WORKING PAPER 2020:8

Family background and the responses to higher SAT scores

Georg Graetz
Oskar Nordström Skans
Björn Öckert

IFAU INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Family background and the responses to higher SAT scores[a]

Georg Graetz[b] Björn Öckert[c] Oskar Nordström Skans[d]

April 22, 2020

Abstract

Using discontinuities within the Swedish SAT system, we show that additional admission opportunities causally affect college choices. Students with high-educated parents change timing, colleges, and fields in ways that appear rational and informed. In contrast, very talented students with low-educated parents react to higher scores by increasing overall enrolment and graduation rates. Remarkably, most of this effect arises from increased participation in college programs and institutions that they could have attended even with a lower score. This suggests that students with low-educated parents face behavioral barriers even in a setting where colleges are tuition-free, student grants are universal and application systems are simple.

KEYWORDS: Educational choice, intergenerational transmission of education, regression discontinuity design

JEL CLASSIFICATION: I21, I23, J62

[a]We thank Susan Dynarski, Edwin Leuven, Guy Michaels, Martin Nybom, Steve Pischke, Alex Solis, and seminar participants at CESifo, LSE, the Stockholm-Uppsala Economics of Education Workshop, the Nordic Summer Institute in Labor Economics 2019, SKILS 2020, and UCLS for helpful suggestions. Milosz Bolibrzuch provided excellent research assistance.

[b]Uppsala University. Email: georg.graetz@nek.uu.se.

[c]IFAU. Corresponding author. Email: bjorn.ockert@ifau.uu.se.

[d]Uppsala University. Email: oskar.nordstrom_skans@nek.uu.se.

Contents

1 Introduction

The determinants of college attendance is a key topic within the economics of education and related social sciences. Much of the focus has been directed towards trying to understand why high-ability students of low socioeconomic status (SES) choose not to attend college.[1] On top of traditional factors such as parental human capital investments, wealth transfers, and access to credit, the recent literature has established an important role for information frictions and behavioral barriers related to the college and grant application processes.[2] In addition, a set of recent studies have highlighted the importance of subjective beliefs for college application rates, while others have noted that low-SES students often fail to take advantage of the most competitive college programs they can enter.[3] Jointly, these studies indicate that the scope and salience of different college opportunities may be important determinants of college choices for low-SES students in particular.

Against this backdrop, this paper investigates the impact of admission opportunities on college choices and completion rates among students with high and low socioeconomic status as measured by parental education. Randomness in opportunities originate from discontinuities within the Swedish scholastic aptitude test (SAT) in a setting where all colleges are required by law to admit at least one third of students purely on the basis of these scores. Achieving a higher SAT score thus entails a marginal but salient expansion of the set of program-college combinations that students can enter, which may affect both *if* and *where* they end up enrolling.[4]

In Sweden, many of the factors traditionally identified as determinants of college attendance in general, and for low-SES individuals in particular, are absent. Colleges are tuition-free and financial aid is both universal and generous. Applying to college is straightforward and admission criteria are simple and transparent. Slots at program-college combinations with an excess number of applicants are allocated solely on the basis of high school GPA and SAT scores within a centralized (national) admission system.[5] Financial aid is administered by the central government through a one-step application system where the *only* requirement for access to grants and loans is college admission, which is verified automatically. But even in the absence of tuition fees and fully automated application systems, there is a sizeable remaining socioeconomic gap in terms of college enrollment and graduation rates between equally able students. The parental education gap in enrolment and graduation rates corresponds to between two and

[1] See for instance Manski and Wise (2013), Landerso and Heckman (2017) and Perna (2006).

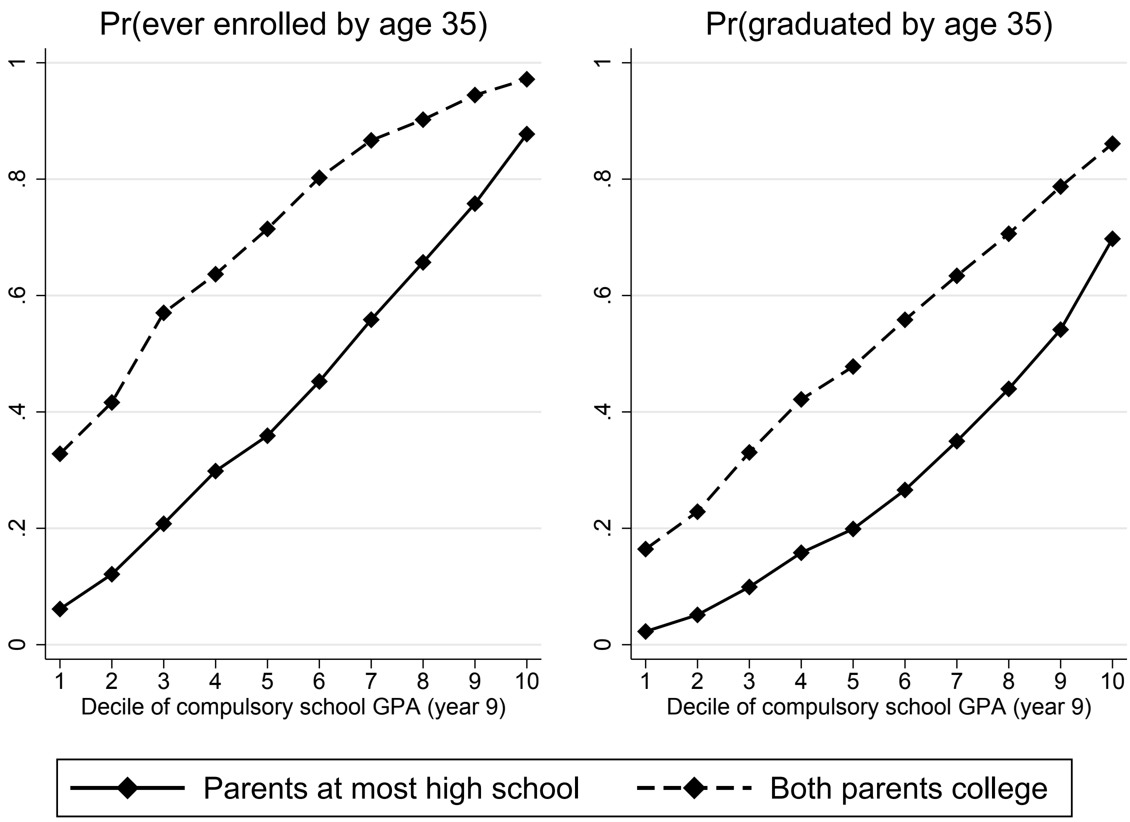
[2] For recent research on the transmission of abilities, see e.g. Grönqvist et al. (2017). On human capital investments and transfers, see e.g. Cameron and Heckman (2001) and Keane and Wolpin (2001). On access to credit, see e.g. Solis (2017) and Lochner and Monge-Naranjo (2012). For further references, see the literature review below.

[3] On the academic mismatch of low-SES students, see Hoxby and Avery (2013), Smith et al. (2013), Dillon and Smith (2017) and Howell and Pender (2016). For further references, see the literature review below.

[4] Students will not know ex ante which programs they can enter, but they know with certainty that the choice set expands with a higher score. Since past admission thresholds are made publicly available, students in principle have easy access to the information needed to make an overall assessment of the extent of these expanded opportunities.

[5] The very few exceptions include creative arts programs where there are specific selection criteria.

four full deciles in the grade distribution as shown in Figure 1.



Notes: The sample includes individuals living in Sweden in 2010 who were born between 1968-1975 and who finished compulsory school in Sweden. Enrollment in and graduation from college programs of any length are considered.

Figure 1: Ability, parental education, and college attendance

We shed new light on origins of these differences by estimating causal responses to expansions of college admission opportunities separately for test takers whose parents are at most high school graduates (low SES) and those with both parents having completed a college education (high SES). To achieve causal identification, we use quasi-random variation in SAT scores that arises when raw scores are converted into reported scores. As shown in Figure 2, reported normalized SAT scores are a step-function of raw scores, allowing us to estimate the causal effects of an incrementally higher score by comparing test takers on either side of a threshold in a sharp regression discontinuity design. The translation matrix mapping raw scores to normalized scores is constructed after the tests are taken, since the distribution of normalized scores is supposed to approximate a truncated normal distribution. The thresholds are therefore unknown at the time of taking the tests. The large number of steps allows us to identify the causal impact of admission opportunities across a wide distribution of initial opportunities and abilities.[6]

[6]Tests take place every semester and the best test (in the past five years) counts. Students can postpone their

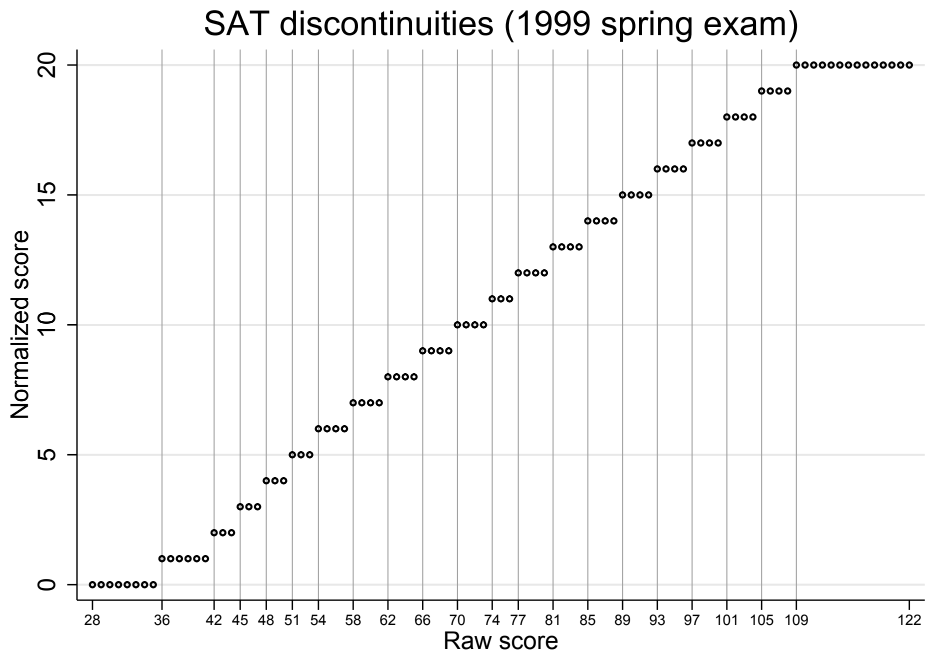


Figure 2: The relationship between raw and normalized SAT scores

We use administrative, population-wide data on raw and normalized scores for all SATs taken during 1994-2006. We link this information to data on parental education and background characteristics such as (pre-college) grades from compulsory school and high school as well as outcome variables measuring college enrollment by field and school, credit production, and graduation.

Starting with the overall patterns, we find that students who by chance receive an incrementally higher SAT score—equivalent to one third of a standard deviation in SAT results—on average are 3 percentage points more likely to be enrolled in college two years after the test. The overall enrollment effect is mirrored by a higher rate of college credit production over 4-5 years after the test. Half of the initial effect on college credits remains in terms of accumulated credits twelve years after the test, and translates into a 3 percentage point higher probability of graduating from college. These effects are strongest for test takers at high thresholds. Remarkably, the extensive margin responses are also visible at the highest threshold which is crossed by less than one percent of all test takers. As these individuals have very attractive college opportunities regardless of the test outcome, we are clearly exploiting variation in the set of possible opportunities and *not* on the margin of whether or not the student would be admitted anywhere.

The differences in effects in relation to parental background are striking. The long-run educational responses are much larger if parents are at most high school graduates (low SES).

studies and retake the test to improve their admission opportunities, but our results show that the crossing of each threshold does have a sizable effect on life-time best scores.

These low-SES individuals have a 4 percentage point higher probability of being enrolled within two years and are 3.5 percentage points more likely to have a college degree twelve years later if they receive the higher score. Corresponding numbers for individuals with college-educated parents (high SES) are 1.5 percent (not statistically significant) on initial enrollment and 1.4 percent on degree probability (also not significant). Instead, high-SES students respond by enrolling in more prestigious colleges and programs, and in that it is mainly timing that is affected, not life-time enrollment.

To help the interpretation of our results, we provide a stylized theoretical framework which mimics our institutional set-up. In the framework, individuals have idiosyncratic valuations of different types of college programs and trade off the best achievable program against the best outside (no college) option. A positive SAT outcome expands the set of possible college opportunities and thus the probability that at least one achievable college option dominates the outside option. The framework suggests that individuals who put a high fixed (that is, regardless of quality) value on college will respond on the intensive margin and enter more competitive programs, whereas individuals who have a low fixed valuation and possibly strong idiosyncratic valuations of educational programs, instead respond on the extensive margin.

The framework naturally excludes the possibility that rational and informed agents increase their enrollment in programs that are attainable even without the expansion of the choice set. These predictions appear valid for high-SES students; a higher score in this group increases enrollment in top institutions and the most competitive fields (law and medicine) and instead reduces their participation at less competitive colleges. In contrast, low-SES students only increase their participation in the *less* competitive institutions without changing enrollment rates in top colleges or law and medicine. The strong extensive margin response we documented at the highest threshold is also fully explained by the low-SES students. Jointly, these results seem to suggest that low-SES students respond to a higher score by increasing their enrollment in programs that they would be eligible for even without the higher score. Using published data on admission thresholds, we directly study the enrollment response for programs where we are certain that the student would be admitted on both sides of the SAT score threshold. Strikingly, we find that almost all of the enrollment effect for low-SES students are within such 'always-attainable' programs.

Our results thus suggest that high-SES students respond to increased admission opportunities by intensive margin adjustments that are consistent with the behavior of rational and informed agents who have a high fixed valuation of college relative to their valuation of an immediate transition to the labor market, and who have a positive marginal valuation of more competitive programs. In contrast, low-SES students primarily respond on the extensive margin relative to programs they are always able to attend. This is inconsistent with our model of rational and informed behavior. We conjecture that the pattern arises because low-SES students are constrained by lack of information or behavioral biases, despite of the streamlined system, and that a higher score provides them with additional incentives to gather information about different aspects of college life in general. A related and possibly reinforcing mechanism is that students may receive additional confidence from the official—that is, normalized—test score. But since test takers also get to see their raw score, as well as the thresholds, they will know if they are on the margin of a better or worse score. The confidence mechanism could therefore only work if there is an additional effect of the normalization on top of the information provided by the raw score.[7]

Overall, our results suggest that parts of the SES-gap among Swedish top-students documented in Figure 1 arise because low-SES students are constrained by information frictions or behavioral barriers, despite of the streamlined institutions. In particular, it appears that some of our very talented low-SES students need a very large opportunity set in order for them to enter college, even if they in the end do not choose to enter into any of the more competitive options.

The paper is structured as follows. First, we provide an overview of the related literature in Section 1.1. Section 2 provides institutional background, in particular it explains the role of the SAT in Swedish college admissions and introduces our data. Section 3 explains our identification strategy. Section 4 introduces a simple theoretical framework that informs our empirical results, which we present in Section 5. Section 6 concludes.

1.1 Related literature

Our study is closely related to the literature on socioeconomic gaps in beliefs and information sets, and how marginal interventions altering these beliefs can affect educational attainment. Hastings et al. (2016) show how beliefs about costs and returns to different college options are related to college choices in Chile. Belfield et al. (2019) and Boneva and Rauh (2019) highlights large socioeconomic gaps in beliefs about non-pecuniary benefits of university education in the UK, which can account for a sizeable proportion of the socioeconomic gaps in students intentions to pursue higher education.

Related experimental studies have found that information, encouragements, simplifications of application processes, and small instant cash rewards have substantial effects on college enrollment of low-SES individuals. Bettinger et al. (2012) showed that helping low-SES parents fill out funding applications had a substantial impact on the rate of applications for college funding in the US. Similarly, Hoxby and Turner (2015) randomly provided students with information about application procedures and costs, as well as no-paperwork application fee waivers, which induced high-achieving low-SES students to increase enrollment at college. Dynarski et al. (2018) randomly provided able low-SES students with a promise of hassle-free tuition

[7] For completeness, we also explore other mechanisms that may explain heterogeneous effects as suggested by our theoretical model. These include SES differences in opportunity costs, returns, as well as in the incidence of repeating the SAT. Only the latter goes some way in accounting for the heterogeneous effects of higher scores, but at most in a proximate sense.

waivers at a prestigious US university, finding a large impact on enrollment rates, which more than doubled in response to the intervention. Other papers in this vein include Hoxby and Avery (2013), Marx and Turner (2019a) and Marx and Turner (2019b). See also the survey by French and Oreopoulos (2017).[8] Although our quasi-experimental variation is very different from the interventions discussed in these paper, they share the feature that random alterations in the perceived choice set can have large impacts on college enrollment rates of low-SES students. Whereas these previous studies find that marginal encouragements or reductions in admission frictions induce enrollment, we show that marginal increases in ex ante (but ex post, unused) college opportunities have sizeable effects on college enrollment in a setting where admission frictions are removed by policy design.

In addition, our paper is related to (but distinct from) a large existing literature using admission cutoffs for identification in regression discontinuity designs. An influential recent contribution is Kirkeboen et al. (2016) studying the effects of program choice in Norway. An earlier example is Öckert (2010) who estimates the returns to college using data on applicant pools and cutoff scores in Sweden.[9]

Papers in this literature typically use data on the ranking of applicants to a given program, and are able to estimate the causal effects of admission by comparing individuals who barely made it across the cutoff score required for admission to those who barely missed the same threshold. Due to the local nature of such comparisons, these papers tend to focus on the intensive margin response (choice of program, or type of school) among students with educational preferences across fields or schools. But the approach used in these papers limits the set of possible responses since the randomization occurs after preferences are elicited. Furthermore, lower-ranked options are removed when students are admitted to higher ranked options, at least in the context of national admissions. In contrast, our identification gives rise to quasi-random variation in the set of programs that individuals are eligible for and our treatment—obtaining a higher SAT score—raises admission probabilities across multiple programs, and over multiple admission cycles, thus allowing preferences and beliefs to change in response to opportunities. Our treatment typically occurs *before* the student applies to college, which has been argued to be an important dividing stage between low-SES and high-SES individuals (Hoxby and Avery, 2013).

