Economics of Education Review xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Who benefits from selective education? Evidence from elite boarding school admissions

Ying Shi

Department of Public Administration and International Affairs, Syracuse University, 215 Eggers Hall, Syracuse, NY 13244-1020, USA

ARTICLE INFO

Keywords:
Selective education
Peer effects
STEM education
Achievement gap

ABSTRACT

Existing research finds minimal gains from attending elite US secondary schools. This paper estimates the causal effect of attending a selective public boarding school, an institutional model increasingly used by states to serve academically gifted students. Regression discontinuity estimates using multiple admissions thresholds show math score gains and college application and enrollment patterns that shift away from less competitive colleges. Effects are concentrated among minorities, students with lower prior individual achievement, from rural neighborhoods, or lower-achieving sending schools. The opportunity to attend selective boarding schools reduces the tendency of disadvantaged or under-represented students to attend a less selective college by at least one-quarter.

1. Introduction

Elite secondary schools are receiving increasing attention as an option for high-achieving students. Proponents of these schools argue that they give talented students from different backgrounds a rigorous education that prepares them for entry into competitive colleges. Yet few studies show that selective schools exert meaningful causal effects on student performance or long-term outcomes. US-based studies find null or small treatment effects in cities such as New York, Boston, and Chicago (Abdulkadiroglu, Angrist, Narita, Pathak, & Zarate, 2017; Abdulkadiroglu, Angrist, & Pathak, 2014; Dobbie & Fryer, 2014).[1] International studies yield mixed results with meaningful achievement gains in Romania and Trinidad and Tobago (Jackson, 2010; Pop-Eleches & Urquiola, 2013), while test score effects are minimal after exposure to selective secondary schools in the United Kingdom, China, and Kenya (Clark, 2010; Lucas & Mbiti, 2014; Zhang, 2016).

A core challenge in evaluating these schools is the non-random selection of students. This paper follows previous regression discontinuity (RD) studies in identifying treatment effects via discontinuities in admissions probabilities. An advantageous feature of the present study is the availability of multiple cutoffs enabling a powerful exploration of effect heterogeneity in the US context. The setting is a selective residential high school in North Carolina with a unique within-district admissions process generated by a legislative mandate to equitably serve all thirteen congressional districts (CD). Applicants compete against peers from the same district. Varying levels of over-subscription and applicant baseline performance generate a wide range of admissions thresholds. Causal effects are then identifiable for students from a range of demographic backgrounds, neighborhoods, and schooling experiences who are exposed to the same academic environment.

Analyses rely on a new dataset linking selective school applicant data with statewide administrative files and National Student Clearinghouse (NSC) records. End of high school outcomes include standardized test scores and postsecondary application behavior. Information on students' college application portfolios permits an evaluation of the correspondence between postsecondary intentions and actual enrollment and completion observed in NSC data. RD results show that students just exceeding the admissions threshold experience a modest 2 percentile point gain in SAT math scores. The same students do not perform better on the SAT verbal exam. Findings on students' college trajectories represent a more striking departure from the preponderance of null or small effects in the selective schools literature. A predicted offer increases the likelihood of college enrollment by 5 percentage points. Students reduce their applications to less competitive colleges by nearly 20% and are significantly less likely to enroll in such a college after high school. As such, a lasting impact of selective school exposure is to shift students away from colleges near the lower end of the selectivity scale. Students' preferred institutions are

E-mail address: yshi78@syr.edu.

[1] Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014) find no consistent evidence that attending exam schools in Boston and New York affects test scores, AP outcomes, and college enrollment in the former, and longer term postsecondary outcomes in the latter. Null achievement effects are also found for selective enrollment schools in Chicago (Abdulkadiroglu et al., 2017).

https://doi.org/10.1016/j.econedurev.2019.07.001
Received 5 February 2019; Received in revised form 27 June 2019; Accepted 1 July 2019
0272-7757/

moreover oriented towards science, technology, engineering, and math (STEM) subjects.

Estimates disguise patterns of effect heterogeneity. I separately evaluate effects by students' individual, neighborhood, and sending school attributes. Math achievement and college selectivity effects are concentrated among disadvantaged students while more privileged students exhibit no meaningful gains. SAT math score gains are 3–5 percentile points among minority students, lower-achievers, and those from rural neighborhoods or lower-achieving sending schools. These students also decrease the share of less selective colleges they apply to by at least one-quarter, which translates to a similarly sized decrease in the likelihood that they enroll in such a college after high school. The selective boarding school experience is in effect bridging existing disparities between disadvantaged students and their better-off peers across these outcomes.[2] Notably, the shift towards more selective colleges has no discernible negative effects on persistence and completion. Disadvantaged and advantaged students are no more or less likely to complete within 4 years or attain a STEM degree within this time period.

To better understand the origins of these patterns, I examine channels through which selective school exposure can affect student outcomes. Applicants exceeding the threshold experience a large drop in relative academic rank and have significantly higher peer achievement.[3] This reflects the reality of attending a high quality selective school as measured not only by student performance but also proxies such as teacher educational attainment. The magnitude of these changes varies by student background. For example, the increase in peer achievement is 0.2-0.3 standard deviations (σ) higher for minority students or those from lower quality sending schools than their more privileged peers. Such differences in part explain observed heterogeneity in treatment effects. A scrutiny of mechanisms also informs our understanding of why this paper finds significant treatment effects in contrast to several other US-based studies. The institutional context and available fallback schools imply that treated students experience a greater jump in peer and school quality than related studies.[4] Another likely reason for the difference is the immersive experience of attending boarding school, which can amplify any positive effects of increased peer and school quality.

This study contributes to three strands of research. Its focus on an elite secondary school places it in the selective schools literature, in particular proximity to a handful of studies relying on multiple admissions cutoffs to investigate effect heterogeneity. These studies find null or small aggregate effects, and evince little evidence that these schools are better at serving particular student groups. Selective school eligibility does not meaningfully improve test scores, and may even exacerbate postsecondary attendance gaps across students from different neighborhood socioeconomic tiers in Chicago.[5] Lucas and Mbiti (2014) find no heterogeneous treatment effects by baseline test scores, gender, or socioeconomic status in Kenyan secondary schools. This paper's findings instead suggest that elite education serving disadvantaged groups with limited access to alternative quality schools has the potential to significantly improve their math scores and reduce the scope for academic undermatch. A closer analogy is available in the gifted and talented programs (GT) literature. While results are mixed overall, one RD study finds that significant benefits accrue to high-performing minority students who miss the IQ-based cutoffs but still enroll in gifted classrooms (Card & Giuliano, 2016).[6] Evidence also exists on the long-term influence of GT programs, from increased high school graduation and college enrollment (Cohodes, 2015) to the choice of a more challenging field of study (Booij, Haan, & Plug, 2016).

Second, this study contributes to a very limited body of research on boarding schools. While these are traditionally cast as private schools serving well-to-do families, public boarding institutions are serving an increasing number of under-represented minorities or economically disadvantaged students. The only two recent studies on this topic exploit admissions lotteries to show that boarding school attendance leads to significant academic achievement effects. In the US, SEED (Schools for Educational Evolution and Development) schools lead to sizable math and reading gains among poor students by combining a "No Excuses" charter school model with a 5-day-a-week boarding component (Curto & Fryer, 2014). A study of French boarding schools document positive effects in math scores only after two years of exposure (Behaghel, de Chaisemartin, & Gurgand, 2017). Both studies stop short of examining longer-term outcomes such as college enrollment and completion. Understanding how boarding schools affect postsecondary trajectories is therefore a key contribution of the present paper.

The finding that selective school eligibility shifts disadvantaged students away from less competitive colleges underscores an important phenomenon: the undermatching of these high-ability students at low-quality colleges (Bowen & Bok, 2016; Dillon & Smith, 2016). Academic undermatch is especially prevalent among low income students, those who live in rural areas, have little access to selective high schools, and are not in a critical mass of similarly high-achieving students (Hoxby & Avery, 2013; Smith, Pender, & Howell, 2013). By embedding students in a rigorous and well-resourced academic environment, the selective boarding school may be providing the necessary conditions for disadvantaged individuals to shift their college application and enrollment behavior towards that of high-income counterparts, or so-called "achievement-typical" as defined in Hoxby and Avery (2013). This potential for selective secondary schools to ameliorate academic undermatch is therefore a significant area of inquiry.

The paper proceeds as follows. Section 2 details the potential advantages and disadvantages of selective boarding schools and provides information on the institutional context of the school in question. Section 3 describes the data sources and construction of the analytic sample. Section 4 on empirical methodology outlines the regression discontinuity approach. I report pooled results and by student background attributes in Section 5 and examine the role of mediating factors in Section 6. I subject the validity of results to several robustness checks in Section 7 before concluding.

[2] Note that gains are not concentrated among disadvantaged students across all outcomes. When it comes to college enrollment, the opportunity to attend the selective high school benefits advantaged students as much as their peers, if not more.

[3] Numerous empirical studies show higher peer achievement to have positive effects (Ammermueller & Pischke, 2009; Hanushek, Kain, Markman, & Rivkin, 2003; Hoxby, 2000; Hoxby & Weingarth, 2005; Imberman, Kugler, & Sacerdote, 2012; Vigdor & Nechyba, 2007). In addition to direct benefits from the spillover effects of elite peers, students tracked to high achieving classrooms may also gain from teachers tailoring instruction to their academic level (Duflo, Dupas, & Kremer, 2011).

[4] The attributes of such fallback schools inform the interpretation of treatment effects, just as they have in the context of oversubscribed charter and traditional public schools (Chabrier, Cohodes, & Oreopoulos, 2016; Deming, Hastings, Kane, & Staiger, 2014).

[5] Barrow, Sartain, and de la Torre (2016) is the only paper I'm aware of in the US that uses multiple admissions thresholds. Chicago uses a serial dictatorship mechanism to allocate selective school applicants from 4 tiers of socioeconomic status (SES), as detailed in Abdulkadiroglu et al. (2017). The North Carolina context is more straightforward in estimation because applicants are not submitting a preference order for multiple high schools, which can introduce omitted variable bias since those with different orderings likely vary along hard-to-observe attributes that matter for outcomes of interest.

[6] Bui, Craig, and Imberman (2014) shows little evidence of higher student achievement with the exception of higher science scores among magnet GT program enrollees.

2. Selective boarding schools

2.1. Conceptual framework

Enrollment in a selective boarding school requires simultaneously changing multiple educational inputs. I consider these inputs underlying the treatment and theorize how each component may positively or negatively affect student outcomes. This facilitates comparisons between the boarding school and other selective day schools while informing the interpretation of key findings.

The move to residential schools could adversely impact student engagement and psychological well-being. Early research on the consequences of attending such a school identified a cluster of potential symptoms known as the Boarding School Syndrome (Schaverien, 2011). Some psychological research shows using qualitative and clinical data that the separation from home and family can manifest in increased detachment and homesickness, although the evidence is mixed. Recent work using Australian data finds mostly parity in boarding and day students' motivation, engagement, and psychological well-being (Martin, Papworth, Ginns, & Liem, 2014), while an evaluation of French boarding schools finds that a negative shock on student well-being and motivation may attenuate benefits from boarding school attendance (Behaghel et al., 2017).

Another factor that can put downward pressure on student outcomes is decreased parental involvement. Parents may help with homework and improve students' engagement with school. Attending a residential school likely leads to reduced parental investment. Suggestive evidence that parental inputs are substitutes with school quality also implies that parents reduce effort when their children enroll in a selective school (Pop-Eleches & Urquiola, 2013). As such the parental channel can mediate any benefits from attending a school with greater resources or better peers.

On the other hand, increased school quality via higher-achieving peers, greater resources, and a more challenging curriculum can lead to improved student performance. While these attributes are present to some degree across selective schools, their impact may intensify in a residential school setting. Students who leave their communities to attend a boarding school are immersed in peer groups both inside and outside of the classroom. To the extent that fellow students can impact classmates' academic interests and aspirations through social interaction and network spillovers (Hanushek et al., 2003; Hoxby, 2000), the prolonged exposure of a residential school can enhance gains from high quality peers. The dearth of "bad apples" decreases opportunities to disrupt learning, a public good that is accompanied by congestion effects (Lazear, 2001).

Outside of peers, the benefits of boarding schools also extend to mentorship opportunities from increased interactions with faculty and improved social capital (Curto & Fryer, 2014). Residential high schools can teach students a set of soft skills such as time management that facilitate the transition to college. While performance outcomes may not capture potential gains along these dimensions, these skills may manifest in improved college access and completion. Taken together, the net effect of boarding schools vis-à-vis selective day schools is ambiguous.

2.2. Institutional context

The context of this study is a selective public boarding school in North Carolina. This institution was among the first in a wave of US public boarding schools established in the past 35 years to serve academically gifted students. Nearly twenty states have created residential schools with selective enrollment.[7] A notable feature of this school is that it only serves grades 11 through 12. The combination of the shortened grade span and the residential nature of the school means that exposure to this academic setting is short but intensive.[8]

A key tenet in the school's mission is to develop future leaders in STEM fields. This focus on mathematics, science, and technology is similar to the curricular specialization of elite public institutions such as the Bronx High School of Science, Brooklyn Technical High School, and Stuyvesant High School.[9] In practice students receive a broader-based education in the arts and humanities in addition to extensive coursework in STEM. The rigorous curriculum has the highest required number of graduation credits in science and mathematics, followed by English. The school also emphasizes project-based learning, with 8–10 day sessions between trimesters that enable students to pursue mini-courses or self-guided projects outside of the traditional classroom context.

Students reside in residential halls constantly supervised by staff to ensure safety and well-being. They can choose to return home on weekends, although depending on distance not all students exercise this option. As such, the extent to which school inputs are substituting for parental investment depends on how far the student is from home. The geographic diversity of the enrolled student body sets it apart from elite day schools in the literature, and owes to a legislative mandate requiring the school to consider each of North Carolina's thirteen congressional districts. The result is a geographically expansive student body drawing on academic high-achievers across the state.

2.3. Admissions

Admission into the high school is highly competitive. The full set of application materials includes a completed form with demographic and sending school information, an academic transcript, recommendations from a counselor and three high school teachers, essays, and listings of extracurricular activities.[10] Applicants must also take the SAT exam between October to January of their 10th grade year. During the annual recruiting cycle, approximately 1000 students apply for fewer than 350 slots.

A defining aspect of the admissions process involves a legislative mandate that obligates the school to equitably serve all 13 congressional districts in North Carolina (Fig. 1). This translates in practice to applications sorted by residential CD at the beginning of the admissions process, such that applicants are only competing against peers from the same district for scarce seats.[11] Three reviewers rate each application using the same admissions rubric. Importantly, the rubric is common to

[7] The state-supported boarding school thus predates institutional models such as SEED schools, which focus more explicitly on disadvantaged populations in

(*footnote continued*)
urban areas.

[8] The shorter grade span is common among selective public boarding schools. While the exposure period is shorter than usual, the application's timing during the middle of students' high school careers discourages selective attrition. Students apply in the fall of their high school sophomore year, with unsuccessful applicants unlikely to switch to private schools or move out of state due to the significant disruptions this would generate during an academically crucial period in which students embark on college applications.

[9] The school is a member of the National Consortium for Secondary STEM Schools, the core STEM school organization in the US comprised of state residential schools, self-contained schools, early college high schools, and schools within schools focused on STEM subjects. Many of residential public high schools have a specialized focus on mathematics and science fields and belong to the same consortium alongside non-residential elite schools such as the institutions examined by Abdulkadiroglu et al. (2014).

[10] Information on the application process is only relevant for application cohorts during the years covered in this study.

[11] Large disparities in the number of applicants exist across districts, such that some are significantly more over-subscribed than others and the probability of admissions varies. The legislative mandate allows for some flexibility in the number of bed spaces allotted per CD. The number of admitted students in a given district can vary from other districts by as much as 2.5% of the full cohort.

