)***	(0.130)***
Male		0.005	0.005	-0.011	-0.874
		(0.004)	(0.004)	(0.006)*	(0.192)***
Covariates	No	Yes	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	No	No	Yes	Yes	Yes
R^2	0.786	0.787	0.863	0.833	0.521
N	43784	43784	34392	34392	33301
Mean Dep. Variable	0.451	0.451	0.430	0.480	81.547

Notes: same as above.

Table B.4: Bounding Selection Effects: School Retention

	Baseline	Selection from Regional Placement		
		Random	Positive	Negative
Panel A. Take SSCE				
Mean BECE of SHS peers	0.047	0.041	-0.001	0.357
	(0.021)**	(0.021)**	(0.020)	(0.025)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.071	0.063	0.068	.
N	34392	38646	38646	38646
Mean Dep. Variable	0.783	0.784	0.807	0.712
Panel B. Take SSCE in three years				
Mean BECE of SHS peers	0.096	0.097	0.015	0.252
	(0.024)***	(0.023)***	(0.023)	(0.025)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.108	0.097	0.103	0.080
N	34392	38646	38646	38646
Mean Dep. Variable	0.640	0.639	0.679	0.569
Panel C. Take SSCE in assigned school				
Mean BECE of SHS peers	0.227	0.246	0.109	0.349
	(0.024)***	(0.024)***	(0.023)***	(0.025)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.167	0.147	0.151	0.134
N	34392	38646	38646	38646
Mean Dep. Variable	0.528	0.520	0.580	0.470

Notes: Table displays results from IV regressions with an indicator for a given school retention measure as the dependent variable. Regressions include cutoff fixed-effects as well as controls for BECE score, BECE score $\times$ 1{BECE $\geq$ cutoff}, gender, and average BECE score in student's JHS. Robust standard errors clustered at the cutoff level are reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

Table B.5: Bounding Selection Effects: Performance on Core SSCE Subjects

	Baseline	Selection from Regional Placement		
		Random	Positive	Negative
Panel A. SSCE Score				
Mean BECE of SHS peers	0.732	1.785	-2.021	3.919
	(0.312)**	(0.336)***	(0.378)***	(0.333)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.466	0.420	0.357	0.402
N	26945	30311	31199	31199
Mean Dep. Variable	15.953	15.611	16.384	14.752
Panel B. SSCE Score [missing=0]				
Mean BECE of SHS peers	1.086	1.878	-1.817	4.856
	(0.415)***	(0.423)***	(0.461)***	(0.454)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.306	0.279	0.253	0.249
N	34392	38646	38646	38646
Mean Dep. Variable	12.499	12.244	13.227	11.174
Panel C. SSCE Passes				
Mean BECE of SHS peers	0.037	0.127	-0.231	0.795
	(0.056)	(0.054)**	(0.060)***	(0.061)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.134	0.126	0.101	0.052
N	26945	30311	31199	31199
Mean Dep. Variable	3.074	3.041	3.200	2.881
Panel D. SSCE Passes [missing=0]				
Mean BECE of SHS peers	0.158	0.215	-0.202	0.752
	(0.081)*	(0.078)***	(0.080)**	(0.084)***
Impute Missing BECE Scores	No	Yes	Yes	Yes
Exclude $(-5, 5)$ Range	Yes	Yes	Yes	Yes
R^2	0.125	0.115	0.110	0.105
N	34392	38646	38646	38646
Mean Dep. Variable	2.408	2.385	2.583	2.143

Notes: Table displays results from IV regressions with a given measure of performance on the core SSCE subjects as the dependent variable. Regressions include cutoff fixed-effects as well as controls for BECE score, BECE score $\times$ 1{BECE $\geq$ cutoff}, gender, and average BECE score in student's JHS. Robust standard errors clustered at the cutoff level are reported in parentheses, *p<0.1, **p<0.05, ***p<0.01.

C Appendix: Moving Schools

Table C.1: Changes in Schools based on Categories

	Category of Assigned School							
	A	B	C	D	Tech	Private	Private Tech.	Total
Category of SSCE School								
A	83.59	4.49	5.65	7.18	13.07	10.58	12.82	22.05
B	3.88	81.23	6.48	8.34	17.15	14.40	17.95	23.44
C	7.70	8.56	76.58	22.11	39.79	31.49	51.28	36.35
D	3.50	3.94	8.52	58.43	22.69	10.44	12.82	14.66
Private	1.32	1.78	2.77	3.94	7.30	33.09	5.13	3.51
Total	22168	24018	38566	16890	2111	3487	39	107279

Notes: Ghana Education Service introduced a categorization scheme for secondary schools in 2009 based on schools' available facilities, with Category A indicating schools with the highest rating facilities.

Table C.2: Changes in Enrollment at Elite Schools

	Assigned to Elite School		
	No	Yes	Total
Take SSCE in Elite School			
No	98.05	13.27	91.22
Yes	1.95	86.73	8.78
Total	98634	8645	107279

Notes: I define elite schools as the 22 schools that are classified as Category A under the GES categorization scheme and were established under the British colonial administration before Ghana gained independence in 1957. These are the oldest schools in the country and they are more selective than the average school.

D Appendix: School Quality

Table D.1: School Summary Statistics

	All	Selectivity[b] Low	High
Profile			
Urban locality	0.518	0.472	0.833
Public[a]	0.965	0.959	1.000
Boys only	0.044	0.013	0.262
Girls only	0.041	0.019	0.198
Number of female students	328.833	295.391	558.144
Number of male students	437.156	397.780	707.164
Total number of enrolled students	765.989	693.171	1265.308
Infrastructure			
Boarding only	0.051	0.037	0.143
Some boarding	0.558	0.503	0.929
Drinking water available	0.869	0.866	0.889
Electricity functional	0.938	0.938	0.944
Girls toilet seats per female student	0.036	0.035	0.047
Boys toilet seats per male student	0.027	0.025	0.040
Pupils per classroom	42.309	42.062	44.003
Temporary classrooms (ratio to permanent)	0.103	0.110	0.059
Share of classrooms needing major repair	0.239	0.242	0.214
Share of classrooms needing minor repair	0.308	0.311	0.290
Management			
School management committee	0.826	0.825	0.833
SMC meetings per year	1.598	1.581	1.714
Circuit supervisor visits per year	1.362	1.429	0.897
In-service trainings in last year	0.947	0.972	0.778
In-service training on HIV/AIDS	0.738	0.742	0.714
HIV/AIDS issues taught in curriculum	0.739	0.741	0.730
Teacher Characteristics			
Male	0.830	0.842	0.752
Median age	37.963	37.500	41.142
Ranked superintendent or higher	0.741	0.743	0.723
Full time	0.952	0.948	0.979
Primary function teaching (not admin)	0.953	0.953	0.958
University graduate	0.675	0.654	0.818
Passed teaching qualification	0.803	0.796	0.846
Pupil teacher ratio	22.159	22.210	21.810
N	495	432	63

Notes: [a]Some schools change ownership between 2005 and 2008, variable indicates median value of public school indicator over the time period. [b]Highly selective schools are those where the average BECE score of admitted students is one standard deviation above the mean BECE score.