Our work is also broadly related to the existing research on the effects of college-relevant

[8] One explanation for the large effects reported in this literature is incomplete development of long-run decision making capabilities among individuals, see Lavecchia et al. (2016), which suggests an important role for parental involvement, which typically increases with parental education, see Guryan et al. (2008).

[9] Other studies in this vein include Hoekstra (2009) and Zimmerman (2014) that estimate the returns to a particular institution in the US., Canaan and Mouganie (forthcoming), Anelli (forthcoming), and Fan et al. (2017) that studies the returns to college in France, Italy, and China, respectively. Other applications include Humlum et al. (2017) that estimate the effect of crossing an admission threshold on the timing of college enrollment and the timing of family formation in Denmark, Zimmerman (2019) who study the returns to elite college acceptance on leadership positions using data from Chile, and Altmejd et al. (2020) that study spillovers from older siblings' college admissions on younger siblings' educational choices using data from Chile, Croatia and Sweden.

test scores. Smith et al. (2017) and Avery et al. (2018) use discontinuities arising from the mapping of raw test scores into reported integer scores to estimate the causal effect of earning college credits in Advanced Placement programs in high school on college choices and completion. Papay et al. (2016) use a similar strategy to estimate the effect of receiving a positive performance label on their state-mandated standardized test in high school on college attendance. We share the general empirical context as well as the econometric strategy with these papers, but the treatments differ. Most importantly, as emphasised in these papers, receiving college credits or a positive performance label does not alter college enrollment opportunities. In contrast, individuals receiving a higher SAT score have the opportunity to enroll to a wider set of colleges and programs, and we document positive effects of higher SAT scores on enrollment even among students of very high ability.

Two other related papers study different features of the SAT in the US. Bond et al. (forthcoming) explore how students' college application portfolios change in response to learning about their SAT score. They make use of a policy that induces students to select colleges prior to taking the exam, and show that students update their choices in terms of college selectivity and tuition when receiving information about their SAT performance. Our paper is similar in terms of the intensive margin analysis, but because our treatment happens much earlier in the process, we are also able to detect extensive margin responses. Finally, Goodman et al. (2018) estimate the effects of retaking the SAT on college enrollment. They exploit discontinuities in the probability to retake the exam at multiples of 100, driven by left-digit bias, to estimate the causal effect of retaking the exam. They find that retaking the exam improves SAT scores and increases the probability to go to four-year colleges, in particular for low-SES students. Also in our setting, passing a SAT threshold reduces the retaking probabilities. In contrast to Goodman et al. (2018), however, passing a SAT threshold in our setting has a direct effect on college enrollment opportunities for all students, irrespective if they retake the exam or not.

2 Background

2.1 College education in Sweden

To be eligible for college, students in Sweden must complete three years of high school, which follow after nine years of compulsory schooling. When starting college, students either enter into programs or into single-subject courses. Teaching follows a two-term system (fall and spring semesters). Programs typically last 3-5 years. Students are awarded a degree on successful completion of a full program. Single courses are usually one semester long, but can be combined to a full degree. On top of high school completion, programs and courses can have special subject-specific eligibility requirements, such as having completed a sufficiently advanced math course in high school.

College in Sweden is always free of tuition, and to cover living costs there are generous study grants and loans available to all students. Access to such financial support is universal and does not depend on any aspect of family background. Since requirements are few, and enrollment can be verified directly by the funding agency, filing a request for financial support is very easy.[10] To continue receiving financial support beyond the first year, the student needs to show adequate progression in terms of credit point production and the number of available terms is capped. Grants are reduced if the student's *own* earnings exceed a certain threshold during enrollment (Högskoleverket, 2000). During our sample period, the total amount (grant plus loan) received for one year of studies corresponds to about 50-60 percent of a full-time full-year employment at the 10th percentile of the wage distribution.

2.2 College admissions

Applicants can apply to many different programs and institutions and admission standards are nationally regulated. Admission to most programs is conducted twice a year by a central agency, with roughly three quarters of places allocated in the fall semester. All programs have a fixed number of slots, and many are oversubscribed. If the number of applicants exceeds the number of slots, candidates are ranked based on high school GPA and SAT scores (details below). In the period studied, applicants with the same formal merits were separated by their preference ordering of the alternative in relation to the other alternatives for which they had applied. Any remaining ties were settled by randomization. By law, colleges are required to fill at least one third of places based on a GPA ranking, and at least one third of places based on an SAT ranking (Edwards et al., 2012). Colleges must decide in advance how many slots are to be filled based on GPA, SAT or other admission instruments (Högskoleverket, 2000).[11]

2.3 The Swedish SAT

The Swedish SAT is a standardized, centralized test administered by the Swedish Council for Higher Education, and designed based on the American SAT (Government of Sweden, 2004). Taking the test is not required for college admission, but many students who wish to proceed to college do take the test as it may improve (and cannot reduce) admission opportunities. The test takes place twice each year, in April and October, and there is a small fee to participate. It includes between 120 and 150 multiple-choice questions testing both numerical skills and language skills. The raw score equals the number of questions answered correctly. From the raw score, a twenty-step normalized score is generated which ranges from 0 to 2 in 0.1 increments. One step on the normalized score corresponds to about one third of a standard deviation. For

[10] See https://www.csn.se/languages/english.html for details.

[11] Other instruments are usually work experience and/or program-specific entry exams, but in most cases schools choose to use GPA for 2/3 of their slots.

simplicity, we multiple the normalized score by ten to obtain an integer scale from 0 to 20. Only the normalized score is reported to the admission process. The raw score, however, is reported to the test taker alongside the normalized score and the threshold values.

The normalized score is a step function of the raw score, as illustrated in Figure 2. The cutoffs are determined such that the distribution of the normalized score mimics a truncated normal distribution. Since the distribution of raw scores is only known after the exam, the cutoffs are set ex post and thus unknown by the student at the time of taking the test. Thus, even if test takers were able to precisely manipulate their raw scores, they would not know which raw score is required to achieve a given normalized score. It is therefore reasonable to assume that the discontinuities in raw SAT scores induce quasi-random variation in normalized scores. This assumption is also corroborated by our analyses in Section 3.

As mentioned above, Swedish colleges are legally required to fill programs based on both high school GPAs and SAT scores. Applicants are placed in different quota groups depending on their relevant merits, and those in the SAT quota group typically simultaneously compete for places in the GPA quota. Thus, passing a SAT-threshold typically opens up more admission opportunities for applicants with lower high school GPA.

Step 1: Assigning first three places by GPA

id	GPA	SAT score
A	2.2	20
B	1.8	17
C	1.7	20
D	1.5	13
E	0.9	19
F	0.6	16
G	0.4	17
H	-0.3	9

Step 2: Assigning remaining two places by SAT

id	SAT score
E	19
G	17
F	16
D	13
H	9

Figure 3: A fictitious applicant pool with five available places

A stylized example for an admission procedure is given in Figure 3. Here, there are eight applicants for five slots. The college assigns the first three places to the top three candidates in the GPA ranking (step 1). It then assigns the two remaining places to the top two remaining candidates in the SAT ranking (step 2). Our empirical strategy generates counter-factual scenarios by essentially randomizing SAT scores via the discontinuities in the raw score. In this example, an SAT score of 17 is sufficient to make the cut. Candidate G's counter-factual of scoring 16 thus would mean the loss of this particular opportunity. Notice however that if candidate B had scored 16 instead of 17 it would not change her opportunities, since her GPA is relatively high and her SAT score was not relevant for admission. Thus, students with a relatively low high school GPA should benefit more from having a higher SAT score, a prediction that is supported by our data. In most of our analysis, we therefore focus on students with a high school GPA below the median among the students near the same SAT threshold. We use data on compulsory school GPA to confirm that the differences in responses we find are due to access to an alternative admission currency (high school GPA) and not due to ability—there are no corresponding differences in responses by compulsory school GPA.

Our empirical strategy will *not* be based on the cutoffs in an admission pool among applicants. In Figure 3, candidate D is not a valid counter-factual for candidate G, since normalized scores are correlated with many potential confounders, as we document below. The admission cutoff from a ranking based on SAT scores would only give credible identification if there were ties that are randomly broken. Our identification strategy instead uses discontinuities in raw SAT scores to induce quasi-random variation in normalized scores, as seen in Figure 2. As will be evident below, the distinction is important for parts of our results since admission cutoffs removes those options that are lower ranked when a student is admitted to a higher ranked alternative. Thus, if using admission cutoffs, it is impossible to detect if students' college attendance respond to irrelevant alternatives, something which turns out to be important in our setting.

To illustrate how SAT scores matter in practice, we draw on data from the 1999 fall admissions. About 80 percent of enrollment was in oversubscribed programs. Half of all students enrolled in a program for which an SAT score of 12 or less sufficed to be admitted, and more than 75 percent enrolled in a program with an SAT cutoff score of 14 or less. Anticipating our empirical results, we see the strongest effects on enrollment and graduation at thresholds 15-20. What were the additional opportunities that became available in 1999 with the crossing of one of these higher thresholds? One important margin was field of study: The cutoff SAT score for Medicine at all colleges was the top score of 20 whereas all economics programs were attainable with a score of 19. Another important margin was institution: To enter the Economics & Business program at Stockholm School of Economics, the cutoff score was 19, while at Umeå University the same program required a score of 14. A third margin was field within institution: At the Royal Institute of Technology, the cutoff scores for Industrial Engineering, Engineering Physics, and Electrical Engineering were 19, 18, and 16, respectively. In sum, students who score at least 14 on the SAT typically have access to a large set (about 75 percent of all slots) of college opportunities. But scoring higher than 14 opens up attractive additional opportunities, in terms of both colleges and programs. More details are presented in the appendix: We show the full distribution of cutoff scores across programs in the fall of 1999 in Figure A1, and provide a list of programs with cutoff scores in the range 14-20 in Table A1.

2.4 Data

Our data come from individual-level administrative registers, covering the entire Swedish population, that have been merged by Statistics Sweden. The data include raw and normalized SAT scores from all tests taken; compulsory school (year 9) and high school (year 12) grade point averages (GPAs); information on college enrollment, credit accumulation, and graduation by institution, program, and course; for individuals older than 16, separate information on highest education attained; annual labor income; year and country of birth; identity of parents. Our sample includes tests 1994-2006 and test takers aged 16-30, hence cohorts born between 1963 and 1990.

We report a variety of descriptive patterns in the appendix, which we briefly summarize here. About half of all high school graduates took the test at least once before age 30 among the more recent birth cohorts in our sample (1975 or later). Compulsory school grades are strongly positively associated with SAT participation, as are high school grades on their own. But conditional on compulsory school grades, high school performance is a *negative*—though weak—predictor of test taking. This suggests that typical test takers are able students without very attractive alternative enrolment opportunities (since high school grades are an 'application currency' similar to SAT scores). Furthermore, female, Swedish-born, and high-SES students are more likely to take the test. Positive predictors of test performance include school grades and parental education. Immigrants and female students do substantially worse on the SAT. The latter is particularly surprising given that female students receive systematically grades in school. This variation in gender gaps is investigated by Graetz and Karimi (2019).

3 Empirical strategy

Our data are at the level of person (i) and test (j). Crossing score threshold s on test date j is indicated by D_{ijs}. The raw score serving as the running variable is denoted by x_{ijs}, and it is expressed in deviations from the threshold s (where zero indicates just having crossed the threshold). As our outcomes, denoted by y_{ij}, we use various measures of educational choices and performance at different time horizons since taking the test. Following standard practices, our RD specification thus regresses the outcomes y_{ij} on a dummy for crossing a threshold, the running variable, their interaction, as well as test date dummies λ_j and a set of predetermined controls $\mathbf{w}_{ij}$. The controls include a dummy for female, foreign born, and age at test dummies. They are not required for consistent estimates of the treatment effects θ_s, but they potentially increase precision. Formally, we estimate by OLS:

$$y_{ij} = \theta_s D_{ijs} + \beta_s x_{ijs} + \gamma_s x_{ijs} D_{ijs} + \lambda_j + \mathbf{w}'_{ij}\delta + \varepsilon_{ij}. \tag{1}$$

For our benchmark specification, we pool observations across all thresholds in the 15-20 interval to improve statistical power (we report non-pooled results in the appendix). When doing so, we estimate the slope of the running variable separately for each threshold and at both sides of the threshold, as well as allowing for separate intercepts at each threshold:

$$y_{ij} = \theta D_{ij} + \sum_{s=15}^{20} \left[\alpha_s + \beta_s x_{ijs} + \gamma_s x_{ijs} D_{ijs}\right] + \lambda_j + \mathbf{w}'_{ij}\delta + u_{ij}. \tag{2}$$

Since individuals may be observed repeatedly, both due to repeated test-taking and due to the stacking of thresholds, we cluster standard errors at the level of individuals.

3.1 Repeaters

Including all tests, that is, using multiple tests for the same individual, increases statistical power and generalizes the results to the overall population of test takers. However, the probability of repeating is itself affected by the normalized score. Estimating (1) with the probability of ever repeating as the outcome (using only each individual's first test) we find that crossing any of the upper six thresholds, which we focus on in most of our analysis, reduces the probability of repeating. The effect is more pronounced at the highest thresholds, as shown in Figure A3.

Importantly, the parameter θ retains its causal interpretation even if repeating is endogenous, since crossing a threshold is quasi-randomly assigned among repeaters and non-repeaters alike. However, endogenous repeating means that θ represents an intention-to-treat rather than an average treatment effect (Cellini et al., 2010). In the extreme, individuals who fail to cross a given threshold may be certain to achieve the relevant score at the next test opportunity, while individuals who cross at the first try may never repeat. In this scenario, both groups would have the same life-time maximum score, and the crossing of a threshold at the first test would not affect long-run outcomes. The intention-to-treat effect on long-run outcomes would then be zero. However, in our setting, crossing a threshold at the first test has large effects on the maximum life-time score, ranging from 0.5 to 0.75 depending on the threshold. For individuals who do not repeat the test—about half of test takers—the effect equals one by necessity. Dropping these individuals from the estimation, we still see positive, if smaller, effects of a higher score in the first test on the maximum life-time score.

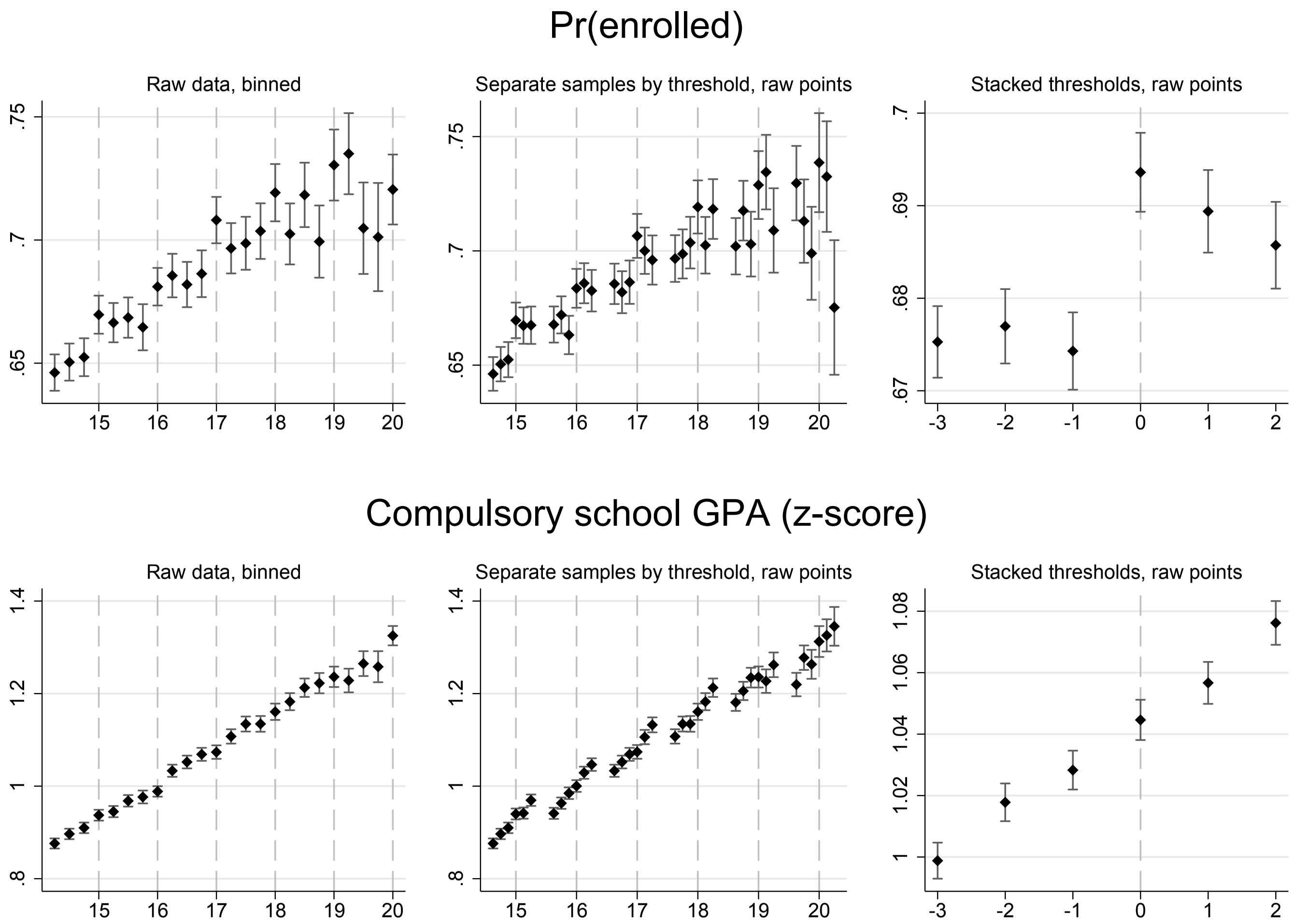
3.2 Validity

Identification in any RD-model relies on comparing individuals very close to the threshold. With a discrete running variable, there is a natural limit of how close one can get and we need data further away from the threshold to model the relationship between the outcome and the running variable. In our setting, we can at most move three points away from a given threshold without crossing one of its neighbouring thresholds. We therefore fit lines on either side of the threshold using these three observations. To illustrate the issues we face, and to motivate our preferred solution, we plot the means of enrollment and the compulsory school GPA (a predetermined ability measure) against various transformations of the raw score in Figure 4. The left-side panels are standard binned scatter plots. The need for binning arises because the number of steps on the raw score between two increments of the normalized score varies across thresholds and test dates. The middle panels group the data separately by threshold, including three points on either side of each threshold. These grouped data are our basis for estimation. The right-side panels collapse the data to the level of raw score distance to a threshold. Most of our results are based on such stacked data. The top right panel of Figure 4 may suggest that we could estimate the effect of a higher score by simply comparing the means of an outcome one point to the left and directly at the threshold. However, the bottom right panel shows that predetermined ability (compulsory school GPA) has a small but statistically significant positive association with the raw score at this frequency which translates into a statistically significant jump across the thresholds if we do not use the controls. To eliminate this potential source of bias, we control for the slope of the running variable linearly, separately for each threshold and on each side of a threshold, as in (2).