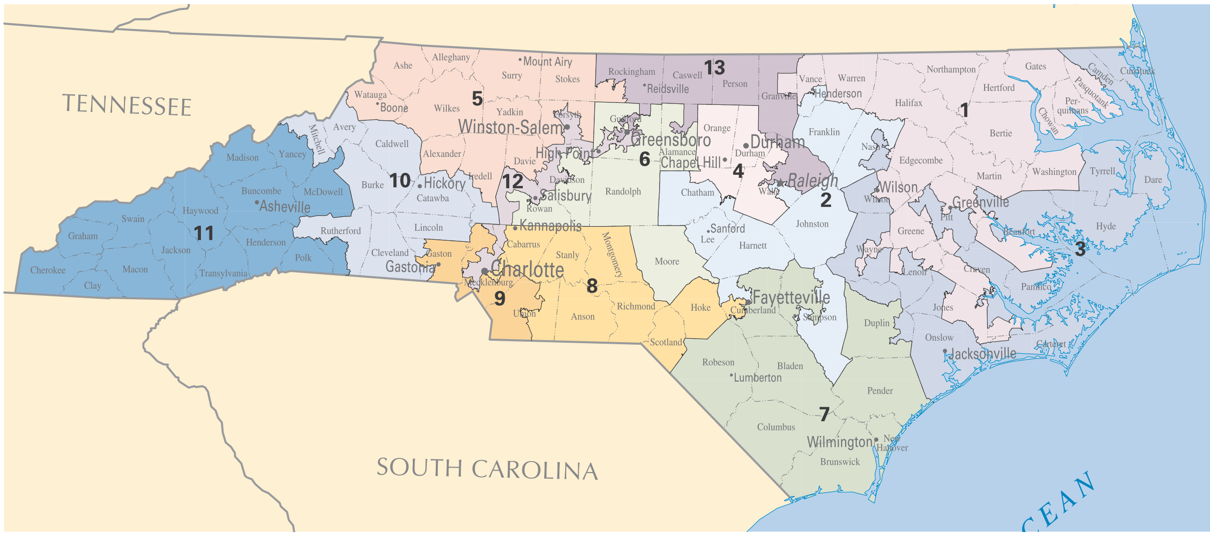

Fig. 1. Congressional district map: 2009 - 2012 application cohorts.

all districts and its contents should have remained constant throughout the course of this study. The criteria for evaluating candidates include academic rigor, quality of sending school, grades, extracurricular involvement, teacher and counselor evaluations, SAT scores, maturity, and interest in math and science. Each item in the rubric has a corresponding weight and reviewers allocate points up to a maximum of 40. The combined rating score across the three reviewers then ranges from 0 to 120. Reviewers use this score to rank all applicants from highest to lowest, admitting students scoring above a certain threshold. Applicants do not know where these cutoffs are set, as thresholds in previous years are not in the public domain, and applicants also do not have sufficient information to predict their placement. The lack of control over cutoff determination and the existence of a continuous admissions score provide the basis for causal identification. An important caveat is that application reviewers are tasked with ensuring diversity in the admitted student population. This means they can exercise discretion in removing high-scoring students or adding lower-scoring students to the final pool.[12] Scrutiny of admissions data and discussions with school personnel suggest that reviewers are more likely to selectively add students whose scores fall below the designated threshold.

3. Data

The analytic sample derives from three sources: applicant data from the selective school, administrative records for all North Carolina public and charter school students, and postsecondary outcomes from the National Student Clearinghouse. Applicant data spans four cohorts from 2009 to 2012. Data elements include the individual's birth month and year, ethnicity, gender, sending school, residential congressional district, county, zip code, and entering SAT scores in 10th grade.[13] In addition to applicant socio-demographics and academic performance, the files also list the combined admissions score from 0 to 120, application status (admitted, waitlisted, rejected), and whether the student enrolled and graduated from the institution.

Administrative records provided by the North Carolina Education Research Data Center (NCERDC) includes detailed information on students, teachers, and schools across all public and charter K-12 institutions. This data helps to contextualize applicants' sending school environments as well as track their outcomes near the end of secondary schooling. Student-level data on socio-demographic characteristics, standardized End-of-Grade (EOG) test scores for grades 3–8, classroom membership, and academic transcripts permit the construction of variables including class rank and average peer achievement.

I examine three sets of outcome variables taken at the end of high school: SAT scores from the latest test administration, student major intentions, and the portfolio of student college applications.[14] SAT testing data includes raw and percentile scores in math and reading. Included alongside these scores are survey responses to a questionnaire on students' preferred majors. It holds particular interest for evaluating whether students are more likely to lean towards STEM subjects as result of selective school attendance.[15] Finally, College Board data also documents all colleges and universities to which students ever sent SAT score reports.[16] I link College Board codes to college characteristics in

[12] To ensure that reviewers' discretion does not affect the admissions score, I regressed the admissions index on 10-point bins of SAT math, reading, and writing scores, minority status, gender, year dummies, and sending school fixed effects. The coefficient on minority status is insignificant, suggesting that reviewers did not manipulate the admissions score on the basis of race or ethnicity.

[13] A redrawing of congressional districts took place in 2011. However, the admissions committee did not adopt the new district boundaries until 2013. So all cohorts in the present study were selected based on the pre-redistricting map.

[14] SAT data comes from 2011 - 2014 College Board files for seniors enrolled in North Carolina public high schools. The data only includes SAT math and verbal performance during the last time students took the SAT, and does not report how many times the student took the exam or scores corresponding to each test date.

[15] STEM is defined as one of the following: (1) Computer and Information Sciences and Support Services, (2) Engineering, (3) Engineering Technologies/ Technicians, (4) Biological and Biomedical Sciences, (5) Mathematics and Statistics, (6) Physical Sciences, or (7) Science Technologies/Technicians. Since the overwhelming share of high school students take the SAT exam during their junior or senior year, this variable captures STEM orientation near the end of students' high school careers. While student responses on this variable may not correspond to actual major choice in subsequent years due to changes in preferences, changes are unlikely to be driven by deliberate misrepresentation of major orientation. Questionnaire answers are made available to colleges, universities, and educational scholarship programs as part of the Student Search Service, so students have an incentive to honestly report major preferences in order to receive relevant information on educational and financial aid opportunities.

[16] Previous evidence shows that this list closely proxies for students' college application behavior. There is high correlation between the number of students who send scores to an institution and the actual applications received by the institution, with the greatest reliability observed among high SAT-achievers applying to large public universities (Card & Krueger, 2005).

the Integrated Postsecondary Education Data System (IPEDS), and construct measures for college selectivity using Barron's Admissions Competitiveness Index, which ranks colleges from non-competitive to most competitive. For ease of presentation I aggregate to three categories of most selective, very selective, and less selective colleges. The first includes Most Competitive and Highly Competitive colleges under the index, the second includes Very Competitive Colleges, while the third includes Competitive, Less Competitive, or Noncompetitive 4-year colleges alongside community colleges. I measure STEM-intensity as the share of 4-year undergraduate degrees awarded in science- and math-specific subjects.[17] The college application data thus enables the juxtaposition of individuals' college portfolios in terms of quantity, selectivity, and curricular focus.

Postsecondary access and completion outcomes are available after matching the full applicant dataset to National Student Clearinghouse records. NSC data includes the timing of student enrollment in each college, whether the student attained a degree, the degree type and date, and students' major at the time of degree completion among participating colleges and universities.[18] As with college application data, I construct the STEM-intensity of the first college students attend using IPEDS data on the share of undergraduate degrees conferred in STEM, and categorize institutional selectivity according to Barron's admissions index.

3.1. Linking applicants to outcome data

Matching procedures depend on the outcome data used. The match to college enrollment and attainment data in the National Student Clearinghouse used identifying information such as full names and birth dates.[19] All resulting matches are unique. In contrast, I did not have full names or exact birth dates for matching to high school outcomes and instead relied on birth month and year, individual demographics, sending schools, and enrollment status to link selective school applicants to the North Carolina administrative database. I significantly reduce the pool of possible matches to approximately 1000 students annually by restricting to students who ever applied to the selective school. These applicants are identifiable in the statewide dataset because they must take the SAT exam and submit a score report using the boarding school's unique 4-digit College Board code. The matching process begins with variables that are least subject to measurement error such as birth year, month, and sex. I further reduce the set of possible matches using additional variables on ethnicity, sending schools, and enrollment status. For all applicants with a unique pairing after each step, I remove the corresponding North Carolina administrative data record from the matching pool to prevent the observation from being linked to a different applicant. The process iterates until the number of matches cannot be reduced for any applicant.[20]

Descriptive statistics for the matched sample of applicants to North Carolina administrative data show a distribution that skews heavily towards single matches. Depending on the cohort, 87% to 89% of applicants match to a unique observation (Table A1). 92% to 96% have three or fewer matches per applicant. The existence of multiple matches per applicant and the possibility of mismatched observations is a source of measurement error that may attenuate estimates. I utilize a number of robustness checks, such as relying on only unique matches and applying bootstrap sampling and estimation methods to ensure results are not sensitive to the matching process. All subsequent analyses using this matched sample employ weights based on the inverse of the total number of matches for each applicant.

3.2. Summary statistics

The full applicant sample covers four enrollment cohorts from 2009 to 2012. It excludes applicants with missing congressional district information or admissions scores. Summary statistics in Table 1 shows a sample of 4085 individuals that skews slightly female, with African Americans taking up 12% and Hispanics taking up 7%. Asians are significantly over-represented at nearly one-quarter of the sample while black students are under-represented relative to public school 10th graders statewide over the same period. The combined entering SAT score during sophomore year is nearly 1200, reflecting a high-achieving applicant pool.[21] Of these, the school admits slightly over one third.

It is worthwhile to note that these numbers disguise substantial heterogeneity across congressional districts. Shares of under-represented minorities are highest in CD 1 in northeastern North Carolina and serpentine CD 12 across the western Piedmont Plateau with its historically high concentrations of minority communities. In contrast, white students dominate applications from CDs 5, 10, and 11 in western North Carolina. The boarding school is closest to CD 4 applicants, who travel an average of 22 miles to reach its campus. The next closest group is CD 13 applicants at 46 miles, while the majority of students live at least 100 miles from the selective boarding school.

In addition to demographic makeup and distance traveled, applicants from these broad geographical areas also diverge in academic preparation. Students applying from CD 1 average 547 in SAT math compared to 658 among CD 4 applicants. Academic performance in this case corresponds closely with economic advantage. Affluent communities are most prevalent in metropolitan areas such as the Triangle (Raleigh-Durham-Cary-Chapel Hill) and Charlotte, or CDs 4 and 9, respectively.[22] CDs also vary significantly in the number and success of applicants. CD 4 has more than five times the application volume of CD 12. Since the school is obligated to admit similar numbers of students across districts, admission rates are significantly lower for the most over-subscribed CD 4 (21%) relative to the least competitive CD 12 at 59%.

[17] This measure characterizes the institution's curricular focus. While STEM-intensity has a positive correlation with selectivity, more selective schools are often not the most STEM-intensive. For example, the only engineering program at the flagship University of North Carolina - Chapel Hill is Environmental Sciences and Engineering, while the less competitive North Carolina State University has more than two and a half times the number of STEM graduates as UNC-Chapel Hill. Note that I define STEM degrees as (1) Computer and Information Sciences and Support Services, (2) Engineering, (3) Engineering Technologies/Technicians, (4) Biological and Biomedical Sciences, (5) Mathematics and Statistics, and (6) Physical Sciences.

[18] Colleges in recent years are increasingly providing information on field of study, and as of 2012 these majors are standardized via Classification of Instructional Programs (CIP) codes (Dynarski, Hemelt, & Hyman, 2015).

[19] While potential sources of measurement error within NSC data documented in the literature include incomplete coverage rates and FERPA blocks, measurement error should not differ systematically from one side of the admissions threshold to the other since information provided by the selective school to the NSC is the same for all applicants. For an extensive discussion of these sources of measurement error, see Dynarski et al. (2015).

[20] A detailed documentation of the order in which supplementary variable are used to narrow down matches and the iteration process is provided by Appendix A.

[21] As another reference point, the average 8th grade math and reading standardized test scores among applicants are 1.8 and 1.4 standard deviations above the statewide mean, respectively. Similarly, their performance in the End-of-Course Algebra I test is also 1.8 standard deviations above average. In terms of demographic characteristics, applicants are less likely to be under-represented minorities and more likely to be Asian compared to the statewide sample of 10th grade students who eventually take the SAT exam. The statewide sample comprises 26% African Americans1, 6% Hispanics, and 4% Asians.

[22] In summary statistics not shown, economic disadvantage is highest among CD 1 applicants at 25%, although rates are also high in western rural communities.

Table 1
Summary statistics - applicants.

	All	CD 1	2	3	4	5	6	7	8	9	10	11	12	13
Socio-demographics														
Female	0.53	0.56	0.55	0.55	0.56	0.46	0.46	0.57	0.58	0.53	0.45	0.51	0.46	0.52
Black	0.12	0.35	0.14	0.08	0.10	0.05	0.08	0.15	0.15	0.12	0.05	0.02	0.23	0.15
Hispanic	0.07	0.07	0.08	0.07	0.05	0.05	0.04	0.10	0.10	0.06	0.08	0.04	0.09	0.08
American Indian	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.01	0.01	0.00
Asian	0.24	0.06	0.08	0.10	0.46	0.15	0.22	0.09	0.29	0.39	0.10	0.09	0.28	0.19
White	0.53	0.48	0.65	0.71	0.36	0.72	0.63	0.60	0.41	0.40	0.73	0.80	0.36	0.53
Distance to school (miles)	127	138	77	192	22	167	84	179	143	201	217	345	154	46
Academic performance														
Entering SAT math	617	547	587	617	658	612	621	575	589	638	620	612	613	604
Entering SAT verbal	579	523	559	579	612	574	574	544	560	593	580	583	575	569
Admissions														
N applied	4085	265	295	321	944	254	273	203	241	374	267	197	168	283
% admitted	0.36	0.45	0.37	0.41	0.21	0.36	0.45	0.50	0.41	0.31	0.38	0.43	0.59	0.36
Average cutoff	98	89	95	97	108	97	95	90	94	99	98	95	88	94

Notes: The applicant sample includes 4085 applicants spanning the 2009–2012 cohorts. The sample size for entering SAT scores and ZIP attributes is 4055. There are 3737 applicants with non-missing sending school quality, and 3741 applicants with non-missing sending school curriculum. Entering SAT scores are taken between October and January of applicants' sophomore year. All distances are computed using latitude and longitude geocode coordinates for the centroid of students' home zip codes.

4. Empirical approach

4.1. Fuzzy regression discontinuity

A challenge to estimating the effects of selective school attendance is endogenous student selection. This paper relies on an admissions process that generates plausible local random assignment into the selective school when accounting for the underlying admissions score. Reviewers base admissions decisions on applicants' score ranking relative to peers in the same CD and cohort. They also consider student demographics to ensure a diverse student body, implying the model should define separate cutoffs by minority status in addition to CD and cohort. Accordingly I fit the following reduced-form specification:

$$Y_{idt} = \pi_0 + \pi_1 D_{idt} + (1 - D_{idt})f_0(s_{idt} - \tilde{s}^m_{dt}) + D_{idt}f_1(s_{idt} - \tilde{s}^m_{dt}) + \phi^m_{dt} + \epsilon_{idt} \tag{1}$$

In this model, i indexes individuals, d the congressional district of residence, t the admissions cohort, and m minority status. Y_{idt} is an outcome variable such as SAT math scores. Treatment status D_{idt} assumes a value of 1 if the student's score s_{idt} surpasses the predicted admissions cutoff for a given district, cohort, and minority status $\tilde{s}^m_{dt}$: $D_{idt} = I(s_{idt} > \tilde{s}^m_{dt})$. The minority status variable m distinguishes between two groups: under-represented minorities who are African-American, Hispanic, or Native American and all remaining individuals. ϕ^m_{dt} is a set of district by cohort by minority intercepts. The choice of fixed effects addresses the possibility of district-specific changes in applicant characteristics over time. The composition of minority or non-minority students applying from a given district can vary for multiple reasons, such as fluctuations in information availability on the selective high school and changes in the curriculum of sending institutions relative to the school.