To test for the validity of our RD design, we estimate (2) with various pre-treatment outcomes on the left-hand side. These include age at test; indicators for female and foreign born; compulsory school GPA and high school GPA;[12] indicators for parental education (at most high school; at least college educated) and the previous SAT score (for repeaters). Table 1 shows that estimated differences in these characteristics around a threshold are of negligible magnitude relative to the means, and statistically insignificant in all cases. In particular, note that the compulsory school GPA now appears to drop by a statistically insignificant 0.002 of a standard deviation (column 4, lower panel). Controlling for the running variable thus eliminates the positive jump we saw in Figure 4.

We also conducted a joint balancing test as follows. We put the crossing of a threshold on the left-hand side of (2), and included all variables considered as outcomes in Table 1 on the right-hand side, while controlling for the running variable as before. The results are supportive of the null hypothesis that all pre-treatment variables have zero effect on the probability of crossing a threshold, once controlling for the running variable (the p-value for the full sample is 0.94 and for the upper thresholds it is 0.89).

[12]Since many students take the SAT before finishing high school, and since the high school GPA is partly determined by performance on courses in the last year, the high school GPA is not a strictly predetermined outcome.



Notes: Means of enrollment two years after the test and of the compulsory school GPA are plotted against various transformations of the raw SAT score. For the left panels, the raw score was re-scaled such that it equals the normalized score at each score threshold, and then rounded such that there remain four steps between each increment of the normalized score (the actual number of steps varies across thresholds and exams). For the middle panels, three points of the raw score to the left and to the right of each threshold are used. For the right panel, the dependent variables are collapsed to the level of the raw score distance to a threshold. Capped bars indicate 95-percent confidence intervals.

Figure 4: Grouping the data around thresholds, and stacking of thresholds

Table 1: Balancing tests

	(1) Age at test	(2) Female	(3) Foreign born	(4) CSGPA	(5) HSGPA	(6) Parents HS	(7) Parents COL	(8) Previous best score
All thresholds								
Above cutoff	-0.0061	-0.0015	0.00071	0.0025	-0.00082	0.0011	0.00053	-0.0074
	(0.012)	(0.0018)	(0.00087)	(0.0026)	(0.0031)	(0.0018)	(0.0013)	(0.011)
Mean of dep. var.	21.3	0.52	0.06	0.59	0.23	0.45	0.14	9.7
Observations	1,656,816	1,656,816	1,656,816	1,475,124	1,499,439	1,625,853	1,625,853	864,566
Thresholds 15-20								
Above cutoff	0.023	0.0055	0.00095	-0.0024	-0.0068	0.0017	-0.0030	-0.020
	(0.029)	(0.0043)	(0.0015)	(0.0060)	(0.0078)	(0.0040)	(0.0038)	(0.019)
Mean of dep. var.	21.7	0.41	0.03	1.03	0.76	0.30	0.26	14.6
Observations	280,566	280,566	280,566	243,940	255,314	278,018	278,018	202,933

Notes: Results from stacked RD specifications are shown. Specifications are as in (2), but with pre-determined outcomes, and without controls $\mathbf{w}_{it}$, some of which instead appear on the left-hand side. The unit of observation is person-test-threshold. Standard errors, clustered by individual, in parentheses.

4 A model of college choice

Here we present a theoretical model of college choice that assumes rational and informed agents and delivers a baseline interpretation of our empirical findings. The model provides conditions under which students respond on the extensive margin, or on intensive margins such as quality and timing, when the choice set randomly expands. Importantly, we highlight what kind of behavior is ruled out by the model.

Individuals choose from menus of non-college jobs and college programs, indexed by j and c, respectively. Each choice has an expected discounted utility V. All parameters and values may differ across individuals, but we omit individual subscripts to avoid clutter.[13] The value of the most attractive of all attainable jobs is denoted by $V^J \equiv \max_j V^J(j)$. Similarly, the value of the most attractive program within the attainable choice set (C_z) is denoted by $V_z^C \equiv \max_c \{V^C(c) : c \in C_z\}$.

The SAT discontinuities in our empirical application are represented by the 'instrumental variable' $z \in \{0, 1\}$, which equals one if a higher SAT score is obtained at a given threshold and zero if the lower score is attained. We use here the conceptual framework of instrumental variables, even though our analysis mainly consists of reduced form results. The framework is useful as it allows us to talk about heterogeneous responses to higher scores in concrete and familiar terms.

The realization of z is positively related to admissions opportunities (the choice set), such that $C_0 \subseteq C_1$. Hence, the utility of the best attainable program is weakly increasing in z, $V_0^C \leq V_1^C$.

Individuals who are rational and informed respond in one of the following ways to the realization of z:

1. If the best program in the expanded set (when $z = 1$) remains dominated by the best job, $V_1^C \leq V^J$, the individual chooses the labor market regardless of the instrument and is thus an *extensive-margin never taker*.[14]

2. If the best program in the smaller set (when $z = 0$) dominates the best job, $V^J < V_0^C$, the individual always chooses a college program and is thus an *extensive-margin always taker*. These come in two forms:

    - If the best choice in the expanded set equals the best choice in the smaller set, $V_0^C = V_1^C$, the individual does not respond to the additional opportunities and chooses the same program regardless of the instrument (*intensive-margin never taker*). Note

[13] Heterogeneity across individuals may be driven by e.g. discount rates, tastes or beliefs.

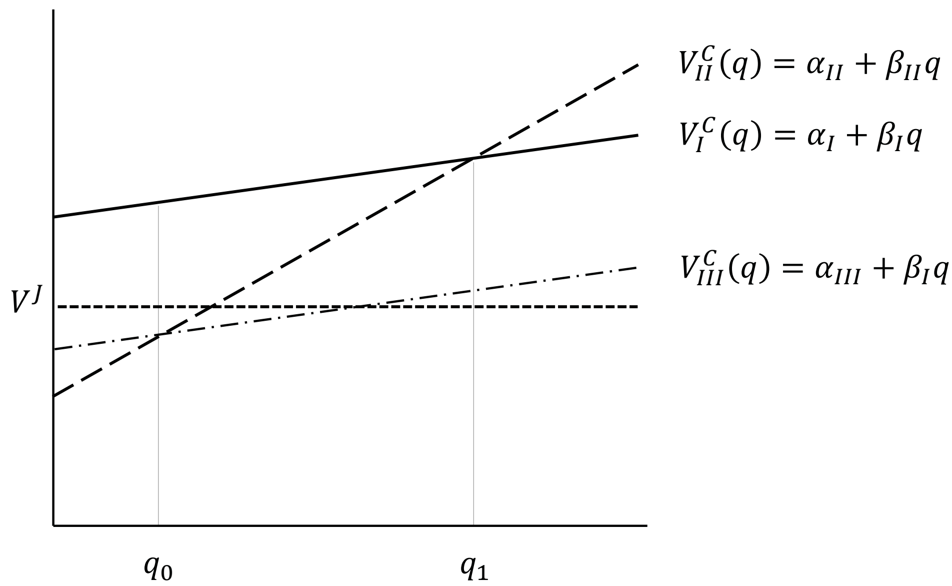
[14] We assume that in case of indifference individuals always choose working over college, and a lower quality college over a higher quality college.

that the case $V_0^C = V_1^C$ also arises if the instrument does not actually provide the individual with additional opportunities ($C_0 = C_1$) say because she has an exceptionally high GPA.

- If the best choice in the expanded set is different from the best choice in the smaller set, if $V^J < V_0^C < V_1^C$, the instrument induces access to a more highly valued, but otherwise unattainable, program and the individual is an *intensive-margin complier* who chooses another program in response to the instrument.

3. If the best job V^J is valued higher than the best choice in the smaller set but lower than the best choice in the expanded set, $V_0^C \leq V^J < V_1^C$, the individual goes to college with a positive realization, but not otherwise (*extensive-margin complier*).

The model implies that intensive-margin compliers place a high *fixed value* on college (relative to work), whereas extensive margin compliers must have a lower fixed valuation, and instead value (perceived) *quality* sufficiently to warrant an extensive-margin reaction.



Notes: Colleges differ by quality q, and the instrument (a higher SAT score) changes the best available option from q_0 to q_1. $V_I^C(q)$ is the value that type I attaches to a college of quality q, and similarly for types II and III.

Figure 5: Valuation of different college programs for three different types

We illustrate the intuition in Figure 5. For simplicity, let colleges vary only along a single-dimensional quality index q, and let college valuations be linear in quality. The instrument changes the best attainable program from q_0 to q_1. We consider three distinct types of individuals with the same ordering of quality and normalize the labor market option to be valued equally at V^J. Type I (solid line) attaches a high fixed value to college and is thus an extensive-margin always taker, and an intensive-margin complier as long as she places any value on quality. Type II (dashed line) has a low fixed valuation of college, but a high valuation of quality. She values the high-quality program as much as Type I but prefers working to attending a low-quality program and is therefore an extensive-margin complier. Finally, we note that an individual with a low valuation of quality can be an extensive margin complier, but only if she values work and a low-quality program relatively equally (Type III, dash-dotted line).

This exercise suggests that if, for instance, individuals with well-educated parents have a higher fixed (potentially non-pecuniary) valuation of college, then they will respond more on the intensive, but not extensive, margin (Type I in Figure 5). And if, for instance, individuals with lower educated parents value high-quality programs, but not college in general, then they will be more likely to respond as types II and III in Figure 5.

4.1 Repeating

A salient feature of our empirical setting is that students may repeat the SAT. The choice of repeating can easily be incorporated into our framework, which leads to the predictions that individuals are more likely to repeat the test if they have a higher probability of succeeding at the next attempt; if they are more patient; if they face lower costs of preparing for the test; and if they enjoy a higher return to a higher-quality college program over their current option (working or being enrolled in a lower-quality program). Parental education may correlate negatively with discount rates and costs of preparing for the SAT. High-SES students will then be more likely to repeat the test, and will exhibit lower intention-to-treat effects of higher scores for long-run educational attainment.[15]

[15] Suppose that individuals have infinite lives and let the life-time utility of the outside option (not repeating) be denoted by V^O. Consider repeating as an option for individuals who work as well as those who are already enrolled, so that $V^O \in \{V^J, V_0^C\}$. Let the discount factor be denoted by β. Suppose that utility flows are constant across periods, except when paying a waiting cost or when switching states. The flow value of the outside option during a single period of waiting can then be written as $(1-\beta)V^O$.

Consider an individual who has not succeeded at the SAT yet (in the sense of achieving the higher score). Let us assume that it is possible to repeat the SAT once, at the beginning of the next period; and that a new enrollment decision can be made immediately after taking the test. Further, we assume that repeating means that a cost κ is incurred, reflecting any study that must be undertaken in preparation of the test. The probability of a higher score at the next test is p. Thus, individuals decide to retake if and only if

$$(1-\beta)V^O - \kappa + \beta\left\{pV_1^C + (1-p)V^O\right\} > V^O,$$

which can be rewritten, also making concrete the two possible outside options, as

$$p\beta\left(V_1^C - V^J\right) > \kappa, \qquad p\beta\left(V_1^C - V_0^C\right) > \kappa. \tag{3}$$

The comparative statics discussed in the text follow directly from (3).

Regarding heterogeneity, note that if low-SES individuals are of Type II, then they could actually be *more* likely to repeat the test, since $V_{II}^C(q_1) - V^J > V_I^C(q_1) - V_I^C(q_0)$. Differences in study costs and discount rates may counteract this, however. If low-SES individuals were of Type III, then based on valuations alone they would be less likely to repeat the test than high-SES individuals since $V_{III}^C(q_1) - V^J < V_I^C(q_1) - V_I^C(q_0)$.

4.2 Beyond the rational and informed

By assuming rational and informed agents our model restricts the set of permissible responses. In particular, the model does not allow a student to respond to a positive draw ($z = 1$) by increasing her participation in a program which already was attainable within the smaller choice set (C_0).

Observing an increase in participation in such 'always-attainable' programs in response to an expansion of the choice set implies that the agent does not act as if fully rational and informed. As noted in the introduction, these types of behavioral responses are typically associated with low-SES students in the existing literature.

One possibility for such a behavioral response in our setting is if an expanded choice set induces students to acquire more information about content, difficulty, or returns of college programs, which may increase enrolment. Similarly, students may not know which programs they are likely to have access to, and an incrementally higher score may induce them to collect information that allows them to better forecast their admission outcomes. Another alternative is that individuals who are uncertain about their own ability may experience a boost in self-confidence if crossing a threshold. But since the empirical model controls for the raw scores this *only* works if confidence is a function of reported (rounded) scores and not of the raw scores directly.

4.3 Differences to admissions discontinuities

The model highlights the contrast between our empirical setting and admission cutoffs used by Kirkeboen et al. (2016) and others. These studies rely on a centralized admission system which can be assumed to elicit a truthful reporting of students' preferences across college programs, and which allocates program slots among applicants based on unpredictable grade cutoffs. In such a setting, the instrument will be a dummy variable for being above the cutoff for a given program c. The complying population for this instrument are individuals who enroll in program c if the instrument is switched on, and in their next-best program if it is switched off (in practice marginal students with cross-field preferences). This setting provides the researcher with variation in program or field of study in isolation. This is crucial for estimating the economic returns to field of study, and for informing policies that change the number of slots in given programs or fields.

In contrast, studying variation in enrollment opportunities arising from SAT-discontinuities allows us to unravel a large set of adjustments in terms of educational choices and outcomes. For example, our strategy allows us to identify effects on college participation that arise if expanded opportunities really do increase the likelihood of participating in 'always-attainable' programs as discussed above. This is not possible with admissions discontinuities since the option of participating in the lower-ranked program is removed when accepted at a higher-ranked program within centralised admissions systems. In addition, our strategy allows us to detect adjustments that arise due to increased returns to information-gathering before students rank their choices, a type of insight that cannot be derived at the later (admissions) stage when the ranking has already been submitted.

5 The effects of SAT scores on college enrollment and graduation

In this section we present our main results. We first present average effects of obtaining a higher SAT score along with extensive robustness checks, before turning to heterogeneity with respect to parental background.

5.1 Average effects

We begin by showing the impact of SAT-scores on enrollment in college for a range of time horizons in Figure 6, top left panel. The results are based on equation (2). For each of the first three years after the test, we find that the probability of being enrolled is increased by 2-3 percentage points for those who obtained a 'randomly' higher score because they just managed to cross a threshold. In the longer run, enrollment rates converges as expected since enrollment rates fade when students leave for the labor market as illustrated in the top right panel.[16]

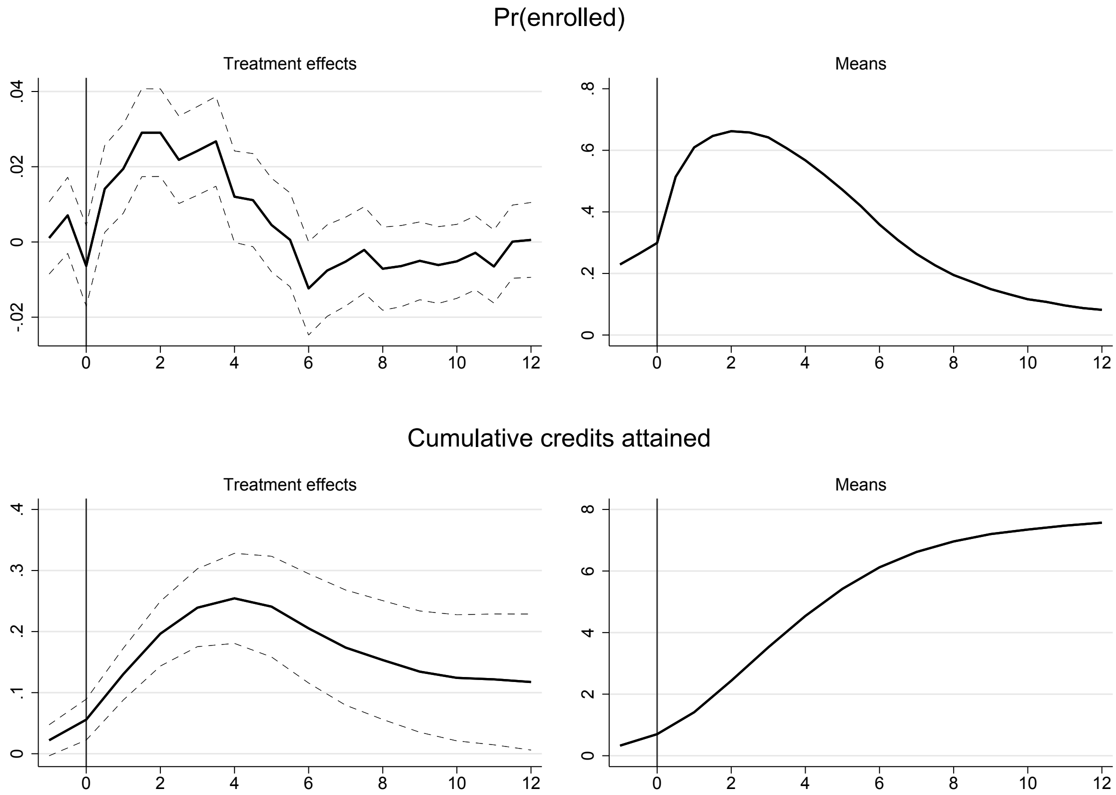
The increased enrollment rates during the first few years translates into a higher number of *accumulated* college credits, both in the short and the long run as shown in the bottom left panel. The maximum impact is seen four years after the test, when higher-scoring students have accumulated the equivalent of a quarter of a full-time term in additional credits. In the long-run (twelve years after the test) the effect is about half of this. Thus, a higher SAT score affects both the timing and the incidence of college education. The magnitudes of the estimates are modest, but of clear economic interest; they amount to about five percent of the respective means as shown in the right-side panels.[17]

For completeness, Figure 6 also displays the estimated effects on enrolling and accumulated credits in the year *prior* to taking the test. Reassuringly, the point estimates are small and statistically indistinguishable from zero. Note that the time-zero effect is positive for credits because the SAT data are by term, while credits are only measured over the full academic year in our data. (See Section 3 above for additional check on the validity using other predetermined variables.)

In Table 2 (Panel A, columns 1, 2, and 3), we summarize the key insights from figure 6 by presenting the impact on enrollment after two years and the impact on accumulated credits

[16]In addition to the extensive margin enrollment effect, there is an intensive margin effect on the choice between 'full programs' and 'stand-alone courses'. In particular, the effect of a higher score on enrolling in a program is slightly higher than the effect on enrollment of any type (which we focus on above). This is explained by the fact that the effect on enrolling in a stand-alone course is negative (though small in magnitude). The results are shown in the top panel of Figure A4.

[17]While effects of a higher score on the probability of enrollment must be short-lived due to the temporary nature of enrollment, we might expect a persistent effect on cumulative terms enrolled. However, the bottom panel of Figure A4 shows that the effects on cumulative enrollment as well as the on the cumulative credit value of courses taken eventually fades away. Individuals appear to continuously enroll in college over the course of their careers. However, this enrollment appears to be of temporary nature (for instance, short evening courses), and does not amount to full degrees, since the long-run effects of higher scores on graduation *are* positive.



Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies.

Figure 6: The effects of a higher score on enrollment and credits

after two and twelve years.[18] The table also presents results on graduation rates. We use two different graduation measures: having accumulated at least six term-equivalent credits, which is a necessary condition for obtaining a degree; and having obtained a formal graduation diploma.[19]

A randomly higher SAT score raises graduation rates by more than 2 percentage points. The conclusion is equally true for both of our measures; accumulated sufficient credits (column 4,

[18]The sample used in Table 2, unlike in Figure 6, is restricted to the students for whom there is information on the education of at least one parent, in order to set a baseline for our analysis of parental background, but this restriction does not affect our results in a meaningful way.

[19]To receive a diploma and appear in the graduation register, Swedish students must show that they have satisfied all the program-specific requirements. It is not uncommon for students to find a job even before leaving college, and then not to submit the paperwork necessary to receive the diploma. Therefore, having accumulated sufficient credits is a potentially more accurate measure of 'graduation', and indeed our estimates using information from the graduation register tend to be less precise.

Panel A) and having obtained a formal degree (column 5, Panel A).[20]

The results presented so far imply that our (high ability) students who are able to enter most competitive programs regardless of the score, respond to a randomly higher score by increasing their participation in higher education. The model of rational and informed agents outlined in the previous section, suggest that this response should come from increased participation in competitive programs. Thus, we next explore the effects of a higher score on college quality, as well as on entering and completing educations within highly competitive fields. We rank institutions into 'top-5', 'top-10', and 'outside top-10' institutions according to the admitted students' average high school GPAs. Overall, the 'top-10' institutions cater to half of the students in our sample. We also show separate estimates for enrolling in a law or medical program (both are long undergraduate programs in Sweden), which are the top-ranking fields in terms of students' high school GPA.

Overall, our results on quality presented in Panel A of Table 3 suggest that a randomly higher SAT scores shift students toward better institutions and programs, as expected from the stylized model. A higher SAT score positively affects enrollment in higher-quality institutions (columns 1-2) as well as in law or medical programs (column 4). In contrast, there appears to be no effect on enrolling in a lower ranked institution (column 3). In the long run, a higher score seems to only have a statistically significant effect on the probability of graduation from a top-10 institution. This effect is quantitatively meaningful at eight percent of the mean (column 6). The effect on obtaining a law or medical degree also appears sizable at eight percent of the mean, but it is imprecisely estimated.

5.1.1 Robustness checks

Here we report an extensive set of robustness checks, all of which can be found in the appendix.

First, we discuss the role of our restriction to only include those with a high school GPA below the median at each threshold. It is illustrative to first focus entirely on individuals with a high school GPA *above* the local median, i.e., the mirror of our main sample. In this group, we find effects on enrollment, sufficient credits, and holding a degree to be small and statistically insignificant. This heterogeneity by local GPA is clearly consistent with the nature of the Swedish admission process since a higher SAT score for this group leaves their admission opportunities unaffected due to the substitutability between high school GPA and SAT scores as discussed in Section 2 and as formalized by the model in Section 4. If we estimate the models on the full sample of students, we find, as expected, effects on enrollment after two years and on the probability of having sufficient credits or obtaining a degree twelve years after the test that are smaller than in the main analysis sample; but estimates remain statistically significant.

[20] The bottom panel of Figure A4 shows the dynamic effects on holding a degree and having sufficient credits for time horizons ranging from one year before the test until twelve years after.

A potential concern may be, however, that the results are instead indicative of heterogeneity by general ability. As a direct test of this hypothesis, we split the sample by *compulsory school* GPA instead since this is a pure ability measure without the competing 'application currency' aspect of high school GPA. We do not find heterogeneous effects here, suggesting that the high school GPA split mainly picks up substitutability between application currencies, and not heterogeneity due to general ability. (See Table A5 for the splits by HSGPA and CSGPA, and A5 for a replication of 6 with the addition of results for the full sample.)

A further set of robustness checks—focusing on our main outcomes of enrollment, sufficient credits, and graduation—include restricting the sample to a *balanced panel*; to *first tests*; and to individuals who are close to a threshold that is sufficiently high so that crossing it would *improve on their previous score* (including all first tests by construction). Across these different sample definitions, the results are very similar. The only exception is that we do not find any statistically significant effect on the graduation outcome among first-time takers, but these estimates are imprecise.[21] (See Figure A6.)

For the three main outcomes we also estimated equation (1) separately by threshold. While there appear to be positive effects even at some lower thresholds, the largest effects are concentrated at the upper thresholds. Remarkably, there is a persistent positive effect on graduation even from crossing the very highest threshold. This is remarkable because the local alternative—a score of 19—provides access to almost all possible choices. This means that a non-trivial fraction of students prefer not going to college to this wide set of alternatives, but prefer going to college if they can chose from an even wider set. We return to this finding, and in particular heterogeneity by SES, in Section 5.3 below. (Results for individual thresholds are shown in Figures A7, A8, and A9.)

[21] As discussed in Section 2, first-time takers tend to score lower, and hence are under-represented among the upper thresholds that we focus on. If we condition on being close to one of the upper thresholds for the first time, our results are nearly identical to the main specification (available on request).

Table 2: The effects of a higher SAT score on enrollment and graduation

	Two years after test		Twelve years after test		
	(1)	(2)	(3)	(4)	(5)
	Enrolled	Cumulative credits	Cumulative credits	Sufficient credits	Degree
A. All (parental education observed)					
Above cutoff	2.91	0.19	0.10	2.12	2.85
	(0.60)	(0.026)	(0.062)	(0.70)	(0.77)
Mean of dep. var.	66.1	2.6	7.5	68.6	50.6
Obs. (total—person-exam—person)	126,542—93,964—59,352		87,689—63,953—42,402		
B. Parents at most high school educated					
Above cutoff	4.32	0.21	0.30	4.17	3.45
	(1.02)	(0.045)	(0.10)	(1.18)	(1.23)
Mean of dep. var.	63.4	2.5	6.8	63.6	47.6
Obs. (total—person-exam—person)	44,867—33,231—22,908		33,480—24,438—17,459		
C. Both parents hold college degree					
Above cutoff	1.53	0.13	-0.27	-0.38	1.45
	(1.27)	(0.055)	(0.14)	(1.48)	(1.75)
Mean of dep. var.	69.2	2.7	8.5	76.4	55.9
Obs. (total—person-exam—person)	26,102—19,334—10,689		17,100—12,433—7,299		
Difference between *B.* & *C.*	2.79	0.08	0.57	4.55	2.00
p-value of difference	0.09	0.23	0.00	0.02	0.35

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Coefficients, standard errors, and means in column (1), (4), and (5) have been multiplied by 100. Standard errors, clustered by individual, in parentheses.

Table 3: Effects on enrollment and graduation by institution and field

	Two years after test: enrollment				Twelve years after test: degree			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 5	Top 10	Not top 10	Law/medicine	Top 5	Top 10	Not top 10	Law/medicine
A. All (parental education observed)								
Above cutoff	1.42	3.34	-0.44	1.30	0.56	2.70	0.15	0.61
	(0.46)	(0.60)	(0.59)	(0.34)	(0.57)	(0.74)	(0.55)	(0.41)
Mean of dep. var.	15.7	34.2	31.9	7.9	16.8	35.7	15.0	7.8
Obs. (total—person-exam—person)	126,542—93,964—59,352				87,689—63,953—42,402			
B. Parents at most high school educated								
Above cutoff	0.23	1.88	2.44	0.38	0.15	2.59	0.86	0.36
	(0.70)	(0.97)	(1.01)	(0.49)	(0.85)	(1.14)	(0.93)	(0.57)
Mean of dep. var.	12.3	29.2	34.2	5.9	13.4	30.1	17.5	5.8
Obs. (total—person-exam—person)	44,867—33,231—22,908				33,480—24,438—17,459			
C. Both parents hold college degree								
Above cutoff	1.94	3.79	-2.26	1.81	0.18	1.37	0.079	-0.27
	(1.11)	(1.36)	(1.28)	(0.87)	(1.45)	(1.73)	(1.14)	(1.13)
Mean of dep. var.	20.5	40.5	28.7	11.0	23.3	44.6	11.3	12.1
Obs. (total—person-exam—person)	26,102—19,334—10,689				17,100—12,433—7,299			
Difference between *B.* & *C.*	-1.70	-1.91	4.70	-1.44	-0.03	1.23	0.78	0.63
p-value of difference	0.19	0.25	0.00	0.15	0.99	0.55	0.60	0.62

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Coefficients, standard errors, and means have been multiplied by 100. Standard errors, clustered by individual, in parentheses.

5.2 Heterogeneity by parental background

We now turn how the effects vary with parental background. To this end, we revisit Table 2 where Panel B reports effects of higher SAT scores on the extensive margins of enrollment, credits and graduation for students whose parents are at most high school educated (low-SES). The estimates suggest that these extensive margin responses are much larger among low-SES students than in the overall sample (Panel A). The estimated impact on enrollment is 4.3 percentage points (instead of 2.9 in the full sample) whereas the impact on long-run credit accumulation is 0.3 (instead if 0.1) and graduation twelve years after the test is 3.5 percentage points (instead of 2.9). The effects for low-SES individuals are even more sizeable (5-10 percent) if compared to the mean outcomes as these tend to be lower than average for this group.

Panel C shows effects for individuals with highly educated parents—college or above (high SES). Short-run effects on enrollment and credits appear positive, but are small and imprecisely estimated (columns 1-2). The probability of enrollment increases by about 1.5 percentage points as a result of crossing a threshold, which is less than half of the impact for low-SES students. Similarly, the effect on the long-run graduation probability for this group is only about 1.5 percentage points, which, however, is estimated even more imprecisely (column 5). The effect on cumulative credits is actually negative and statistically significant for high-SES individuals (columns 3). A potential explanation, consistent with results presented in the next sub-section, could be that high-SES students who miss out on a higher score study something else until they gain access to their preferred program in later admissions, and therefore accumulate more credits.

To sum up, among low-SES students, crossing a threshold has sizeable effects on long-run credit accumulation and the probability of holding a degree, with 95-percent confidence intervals clearly excluding zero in each case. In contrast, among high-SES students these effects are estimated to be much smaller, and indistinguishable from zero by conventional statistical criteria. The difference across the two groups is statistically significant at the 5 or 1 percent level (see the table foot) in the case of short-run credits and sufficient credits for a degree, and at the 10 percent-level for short-run enrollment. For formal degrees and cumulative credits in the long run, estimates are noisier and differences are therefore not statistically significant at conventional levels.

The relatively large responses on the external margins among the low-SES students is what we should expect if they have a low valuation of college in general, but a high valuation of quality. However, when we revisit the relationship to our quality indicators in Table 3, a very different picture emerges. The results show that even though students with low-educated parents are more likely to respond on the overall enrollment margins, they are in fact *not* more likely to enter a top-5 institution, *nor* a law or medical program, if obtaining a higher SAT score. Instead, they *are* are more likely to enter, and graduate from, institutions outside of the very top (columns 1-2, 4; column 3). Thus, these results do not indicate that low-SES students value quality more than others.[22] This is in sharp contrast to high-SES students, as seen in panel C. For them, a higher score positively affects the probability of enrolling in a high quality institution or program (columns 1-2, 4), but it *negatively* affects the probability of enrolling in a lower quality institution (column 3). This is the expected response of rational and informed students who have a high relative valuation of college, and some (possibly high) valuation of quality. None of the graduation outcomes appear to be affected (columns 5-8; admittedly, these estimates are imprecise). Thus, higher SAT scores among high-SES students appear to affect the intensive margins of college quality and timing of enrolment, but not the extensive margin.

Before probing deeper into potential explanations for this heterogeneity we first rule out a mechanical explanation related to the pooling of thresholds: Effects may be heterogeneous across thresholds, and low-SES test takers may be under-represented at the higher thresholds (and over-represented at lower ones). In fact, this turns out not to be a viable explanation for three reasons. First, there are no clear patterns of heterogeneity *by threshold* across the upper thresholds, with extensive margin effects even at the highest threshold (see Section 5.1.1). Second, the effects of crossing the highest threshold exhibit a very similar pattern of heterogeneity by parental education as the results from pooled specifications—see more on this in the next sub-section. Third, we re-estimated the specifications reported in Tables 2 and 3 on re-weighted samples in which the distribution across thresholds is weighted to be exactly the same for each parental education group as in the full sample. This does not alter the main results (see Tables A6 and A7).

[22] Of the graduation outcomes, only obtaining a degree from a top-10 institution seems to be affected in this group (column 6). The means in Table 3 show that enrolling in lower ranked institution is much more common than graduating from one, perhaps suggesting that these institutions are partly used as stepping stones.

5.3 Behavior beyond the rational and informed

The finding of such a large extensive margin response by low SES students is quite surprising. In particular, since the reason seems to be that a higher SAT score among low-SES students primarily tend to increase enrollment at institutions that falls *outside* of the top 10, but *not* in the more competitive colleges or fields. To make this point even more forcefully, it is useful to zoom in on the uppermost threshold, which is crossed by less than 1 percent of test-takers. Achieving the highest possible score—as opposed to the second-highest—increases the probability of enrolment by 18 percent of the mean among these extremely talented low-SES students. Furthermore, the probability of having graduated from college twelve years after the test increases by 25 percent in the low-SES sample, whereas for high-SES the corresponding point estimate is negative and very imprecise (see Table A11 for details).

Since almost all programs are attainable with a score of 19, these results seem to suggest that attaining a higher score may cause some students to increase their enrollment into programs that they could have entered even if they had failed to cross the thresholds. To more thoroughly explore this possibility, we use data on actual SAT admission requirements that we could find for the years 1999-2006. We define as a new outcome an indicator for enrolling in any program that one would have had access to regardless of crossing the threshold, again focusing on the six upper thresholds 15-20. For instance, if an individual enrols in a program with a required SAT score of 16, then this indicator takes value one if she is close to the threshold that separates scores of 17 and 18. For she would be eligible for the program regardless if the score is 17 or 18 (if scoring 16 she is not guaranteed to be eligible, because of ties). We also define similar outcomes separately for top 10 and top 5 schools.

Table 4 reveals three important findings. First, columns (1)-(3) in Panel A show that crossing one of the upper thresholds has a positive effect on enrolling in 'always-attainable' programs, i.e. programs that the student is eligible for even if failing to cross the threshold. This holds for enrolment at any institution, as well as when focusing on top-10 institutions, but to a lesser extent for top 5 colleges, possibly because there are fewer 'always-attainable' programs there. Second, these effects are very large for low-SES individuals, but indistinguishable from zero for high-SES (columns (1)-(3) in panels B and C). Third, the effects on any enrolment—including where crossing a threshold indeed opens up new opportunities—are very similar in magnitude to the effects on enrolling in 'always-attainable' programs in the case of low-SES individuals (panel B, columns (4)-(5)) implying that most of the responses to a randomly higher SAT score for this group is through increased enrollment in programs where we know that the student is able to enroll with scores from both sides of the threshold.

Thus, overall, the results for individuals whose parents are college-educated remain consistent with our stylized model of how students with a good understanding of their enrolment opportunities should behave, including an understanding of which additional opportunities arise from achieving a higher SAT score. In contrast, the behaviour of individuals whose parents did not attend college do not concur with this model. This seems to suggest that low-SES students optimise under limited information regarding their enrolment opportunities, limited information about the utility of the various programs, or limited information about their own ability. Achieving a higher score may induce them to collect information on which courses they are eligible for, about their own valuation of the programs, or, possibly, that a better score may be perceived as a positive signal of own ability. The last explanation would however require that they themselves draw a stronger signal from the public score beyond what they can infer from the raw score since the students (but not the schools) also see the raw scores.

5.4 Other sources of heterogeneous effects

As highlighted above, the overall patterns for low-SES students are difficult to align with a presumption that they are rational and informed. But our theory section discussed some proximate explanations within the framework of the rational and informed that that are interesting in their own right. In particular the model highlights the potential role of costs and benefits of college enrollments, and the role of discount rates and patience for explaining the incidence of repeating. These issues are analyzed in detail below.

5.4.1 Costs and benefits of college

An obvious candidate explanation for heterogeneous effects is differences in opportunity costs or the returns to college by socio-economic background (recall that the direct financial costs of college are the same for everyone, given lack of tuition fees and generous, universal financial support).