The specification flexibly controls for the running variable of $s_{idt} - \tilde{s}^m_{dt}$, the distance between a student's own admission score and the predicted cutoff. Interactions with D_{idt} permit the slope to vary on either side of the cutoff. I rely on nonparametric and parametric specifications as complementary estimates. The nonparametric model controls for a linear function of the running variable, computes optimal bandwidths following the methods of Imbens and Kalyanaraman (2012), and uses a tent-shaped edge kernel centered around the admissions cutoff denoted by $K(s_{idt} - \tilde{s}^m_{dt})$:

$$K(s_{idt} - \tilde{s}^m_{dt}) = \mathrm{I}\left\{\left|\frac{s_{idt} - \tilde{s}^m_{dt}}{h}\right| \le 1\right\}\cdot\left(1 - \left|\frac{s_{idt} - \tilde{s}^m_{dt}}{h}\right|\right) \tag{2}$$

h represents the optimal computed bandwidth. Parametric specifications use the cubic polynomial $\alpha_{k1}(s_{idt} - \tilde{s}^m_{dt}) + \alpha_{k2}(s_{idt} - \tilde{s}^m_{dt})^2 + \alpha_{k3}(s_{idt} - \tilde{s}^m_{dt})^3$ with $k \in \{0, 1\}$. Results are presented using different degrees of polynomials to ensure that π_1 is robust to alternative functional form assumptions.

The use of D_{idt} underscores the 'fuzzy' part of the regression discontinuity. Scores and CD-cohort intercepts do not predict admissions with certainty. Selection and sorting behavior leads to possible correlations between enrollment eligibility and the error term. Furthermore, students who were accepted may not always enroll, and those who decide not to attend the school may differ systematically from enrollees in their individual characteristics, counterfactual high schools, and family backgrounds. Given these considerations, I rely on an indicator for exceeding score cutoffs to randomize at the intent-to-treat level. Estimates quantify the effect of becoming *eligible for enrollment* in the selective school. I complement these core findings with 2SLS estimates using an indicator for exceeding the score threshold as an instrument for admissions. This quantifies the causal impact of admissions to the selective boarding school.

4.2. Predicted cutoffs $\tilde{s}^m_{dt}$ and covariate smoothness

Eligibility status depends on exact cutoff scores which are not observed. However, it is possible to create assignment rules and empirically test for their accuracy in predicting admissions outcomes. Discussions with school personnel reveal that the selection committee ranks applicants from each congressional district from highest to lowest. The committee then admits the highest scoring applicants, subject to additional considerations of demographic characteristics. To the extent that racial and ethnic diversity is an objective, this translates effectively to different admissions thresholds for minority and non-minority applicants for a given district and cohort. I define cutoffs as the admissions score that maximizes the number of applicants that are correctly classified as admitted if their scores exceed the threshold, or not admitted for those at or below the threshold. This approach minimizes the combined probability of type I and II errors and yields a cutoff for each cohort-CD-ethnic group.

Table 2 shows that this set of predicted cutoffs correctly classify the

Table 2
Accuracy of predicted cutoffs.

	Cohort				
Congressional District	2009	2010	2011	2012	Total
1	0.98	0.92	0.98	0.93	0.95
2	0.88	0.91	0.89	0.93	0.91
3	0.89	0.98	0.97	0.92	0.94
4	0.87	0.95	0.94	0.96	0.93
5	0.96	0.93	0.98	0.87	0.93
6	0.96	0.96	0.97	0.95	0.96
7	0.90	0.89	0.98	0.96	0.93
8	0.94	0.96	1.00	0.95	0.96
9	0.89	0.96	0.95	0.94	0.94
10	0.85	0.97	0.97	0.87	0.92
11	0.91	0.98	0.98	0.93	0.95
12	0.95	0.90	0.93	0.86	0.91
13	0.94	0.91	0.93	0.96	0.93
Total	0.91	0.94	0.96	0.93	0.93

Notes: Admissions status is correctly predicted if distance from the predicted cutoff is equal or less than 0 and the applicant is not admitted, or if distance from the predicted cutoff is positive and the applicant is admitted.

admissions status of 93% of all applicants.[23] In a visual example, Fig. 2 shows the classification of 2010 non-minority applicants across districts. Each dot represents an applicant and red dots denote misclassified observations. Applications above the threshold are much more likely to be admitted than students below the threshold, such that there is a jump in treatment probability at $\tilde{s}^m_{dt}$. I then pool all applicants across districts, cohorts, and minority status and examine the relationship between the running variable and admission probabilities. Fig. 3 shows a discontinuous increase of over 60 percentage points in the likelihood of admissions when an applicant exceeds the cutoff. One reason for the fuzzy discontinuity is that predicted cutoffs do not perfectly measure the true thresholds. Another reason is discretion on the part of admissions officers. They can add (remove) individuals to the acceptance list even when students are below (above) the cutoff.[24]

If treatment is as good as randomized in the neighborhood of the admission score cutoff, then the distribution of observed baseline characteristics should be smooth around the threshold. Graphs of baseline covariates enable a direct test of the identifying assumptions implicit in a regression discontinuity framework. Fig. 4 shows student characteristics relative to the admissions score, with a linear fit line on either side of the pooled sample. Covariates include student ethnicity, SAT math and verbal scores at the time of application, and characteristics of residential zip codes and sending high schools.

Students' entering SAT math and verbal scores are predictably increasing in admission scores, with no observable discontinuities around predicted thresholds. The shares of applicants belonging to an underrepresented minority group are smooth across the admissions cutoffs, although the covariate exhibits more variation relative to earlier measures of academic performance. There are also no visual discontinuities by the share of households in applicants' residential zip codes that are rural, or in the quality of sending schools as measured by average student math scores or pupil-teacher ratios. Regressions in Appendix Table B1 similarly show no evidence of statistically significant differences around the threshold across individual characteristics. To further verify covariate balance and rule out manipulation in the admissions process, I use the density test outlined in McCrary (2008). The density distribution shown in Fig. B1 is continuous across the threshold.

5. Results

5.1. End of high school outcomes

Fig. 5 depicts the relationship between the running variable and multiple outcomes measured near the end of students' high school careers. The first two graphs show SAT math and verbal scores, while the remainder examines a relatively under-studied set of outcomes in the literature: students' major and postsecondary intentions as realized via college applications.[25] A discontinuous increase is apparent at the cutoff for SAT math percentile. In contrast, graphs corresponding to SAT verbal percentiles and a consistent intention to major in STEM show continuity through fitted lines.[26] The math score increase does not appear to extend to verbal ability or translate to greater interest in pursuing STEM subjects for the pooled applicant sample.

College application outcomes focus on the total number of postsecondary institutions the student applied to, their selectivity, and STEM-intensity as measured by the share of all Bachelor's degrees awarded in STEM fields. Evidence suggests that selective school eligibility did not have a measurable effect on the total number of applications. Next I examine college selectivity outcomes documenting the share of applications to (1) the most selective, (2) very selective, and (3) less selective institutions. The first group includes Ivy League universities, elite private non-for-profit colleges, state public flagships, and top liberal arts colleges. Only five universities in North Carolina meet the ranking criteria alongside 95 universities elsewhere in the U.S.: Duke University, Davidson College, Elon University, the University of North Carolina - Chapel Hill, and Wake Forest University. The second tier of institutions includes selective four-year institutions such as Purdue University, Brigham Young University, Spelman College, and the former land-grant North Carolina State University. Remaining 4-year institutions and open enrollment community colleges are classified under the final tier. Visual inspection suggests a possible increase in applications to the most selective institutions and a discontinuous drop in applications to less selective institutions.[27] Finally, I examine

[23] This method of predicting admissions cutoffs yields very similar results compared to other data-driven approaches. For example, I follow Hoekstra (2009) by partitioning into district, cohort, and minority status-specific categories and within each subgroup regress admission status on an indicator for whether the applicant's admissions score is higher than a given cutoff. After repeating this procedure across the range of possible cutoffs, I select the cutoff for which the R^2 is the highest. The correlation between the resulting cutoffs and original estimates across the full sample is 0.98. The new approach also correctly classifies 93% of all applicants' admissions statuses.

[24] Very few students who are not admitted outright eventually enroll. The selective school maintains an admissions waitlist and students can come off the waitlist if spots become available before Labor Day weekend during the fall semester of 11th grade. Only 4 students on the waitlist eventually enroll in the full applicant sample. Since their admission scores are at or below the cutoffs, they are classified as not treated in the analyses.

[25] Test scores and major intentions are measured at the time of the latest SAT administration. The majority of high school students take the SAT exam during their senior year in the analytic sample, while the remaining take the exam during their junior year. As such these outcomes capture test scores and major orientation near the end of students' high school careers.

[26] STEM major intention directly tests for the effect of attending a science- and math-intensive school on students' predilections for these fields. One justification for the existence of selective institutions is that they increase the ranks of science, technology, engineering, and math (STEM) professionals vital to a knowledge economy. STEM academic tracks also bolster students' future economic well-being. Limited evidence exists, however, on the causal effect of science- and math-oriented schools on students' STEM trajectories. By estimating the causal impact on major and postsecondary STEM orientation, this study is well suited to comment on one of the primary goals of STEM institutions.

[27] To ensure that these results are not sensitive to the way college selectivity is defined, I use alternative classifications based on the distribution of SAT scores among enrolled students. Results under this definition are consistent with original findings.

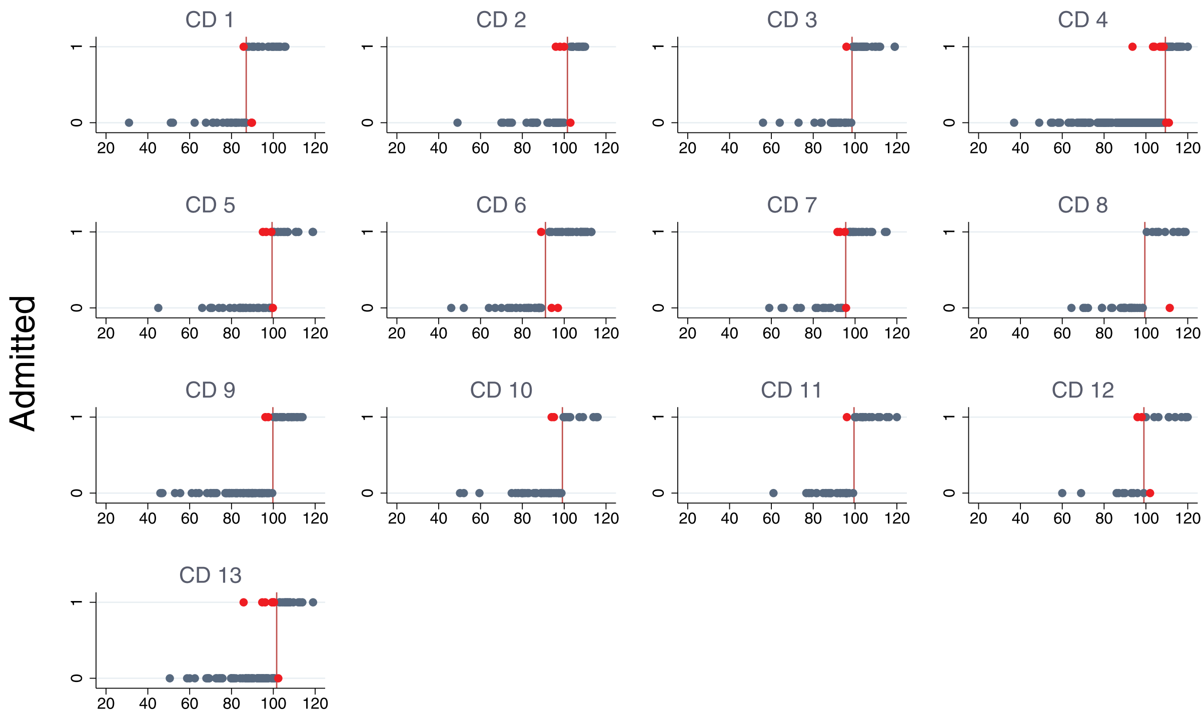

Fig. 2. Predicted cutoffs for non-minority applicants in 2010.

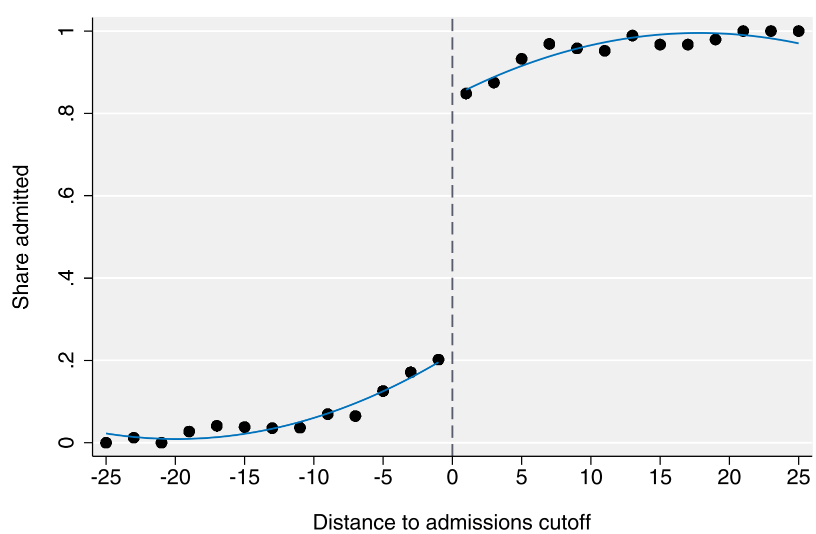

Fig. 3. First stage relationship *Note:* This figure pools all applicants across districts, cohorts, and minority status to examine the relationship between the running variable and admission probabilities.

whether an individual is more likely to apply to STEM-intensive institutions as measured by the share of undergraduate degrees conferred in these majors. The scatter plot strongly suggests a discontinuous increase in STEM-intensity among institutions to which the students applied.

Regression results in Table 3 confirm graphical evidence using both parametric and nonparametric specifications. Students who exceed the admissions cutoff increase their SAT math performance by 2 percentile points, while the coefficients on SAT reading percentile are insignificant. The analogous 2SLS estimate on the causal effect of selective boarding school admissions induced by surpassing the cutoff is 3 percentile points (Table D1). While modest, the math score effect still contrasts with US-based selective school studies finding null effects across types of standardized tests (Abdulkadiroglu et al., 2014; Barrow et al., 2016).[28] An exception is the sizable 0.2σ impact in math scores of attending boarding schools in the US and France (Behaghel et al., 2017; Curto & Fryer, 2014).

Consistent with graphical evidence, students are not more likely to track into STEM fields as the result of enrolling in the selective school. Relatively large standard errors accompany the measurement of major

[28] In Abdulkadiroglu et al. (2014), few test score effects under the Massachusetts Comprehensive Assessment System and the New York Regents exam are significantly different from zero, while the pooled sample across years and grades are more precisely measured and close to zero or negative. Estimates of the effect of being admitted to a selective Chicago-area secondary school on 11th grade ACT test scores are statistically insignificant (Barrow et al., 2016).