First, consider opportunity costs. In the simple model of Section 4, individuals are more likely to respond to additional opportunities if their outside option is unattractive. The value of the outside option is partly determined by foregone earnings while in college. We can estimate the monetary opportunity costs of enrolling in college by instrumenting enrollment with the dummy for crossing a threshold.[23] We predict college enrollment using (2) as the first stage. In the second stage, we use income during the second year after the test as an outcome in (2), and replace the dummy for crossing a threshold with the fitted values for college enrollment. We find that foregone earnings are of similar magnitude among low-SES students as in the full sample (see the more detailed discussion in the appendix, and Table A1).[24]

[23]Unfortunately, we cannot estimate non-monetary aspects of opportunity costs, for example, the value of foregone leisure. For wealthy individuals, the alternative to college may not be working but travelling around the world. We are not able to capture such aspects with our data.

[24]Ideally, we would like to estimate the opportunity costs for high-SES individuals, too. However, given that enrollment effects are close to zero for this group, we lack a first stage, and hence the IV estimates are not informative.

Foregone earnings while in college are of course only part of the life-time earnings comparison that enrollment decisions are based on. Thus, we also explore the returns to college, as well as the returns to college quality. We employ a selection-on-observables strategy similar to Dale and Krueger (2002), who estimate the return to college quality by comparing students who applied to and were admitted by the same colleges.[25]

While we do not have data on applications, we proxy admission opportunities by a full set of interactions of high school GPA ventiles and normalized SAT scores (including a separate category for missing SAT scores), motivated by our institutional setting as described in Section 2.2. We thus compare students who had access to the same college programs. We use the compulsory school GPA to assess to what extent this approach succeeds in reducing bias. Our sample includes Swedish residents born between 1972-1983, and earnings are measured in 2013.

We estimate a sizeable return to graduating from a top 10 institution—compared to graduating from an institution below the top-10—in the full sample of about ten percent of mean earnings. The additional return to a degree from a top-5 institution is estimated to be economically and statistically insignificant. The returns to a top-10 degree (as well as to any college) are larger for high-SES than low-SES individuals, while the opposite holds for a top-5 degree (though even among the low-SES, returns to a top-5 degree are very modest at about 1-2 percent of mean earnings). Our selection-on-observables approach eliminates a large part but not all of the CSGPA differences by degree quality, and estimated returns are thus somewhat more modest when controlling for the CSGPA. (See the more detailed discussion in the appendix, and Table A9.)

To sum up, there is no evidence that low-SES students face substantially different opportunity costs of college. While there is some heterogeneity in returns, the pattern is inconsistent with the larger enrolment and graduation effects concerning top-5 colleges for high-SES than low-SES students documented in Section 5.2. Thus, differences in opportunity costs and returns do not seem to be a viable explanation for the heterogeneous effects of higher SAT scores that we document.

5.4.2 The incidence of repeating the SAT

A possible, *proximate* explanation for heterogeneous effects on the extensive margin by parental education is differences in the incidence of repeating. As argued in Section 3, the possibility of repeating could lead to a zero effect of a higher score in the first test. This would be true in an extreme scenario where individuals who fail to cross a given threshold are certain to

[25] Our RD strategy is not suitable for estimating the life-time returns to college, given that higher SAT scores have both extensive and intensive margin effects. This is not a concern for the opportunity cost estimates above, since choice of program or field should not affect *foregone* earnings. We report reduced form estimates for income over different horizons in the appendix, see Figure A10. We lack statistical power to precisely estimate long-term income effects of a higher score.

achieve the relevant score at the next test opportunity, while individuals who cross at the first try never repeat. In this scenario, both groups would have the same life-time maximum score. We have seen in Section 3 that this extreme scenario does not prevail among the population of test takers. Here we explore whether differences in repeating by parental background could nevertheless explain the heterogeneity in the effects of a higher score.

Recall that, in the model of Section 4, students are more likely to repeat the test (conditional on having failed to cross the threshold at the first opportunity) if they have a higher probability of succeeding; if they are more patient; if they enjoy a higher return to college; or if they face lower costs of preparing for the test. Indeed, we find that high-SES students are more likely to repeat the test, especially if scoring in the upper part of the first-test score distribution; and they are more likely to improve their score, unconditionally and also conditional on repeating the test (see Figure A11). The effect of a higher score on repeating also varies with parental education as expected. The probability of repeating the test declines by 20 percent of the mean for low-SES students, compared to 15 percent for high-SES. Consequently, the effect of a higher score in the first test on the life-time maximum score is larger for the low-SES: 0.7 points, compared to 0.5 points (see Table A10).

The extensive margin response in enrollment and graduation that we observe for low-SES students could thus be in part due to the SAT score 'sticking' with these students more, since they repeat relatively less often, are less likely to improve their score conditional on repeating, and exhibit larger negative effects—relative to the mean—of a higher score on repeating. High-SES students, on the other hand, are more likely to erase a bad score by repeating and improving. However, the evidence on repeating does not speak to our findings regarding 'always-attainable' programs discussed in Section 5.3.

6 Conclusions

In this paper we have explored the causal impact of college admission opportunities on college attendance and graduation. For identification, we use discontinuities generated by aggregation procedures mapping raw scores to normalized reported SAT scores. We find that admission opportunities on average have positive effects on college attendance, which translate into long-run differences in college graduation. The effects arise even at thresholds where most of attractive college educations are available for both winners and losers. The magnitudes are of clear economic interest, about 5 percent of the mean outcome, and the results thus suggest that preferences over schools and fields is a crucial determinant of college attendance.

Our Swedish setting is one where all colleges are tuition free and all students are granted generous financial aid while enrolled. We are therefore able to study how differences in admission opportunities affect college enrollment and graduation for children of high- and low-educated parents without concerns about how these opportunities interact with access to financial support. Our results show that low-SES individuals are responsible for most of the extensive margin responses, whereas high-SES individuals instead respond by adjusting the quality and timing of their college education.

Interpreted through the lens of our stylized theoretical framework, the results suggest that individuals whose parents are college-educated have a high valuation of college relative to market work, but with a preference for more competitive programs. They therefore attend college even if the SAT score is randomly lower, but they enter faster and into more competitive programs if the score is higher. In contrast, the behavior of individuals whose parents did not attend college is not easily understood within a framework of rational and informed agents. The reason is that a higher score primarily tends to increase their enrollment into programs that they could enter even with a lower score. One plausible explanation for this behavior is that these low-SES students make their choices under limited information regarding their enrolment opportunities and/or their subjective valuation of the various programs. Achieving a higher score, which obviously and transparently increases their choice set, may induce them to collect information on the various aspects that affect their subjective valuation of different available programs. A related alternative explanation for why these students may deviate from the rational and informed baseline is if they have limited information about their own ability and that they perceive a better score as a positive ability signal. However, since the students—in contrast to the admission authorities—also receive their raw scores on their report cards, this explanation would require that the students themselves perceive a stronger signal from the public score beyond what they can infer from the raw score.

The fact that we find substantial extensive margin responses sets us apart from the stream of studies using admission cut-offs in national admission systems. This difference in results appears to arise for two interrelated reasons. First, variations around admission cut-offs rules out, by construction, the probability that the admitted students increase their probability of entering a lower-ranked option if being admitted to the higher ranked option (the lower ranked options are mechanically dropped). Second, the impact of admission discontinuities is, again by construction, estimated within the sample of students who have already made their applications. In contrast, our estimates are based on variation that arises at an earlier stage in the decision process and the SAT score can thus affect both *if* students apply and *where* they apply, including the option of applying to a low ranked program.

Overall, our results suggest that a partial explanation for the SES enrollment gap among top-performing Swedish students, which we documented in the introduction, are admission-related information frictions and behavioral barriers, despite the streamlined application systems used for both enrollment and funding. It appears that low-SES students require a very broad set of options, or a public-use signal of their own ability, before being willing to take the leap into higher education, even though they are unlikely to actually enter into one of their most competitive options. Our results thus highlight that behavioral barriers and information frictions may cause intergenerational persistence in college graduation even in settings where financial constraints are unlikely to bind, application procedures are simple, and where at least some college programs always have open slots.

Table 4: Effects on enrollment where eligible regardless

	Eligible even if below cutoff			Any eligibility		
	(1)	(2)	(3)	(4)	(5)	(6)
	Enrolled	Top 5	Top 10	Enrolled	Top 5	Top 10
A. All (parental education observed, tests from 1999-2006)						
Above cutoff	1.85	0.88	1.49	3.05	2.14	3.39
	(0.73)	(0.43)	(0.60)	(0.82)	(0.63)	(0.82)
Mean of dep. var.	23.8	6.7	14.2	65.3	16.0	35.0
Obs. (total—person-exam—person)			68,581—52,027—36,005			
B. Parents at most high school educated						
Above cutoff	3.69	1.37	2.00	4.75	1.43	1.73
	(1.28)	(0.67)	(0.98)	(1.46)	(0.97)	(1.37)
Mean of dep. var.	24.3	5.4	12.2	62.0	12.1	29.8
Obs. (total—person-exam—person)			22,721—17,293—13,063			
C. Both parents hold college degree						
Above cutoff	0.087	-0.0097	-0.69	0.33	2.56	2.77
	(1.54)	(1.03)	(1.37)	(1.70)	(1.48)	(1.81)
Mean of dep. var.	21.6	7.9	15.6	69.0	20.5	40.8
Obs. (total—person-exam—person)			14,871—11,206—6,808			
Difference between *B.* & *C.*	3.60	1.38	2.69	4.42	-1.13	-1.04
p-value of difference	0.07	0.26	0.11	0.05	0.52	0.65

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Coefficients, standard errors, and means have been multiplied by 100. Standard errors, clustered by individual, in parentheses.

References

ALTMEJD, A., A. BARRIOS-FERNANDEZ, M. DRLJE, D. KOVAC, AND C. NEILSON (2020): "Siblings Spillover Effects on College and Major Choice: Evidence from Chile, Croatia and Sweden," *Princeton Industrial Relations Section Working Paper*.

ANELLI, M. (forthcoming): "Returns to elite university education: a quasi-experimental analysis," *Journal of the European Economic Association*.

AVERY, C., O. GURANTZ, M. HURWITZ, AND J. SMITH (2018): "Shifting College Majors in Response to Advanced Placement Exam Scores," *Journal of Human Resources*, 53, 918–956.

BELFIELD, C., T. BONEVA, C. RAUH, AND J. SHAW (2019): "What drives enrolment gaps in further education? the role of beliefs in sequential schooling decisions," *Economica*, 87, 1–4.

BETTINGER, E. P., B. T. LONG, P. OREOPOULOS, AND L. SANBONMATSU (2012): "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment," *The Quarterly Journal of Economics*, 127, 1205–1242.

BOND, T. N., G. BULMAN, X. LI, AND J. SMITH (forthcoming): "Updating Human Capital Decisions: Evidence from SAT Score Shocks and College Applications," *Journal of Labor Economics*.

BONEVA, T. AND C. RAUH (2019): "Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns," *Unpublished manuscript, University of Oxford*.

CAMERON, S. V. AND J. J. HECKMAN (2001): "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 109, 455–499.

CANAAN, S. AND P. MOUGANIE (forthcoming): "Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity," *Journal of Labor Economics*.

CELLINI, S. R., F. FERREIRA, AND J. ROTHSTEIN (2010): "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 125, 215–261.

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

DILLON, E. W. AND J. A. SMITH (2017): "Determinants of the Match between Student Ability and College Quality," *Journal of Labor Economics*, 35, 45–66.

DYNARSKI, S., C. LIBASSI, K. MICHELMORE, AND S. OWEN (2018): "Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students," NBER Working Papers 25349, National Bureau of Economic Research, Inc.

EDWARDS, D., H. COATES, AND T. FRIEDMAN (2012): "A Survey of International Practice in University Admissions Testing," *Higher Education Management and Policy*, 24, 1–18.

FAN, E., X. MENG, Z. WEI, AND G. ZHAO (2017): "Rates of Return to Four-Year University Education: An Application of Regression Discontinuity Design," *The Scandinavian Journal of Economics*, n/a–n/a.

French, R. and P. Oreopoulos (2017): "Behavioral barriers transitioning to college," *Labour Economics*, 47, 48–63.

Goodman, J., O. Gurantz, and J. Smith (2018): "Take Two! SAT Retaking and College Enrollment Gaps," Working Paper 24945, National Bureau of Economic Research.

Government of Sweden (2004): "Tre vägar till den öppna högskolan—Betänkande av Tillträdesutredningen," *Statens Offentliga Utredningar*, 29.

Graetz, G. and A. Karimi (2019): "Explaining gender gap variation across assessment forms," Working Paper Series 2019:8, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Grönqvist, E., B. Öckert, and J. Vlachos (2017): "The Intergenerational Transmission of Cognitive and Noncognitive Abilities," *Journal of Human Resources*, 52, 887–918.

Guryan, J., E. Hurst, and M. Kearney (2008): "Parental Education and Parental Time with Children," *Journal of Economic Perspectives*, 22, 23–46.

Hastings, J. S., C. A. Neilson, A. Ramirez, and S. D. Zimmerman (2016): "(Un)informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 51, 136 – 151, access to Higher Education.

Hoekstra, M. (2009): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, 91, 717–724.

Högskoleverket (2000): "Högskoleprovet. Gårdagens mål och framtida inriktning," *Högskoleverkets rapport-serie*, 2000:12R.

Howell, J. S. and M. Pender (2016): "The costs and benefits of enrolling in an academically matched college," *Economics of Education Review*, 51, 152–168.

Hoxby, C. and C. Avery (2013): "The Missing 'One-Offs': The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 44, 1–65.

Hoxby, C. M. and S. Turner (2015): "What High-Achieving Low-Income Students Know about College," *American Economic Review*, 105, 514–517.

Humlum, M. K., J. H. Kristoffersen, and R. Vejlin (2017): "College admissions decisions, educational outcomes, and family formation," *Labour Economics*, 48, 215 – 230.

Keane, M. P. and K. I. Wolpin (2001): "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42, 1051–1103.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016): "Field of Study, Earnings, and Self-Selection," *The Quarterly Journal of Economics*, 131, 1057–1111.

Landerso, R. and J. J. Heckman (2017): "The Scandinavian Fantasy: Sources of Intergenerational Mobility in Denmark and the US," *Scandinavian Journal of Economics*, 119, 178–230.

Lavecchia, A., H. Liu, and P. Oreopoulos (2016): "Chapter 1 - Behavioral Economics of Education: Progress and Possibilities," Elsevier, vol. 5 of *Handbook of the Economics of Education*, 1 – 74.

LOCHNER, L. AND A. MONGE-NARANJO (2012): "Credit Constraints in Education," *Annual Review of Economics*, 4, 225–256.

MANSKI, C. F. AND D. A. WISE (2013): *College Choice in America*, Harvard University Press.

MARX, B. M. AND L. J. TURNER (2019a): "Student Loan Choice Overload," NBER Working Papers 25905, National Bureau of Economic Research, Inc.

——— (2019b): "Student Loan Nudges: Experimental Evidence on Borrowing and Educational Attainment," *American Economic Journal: Economic Policy*, 11, 108–141.

ÖCKERT, B. (2010): "What's the Value of an Acceptance Letter? Using Admissions Data to Estimate the Return to College," *Economics of Education Review*, 29, 504–516.

PAPAY, J., R. MURNANE, AND J. WILLETT (2016): "The Impact of Test Score Labels on Human-Capital Investment Decisions," *Journal of Human Resources*, 51, 357–388.

PERNA, L. W. (2006): "Studying College Access and Choice: A Proposed Conceptual Model," Springer, vol. XXI of *Higher Education: Handbook of Theory and Research*, 99 – 157.

SMITH, J., M. HURWITZ, AND C. AVERY (2017): "Giving College Credit Where It Is Due: Advanced Placement Exam Scores and College Outcomes," *Journal of Labor Economics*, 35, 67–147.

SMITH, J., M. PENDER, AND J. S. HOWELL (2013): "The full extent of student-college academic undermatch," *Economics of Education Review*, 32, 247–261.

SOLIS, A. (2017): "Credit Access and College Enrollment," *Journal of Political Economy*, 125, 562–622.

ZIMMERMAN, S. D. (2014): "The Returns to College Admission for Academically Marginal Students," *Journal of Labor Economics*, 32, 711–754.

——— (2019): "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes," *American Economic Review*, 109, 1–47.

A1 Descriptive evidence on SAT takers

Here, we briefly discuss a set of descriptive patterns regarding SAT participation and performance as a general background.[26] The Swedish SAT is not compulsory and it is possible to apply for college using only the high school GPA. Nevertheless, within the most recent cohorts in our data about half of high school graduates take the test at least once before age 30 and a quarter participate more than once. Students typically take the test for the first time in the fall of the year they graduate from high school. It is, however, not uncommon to take the test later on, even for the first time, as is reflected by the age distribution of test takers.

We report predictors of SAT participation in Table A2. As expected, raw correlations between ability measures (compulsory school GPA and high school GPA) and the probability of ever taking the SAT are positive. Notably, however, the high school GPA has two potentially counteracting effects on the incentives to take the test. It is correlated with ability, which should increase the usefulness of the test, but it is also an alternative 'application currency' which should reduce the incentives to take the test (conditional on ability). Along these lines, the impact of high school GPA is found to be *negative* when we control for compulsory school GPA (which only measures ability, not application currency). Thus, somewhat simplified, the results suggest that the typical test takers are able students without very attractive alternative admission opportunities. Notably, compulsory school GPA remains a very strong predictor of SAT participation even conditional on high school GPA, gender, immigration status, and parental education: a one-standard-deviation increase in the compulsory school GPA is associated with a 20 percentage point increase in the probability of taking the SAT.

The relationships between grades and the probability of repeating the test are similar to those for incidence. Compulsory school GPA (that is, ability) is positively associated with repeating, whereas high school GPA is negatively associated once we condition on compulsory school GPA.

Females are more likely to take the test, but less likely to repeat. Natives and individuals with high-educated parents are more likely to ever take the test and to repeat. These patterns remain (although a bit muted) after controlling for grades. However, after controlling for the SAT score in the first attempt, the association between female and repeating turns positive (we return to the *causal* effect of a higher score on repeating below).

Turning to the predictors of performance on the SAT, we find large differences related to ability indicators and demographic background. Females and foreign born score worse on the SAT by about 50 and 80 percent of a standard deviation, respectively. As expected, GPAs are positive predictors of SAT performance, and when controlling for the GPAs, as well as parental education, the coefficient on female becomes even more negative. First-time takers and older

[26]For incidence, frequency, and timing of test taking, see Figure A2. See Tables A3 and A4 for results on repeating and performance, respectively.

test takers perform worse, although the association with age is positive after controlling for GPA. Controlling for individual fixed effects, first-time takers still do worse, and older test takers do better, both of which are likely to reflect learning.