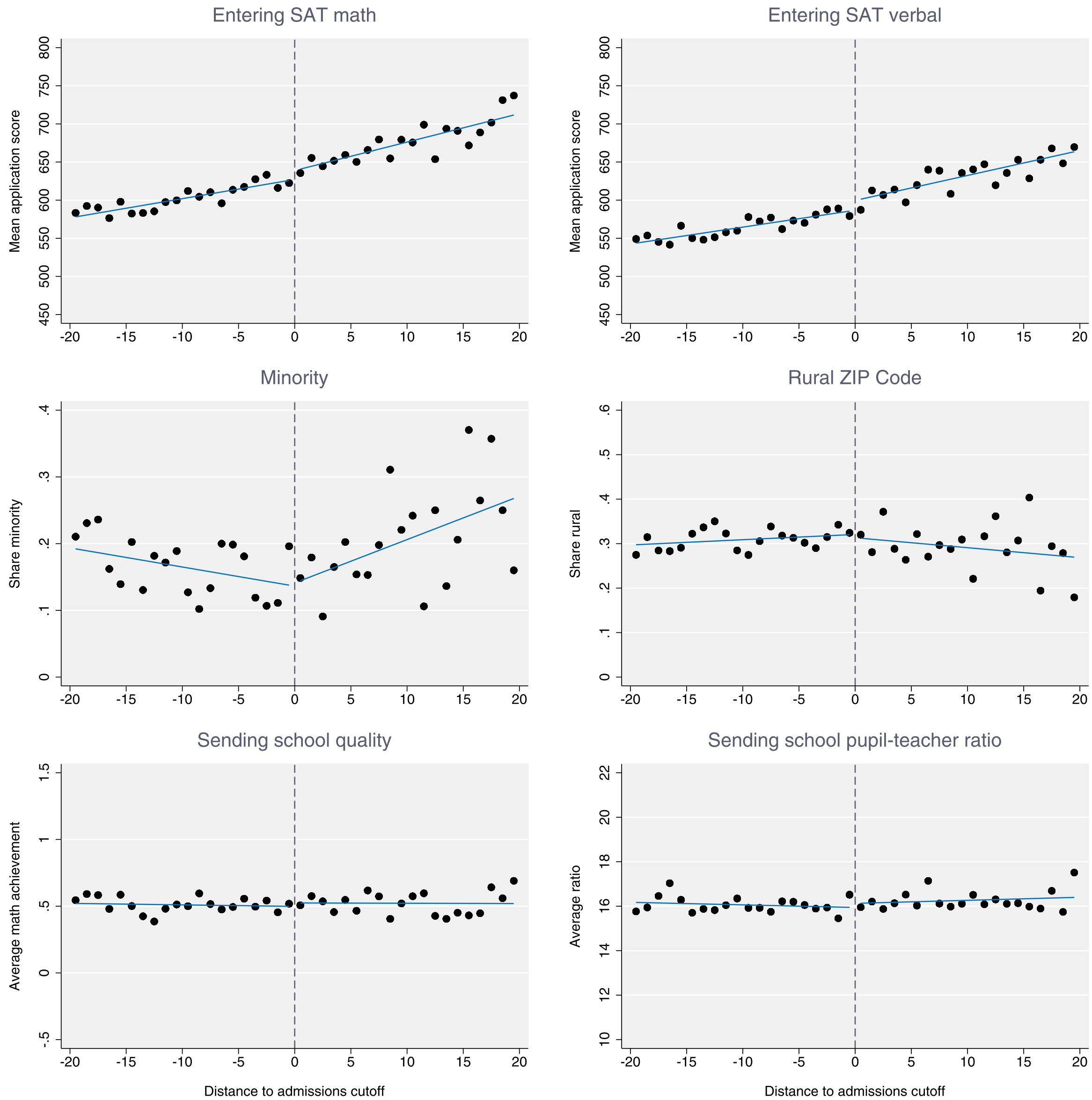

Fig. 4. Covariate smoothness *Note:* This figure shows the existence of possible discontinuities across the admissions threshold, using a pooled sample of applicants across cohorts and congressional districts. SAT math and verbal scores refer to student performance at the time of application to the selective school. Minority status includes African-American, Hispanic, and Native American applicants. Rural zip code is defined as the share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores.

intentions, and I cannot exclude the possibility of measurement error leading to attenuated results. Eligibility for selective school enrollment also does not induce students to increase the volume of college applications. Meanwhile, there is a significant decrease in applications to less selective colleges of over 3 percentage points. Relative to a baseline of 17%, students predicted to enroll in the selective school are applying to nearly one-fifth fewer less selective colleges (Table C1). An alternative interpretation in terms of total applications is that students apply to 0.14-0.15 fewer colleges in this tier. Curricular emphasis of the college application portfolio also shifts, as both parametric and nonparametric specifications point to an increase in the institutions' STEM intensity. The effect of 1.7–1.8 percentage points is equivalent to an increase of 8%. The magnitudes of these changes are sizable considering a condensed exposure window of no more than two years.

5.2. Postsecondary outcomes

In light of measurable test score and college application effects, the question remains of whether these changes carry over to postsecondary outcomes. Fig. 6 turns to National Student Clearinghouse data to address the longer-term ramifications of selective school exposure. Students above the threshold are more likely to enroll in any 2-year or 4-year institution as identified in the NSC database.[29] Conditional on

[29] Applicants from the 2009–2012 matriculating cohorts are matched to NSC data through Fall 2017. To ensure that these results are not driven by differential time frames for enrollment across cohorts, I rerun the analyses after restricting the postsecondary enrollment window to within one year of expected high school graduation. Results are very consistent using the alternative outcome definition.

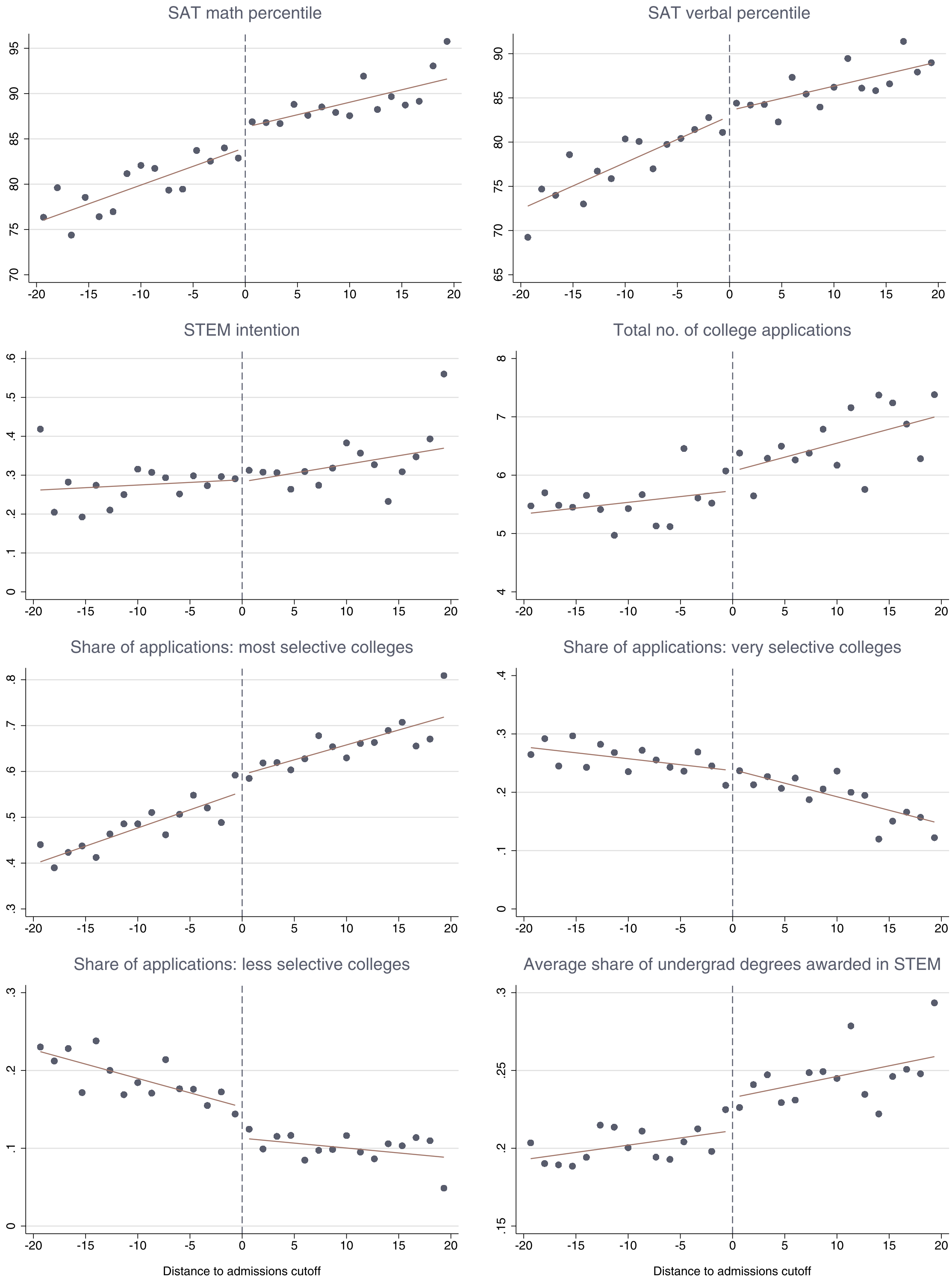

Fig. 5. End of high school outcomes *Note:* This figure shows outcomes relative to distance to the admissions cutoff. STEM intention takes a value of 1 if the individual listed only STEM majors out of five possible major choices. Most selective colleges are ranked as Most Competitive or Highly Competitive under Barron's admissions index. Very selective colleges correspond to Very Competitive colleges, while less selective institutions include Competitive, Less Competitive, and Noncompetitive colleges as well as community colleges.

enrollment, the opportunity to attend the selective school visibly alters the type of college initially chosen by the student. I examine institutional selectivity defined using the same parameters from Barron's admissions index, and find a marked decrease in the likelihood of enrolling in a less selective college after high school graduation. This suggests that the selectivity of students' college application portfolios carried over to their eventual enrollment decisions. Upon closer inspection, one factor driving the decrease is a discontinuous drop in the share of students who *ever* enrolled in a 2-year institution during the observed period.

Table 3
Reduced form estimates - end of high school outcomes.

	Parametric	Nonparametric
Dependent variable	(1)	(2)
SAT math percentile	2.119*	1.708**
	(1.086)	(0.757)
	⟨ 3441 ⟩	⟨ 1580 ⟩
SAT reading percentile	0.715	0.236
	(1.353)	(0.980)
	⟨ 3442 ⟩	⟨ 1541 ⟩
STEM intention	0.019	0.004
	(0.047)	(0.031)
	⟨ 3442 ⟩	⟨ 1753 ⟩
Total no. of college applications	0.048	0.238
	(0.351)	(0.213)
	⟨ 3366 ⟩	⟨ 1929 ⟩
Share of applications: most selective colleges	0.023	0.020
	(0.026)	(0.019)
	⟨ 3226 ⟩	⟨ 1755 ⟩
Share of applications: very selective colleges	0.009	0.010
	(0.022)	(0.016)
	⟨ 3226 ⟩	⟨ 1871 ⟩
Share of applications: less selective colleges	−0.030*	−0.034***
	(0.016)	(0.011)
	⟨ 3226 ⟩	⟨ 1634 ⟩
Average share of BAs awarded in STEM	0.017*	0.018**
	(0.010)	(0.007)
	⟨ 3226 ⟩	⟨ 1588 ⟩

Notes: Table shows the effects of becoming eligible for selective school enrollment for the pooled sample. Column (1) reports estimates from the parametric model that flexibly allows for different cubic functions of distance to the cutoff on either side of the cutoff. Column (2) reports nonparametric estimates using optimal bandwidths and a linear control of distance to the cutoff. The outcomes of SAT math and verbal percentile refer to student performance at the latest SAT administration near the end of high school. STEM intention takes a value of 1 if the individual listed only STEM majors out of five possible major choices. Total institutions refer to the number of postsecondary institutions receiving a SAT score report by the end of the individual's high school career. Most selective colleges are ranked as Most Competitive or Highly Competitive under Barron's admissions index. Very selective colleges include Very Competitive college, while less selective institutions include colleges labeled as Competitive, Less Competitive, and Noncompetitive as well as community colleges. Average share of STEM degrees refers to the mean share of bachelor's degrees awarded that belong to a STEM field among all institutions receiving score reports. All results are weighted by the inverse of the number of matches. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. Sample sizes in brackets. Applicants with admission scores more than 40 points removed from the cutoff are excluded. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Another postsecondary attribute that echoes patterns from college applications is curricular focus. Selective school eligibility appears to increase the STEM-intensity of colleges that students attend. I explore whether this affects time to graduation and major choice. The final two graphs suggest no effects on students' likelihood of graduating in four years and graduating with a science- and math-focused Bachelor's during the same period.[30]

Table 4 presents coefficients corresponding to each graph. Applicants who clear the admissions threshold are 5–6 percentage points more likely to enroll in any college.[31] Given that 92% of non-admits in the neighborhood of the thresholds are enrolling in postsecondary institutions, the treatment effect pushes college enrollment close to the upper limit. In addition to increased enrollment, applicants with scores exceeding the cutoff are between 5 and 8 percentage points less likely to attend an open enrollment or less selective college. Even the more conservative nonparametric estimate corresponds to a one-fifth decrease from the baseline rate (Table C1). Strikingly, some of this effect is driven by reduced enrollment in open enrollment community colleges of 6–8 percentage points. This suggests that a notable and possibly overlooked benefit of selective schools may not be shepherding academically gifted students into elite Ivy League schools, but rather preventing them from undermatching at less selective 4-year institutions and even community colleges.

Finally, selective school eligibility leads to a 2–3 percentage point increase in the STEM-intensity of students' first colleges, which does not appear to translate into changes in college graduation broadly or increases in the share that graduates with a STEM degree. Note that caution is advisable when interpreting graduation outcomes due to limitations in National Student Clearinghouse data availability and coverage. The analyses are somewhat under-powered due to 4-year attainment outcomes only being available for the earliest three cohorts. Degree attainment reporting in the NSC is also low compared to the comprehensive coverage of enrollment records (Dynarski et al., 2015). This source of measurement error can in part explain why enrollment outcomes are not yet translating to higher graduation rates.

5.3. Heterogeneous treatment effects

A unique advantage of this study is the relatively heterogeneous population of students generated by the district-based admissions rule. Multiple district-based cutoffs ensure that marginal students enroll not only near the bottom of the selective school class but across the ability distribution. This type of admissions context permits evaluating the selective school's impact for a broader range of students. I measure student background along several dimensions: individual demographics, neighborhood characteristics, sending high school quality, and academic achievement. Variables corresponding to the first category include gender and ethnicity. The focus on urban vs. rural neighborhoods at the zip code level reflects the possibility that denser neighborhoods have more high-quality alternative schools, such that enrolling in the selective school is a smaller 'treatment' relative to students with fewer schooling options. Sending school quality is measured by average student math scores, while achievement grouping depends on the individual's SAT math performance at the time of selective school application.[32] Separate models are run for each set of variables, interacting indicators such as low and high sending school quality with treatment status.

Table 5 shows that SAT math gains are most prominent among students with lower baseline math achievement or relatively disadvantaged groups. The opportunity to enroll in a selective school yields a 2–3 percentile point gain in SAT math scores among female students and those from more rural neighborhoods or with lower baseline achievement, while males, urban, and higher-achieving peers are no better off. The effect is particularly salient for minority students at 5 percentile points while non-minorities observe no discernible effects. Formal tests of differences between paired subgroups show that

[30] I restrict analyses of the final two outcomes to the 2009–2011 matriculating cohorts. They are expected to graduate from high school and enroll in college in 2011–2013, while the 2012 cohort only has 3 years of NSC data.

[31] By comparison, Abdulkadiroglu et al. (2014) find no effect of Boston exam school offers on college attendance, 4-year college attendance, or attendance at any competitive or highly competitive college. Similarly, Dobbie and Fryer (2014) find little evidence of exam school offers in New York City on college enrollment and attainment.

[32] The sample is by construction evenly divided by neighborhood rurality, sending school quality, and achievement sores. Rural status is defined by the share of all households within a zip code located in rural areas. So-called urban zip codes contain 0 - 16% of rural households. School quality is based on average 8th grade EOG math scores of 10th grade students at the sending school. The lower half of sending schools by quality has average scores that are up to 0.5σ above the statewide average, while the upper half ranges from 0.5 to 2.3σ. The lower half of entering 10th grade SAT math scores averages 550 while the upper half averages 700.

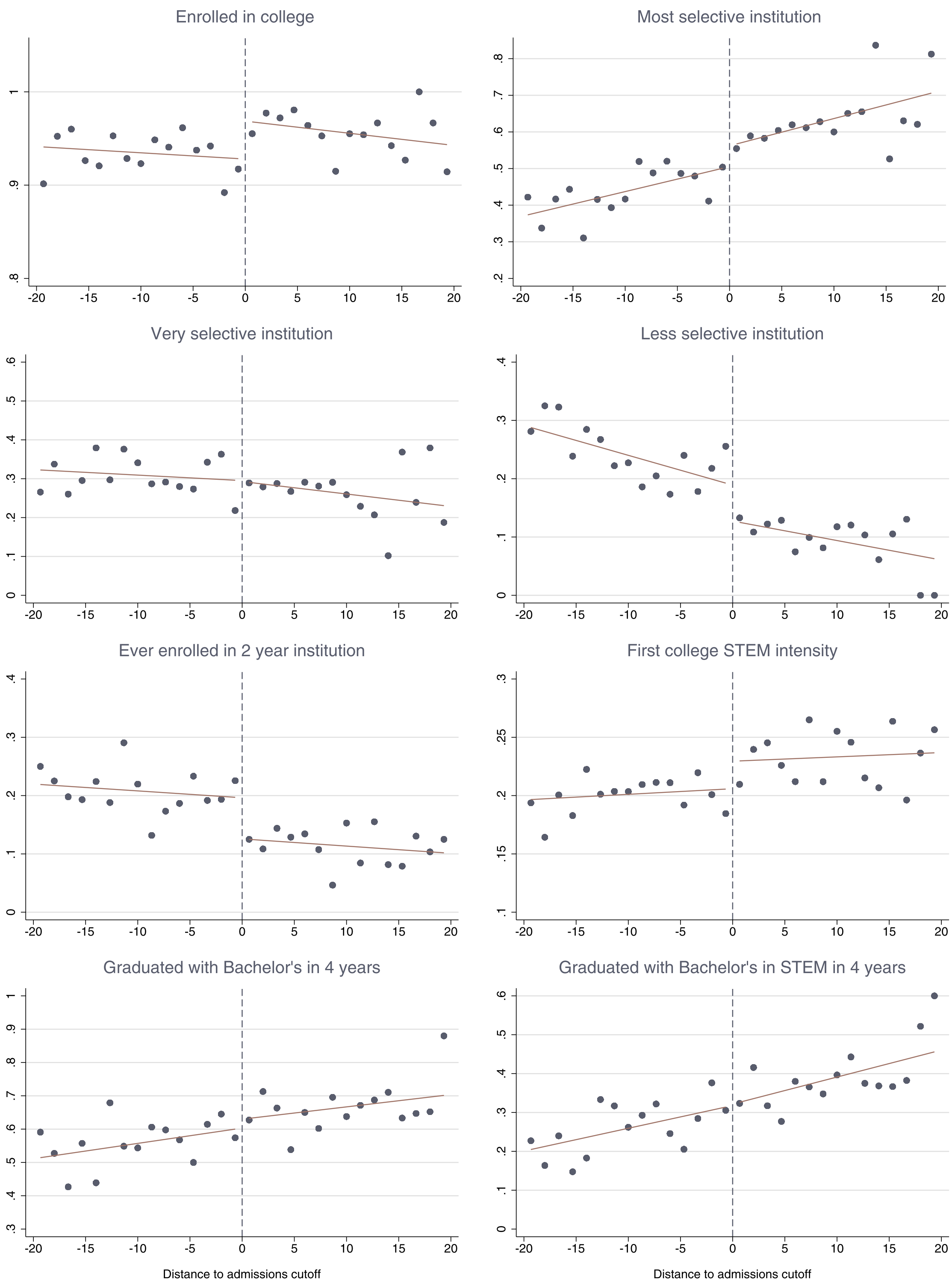

Fig. 6. Postsecondary outcomes *Note:* This figure shows outcomes relative to distance to the admissions cutoff. College enrollment takes a value of 1 if the individual is matched to NSC records up to Fall 2017. Most selective colleges are ranked as Most Competitive or Highly Competitive under Barron's admissions index. Very selective colleges correspond to Very Competitive colleges, while less selective institutions include Competitive, Less Competitive, and Noncompetitive colleges as well as community colleges. STEM intensity refers to average share of Bachelor's degrees awarded that belong to a STEM field.

effects among more disadvantaged groups are 2–4 percentile points significantly higher than peers. The magnitudes translate to approximately one-quarter up to over one-third of existing disparities, suggesting that selective school exposure bridges existing gaps (Table C1). In contrast to gains for SAT math, results show only null effects for SAT verbal scores. These patterns of concentrated math gains among high-achieving disadvantaged students that are not replicated for reading are consistent with findings of the effect of French boarding schools and other studies on secondary education (Behaghel et al., 2017). In another boarding school context, poor minority children in US-based

Table 4
Reduced form estimates – postsecondary outcomes.

	Parametric	Nonparametric
Dependent variable	(1)	(2)
Enrolled in college	0.057*	0.048**
	(0.032)	(0.022)
	⟨ 3929 ⟩	⟨ 3403 ⟩
First college was most selective	0.022	0.030
	(0.048)	(0.031)
	⟨ 3705 ⟩	⟨ 3304 ⟩
First college was very selective	0.046	0.020
	(0.045)	(0.026)
	⟨ 3705 ⟩	⟨ 3276 ⟩
First college was less selective	−0.077**	−0.047**
	(0.035)	(0.023)
	⟨ 3705 ⟩	⟨ 3682 ⟩
Ever attended 2-year college	−0.082*	−0.064**
	(0.043)	(0.026)
	⟨ 3705 ⟩	⟨ 3507 ⟩
Share of BAs awarded in STEM in first college	0.029*	0.018**
	(0.015)	(0.009)
	⟨ 3488 ⟩	⟨ 3287 ⟩
Graduated with Bachelor's in 4 years	0.039	0.013
	(0.067)	(0.042)
	⟨ 2755 ⟩	⟨ 2683 ⟩
Graduated with Bachelor's in STEM in 4 years	−0.002	0.003
	(0.049)	(0.032)
	⟨ 2755 ⟩	⟨ 2732 ⟩

Notes: Table shows the effects of becoming eligible for selective school enrollment for the pooled sample. Column (1) reports estimates from the parametric model that flexibly allows for different cubic functions of distance to the cutoff on either side of the cutoff. Column (2) reports nonparametric estimates using optimal bandwidths and a linear control of distance to the cutoff. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. Sample sizes in brackets. Applicants with admission scores more than 40 points removed from the cutoff are excluded. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

SEED schools improved significantly in both math and reading (Curto & Fryer, 2014). The differential effects by math and verbal skills seen here may reflect the curricular focus of a science- and math-focused institution.

Results on college application behavior parallel the findings on SAT math scores. Females, rural applicants, lower-achieving students and those from lower quality sending schools apply to 5–6 percentage point fewer less selective colleges. This represents a one-quarter decrease from average application rates of 20–24% in this selectivity bracket. The relative magnitude is largest for minority students at 10 percentage points. As before, these effects depart significantly from the mostly null effects among advantaged peers.[33] Significant treatment effects are most apparent for groups identified in the literature to be especially susceptible to academic undermatch (Hoxby & Avery, 2013; Smith et al., 2013).

Table 6 shifts from college application behavior to actual enrollment.[34] Notably, gains in college enrollment appear to accrue to students exceeding the admissions cutoff who are non-minorities, have higher prior achievement, or from higher-achieving sending schools. Formal tests show that in most cases there are no statistically significant differences across subgroups. In contrast to the results on college enrollment, only students from more disadvantaged groups meaningfully changed the selectivity profile of their first college attended. For one, minority applicants surpassing the admissions cutoff are 15 percentage points less likely to attend a less selective college directly after high school. This decrease is matched by a comparable jump in the likelihood of attending a very selective 4-year institution, with in-state examples such as North Carolina State University and UNC - Asheville. Analogous decreases for rural applicants and those from lower quality sending schools range from one-quarter to one-third of the baseline. A clue into the nature of these decreases is evident in sizable drops in the likelihood of ever attending a community college during students' observed college careers. Around one-quarter of students in disadvantaged groups fall into this category despite their records of academic achievement. Evidence suggests that the elite residential high school eases the likelihood that such students will eventually settle for a 2-year college.

While college selectivity effects are concentrated among certain groups, the same cannot be said for major orientation, curricular focus, and college persistence. Table 5 finds no evidence that attending the selective school affects STEM major orientation for any subgroup. Meanwhile, significant coefficients of similar magnitude are found when examining individuals' propensity to apply to STEM-oriented universities. Institutional STEM intensity increased 1.6 to 1.9 percentage points across almost all subgroups. Males and minority students were the only groups that did not experience a significant increase, although the possibility of a treatment effect cannot be ruled out given the standard error sizes. Similarly, increases in the STEM-intensity of students' undergraduate institutions are statistically indistinguishable from one group to another. The majority of treatment effects are around 2 percentage points. While students may be more likely to attend a rigorous science- and math-focused campus such as North Carolina State University over a less competitive or STEM-intensive institution, this shift has no apparent effect on 4-year college graduation rates. Greater availability and coverage of NSC attainment data may be necessary to ensure that observed outcomes are driven by actual persistence patterns and not limited coverage of graduation data.

6. Mediating factors

Patterns of heterogeneous treatment effects give rise to the question of why some benefit more than others. This prompts a deeper inquiry into the nature of the treatment and its components. Surpassing the score threshold induces an increase in school quality relative to the counterfactual. The magnitude of this shift and interpretation of treatment effects then depend on the nature and quality of students' fallback schools (Chabrier et al., 2016; Deming et al., 2014). In Deming et al. (2014), gains are concentrated among school lottery winners from lower performing neighborhood schools who experience the largest increases in school quality. Similarly, the performance of fallback schools explains a substantial share of the superior performance of No Excuses urban schools, suggesting that the counterfactual school environment is key to interpretations of findings (Chabrier et al., 2016).

Interpretation is challenged by simultaneous changes along multiple dimensions of school quality. The selective schools literature advances several channels through which admissions to such schools can shape student outcomes. Foremost is peer achievement (Abdulkadiroglu et al., 2014; Jackson, 2010). Higher quality peers have been shown to positively affect performance (Ammermueller & Pischke, 2009; Hanushek et al., 2003; Hoxby, 2000; Hoxby & Weingarth, 2005; Imberman et al., 2012; Vigdor & Nechyba, 2007). This may induce higher scores through teamwork-based classroom learning or induce teachers to tailor curriculum for a more gifted population. Meanwhile, large increases in peer quality are usually accompanied by sizable decreases in relative class rank. The positive effects from high ability peers may be mediated by

[33] Tests of differences show that effects are 3–4 percentage points higher for disadvantaged groups in all categories except for race, for which minority students shift away from less competitive colleges by 7 additional percentage points relative to non-minorities.

[34] Patterns of differential results by student background remain robust to alternative specifications including parametric regressions using cubic polynomials of the running variable on both sides of the cutoff. Estimates using 2SLS furthermore replicate these results, suggesting that differential patterns of selection or compliance rates do not play an important role (Tables D2 and D3).

Table 5
Heterogeneous treatment effects - end of high school outcomes.

	Gender		Minority		Home ZIP		HS quality		Achievement	
	Female	Male	Yes	No	Rural	Urban	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SAT math percentile	2.446***	−0.023	4.957***	1.174	2.999***	0.325	3.570***	−0.624	3.171***	−0.511
	(0.793)	(0.887)	(1.508)	(0.785)	(0.886)	(0.882)	(0.801)	(0.737)	(0.828)	(0.738)
SAT reading percentile	0.330	−0.110	2.714	−0.174	1.382	−0.701	1.202	−1.017	1.140	−0.822
	(1.007)	(1.366)	(2.180)	(0.943)	(1.175)	(1.153)	(1.206)	(1.059)	(1.372)	(1.063)
STEM intention	0.009	−0.024	−0.001	0.005	0.006	0.006	−0.011	−0.008	−0.023	0.019
	(0.032)	(0.037)	(0.059)	(0.032)	(0.038)	(0.036)	(0.039)	(0.035)	(0.039)	(0.032)
Total no. of applications	0.253	0.209	0.232	0.239	0.036	0.466*	0.083	0.281	0.420	0.049
	(0.255)	(0.258)	(0.460)	(0.209)	(0.243)	(0.243)	(0.252)	(0.257)	(0.308)	(0.218)
Share of applications: most selective	0.038**	−0.007	−0.005	0.024	0.036	0.004	0.026	0.019	0.035	−0.001
	(0.018)	(0.026)	(0.032)	(0.020)	(0.022)	(0.022)	(0.022)	(0.020)	(0.022)	(0.024)
Share of applications: very selective	0.007	0.014	0.097***	−0.003	0.013	0.010	0.022	−0.005	0.025	0.001
	(0.016)	(0.020)	(0.024)	(0.016)	(0.017)	(0.018)	(0.019)	(0.016)	(0.019)	(0.018)
Share of applications: less selective	−0.049***	−0.009	−0.097***	−0.024**	−0.052***	−0.018	−0.049***	−0.018	−0.064***	−0.002
	(0.014)	(0.013)	(0.031)	(0.010)	(0.013)	(0.013)	(0.012)	(0.011)	(0.016)	(0.013)
Average share of STEM BAs	0.017**	0.012	0.016	0.018**	0.019**	0.019**	0.016*	0.018**	0.014	0.018**
	(0.008)	(0.008)	(0.012)	(0.007)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.007)

Notes: Reported coefficients are the interactions between the individual or HS attributes (e.g. female) and an indicator for exceeding the score cutoff. All specifications include CD X cohort X minority fixed effects and use optimal bandwidths and local linear regressions. Rural ZIP is defined as the upper half in terms of share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores and divided into halves. Achievement is the 10th grade SAT math score of applicants divided into halves. All results are weighted by the inverse of the number of matches. Robust standard errors are clustered at the cohort and CD level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6
Heterogeneous treatment effects - postsecondary outcomes.

	Gender		Minority		Home ZIP		HS quality		Achievement	
	Female	Male	Yes	No	Rural	Urban	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enrolled in college	0.056**	0.039	0.026	0.052**	0.052**	0.044	0.026	0.077***	0.033	0.062**
	(0.021)	(0.027)	(0.028)	(0.023)	(0.022)	(0.026)	(0.020)	(0.027)	(0.020)	(0.026)
First college was most selective	0.054	0.000	−0.004	0.036	0.048	0.013	0.020	0.033	0.085**	−0.021
	(0.032)	(0.041)	(0.053)	(0.031)	(0.031)	(0.037)	(0.035)	(0.039)	(0.033)	(0.038)
First college was very selective	0.017	0.022	0.150***	0.000	0.039	0.004	0.046	−0.032	−0.007	0.041
	(0.028)	(0.037)	(0.050)	(0.028)	(0.029)	(0.033)	(0.034)	(0.032)	(0.030)	(0.030)
First college was less selective	−0.056**	−0.032	−0.146***	−0.032	−0.081***	−0.016	−0.068**	0.000	−0.072**	−0.019
	(0.027)	(0.025)	(0.042)	(0.022)	(0.025)	(0.027)	(0.029)	(0.025)	(0.029)	(0.023)
Ever attended 2−year college	−0.079**	−0.043	−0.091*	−0.060**	−0.095***	−0.029	−0.076**	−0.058*	−0.066*	-0.058**
	(0.031)	(0.026)	(0.050)	(0.025)	(0.034)	(0.027)	(0.033)	(0.032)	(0.036)	(0.026)
Share of BAs awarded in STEM in first college	0.023**	0.012	0.026	0.017*	0.016*	0.021*	0.023**	0.008	0.012	0.023**
	(0.009)	(0.013)	(0.019)	(0.009)	(0.010)	(0.012)	(0.011)	(0.012)	(0.011)	(0.010)
Graduated with Bachelor's in 4 years	0.010	0.023	−0.011	0.016	0.017	0.016	0.004	0.050	0.031	0.002
	(0.049)	(0.044)	(0.076)	(0.040)	(0.046)	(0.053)	(0.051)	(0.049)	(0.044)	(0.047)
Graduated with Bachelor's in STEM in 4 years	0.005	−0.001	−0.024	0.007	−0.012	0.017	−0.013	0.021	0.002	−0.007
	(0.036)	(0.047)	(0.061)	(0.031)	(0.038)	(0.043)	(0.041)	(0.039)	(0.038)	(0.040)

Notes: Reported coefficients are the interactions between the individual or HS attributes (e.g. female) and an indicator for exceeding the score cutoff. All specifications include CD X cohort X minority fixed effects and use optimal bandwidths and local linear regressions. Rural ZIP is defined as the upper half in terms of share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores and divided into halves. Achievement is the 10th grade SAT math score of applicants divided into halves. Robust standard errors are clustered at the cohort and CD level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

negative consequences of being in a more competitive academic environment (Cullen, Jacob, & Levitt, 2006). Finally, exposure to elite education can also lead to discontinuous changes in peer composition (Abdulkadiroglu et al., 2014). Changes in the racial distribution mark an important difference in the peer environment.

Fig. 7 plots peer composition, quality, and class rank measured during the fall of junior year against the running variable for the pooled sample.[35] Applicants are predicted to have a discontinuous increase in

[35] Let R_{ist} denote percentile rank computed from high school peers' 8th grade EOG math scores, where i indexes individuals, s the high school, and t the

(*footnote continued*)
cohort:

$$R_{ist} = 100 \times \left(\frac{n_{ist} - 1}{N_{st} - 1} \right)$$

Individual ordinal rank n_{ist} ranges from 1 at the bottom of the class to grade size N_{st} corresponding to the highest achiever. As such, the lowest ranking student has $R_{ist} = 0$ while the highest has $R_{ist} = 100$. Rank is defined relative to peers in the same school and grade.