A2 The estimation of costs of and returns to college

A2.1 Costs

Here we provide more detail about the estimation of costs of and returns to college discussed in Section 5.4.1.

We estimate the monetary opportunity costs of enrolling in college by instrumenting enrollment with the dummy for crossing a threshold. We predict college enrollment using (2) as the first stage. In the second stage, we use income during the second year after the test as an outcome in (2), and replace the dummy for crossing a threshold with the fitted values for college enrollment. We also estimate a reduced form where we simply put income on the left-hand side of (2).

We rely on a normalized earnings measure constructed as annual earnings divided by 12 times the monthly minimum wage as proxied by the tenth percentile of the wage distribution.[27] We refer to this variable as *10th-percentile earnings*. The reason for the normalization is the following. Earnings grow over time due to productivity growth and inflation. Many students have zero earnings. Thus, the earnings difference between a young worker earning say the 10th-percentile wage and a student earning zero, increases over time, even if adjusting earnings for inflation. Our normalization helps us circumvent this issue.

Given that earnings are measured over a full calendar year, but enrollment data are by term, we set the enrollment variable equal to zero, one-half, and one if an individual is not enrolled, enrolled for one term, or enrolled for two terms, respectively, during the second year after the test.

Column 1 of Table A8 shows the first-stage results, which are similar to the results in column 1 of Table 2. The reduced form is shown in column 2. We see a drop in 10th-percentile earnings in the full sample and for individuals with low-educated parents as a result of a higher score. There is no effect on the earnings of individuals with high-educated parents—this is not surprising given that the enrollment effect is close to zero for this group. Column 3 in panel A shows that the opportunity cost of enrolling in college for one year amount to half of 10th-percentile earnings in the full sample. Column 3 in panel B shows that opportunity costs are essentially of that same magnitude for individuals with low-educated parents, thus suggesting that differences in (short-run monetary) opportunity costs are unlikely to explain the differences

[27] Minimum wages in Sweden are not imposed by the state, but are the result of collective agreements between trade unions and employer associations. The lowest agreed wages tend to be near the 10th percentile, as some workers are not covered by collective agreements.

in enrollment effects.

Ideally, we would like to estimate the opportunity costs for high-SES individuals, too. However, given that enrollment effects are close to zero for this group, we lack a first stage, and hence the IV estimates in column 3 of panel C are not informative.

A2.2 Returns

Next, we explore the returns to college, as well as the returns to college quality. We employ a selection-on-observables strategy similar to Dale and Krueger (2002), who estimate the return to college quality by comparing students who applied to and were admitted by the same colleges.[28]

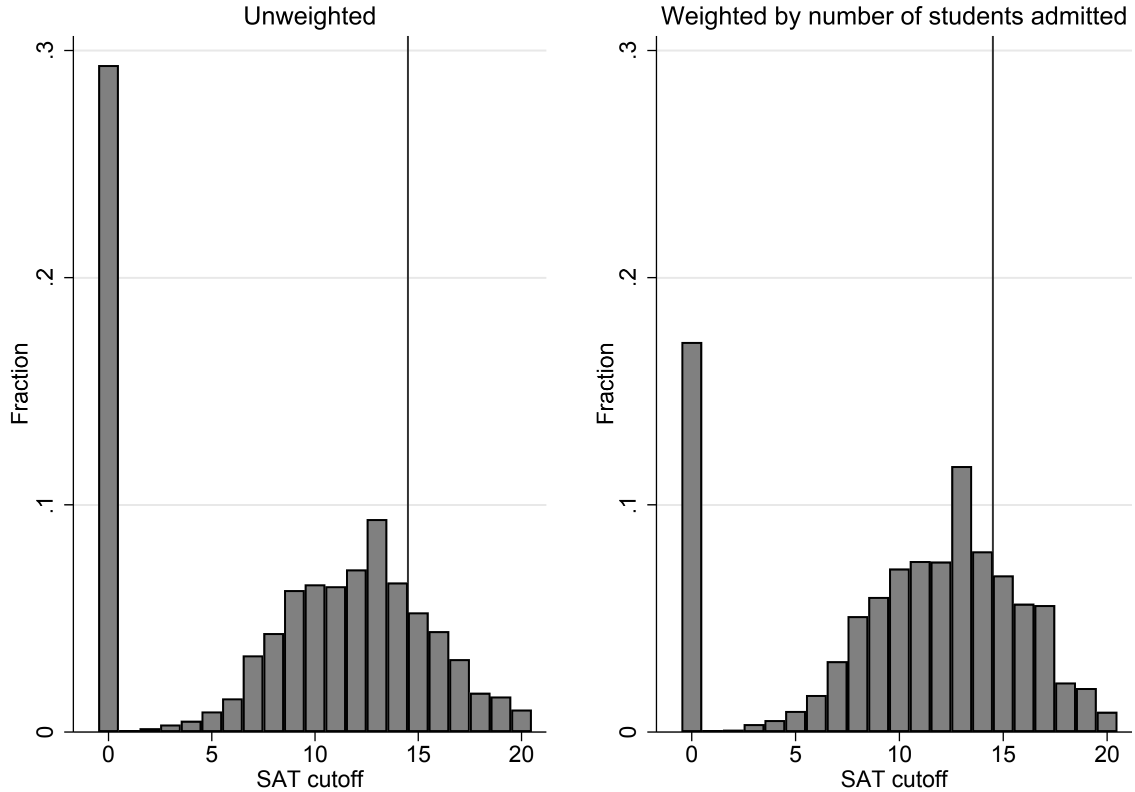
While we do not have data on applications, we proxy admission opportunities by a full set of interactions of high school GPA ventiles and normalized SAT scores (including a separate category for missing SAT scores), motivated by our institutional setting as described in Section 2.2. We thus compare students who had access to the same college programs. Baseline controls include dummies for female, immigrant, cohort by graduation year interactions, as well as parental education. We use the compulsory school GPA to assess to what extent this approach succeeds in reducing bias. Our sample includes Swedish residents born between 1972-1983, and earnings are measured in 2013.

Column (1) of Table A9 shows that conditional on baseline controls, individuals who graduated college have an about one standard deviation higher compulsory school GPA, and this difference increases by an additional one fourth (and another one tenth) of a standard deviation for those who graduated from a top 10 (top 5) institution (panel A). These differences are similar in the low- and high-SES samples (panels B and C). Controlling for admission opportunities reduces the selection in terms of compulsory school GPA by more than 50 percent in each case, and the reduction is especially large when it comes to differences by college quality (column (2)). Nonetheless, some selection appears to remain, so we report earnings regressions that control for the compulsory GPA, as well.

Column (4) of Table A9 suggests that graduating from college yields a large earnings return of about 30 percent of mean earnings, and an additional ten percent of mean earnings for graduating from a top 10 institution; there appears to be no additional return to graduating from a top 5 college (panel A). High-SES individuals appear to have larger (smaller) returns in the case of any college and top 10 institutions (top 5) than low-SES individuals, as shown in panels B and C. Controlling for the compulsory school GPA reduces these estimates slightly (column (5)). Results for earnings rank are qualitatively similar (columns (7) and (8)).

[28] Our RD strategy is not suitable for estimating the life-time returns to college, given that higher SAT scores have both extensive and intensive margin effects. This is not a concern for the opportunity cost estimates above, since choice of program or field should not affect *foregone* earnings. We report reduced form estimates for income over different horizons in the appendix, see Figure A10. We lack statistical power to precisely estimate long-term income effects of a higher score.

A3 Appendix figures and tables



Notes: A cutoff score of zero indicates that the program was undersubscribed. Most of our results focus on the thresholds to the right of the vertical lines.

Figure A1: Distribution of SAT cutoffs across programs from 1999 fall admissions

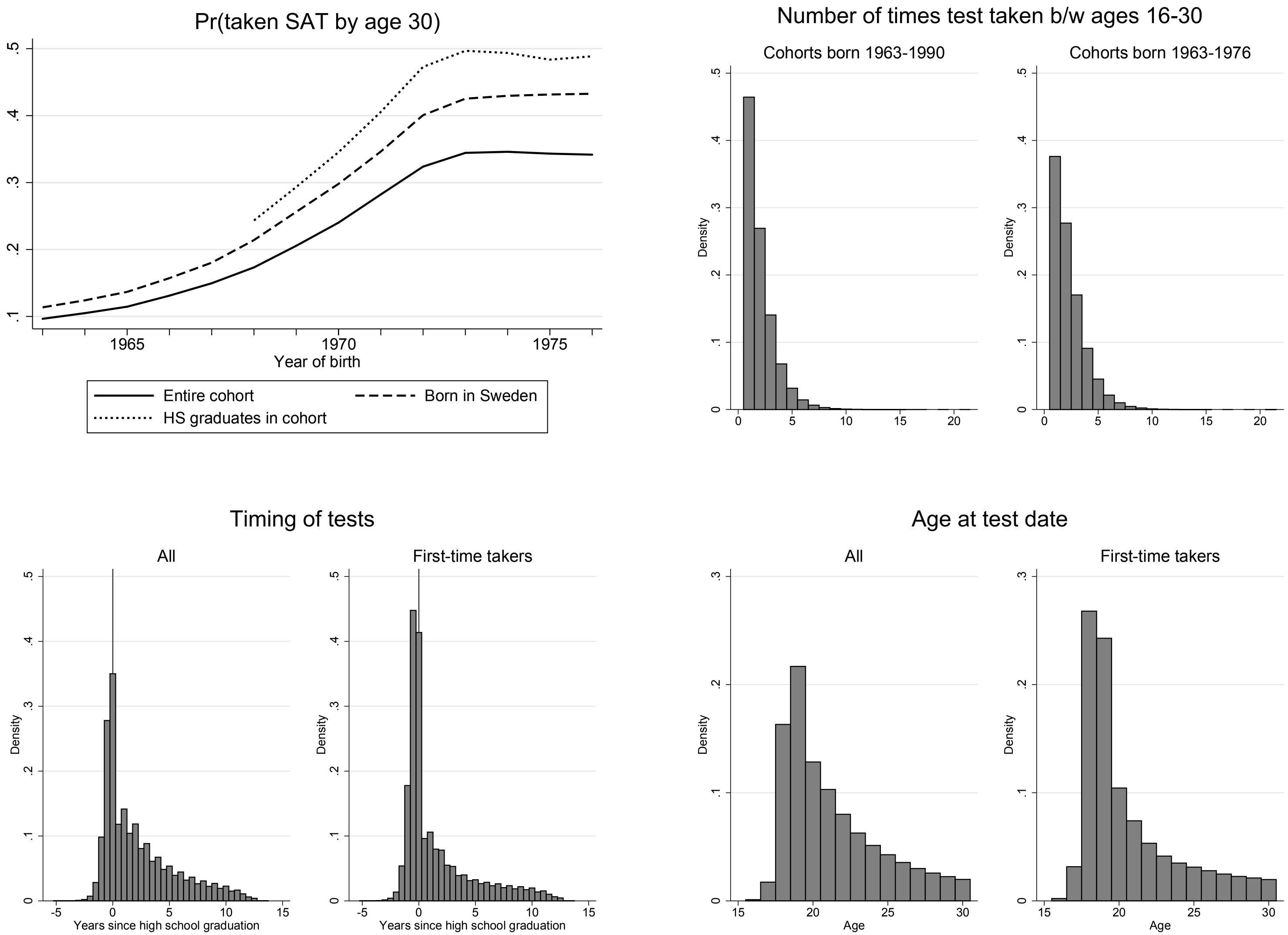
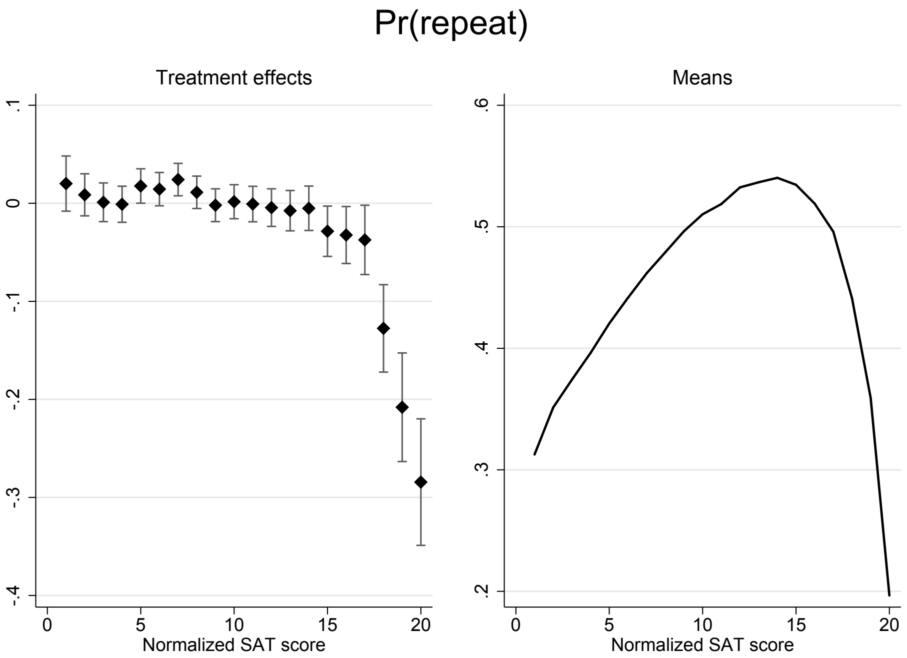
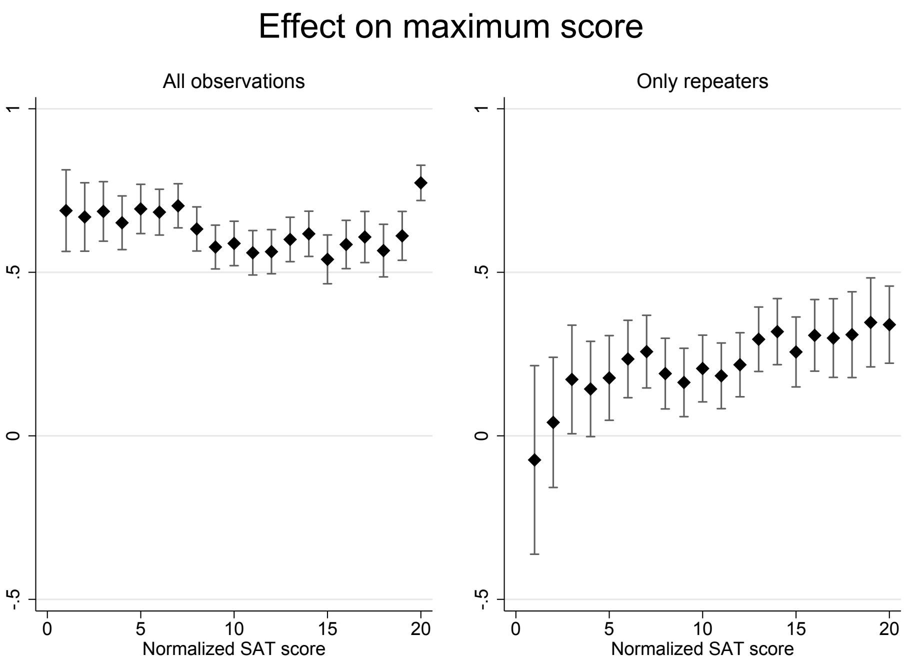


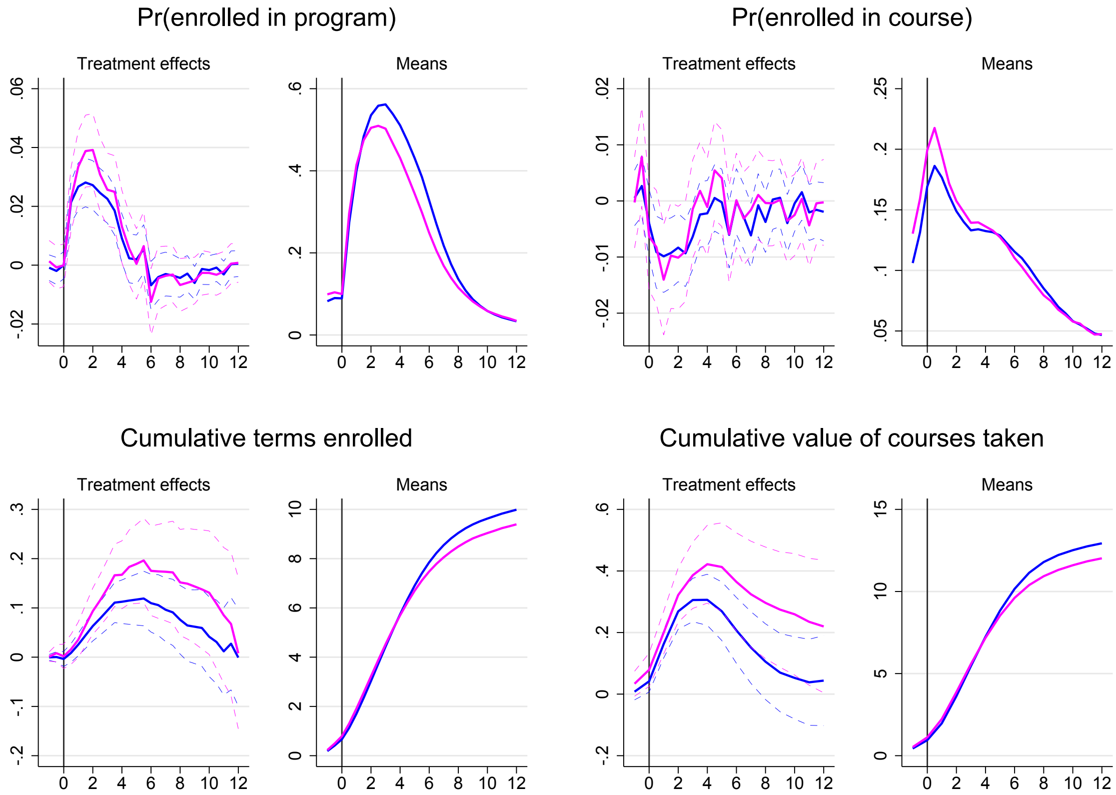
Figure A2: The incidence, frequency, and timing of taking the SAT





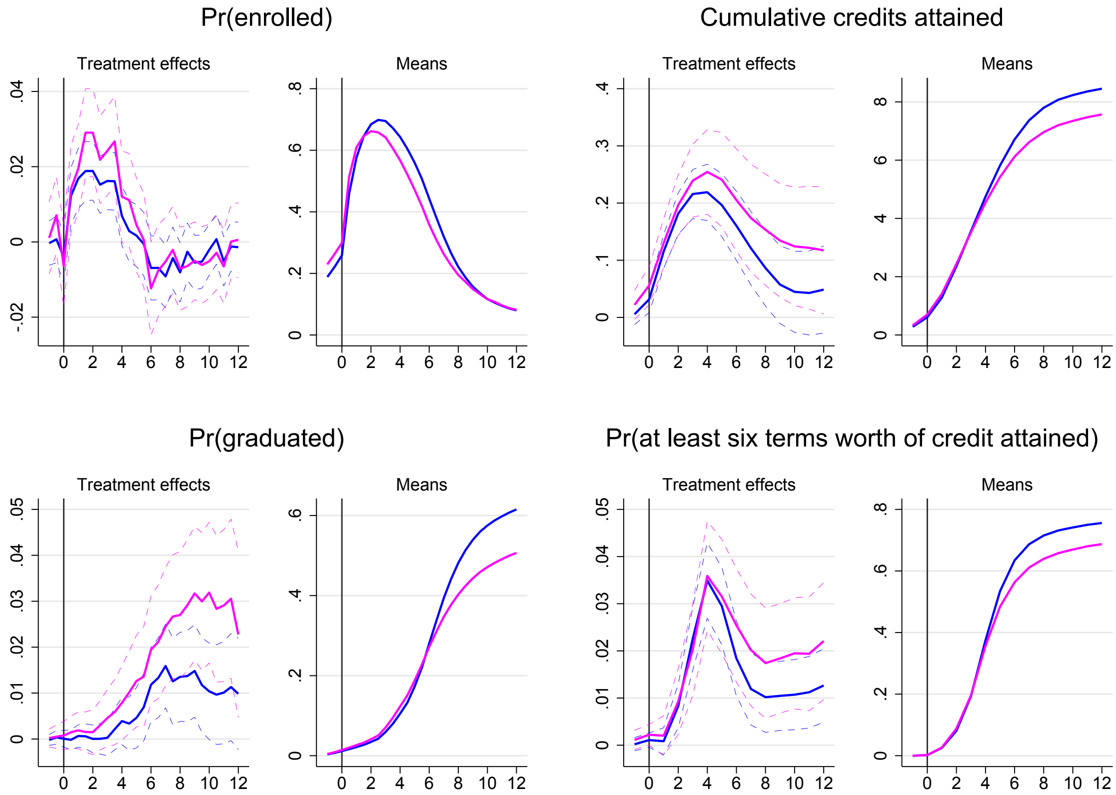
Notes: The estimated effects of crossing the indicated threshold on repeating the test and on the maximum life-time score are displayed together with 95-percent confidence intervals. Effects are estimated by RD, implemented via local linear regressions with a bandwidth of three raw marks on each side. Only first tests are included. The top right panel shows the means of repeating by normalized score.

Figure A3: The effects of a higher score on the probability of repeating and on the maximum life-time score



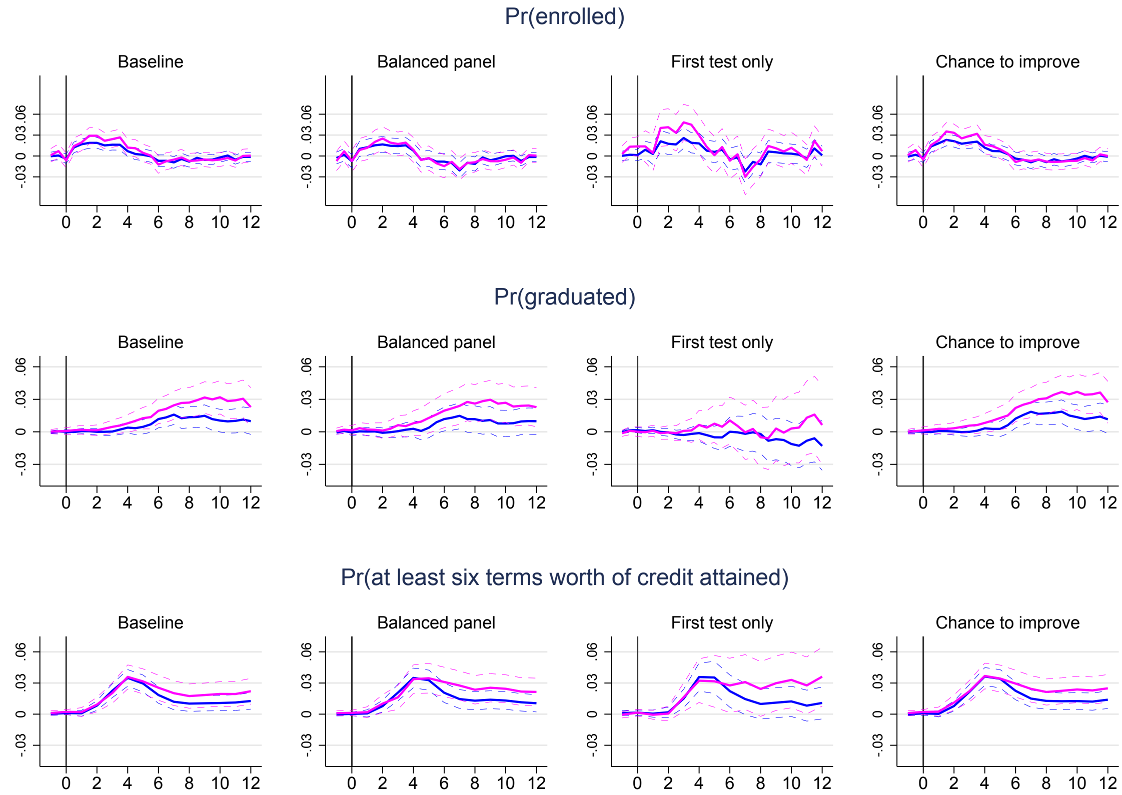
Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A4: The effects of a higher score on type of enrollment and cumulative enrollment



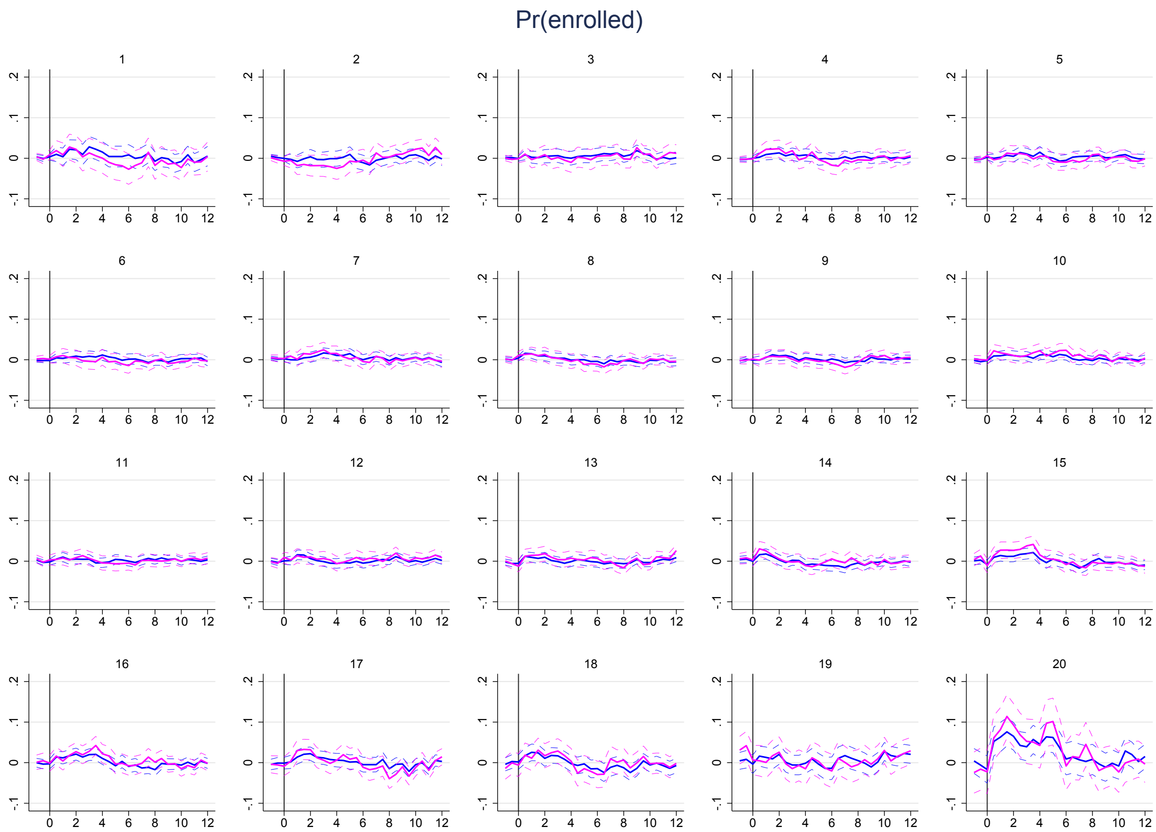
Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A5: The effects of a higher score on enrollment, credits, and graduation



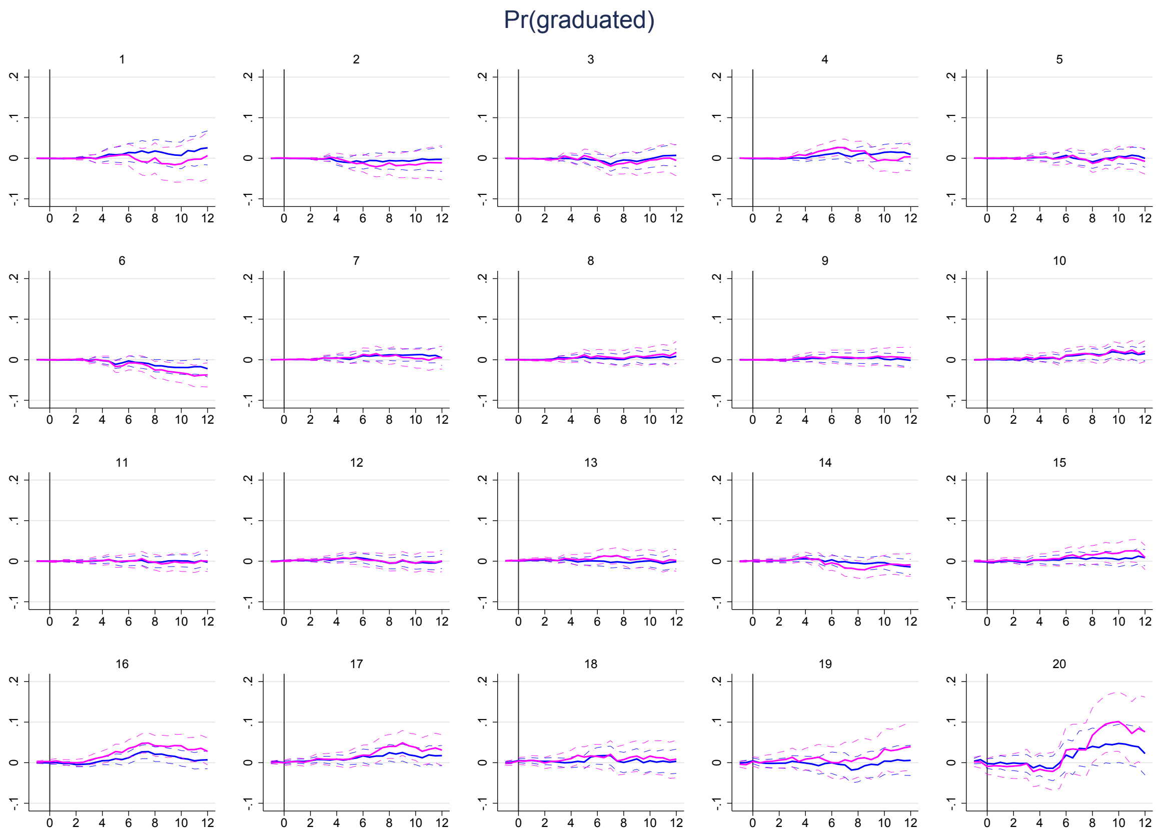
Notes: Point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (2) are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A6: The effects of a higher score on enrollment, graduation, and credits—robustness checks



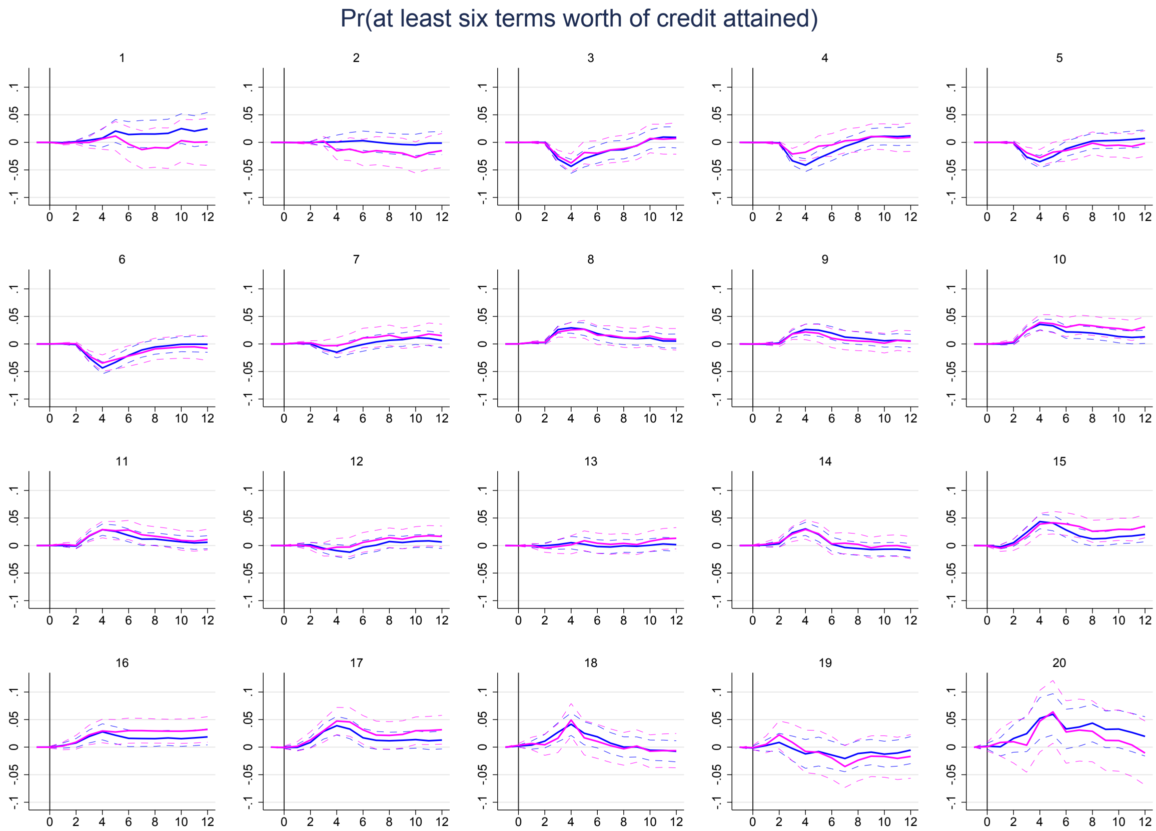
Notes: Point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (1) are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A7: The effects of a higher score on enrollment, by threshold



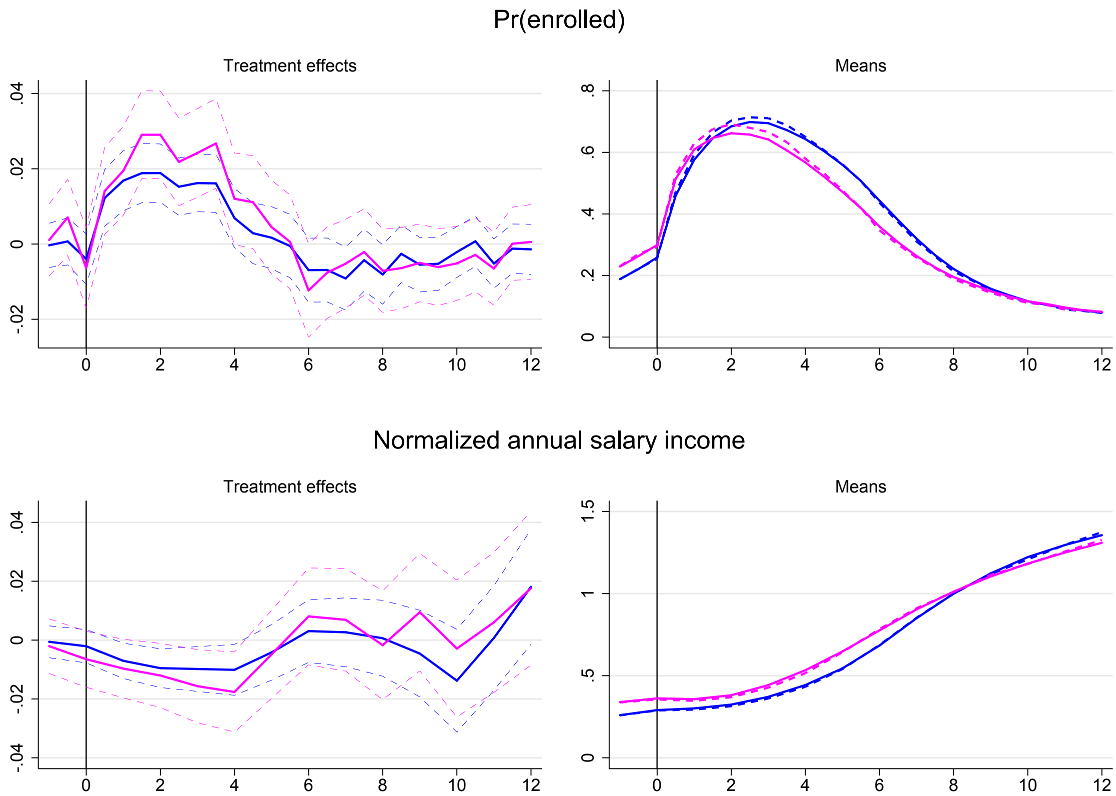
Notes: Point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (1) are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A8: The effects of a higher score on graduation, by threshold



Notes: Point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (1) are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively.