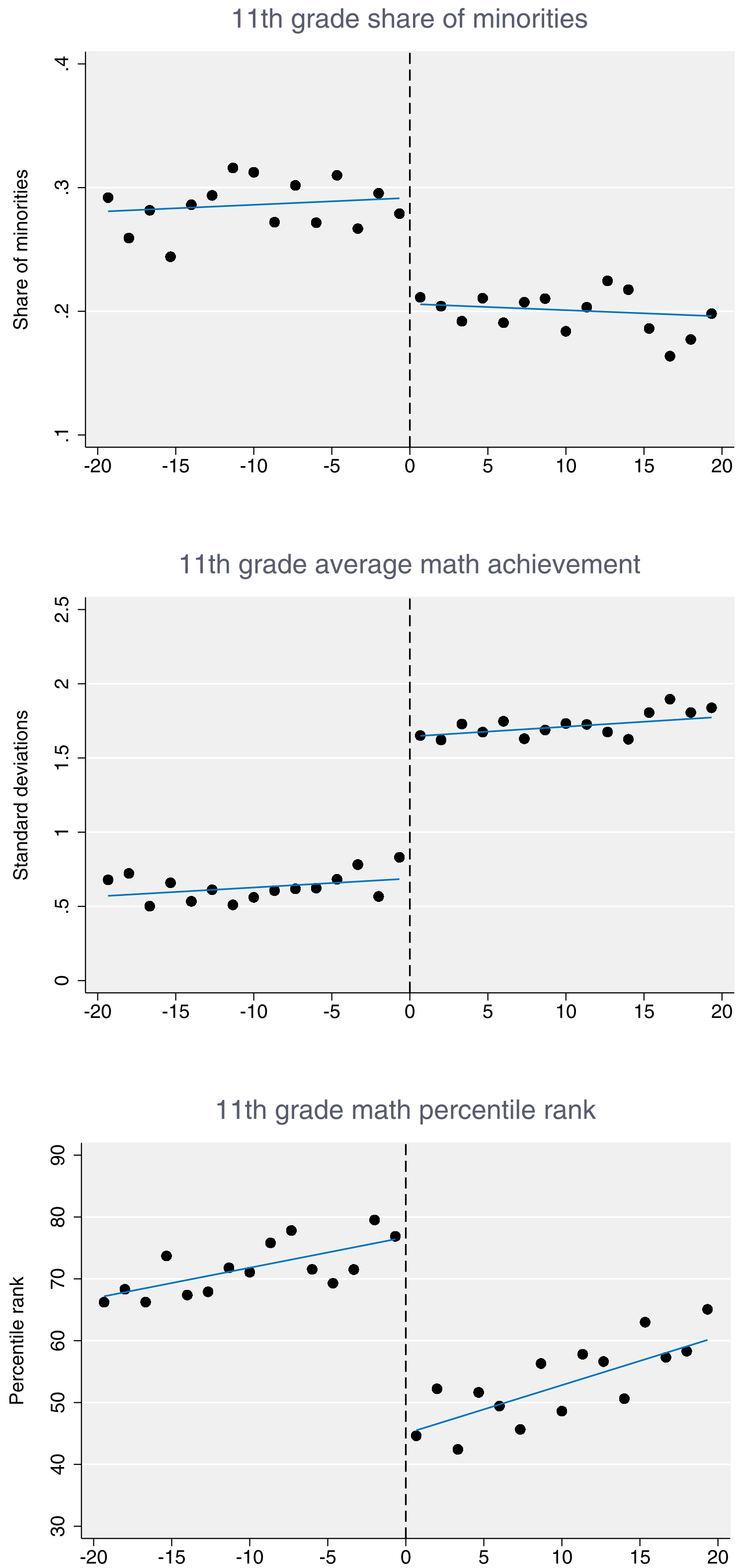

Fig. 7. Discontinuities in school characteristics *Note:* This figure shows high school characteristics relative to distance to the admissions cutoff. The share of minority students describes the proportion of the 11th grade class that is African-American, Hispanic, or Native American. Average math score computes the mean 8th grade End-of-Grade math score of the 11th grade class, while the percentile rank outcome denotes relative ordering from highest (100) to lowest (0).

Table 7
Mediating factors.

		Minority		Sending school quality		Distance to school	
	All	Yes	No	Low	High	Far	Near
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High school characteristics							
Share of minority students	−0.070***	−0.173***	−0.054***	−0.114***	−0.014	−0.074***	−0.056
	(0.017)	(0.036)	(0.015)	(0.024)	(0.020)	(0.019)	(0.035)
Average math score	0.792***	0.987***	0.761***	0.924***	0.612***	0.836***	0.635***
	(0.072)	(0.132)	(0.070)	(0.082)	(0.071)	(0.077)	(0.055)
Percentile rank by math score	−24.735***	−39.070***	−22.710***	−30.134***	−20.290***	−27.593***	−14.312***
	(3.312)	(5.119)	(3.140)	(3.516)	(3.318)	(3.344)	(4.371)
College application portfolio characteristics							
Average share of black students	−0.330	−3.710**	0.209	−0.984**	0.139	−0.446	0.020
	(0.370)	(1.474)	(0.290)	(0.438)	(0.318)	(0.393)	(0.551)
Average 75th percentile SAT math score	7.761**	10.369	7.352**	7.305**	7.791**	10.365***	−1.039
	(3.309)	(7.161)	(3.203)	(3.555)	(3.562)	(3.519)	(4.720)
Share of colleges that are out of state	0.038**	0.034	0.039**	0.036	0.033*	0.046**	0.006
	(0.022)	(0.029)	(0.019)	(0.024)	(0.019)	(0.020)	(0.026)

Notes: Table shows the effects of selective school eligibility on 11th grade school characteristics and analogous attributes of college application portfolios. It reports nonparametric estimates using optimal bandwidths and a linear control of distance to the cutoff. Columns 2–7 interacts attributes such as student ethnicity with an indicator for exceeding the score cutoff. The share of minority students describes the proportion of the 11th grade class that is African-American, Hispanic, or Native American. Average math score computes the mean 8th grade End-of-Grade math score of the 11th grade class, while the percentile rank outcome denotes relative ordering from highest (100) to lowest (0). Among college application mediators, the share of black students is averaged across all institutions that the student applied to, as proxied by SAT score reports. The same goes for institutions' 75th percentile SAT math scores. The last mediator captures the share of all colleges a student applied to that are located out-of-state. All results are weighted by the inverse of the number of matches, and include CD X cohort X minority fixed effects. Robust standard errors are clustered at the cohort and CD level.

the likelihood of having selective school peers upon crossing the threshold. Applicants who barely miss the cutoff have over 7 percentage point fewer minority peers in their 11th grade classroom relative to the counterfactual of nearly 30%. The accompanying increase in peer math ability is 0.8σ, up from an already high-achieving 0.8σ above the statewide average. The final graph shows a discontinuous drop in math percentile rank among 11th grade classmates, reflecting the reality of enrolling in a more competitive institution in which everyone excels academically.

Examining the magnitude of these changes by student background further informs an understanding of heterogeneous treatment effects. Table 7 separately estimates these discontinuities by minority status, sending school quality, and distance to the selective boarding school. I distinguish between students who live proximal to campus ($\leq$ 30 miles) to those who live farther away for two reasons. The first is that students living close by can go home with greater frequency, such that the treatment entails a smaller shift in parental inputs and the home environment. The second is that this range includes the majority of students applying from CD 4, which has an unusually large number of well-qualified applicants compared to the rest of the sample.

Increases in school quality are between 0.2 and 0.3σ larger for minority students, those from lower quality sending schools, or those residing further away from the school who surpass the admissions threshold.[36] Meanwhile, minority students and those from lower quality sending schools exceeding the cutoff experience a statistically larger drop in the share of minority peers. Changes in racial composition are accompanied by sizable decreases in math percentile rank across all groups that are particularly pronounced among disadvantaged students.

The rigorous academic and peer environment of the selective boarding school has been described by some alumni as mimicking the college experience. I explore how students exposed to different magnitudes of peer treatments respond in terms of college application portfolios. The bottom panel of Table 7 shows that minorities and students from lower quality sending schools apply to colleges with fewer black students on average if they surpass the admissions threshold. With the exception of historically black colleges and universities, competitive colleges are generally more likely to resemble the racial composition of the selective school. Notably, students residing more than 30 miles from the selective school who exceed the admissions threshold apply to higher-performing colleges as measured by SAT math score distributions as well as a greater share of out-of-state schools. No comparable effects are observed for students residing close to the boarding school. Part of this is driven by lower-scoring counterfactual schools among the far group, in contrast to high quality options for those residing within the 30-mile radius (Table C2).

Findings of significant treatment effects in this paper stand in contrast to other domestic evaluations of selective schools as well as several other international studies. I attribute these differences to a treatment that is larger in magnitude than the majority of other studies and somewhat different in content due to the residential nature of the program, which may amplify the positive effects of increased peer and school quality. An applicant in this study who just exceeds the cutoff ends up in a classroom with average peer math scores of over 1.6σ above the state mean, putting it on par with exam schools such as Bronx Science and Stuyvesant in New York City or Latin School in Boston (Abdulkadiroglu et al., 2014; Dobbie & Fryer, 2014). Those who miss the cutoff face a drop in average peer achievement of 0.8σ. The size of this selective school treatment is as large, if not substantially larger, than all six discontinuities documented in Abdulkadiroglu et al. (2014). Peer achievement means also disguise differences in the nature of counterfactual schools. Students not admitted to the most elite Boston or NYC schools are likely to proceed to the second most selective institution in the hierarchy of exam schools, even when the district allows for considerable choice across its public school offerings. The counterfactual case in North Carolina includes a wider range of public and charter school models across dozens of administrative districts that can vary widely in terms of resources and curriculum.

[36] Table C2 documents increases in average math scores of this magnitude for minority students and those from rural neighborhoods, focusing on compliers.

To be sure, peer achievement is only one proxy for school quality. It seems reasonable to assume that an academic environment enrolling higher-performing peers will also have greater resources at its disposal that manifest in more experienced teachers, college counselors, or advanced curricular exposure that shape students' propensity for competitive universities. While I do not have access to administrative staff data, publicly available documents show that over 96% of current instructors at the selective boarding school have at least a Master's degree.[37] The analogous proportions in the top three New York City public high schools are 59–64% among compliers. This suggests the distribution of educational levels in this institution is well to the right of many other elite secondary schools in North Carolina and beyond.

The residential nature of this school is another key dimension that can explain differential treatment effects between this study and other US-based exam school contexts. The experience of living on campus and persistent exposure to high-performing peers in both academic and social settings comprise a stronger treatment than regular day schools. Those with more difficult home environments can benefit from reduced academic disruptions on campus. There is also potential for greater faculty interaction via research opportunities and mentorship programs. Lastly, the school's leadership aims to develop skills that ease the transition into college. The experience of living independently may facilitate the acquisition of these soft skills in ways that pay dividends in postsecondary access and completion outcomes. Taken together, the evidence suggests that mechanisms specific to boarding schools that can adversely affect students' performance and well-being do not dominate in this case.

7. Robustness checks

To ensure results have an unbiased causal interpretation, I present several sets of analyses that test the robustness of the study's findings. I begin by restricting the sample to applicants that fall within increasingly narrower bandwidths around admissions cutoffs. Table E1 presents the effect of selective school eligibility across all high school and postsecondary outcomes under bandwidths that are 90%, 75%, and 50% of the original. These checks reaffirm earlier findings. Reducing bandwidths by a quarter yields similarly sized coefficients as before. It's only when bandwidths are halved that outcomes on SAT math begins to lose significance. Despite larger standard errors, the effects on postsecondary outcomes remain significant and even appear to increase in magnitude.

Column 4 of Table E1 takes the original sample and excludes CD 4, a district with unique properties such as its proximity to the boarding school, relatively affluent status, and a surplus of demand leading to high cutoff scores and low admissions rates. After dropping this district, estimates of the effect of selective school eligibility on both high school and college outcomes are larger in magnitude relative to results using the full sample, suggesting that if anything, treatment effects are more pronounced outside of CD 4 (Table E1). The nonparametric specification shows changes in applications to and enrollment in less selective colleges of -4.4 and -5.9 percentage points, respectively.

Another specification check varies the functional form assumptions of the parametric model. A consideration is that high-order polynomials may lead to noisy estimates and poorer inference (Gelman & Imbens, 2017). Table E2 juxtaposes results using cubic polynomials with those that rely on quadratic functions of the distance to the cutoff. Estimates using quadratic functions produce more precise estimates. Coefficients are mostly of similar magnitudes, with the exception of slightly attenuated SAT math estimates and postsecondary outcomes on attending less selective colleges. The significant result on the latter is closer in magnitude to nonparametric models.

One consideration that is specific to high school outcomes is the role of the matching process on estimated treatment effects. Applicants are linked to potential matches in statewide administrative data using individual characteristics, sending school codes, and enrollment status. Any measurement error introduced by the process can affect causal inference. I undertake two approaches to gauge the sensitivity of estimated effects to the matching process. The first re-estimates the models using only applicants with an unique match from North Carolina administrative data while the second randomly samples with replacement one observation from the set of matches and computes bootstrapped coefficients and standard errors for 1000 repetitions. Estimates using the restricted sample of unique matches in Table E3 maintain the statistical significance of the original results and adhere closely to the estimates' magnitudes. Bootstrapped coefficients provide further evidence that the estimates are not sensitive to the choice of matched observation (Fig. E1).

8. Conclusion

This paper uses discontinuous admissions probabilities to estimate the impact of selective secondary schools on students' test scores, college application behavior, enrollment, and completion. I rely on a district-based admissions rule and regression discontinuity design with multiple cutoffs to evaluate how a diverse group of students fare from exposure to selective education. The study contributes to an exam schools literature that has limited evidence on effect heterogeneity, due to most regression discontinuity studies' reliance on variation in the neighborhood of a singular school cutoff. The few studies employing multiple thresholds find no differential test score effects by individual SES, baseline ability, or gender (Barrow et al., 2016; Lucas & Mbiti, 2014). This paper provides new evidence in the US context that high achieving students from disadvantaged backgrounds benefit in meaningful ways from attending a selective secondary school. In addition to SAT math gains of 3–5 percentile points, under-represented minorities, lower achieving students, and those from more rural neighborhoods or lower quality sending schools decrease their applications to and enrollment in less competitive colleges by at least one-quarter. Matching to a higher quality university does not reduce 4-year graduation rates for these groups, suggesting that lower rates of academic undermatch are not offset by negative consequences from being in a more academically rigorous environment.

The findings on test scores are qualitatively in line with treatment effects identified for disadvantaged students in gifted and talented programs or other public boarding schools (Behaghel et al., 2017; Card & Giuliano, 2016; Curto & Fryer, 2014). These studies consistently document the sensitivity of minority or low-income students to increases in education quality. What remain unknown are longer-term consequences beyond test scores. This paper addresses the gap by tracking students through postsecondary education and finding benefits that persist from the college application stage through actual enrollment.

Results inform the relevance of an institutional model that states increasingly rely on to serve academically gifted students. The study provides causal evidence that public boarding schools, particularly those with a science and math curricular focus, are benefiting disadvantaged students while treatment effects in test scores and postsecondary outcomes are mostly null among their more advantaged peers. These patterns of heterogeneous effects advance an efficiency-based argument for more inclusive selective institutions.

A persistent challenge in this work is parsing out the channels of influence when the treatment entails simultaneous changes along multiple inputs. A lingering question is how the magnitude of benefits would change if students did not board at the selective school. One potential path of inquiry is evaluating more parsimonious modes of

[37] 49% of teachers in the selective school have a Ph.D., with some variation across subjects in the proportion of doctoral degrees attained. For example, 31% and 39% of math and humanities instructors have a Ph.D. respectively, compared to 68% in science and engineering.

delivery including day schools and online programs to determine the minimum set of inputs leading to improved student outcomes.

Acknowledgments

The author is grateful to Atila Abdulkadiroglu, Elizabeth Ananat, Peter Arcidiacono, Charles Clotfelter, Philip Cook, Tom Dee, Robert Garlick, V. Joseph Hotz, Helen Ladd, Susanna Loeb, Hugh Macartney, Manoj Mohanan, Marcos A. Rangel, Sean Reardon, Seth Sanders, Jeffrey Smith, Jacob Vigdor, and seminar participants at Duke University, IZA Brown Bag Seminar, and the 2017 SOLE annual conference for their helpful feedback. This work was supported by the Institute for Education Sciences (Award R305B130017).

Appendix A. Matching procedure

I link the 2009–2012 applicant sample to North Carolina administrative data in several steps. First I restrict the pool of possible matches in K-12 statewide data to students who ever submitted a SAT score report to the selective school using its unique 4-digit College Board code. This restricts the sample to approximately 1000 students annually. Next I construct a match ID using variables subject to minimum measurement error: birth date and sex. Other variables, such as residential zip code and ethnicity, are likely to have higher incidences of mis-reporting. Additional attributes are incorporated in sequence to narrow down the set of possible matches.

Reliance on birth date and sex alone leads almost all applicants to have multiple matches in the North Carolina dataset. The next step uses variables on enrollment status, race/ethnicity, and sending high school. Race is defined as black, Asian, or other because distinguishing further by Hispanic origin or multiple races leads to higher rates of unmatched observations. A likely match occurs if applicant records indicate she was enrolled during 10th grade in the same sending high school as given on the application, belongs to a race/ethnicity category that matches the one on her application, and graduated from the selective high school while being marked as being enrolled in the selective school at the time she last took the SAT. If there is at least one match on these variables, I drop all other matches for the applicant. In some cases only a single match remains for an applicant. I then remove the North Carolina student ID corresponding to the so-called likely match from the match sets of other applicants with multiple pairings. This further limits the match sets for some applicants and creates additional unique matches. The process iterates until no further changes can be made.

The following step matches on the same variables but focuses on those who did not enroll or graduate from the selective school. An observation is a likely match if applicant records indicate she was enrolled during 10th grade in the same sending high school as given on the application, belongs to a race/ethnicity category that matches the one on her application, and did not graduate from the selective high school while being marked as being enrolled in any school except for the selective school at the time she last took the SAT. As before, I remove student IDs corresponding to likely matches from the list of possible matches in North Carolina data.

I narrow the number of remaining matches using (1) location, (2) sending school, and (3) detailed ethnicity variables. First I check if the residential zip code and county given on the applicant corresponds with the student's 10th grade residential zip code and county as shown in statewide administrative geocode data. Then I identify applicants with at least one sending school matched pairing, and drop all other mismatched observations. Finally, I drop mismatches based on detailed ethnicity categories.

Table A1
Summary statistics - matching procedure.

	2009		2010		2011		2012	
Group	Unique	Total	Unique	Total	Unique	Total	Unique	Total
1 match	812	812	803	803	714	714	814	814
2 matches	49	98	37	74	28	56	47	94
3 matches	10	30	13	39	3	9	14	42
4 matches	9	36	20	80	15	60	6	24
5 matches	5	25	15	75	8	40	9	45
6 matches	9	54	15	90	12	72	9	54
7 matches	10	70	7	49	4	28	4	28
8 matches	6	48	9	72	3	24	2	16
9 matches	4	36	2	18	11	99	4	36
10 matches	6	60	2	20	6	60	3	30
11 matches	2	22	0	0	2	22	0	0
12 matches	1	12	0	0	0	0	0	0
13 matches	0	0	0	0	0	0	0	0
14 matches	1	14	0	0	1	14	0	0
Total	924	1499	923	1405	807	1316	912	1324

Notes: Table summarizes the number of matches corresponding to 3566 unique selective school applicants spanning the 2009 - 2012 cohorts. Applicants in the analytical sample have non-missing congressional district information and positive admissions scores. Matched observations derive from NCERDC records.

Appendix B. Covariate smoothness

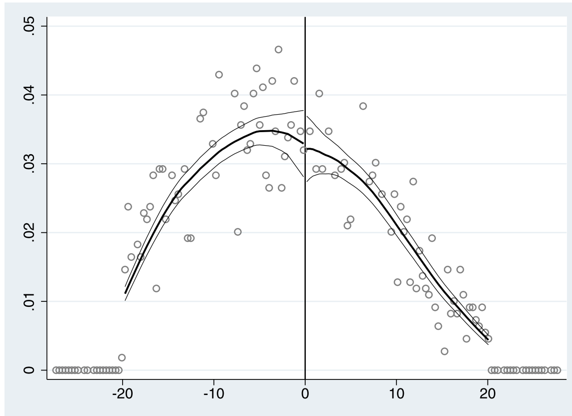

Fig. B1. Density of the admission score variable around the cutoff *Notes*: X-axis shows the distance from cutoff in the pooled sample. The approach computes frequency counts within equally-spaced bins and smooths the histogram separately on either side of the cutoff using local linear regressions.

Table B1
Covariate balance.

Covariate	Parametric (1)	Nonparametric (2)
Entering SAT math	3.260	6.034
	(5.693)	(4.143)
Entering SAT verbal	3.523	5.740
	(6.380)	(4.242)
Black	0.002	0.019
	(0.020)	(0.012)
Rural zip code	0.019	0.010
	(0.031)	(0.022)
Sending school quality	0.025	0.029
	(0.036)	(0.023)
Sending school pupil-to-teacher ratio	-0.000	0.136
	(0.300)	(0.198)

Notes: This table tests for discontinuities across the admissions threshold. Models regress each covariate on an indicator for exceeding the cutoff and report the coefficients on treatment status. Column (1) shows estimates from the parametric model that flexibly allows for different cubic functions of distance to the cutoff on either side of the cutoff. Applicants with admission scores more than 40 points removed from the cutoff are excluded. Column (2) reports nonparametric estimates using optimal bandwidths and a linear control of distance to the cutoff. SAT math and verbal scores refer to student performance at the time of application to the selective school. Rural zip code is defined as the share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix C. Baseline outcomes and complier characteristics

Table C1
Baseline outcomes.

		Gender		Race/ethnicity		Urban vs. rural		HS quality		Achievement	
	All	Female	Male	Minority	Non-min	Rural	Urban	Low	High	Low	High
Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
High school outcomes											
SAT math percentile	83.13	80.47	86.79	69.79	85.34	78.76	87.84	78.22	88.93	75.62	91.57
SAT verbal percentile	81.63	80.28	83.48	70.26	83.51	77.21	86.23	77.54	86.45	76.17	87.59
STEM intention	0.29	0.20	0.43	0.23	0.31	0.30	0.29	0.28	0.32	0.27	0.32
Total no. of applications	5.94	5.89	6.02	5.45	6.02	5.24	6.70	5.34	6.65	5.24	6.72
Share of applications: most selective	0.57	0.54	0.61	0.53	0.58	0.47	0.67	0.48	0.67	0.46	0.69
Share of applications: very selective	0.26	0.26	0.26	0.17	0.27	0.30	0.21	0.29	0.22	0.29	0.22
Share of applications: less selective	0.17	0.20	0.13	0.30	0.14	0.22	0.11	0.22	0.11	0.24	0.09
Avg share of STEM BAs	0.22	0.20	0.26	0.20	0.23	0.21	0.24	0.21	0.24	0.20	0.25
Postsecondary outcomes											
Enrolled in college	0.92	0.92	0.92	0.94	0.92	0.95	0.89	0.95	0.87	0.96	0.88
First college was most selective	0.47	0.49	0.44	0.43	0.47	0.39	0.57	0.39	0.57	0.36	0.59
First college was very selective	0.30	0.23	0.39	0.19	0.32	0.36	0.23	0.32	0.28	0.35	0.25
First college was less selective	0.22	0.27	0.16	0.38	0.20	0.25	0.19	0.28	0.14	0.28	0.16
Ever attended 2 year	0.22	0.24	0.19	0.29	0.20	0.26	0.17	0.24	0.19	0.25	0.18
Avg share of STEM BAs	0.20	0.18	0.22	0.18	0.20	0.19	0.21	0.18	0.22	0.19	0.22
Graduate in 4 years	0.45	0.49	0.40	0.36	0.47	0.42	0.48	0.49	0.41	0.42	0.48
Graduate in 4 with STEM degree	0.23	0.23	0.24	0.17	0.24	0.19	0.27	0.22	0.25	0.16	0.32

Notes: Table shows average outcomes for applicants with admissions scores 5 points or less below the cutoff.

Table C2
Complier characteristics.

	All			
Full sample	D = 0	D = 1		
Share of minority students	0.31	0.20		
Average math score	0.55	1.89		
Percentile rank by math score	89.29	53.99		
Rural high school	0.47	0.04		
	Minority		Non-minority	
Minority vs. non-minority	D = 0	D = 1	D = 0	D = 1
Share of minority students	0.47	0.22	0.29	0.19
Average math score	0.33	1.92	0.58	1.88
Percentile rank by math score	80.80	37.17	90.27	57.49
Rural high school	0.44	0.01	0.48	0.05
	Lower		Higher	
Lower vs. higher quality sending school	D = 0	D = 1	D = 0	D = 1
Share of minority students	0.39	0.21	0.23	0.18
Average math score	0.27	1.83	0.83	1.98
Percentile rank by math score	90.10	47.85	87.61	59.80
Rural high school	0.55	0.05	0.37	0.02
	Far		Near	
Far vs. near to selective school	D = 0	D = 1	D = 0	D = 1
Share of minority students	0.32	0.20	0.31	0.20
Average math score	0.48	1.89	0.76	1.95
Percentile rank by math score	89.39	51.46	87.93	65.36
Rural high school	0.52	0.04	0.29	0.02

Notes: This table shows summary statistics for compliers to the left (D = 0) and right (D = 1) of the admissions cutoff. Estimates are computed using the IV strategy defined in Abadie (2003) with optimal bandwidths and local linear regressions. Minority status includes African-American, Hispanic, and Native American students. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores and divided into halves. Far from selective school is defined as residing in a home zip code that is more than 30 miles away. The share of minority students describes the proportion of the 11th grade class that is African-American, Hispanic, or Native American. Average math score computes the mean 8th grade End-of-Grade math score of the 11th grade class, while the percentile rank outcome denotes relative ordering from highest (100) to lowest (0). Rural high school is an indicator of whether the student is enrolled in a school located in a rural area. All results are weighted by the inverse of the number of matches.

Appendix D. 2SLS Estimates

Table D1
Selective boarding school effects.

	Reduced form	2SLS
	(1)	(2)
High school outcomes		
SAT math percentile	1.708**	2.651**
	(0.757)	(1.188)
SAT reading percentile	0.236	0.367
	(0.980)	(1.532)
STEM intention	0.004	0.006
	(0.031)	(0.048)
Total no. of college applications	0.238	0.355
	(0.213)	(0.318)
Share of applications: most selective	0.020	0.031
	(0.019)	(0.029)
Share of applications: very selective	0.010	0.015
	(0.016)	(0.023)
Share of applications: less selective	-0.034***	-0.051***
	(0.011)	(0.018)
Average share of BAs awarded in STEM	0.018**	0.028**
	(0.007)	(0.011)
Postsecondary outcomes		
Enrolled in college	0.048**	0.073**
	(0.022)	(0.033)
First college was most selective	0.030	0.046
	(0.031)	(0.047)
First college was very selective	0.020	0.030
	(0.026)	(0.039)
First college was less selective	-0.047**	-0.067**
	(0.023)	(0.033)
Ever attended 2-year college	-0.064**	-0.094**
	(0.026)	(0.037)
Share of BAs awarded in STEM in first college	0.018**	0.027**
	(0.009)	(0.013)
Graduated with Bachelor's in 4 years	0.013	0.018
	(0.042)	(0.058)
Graduated with Bachelor's in STEM in 4 years	0.003	0.005
	(0.032)	(0.046)

Notes: Table shows the effects of selective school admissions using indicators for exceeding the threshold as an instrument for being admitted. Column (1) reports reduced form estimates from the nonparametric estimates using optimal bandwidths and a linear control of distance to the cutoff. Column (2) reports 2SLS estimates. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. Applicants with admission scores more than 40 points removed from the cutoff are excluded. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D2
Heterogeneous 2SLS treatment effects - high school outcomes.

	Gender		Minority		Home ZIP		HS quality		Achievement	
	Female	Male	Yes	No	Rural	Urban	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SAT math percentile	3.281***	0.294	6.840***	1.924	4.196***	0.892	4.911***	-0.353	4.279***	-0.315
	(1.124)	(1.408)	(2.171)	(1.221)	(1.313)	(1.345)	(1.206)	(1.114)	(1.225)	(1.160)
SAT reading percentile	0.427	-0.123	3.517	-0.188	1.809	-0.822	1.538	-1.267	1.452	-1.049
	(1.466)	(2.144)	(3.072)	(1.473)	(1.734)	(1.761)	(1.779)	(1.593)	(1.958)	(1.645)
STEM intention	0.009	-0.034	-0.001	0.007	0.009	0.009	-0.016	-0.012	-0.029	0.025
	(0.044)	(0.056)	(0.080)	(0.048)	(0.054)	(0.054)	(0.056)	(0.051)	(0.055)	(0.049)

(*continued on next page*)

Table D2 (*continued*)

	Gender		Minority		Home ZIP		HS quality		Achievement	
Total no. of applications	0.362	0.330	0.363	0.353	0.112	0.654*	0.149	0.396	0.565	0.119
	(0.348)	(0.386)	(0.619)	(0.309)	(0.343)	(0.356)	(0.358)	(0.373)	(0.422)	(0.322)
Share of applications: most selective	0.048*	-0.006	0.000	0.035	0.048	0.010	0.037	0.029	0.046	0.002
	(0.026)	(0.039)	(0.044)	(0.030)	(0.031)	(0.032)	(0.031)	(0.029)	(0.031)	(0.036)
Share of applications: very selective	0.012	0.020	0.125***	-0.000	0.018	0.015	0.028	-0.004	0.034	0.004
	(0.023)	(0.030)	(0.035)	(0.024)	(0.025)	(0.027)	(0.027)	(0.023)	(0.027)	(0.027)
Share of applications: less selective	-0.066***	-0.020	-0.135***	-0.039**	-0.073***	-0.033	-0.070***	-0.033*	-0.087***	-0.011
	(0.020)	(0.020)	(0.045)	(0.016)	(0.019)	(0.020)	(0.018)	(0.017)	(0.023)	(0.020)
Average share of STEM BAs	0.025**	0.019	0.026	0.028**	0.029**	0.030**	0.024*	0.028**	0.022*	0.027**
	(0.011)	(0.012)	(0.018)	(0.011)	(0.011)	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)

Notes: Reported coefficients are the interactions between the individual or HS attributes (e.g. female) and the instrument for being admitted. All specifications include CD X cohort X minority fixed effects and use optimal bandwidths and local linear regressions. Rural ZIP is defined as the upper half in terms of share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores and divided into halves. Achievement is the 10th grade SAT math score of applicants divided into halves. All results are weighted by the inverse of the number of matches. Robust standard errors are clustered at the cohort and CD level. * p < 0.1, ** p < 0.05, *** p < 0.01

Table D3
Heterogeneous 2SLS treatment effects - postsecondary outcomes.

	Gender		Minority		Home ZIP		HS quality		Achievement	
	Female	Male	Yes	No	Rural	Urban	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enrolled in college	0.079**	0.063	0.048	0.076**	0.077**	0.067*	0.046	0.108***	0.052*	0.092**
	(0.031)	(0.041)	(0.039)	(0.034)	(0.033)	(0.038)	(0.031)	(0.038)	(0.029)	(0.038)
First college was most selective	0.069	0.007	0.005	0.052	0.068	0.025	0.032	0.049	0.110**	-0.019
	(0.045)	(0.062)	(0.072)	(0.047)	(0.046)	(0.052)	(0.051)	(0.055)	(0.047)	(0.057)
First college was very selective	0.025	0.034	0.191***	0.005	0.055	0.012	0.059	-0.037	-0.004	0.058
	(0.038)	(0.056)	(0.066)	(0.041)	(0.043)	(0.048)	(0.048)	(0.046)	(0.041)	(0.046)
First college was less selective	-0.074**	-0.050	-0.190***	-0.050	-0.109***	-0.031	-0.089**	-0.008	-0.093**	-0.033
	(0.035)	(0.037)	(0.057)	(0.032)	(0.036)	(0.037)	(0.040)	(0.035)	(0.040)	(0.034)
Ever attended 2-year college	-0.107**	-0.071*	-0.131*	-0.089**	-0.132***	-0.053	-0.109**	-0.088*	-0.093*	-0.089**
	(0.042)	(0.039)	(0.066)	(0.037)	(0.047)	(0.038)	(0.045)	(0.044)	(0.048)	(0.038)
Share of BAs awarded in STEM in first college	0.030**	0.020	0.038	0.026**	0.024*	0.031*	0.032**	0.013	0.018	0.034**
	(0.012)	(0.019)	(0.025)	(0.013)	(0.013)	(0.016)	(0.016)	(0.017)	(0.014)	(0.015)
Graduated with Bachelor's in 4 years	0.014	0.033	-0.010	0.022	0.023	0.023	0.010	0.064	0.039	0.006
	(0.063)	(0.062)	(0.097)	(0.056)	(0.062)	(0.072)	(0.068)	(0.068)	(0.058)	(0.067)
Graduated with Bachelor's in STEM in 4 years	0.006	-0.001	-0.028	0.009	-0.014	0.021	-0.015	0.026	0.002	-0.009
	(0.047)	(0.066)	(0.079)	(0.044)	(0.051)	(0.059)	(0.056)	(0.054)	(0.049)	(0.057)

Notes: Reported coefficients are the interactions between the individual or HS attributes (e.g. female) and the instrument for being admitted. All specifications include CD X cohort X minority fixed effects and use optimal bandwidths and local linear regressions. Rural ZIP is defined as the upper half in terms of share of all households within one's residential zip code located in rural areas. Sending school quality is constructed using average 8th grade End-of-Grade math standardized scores and divided into halves. Achievement is the 10th grade SAT math score of applicants divided into halves. Robust standard errors are clustered at the cohort and CD level. * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix E. Robustness checks

Table E1
Robustness to Varying Bandwidths and Excluding CD 4.

	Optimal bandwidth h			Exclude
	$h \times 0.9$	$h \times 0.75$	$h \times 0.5$	CD 4
	(1)	(2)	(3)	(4)
High school outcomes				
SAT math percentile	1.753**	1.928**	1.735	2.302**
	(0.812)	(0.903)	(1.162)	(0.875)
SAT reading percentile	0.218	0.323	0.479	0.334
	(1.015)	(1.093)	(1.205)	(1.037)
STEM intention	0.006	0.012	0.025	-0.002
	(0.033)	(0.036)	(0.043)	(0.038)
Total no. of college applications	0.183	0.106	0.033	0.305
	(0.219)	(0.238)	(0.295)	(0.248)
Share of applications: most selective	0.019	0.017	0.015	0.030
	(0.020)	(0.020)	(0.021)	(0.022)
Share of applications: very selective	0.012	0.015	0.015	0.011
	(0.016)	(0.017)	(0.018)	(0.018)
Share of applications: less selective	-0.033***	-0.031**	-0.028*	-0.044***
	(0.011)	(0.013)	(0.015)	(0.012)
Average share of BAs awarded in STEM	0.018**	0.016*	0.010	0.018**
	(0.008)	(0.008)	(0.010)	(0.008)
Postsecondary outcomes				
Enrolled in college	0.049**	0.052**	0.060**	0.020
	(0.023)	(0.025)	(0.030)	(0.016)
First college was most selective	0.030	0.026	0.048	0.041
	(0.032)	(0.035)	(0.043)	(0.036)
First college was very selective	0.025	0.033	0.024	0.025
	(0.027)	(0.031)	(0.041)	(0.028)
First college was less selective	-0.048**	-0.051**	-0.063**	-0.059**
	(0.023)	(0.024)	(0.026)	(0.026)
Ever attended 2-year college	-0.067**	-0.072**	-0.082**	-0.064**
	(0.027)	(0.030)	(0.037)	(0.031)
Share of BAs awarded in STEM in first college	0.020**	0.022**	0.024*	0.021**
	(0.009)	(0.009)	(0.012)	(0.010)
Graduated with Bachelor's in 4 years	0.013	0.014	0.019	-0.007
	(0.041)	(0.043)	(0.048)	(0.048)
Graduated with Bachelor's in STEM in 4 years	0.003	0.003	0.004	-0.008
	(0.033)	(0.034)	(0.037)	(0.037)

Notes: This table shows the effects of selective school eligibility for samples that use different optimal bandwidths or exclude applicants from CD 4. The first three columns use increasingly narrower bandwidths and a linear control of distance to the cutoff. High school results are weighted by the inverse of the number of matches. All models include CD X cohort X minority fixed effects. Robust standard errors are clustered at the cohort and CD level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table E2
Robustness to functional form assumptions.

	Cubic	Quadratic
	(1)	(2)
High school outcomes		
SAT math percentile	2.119*	1.205
	(1.086)	(0.869)
SAT reading percentile	0.715	-0.237
	(1.353)	(0.986)
STEM intention	0.019	0.006
	(0.047)	(0.037)
Total no. of college applications	0.048	0.065
	(0.351)	(0.257)
Share of applications: most selective	0.023	0.007
	(0.026)	(0.022)
Share of applications: very selective	0.009	0.024
	(0.022)	(0.019)
Share of applications: less selective	-0.030*	-0.033**
	(0.016)	(0.014)
Average share of BAs awarded in STEM	0.017*	0.017**
	(0.010)	(0.008)
Postsecondary outcomes		
Enrolled in college	0.057*	0.056**
	(0.032)	(0.021)
First college was most selective	0.022	0.017
	(0.048)	(0.037)
First college was very selective	0.046	0.022
	(0.045)	(0.032)
First college was less selective	-0.077**	-0.048*
	(0.035)	(0.028)
Ever attended 2-year college	-0.082*	-0.056*
	(0.043)	(0.032)
Share of BAs awarded in STEM in first college	0.029*	0.019*
	(0.015)	(0.011)
Graduated with Bachelor's in 4 years	0.039	0.010
	(0.067)	(0.054)
Graduated with Bachelor's in STEM in 4 years	-0.002	0.002
	(0.049)	(0.038)

Notes: This table shows the estimates from parametric models using a cubic or quadratic function of distance to the cutoff. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. Applicants with admission scores more than 40 points removed from the cutoff are excluded. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table E3
Estimates using unique matches only.

SAT math percentile	1.912**
	(0.732)
SAT reading percentile	0.353
	(1.108)
STEM intention	0.002
	(0.034)
Total no. of college applications	0.305
	(0.224)
Share of applications: most selective	0.016
	(0.020)
Share of applications: very selective	0.012
	(0.017)
Share of applications: less selective	-0.032**
	(0.012)
Average share of BAs awarded in STEM	0.018**
	(0.008)

Notes: This table shows the effects of selective school eligibility for the pooled sample. Nonparametric estimates use optimal bandwidths and a linear control of distance to the cutoff. All models include CD X cohort X minority fixed effects and robust standard errors are clustered at the cohort and CD level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

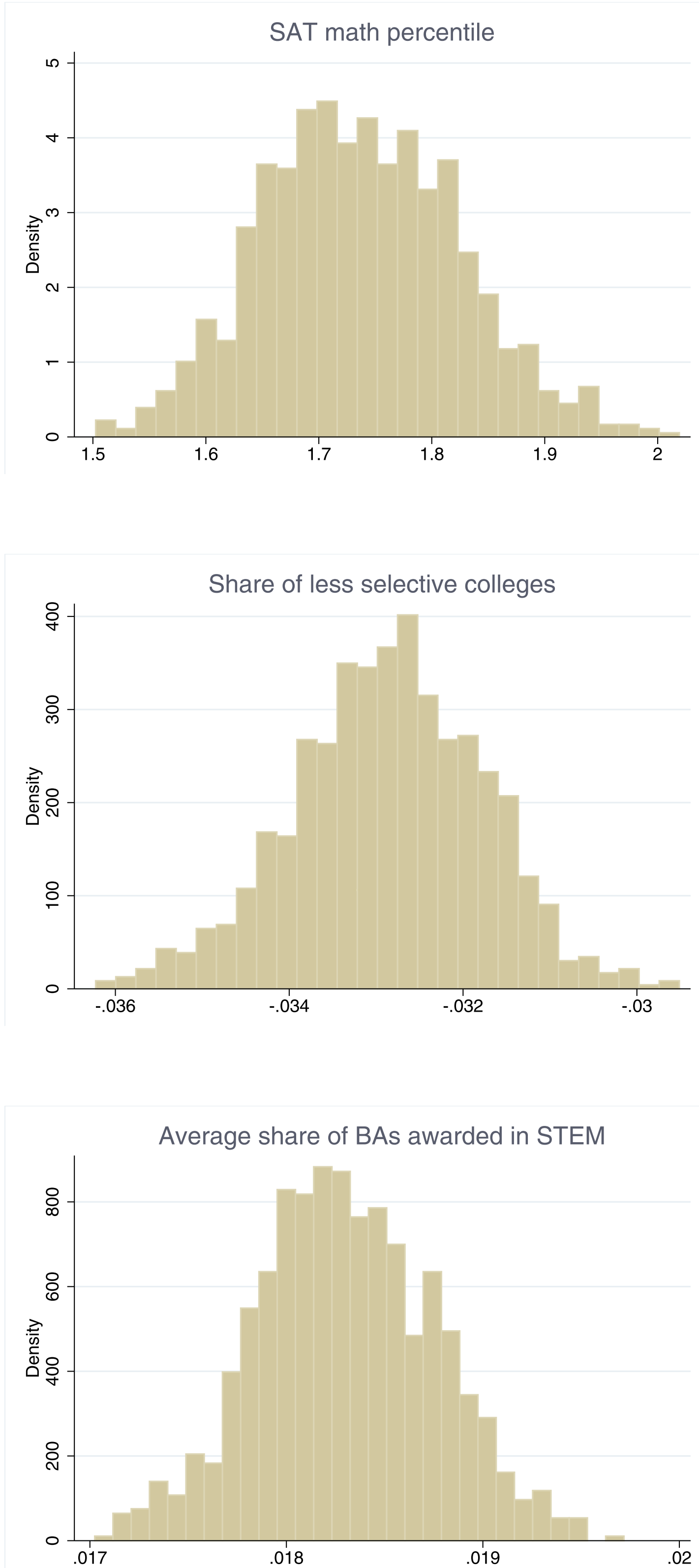

Fig. E1. Sensitivity of coefficients to matching process *Note:* Distributions of coefficients computed by randomly sampling one observation from the full set of matches for each applicant.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.econedurev.2019.07.001 .

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics, 113*(2), 231–263.

Abdulkadiroglu, A., Angrist, J. D., Narita, Y., Pathak, P. A., & Zarate, R. A. (2017). Regression discontinuity in serial dictatorship: Achievement effects at Chicago's exam schools. *American Economic Review, 107*(5), 240–245.

Abdulkadiroglu, A., Angrist, J. D., & Pathak, P. (2014). The elite illusion: Achievement effects at Boston and New York exam schools. *Econometrica, 82*(1), 137–196. https://doi.org/10.3982/ECTA10266.

Ammermueller, A., & Pischke, J.-S. (2009). Peer effects in european primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics, 27*(3), 315–348.

Barrow, L., Sartain, L., & de la Torre, M. (2016). *The role of selective high schools in equalizing educational outcomes: heterogeneous effects by neighborhood socioeconomic status*.

Behaghel, L., de Chaisemartin, C., & Gurgand, M. (2017). Ready for boarding? The effects of a boarding school for disadvantaged students. *American Economic Journal: Applied Economics, 9*(1), 140–164. https://doi.org/10.1257/app.20150090.

Booij, A. S., Haan, F., & Plug, E. (2016). Enriching students pays off: Evidence from an individualized gifted and talented program in secondary education.

Bowen, W. G., & Bok, D. (2016). *The shape of the river: Long-term consequences of considering race in college and university admissions.* Princeton University Press.

Bui, S. A., Craig, S. G., & Imberman, S. A. (2014). Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students. *American Economic Journal: Economic Policy, 6*(3), 30–62. https://doi.org/10.1257/pol.6.3.30.

Card, D., & Giuliano, L. (2016). Can tracking raise the test scores of high-ability minority students? *American Economic Review, 106*(10), 2783–2816. https://doi.org/10.1257/aer.20150484.

Card, D., & Krueger, A. B. (2005). Would the elimination of affirmative action affect highly qualified minority applicants? evidence from california and texas. *Industrial & Labor Relations Review, 58*(3), 416–434.

Chabrier, J., Cohodes, S., & Oreopoulos, P. (2016). What can we learn from charter school lotteries? *Journal of Economic Perspectives, 30*(3), 57–84. https://doi.org/10.1257/jep.30.3.57.

Clark, D. (2010). Selective schools and academic achievement. *The B.E. Journal of Economic Analysis & Policy, 10*(1), 1–40.

Cohodes, S. (2015). *The long-run impacts of tracking high-achieving students: evidence from Boston's Advanced Work Class*.

Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on participants: evidence from randomized lotteries. *Econometrica, 74*(5), 1191–1230. https://doi.org/10.1111/j.1468-0262.2006.00702.x.

Curto, V. E., & Fryer, R. G. (2014). The potential of urban boarding schools for the poor: Evidence from SEED. *Journal of Labor Economics, 32*(1), 65–93. https://doi.org/10.1086/671798.

Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *The American Economic Review, 104*(3), 991–1013.

Dillon, E. W., & Smith, J. A. (2016). Determinants of the match between student ability and college quality. *Journal of Labor Economics, 35*(1), 45–66. https://doi.org/10.1086/687523.

Dobbie, W., & Fryer, R. G. Jr. (2014). The impact of attending a school with high-achieving peers: Evidence from the new york city exam schools. *American Economic Journal: Applied Economics, 6*(3), 58–75 http://www.aeaweb.org/aej-applied/

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739–1774. https://doi.org/10.1257/aer.101.5.1739.

Dynarski, S. M., Hemelt, S. W., & Hyman, J. M. (2015). The missing manual: Using National Student Clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis, 37*(1_suppl), 53S–79S. https://doi.org/10.3102/0162373715576078.

Gelman, A., & Imbens, G. (2017). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics, 0*(0), 1–10. https://doi.org/10.1080/07350015.2017.1366909.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18*(5), 527–544. https://doi.org/10.1002/jae.741.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *Review of Economics and Statistics, 91*(4), 717–724. https://doi.org/10.1162/rest.91.4.717.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variationWorking Paper*. National Bureau of Economic Research.

Hoxby, C., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity, 2013*(1), 1–65. https://doi.org/10.1353/eca.2013.0000.

Hoxby, C. M., & Weingarth, G. (2005). Taking race out of the equation: School reassignment and the structure of peer effects. *Working paper*.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies, 79*(3), 933–959. https://doi.org/10.1093/restud/rdr043.

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review, 102*(5), 2048–2082. https://doi.org/10.1257/aer.102.5.2048.

Jackson, K. C. (2010). Do students benefit from attending better schools? evidence from rule-based student assignments in Trinidad and Tobago. *The Economic Journal, 120*(549), 1399–1429. https://doi.org/10.1111/j.1468-0297.2010.02371.x.

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics, 116*(3), 777–803. https://doi.org/10.2307/2696418.

Lucas, A. M., & Mbiti, I. M. (2014). Effects of school quality on student achievement: Discontinuity evidence from Kenya. *American Economic Journal: Applied Economics, 6*(3), 234–263. https://doi.org/10.1257/app.6.3.234.

Martin, A. J., Papworth, B., Ginns, P., & Liem, G. A. D. (2014). Boarding school, academic motivation and engagement, and psychological well-being: A large-scale investigation. *American Educational Research Journal, 51*(5), 1007–1049. https://doi.org/10.3102/0002831214532164.

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

Pop-Eleches, C., & Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review, 103*(4), 1289–1324 http://www.aeaweb.org/aer/

Schaverien, J. (2011). Boarding school syndrome: Broken attachments a hidden trauma. *British Journal of Psychotherapy, 27*(2), 138–155. https://doi.org/10.1111/j.1752-0118.2011.01229.x.

Smith, J., Pender, M., & Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review, 32*, 247–261. https://doi.org/10.1016/j.econedurev.2012.11.001.

Vigdor, J., & Nechyba, T. (2007). Peer effects in North Carolina public schools. In W. Ludger, & P. E. Paul (Eds.). *Schools and the Equal Opportunity Problem* (pp. 73–102). Cambridge: MIT Press.

Zhang, H. (2016). Identification of treatment effects under imperfect matching with an application to Chinese elite schools. *Journal of Public Economics, 142*, 56–82. https://doi.org/10.1016/j.jpubeco.2016.03.004.