Figure A9: The effects of a higher score on the probability of having sufficient credits, by threshold



Notes: In the panels titled 'Treatment effects', point estimates (solid) and cluster-robust 95-percent confidence intervals (dashed) from estimating (2) are plotted against years since taking the test. In the panels titled 'Means', means of the dependent variable within the estimation sample are plotted against years since taking the test. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Results from the full and locally low GPA samples are shown in blue and magenta, respectively. Normalized annual salary income refers to nominal labor income during the relevant year, divided by the full-time-full-year nominal earnings of a worker at the tenth percentile of the wage distribution in that year.

Figure A10: The effects of a higher score on enrollment and income

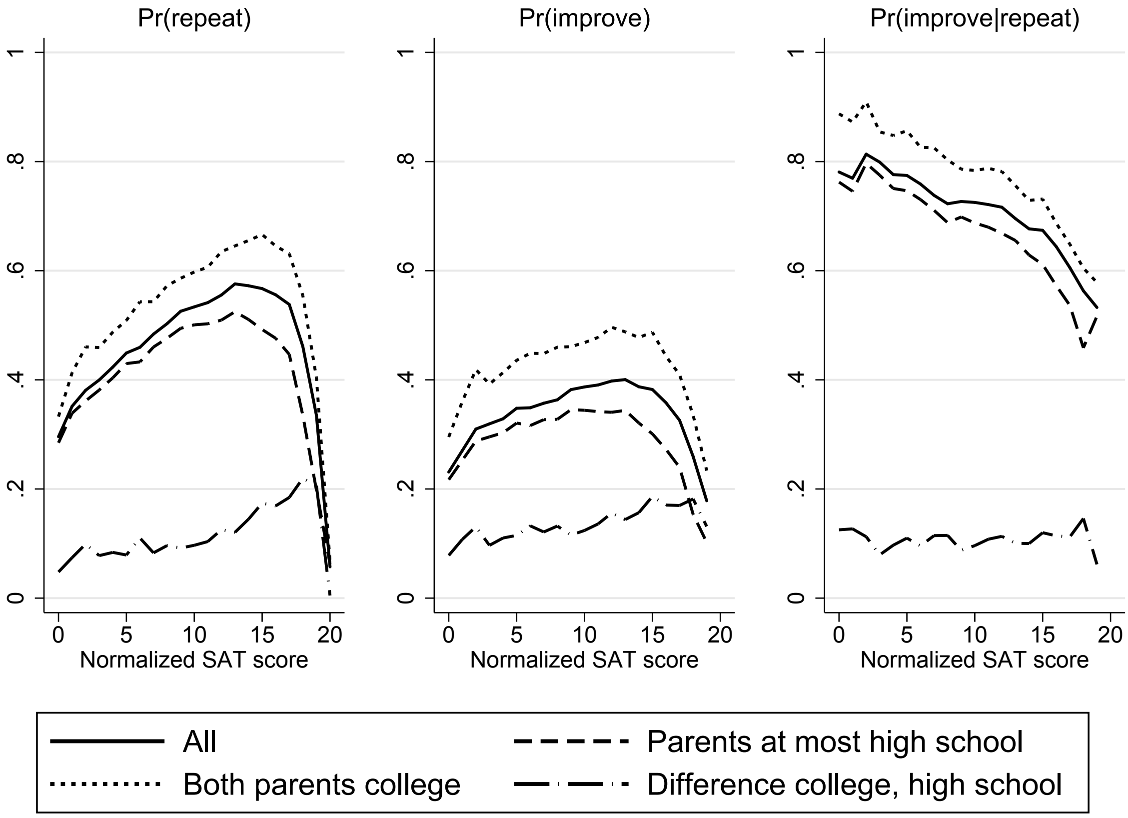


Figure A11: Repeated test taking and improvements in maximum life-time scores

Table A1: SAT and GPA cutoffs from 1999 fall admissions

College	Program	SAT cutoff	GPA cutoff
Linköping University	Medicine	20	2.4
Karolinska Institute	Medicine	20	2.4
Linköping University	Industrial Engineering	20	2.3
Uppsala University	Medicine	20	2.3
University of Gothenburg	Medicine	20	2.3
Lund University	Medicine	20	2.3
Karolinska Institute	Speech-Language Pathology	20	2.3
University of Gothenburg	European Studies	20	2.2
University of Gothenburg	Psychology	20	2.2
Swedish University of Agricultural Sciences	Veterinary Medicine	19	2.2
Stockholm School of Economics	Economics & Business	19	2.2
Royal Institute of Technology	Industrial Engineering	19	2.0
Royal Institute of Technology	Architecture	19	1.9
Royal Institute of Technology	Engineering Physics	18	2.0
Royal Institute of Technology	Computer Engineering	18	1.6
Stockholm University	Information Systems	18	1.4
University of Gothenburg	Information Systems	18	1.4
University of Gothenburg	Law	17	1.8
Stockholm University	Economics & Business	17	1.6
Stockholm University	Law	17	1.5
Chalmers University of Technology	Computer Engineering	17	1.4
Uppsala University	Economics & Business	16	1.7
Lund University	Law	16	1.6
Lund University	Economics & Business	16	1.5
Royal Institute of Technology	Electrical Engineering	16	1.1
Uppsala University	Law	15	1.5
University of Gothenburg	Social Services	15	1.4
Linköping University	Economics & Business	15	1.2
Chalmers University of Technology	Electrical Engineering	15	1.0
Stockholm University	Multimedia Pedagogy	14	1.2
Umeå University	Economics & Business	14	1.0
Lund University	Electrical Engineering	14	0.8
Royal Institute of Technology	Mechanical Engineering	14	0.7

Notes: 'SAT cutoff' is the lowest SAT score among the students admitted via the SAT quota, and 'GPA cutoff' is the lowest high school GPA (standardized by the mean and standard deviation of the 1999 high school cohort) among the students admitted via the GPA quota. For values of the SAT cutoff below 20, only the four largest programs (in terms of total number of students admitted) are listed for each cutoff, to save space. The data on cutoffs were obtained from `https://www.uhr.se/studier-och-antagning/Antagningsstatistik/Tidigare-terminer/Antagningsstatistik-urval-2-ar-2010-och-tidigare/` on 16 Feb 2018.

Table A2: Predictors of ever taking the SAT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.048	0.062	0.064	0.063	-0.0084	0.021	-0.0071	0.0020
	(0.00039)	(0.00071)	(0.00070)	(0.00071)	(0.00068)	(0.00071)	(0.00068)	(0.00068)
Foreign born	-0.22	-0.051	-0.042	-0.056	-0.0070	-0.012	-0.0096	-0.013
	(0.00038)	(0.0016)	(0.0016)	(0.0016)	(0.0015)	(0.0016)	(0.0015)	(0.0015)
Parents at most high school			-0.20					-0.092
			(0.00071)					(0.00076)
Both parents hold college degree				0.24				0.059
				(0.0013)				(0.0013)
CSGPA (z-score)					0.20		0.21	0.20
					(0.00037)		(0.00054)	(0.00055)
HSGPA (z-score)						0.11	-0.015	-0.021
						(0.00034)	(0.00047)	(0.00047)
Restricted sample		✓	✓	✓	✓	✓	✓	✓
R-squared	0.12	0.14	0.18	0.16	0.27	0.19	0.27	0.28
Obs	3,870,140	1,539,617	1,539,617	1,539,617	1,539,617	1,539,617	1,539,617	1,539,617

Notes: Results from OLS regressions are shown. The dependent variable is an indicator for ever having taken the test. All regressions control for year-of-birth fixed effects. Samples include all individuals born between 1963 and 1990. The restricted sample consists of individuals whose compulsory school GPA, high school GPA, and parents' education are observed. Robust standard errors in parentheses.

Table A3: Predictors of repeating the SAT

	(1)	(2)	(3)	(4)	(5)
Female	-0.032	-0.027	-0.032	-0.024	0.010
	(0.0014)	(0.0014)	(0.0015)	(0.0015)	(0.0015)
Foreign born	-0.064	-0.061	-0.069	-0.070	-0.031
	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Parents at most high school		-0.048		-0.049	-0.036
		(0.0016)		(0.0016)	(0.0016)
Both parents hold college degree		0.056		0.063	0.056
		(0.0023)		(0.0023)	(0.0023)
CSGPA (z-score)			0.063	0.055	0.018
			(0.0014)	(0.0014)	(0.0015)
HSGPA (z-score)			-0.054	-0.059	-0.075
			(0.0011)	(0.0011)	(0.0011)
Score fixed effects					✓
R-squared	0.13	0.13	0.13	0.14	0.16
Obs	419,368	419,368	419,368	419,368	419,368

Notes: Results from OLS regressions are shown. The dependent variable is an indicator for repeating the test at least once within five years of the first test. The sample includes only first tests. All regressions control for date-of-test and cohort dummies. Robust standard errors in parentheses.

Table A4: Predictors of performance in the SAT

	(1)	(2)	(3)	(4)	(5)
Female	-1.54	-2.17	-1.46	-2.17	
	(0.015)	(0.012)	(0.014)	(0.012)	
Foreign born	-2.49	-1.77	-2.35	-1.66	
	(0.041)	(0.033)	(0.038)	(0.031)	
Parents at most high school		-0.88		-0.80	
		(0.013)		(0.013)	
Both parents hold college degree		1.03		0.90	
		(0.019)		(0.018)	
CSGPA (z-score)		2.17		2.19	
		(0.012)		(0.012)	
HSGPA (z-score)		1.11		1.31	
		(0.010)		(0.0098)	
First time			-2.80	-1.86	-0.66
			(0.0093)	(0.0082)	(0.0081)
Age at test			-0.024	0.36	0.57
			(0.0028)	(0.0024)	(0.0031)
Individual fixed effects					✓
R-squared	0.04	0.32	0.13	0.40	0.94
Obs	880,018	880,018	880,018	880,018	880,018

Notes: Results from OLS regressions are shown. The dependent variable is the normalized SAT score ranging from 0 to 20 in integer steps. One step corresponds to about one third of a standard deviation. All regressions control for cohort dummies. Standard errors, clustered by individual, in parentheses.

Table A5: Effects on enrollment and graduation by various GPA groupings

	(1) Enrolled	(2) Sufficient credits	(3) Degree
A. Full sample			
All	1.89	1.16	1.21
	(0.40)	(0.44)	(0.51)
Mean of dep. var.	68.0	75.1	60.5
Observations (person-test)	120,808	84,646	84,646
B. HSGPA non-missing			
All	1.97	1.15	1.39
	(0.41)	(0.45)	(0.53)
Mean of dep. var.	69.0	76.3	61.8
Observations (person-test)	109,094	77,338	77,338
HSGPA below local median	2.91	2.15	2.92
	(0.59)	(0.70)	(0.77)
Mean of dep. var.	66.0	68.6	50.6
Observations (person-test)	59,797	42,722	42,722
HSGPA above local median	1.08	0.31	0.16
	(0.56)	(0.54)	(0.69)
Mean of dep. var.	72.0	84.4	73.5
Observations (person-test)	55,581	38,970	38,970
C. HSGPA & CSGPA non-missing			
All	1.96	0.89	1.06
	(0.43)	(0.47)	(0.57)
Mean of dep. var.	70.3	80.2	63.9
Observations (person-test)	93,456	62,356	62,356
HSGPA below local median	3.21	2.14	2.75
	(0.64)	(0.76)	(0.86)
Mean of dep. var.	67.5	72.9	51.8
Observations (person-test)	48,456	31,746	31,746
HSGPA above local median	0.91	0.018	-0.012
	(0.58)	(0.54)	(0.71)
Mean of dep. var.	72.8	86.9	74.9
Observations (person-test)	50,698	34,421	34,421
CSGPA below local median	2.15	0.90	1.32
	(0.63)	(0.72)	(0.84)
Mean of dep. var.	68.6	75.0	55.0
Observations (person-test)	48,866	32,938	32,938
CSGPA above local median	1.82	0.96	0.94
	(0.59)	(0.58)	(0.74)
Mean of dep. var.	71.9	85.4	72.8
Observations (person-test)	49,642	32,805	32,805

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Coefficients, standard errors, and means have been multiplied by 100. Standard errors, clustered by individual, in parentheses.

Table A6: The effects of a higher SAT score on enrollment and graduation—re-weighted by threshold

	Two years after test		Twelve years after test		
	(1)	(2)	(3)	(4)	(5)
	Enrolled	Cumulative credits	Cumulative credits	Sufficient credits	Degree
A. All (parental education observed)					
Above cutoff	2.91	0.19	0.10	2.12	2.85
	(0.60)	(0.026)	(0.062)	(0.70)	(0.77)
Mean of dep. var.	66.1	2.6	7.5	68.6	50.6
Obs. (total—person-exam—person)	126,542—93,964—59,352		87,689—63,953—42,402		
B. Parents at most high school educated					
Above cutoff	4.37	0.20	0.29	4.05	3.44
	(1.03)	(0.045)	(0.10)	(1.18)	(1.24)
Mean of dep. var.	63.4	2.5	6.8	63.6	47.6
Obs. (total—person-exam—person)	44,867—33,231—22,908		33,480—24,438—17,459		
C. Both parents hold college degree					
Above cutoff	1.48	0.12	-0.21	0.37	1.83
	(1.30)	(0.055)	(0.14)	(1.49)	(1.76)
Mean of dep. var.	69.2	2.7	8.5	76.4	55.9
Obs. (total—person-exam—person)	26,102—19,334—10,689		17,100—12,433—7,299		
Difference between *B.* & *C.*	2.89	0.09	0.50	3.68	1.60
p-value of difference	0.08	0.22	0.00	0.05	0.46

Notes: Results from the same specifications as in Table 2 are shown, except that the samples in panels B and C have been re-weighted such that the distribution of observations across thresholds is the same as that in the full sample (which the results in panel A are from). See also the notes to Table 2.

Table A7: Effects on enrollment and graduation by institution and field—re-weighted by threshold

	Two years after test: enrollment				Twelve years after test: degree			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 5	Top 10	Not top 10	Law/medicine	Top 5	Top 10	Not top 10	Law/medicine
A. All (parental education observed)								
Above cutoff	1.42	3.34	-0.44	1.30	0.56	2.70	0.15	0.61
	(0.46)	(0.60)	(0.59)	(0.34)	(0.57)	(0.74)	(0.55)	(0.41)
Mean of dep. var.	15.7	34.2	31.9	7.9	16.8	35.7	15.0	7.8
Obs. (total—person-exam—person)	126,542—93,964—59,352				87,689—63,953—42,402			
B. Parents at most high school educated								
Above cutoff	0.19	1.75	2.62	0.31	0.21	2.51	0.93	0.28
	(0.72)	(0.98)	(1.01)	(0.52)	(0.87)	(1.15)	(0.93)	(0.60)
Mean of dep. var.	12.3	29.2	34.2	5.9	13.4	30.1	17.5	5.8
Obs. (total—person-exam—person)	44,867—33,231—22,908				33,480—24,438—17,459			
C. Both parents hold college degree								
Above cutoff	1.71	3.65	-2.17	1.82	-0.17	1.51	0.32	0.0041
	(1.10)	(1.37)	(1.31)	(0.82)	(1.45)	(1.75)	(1.16)	(1.10)
Mean of dep. var.	20.5	40.5	28.7	11.0	23.3	44.6	11.3	12.1
Obs. (total—person-exam—person)	26,102—19,334—10,689				17,100—12,433—7,299			
Difference between *B.* & *C.*	-1.52	-1.90	4.79	-1.52	0.37	0.99	0.61	0.28
p-value of difference	0.25	0.26	0.00	0.12	0.83	0.63	0.68	0.82

Notes: Results from the same specifications as in Table 3 are shown, except that the samples in panels B and C have been re-weighted such that the distribution of observations across thresholds is the same as that in the full sample (which the results in panel A are from). See also the notes to Table 3.

Table A8: The opportunity cost of college

	First stage (1) Enrolled	Reduced form (2) Income	IV (3) Income
A. All (parental education observed)			
Estimate	0.024	-0.012	-0.51
	(0.0052)	(0.0056)	(0.21)
Mean of dependent variable	0.65	0.38	
F-statistic			20.6
Obs. (total—person-exam—person)	125,661—93,288—58,963		
B. Parents at most high-school educated			
Estimate	0.033	-0.018	-0.54
	(0.0091)	(0.010)	(0.26)
Mean of dependent variable	0.63	0.43	
F-statistic			13.3
Obs. (total—person-exam—person)	44,537—32,984—22,752		
C. Both parents hold college degree			
Estimate	0.017	-0.0036	-0.21
	(0.011)	(0.011)	(0.59)
Mean of dependent variable	0.69	0.31	
F-statistic			2.5
Obs. (total—person-exam—person)	25,936—19,207—10,631		

Notes: Samples are restricted to individuals whose GPA is below the local median, and to observations two years after the test. Enrollment equals one (one-half, zero) if the individual is enrolled both terms (only one, none) in the year coming two years after the test. Income refers to nominal labor income over the year coming two years after the test, divided by the full-time-full-year nominal earnings of a worker at the tenth percentile of the wage distribution in that year. Results in columns (1)-(2) refer to the coefficient on the indicator for being above the cutoff form estimating (2). Results in column (3) are from instrumental variable regressions where enrollment is instrumented by the indicator for being above the cutoff, and the running variable is controlled for as in (2). The unit of observation is person-test-threshold. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Standard errors, clustered by individual, in parentheses.

Table A9: The returns to college quality

	Compulsory school GPA		Earnings (SEK'000)			Earnings rank		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All								
Graduated: any	0.94	0.41	102.7	78.3	68.8	17.0	13.6	12.2
	(0.012)	(0.0100)	(2.15)	(2.14)	(2.13)	(0.34)	(0.34)	(0.34)
Graduated: top 10	0.26	0.059	37.8	27.9	26.5	3.57	2.63	2.43
	(0.0030)	(0.0024)	(0.79)	(0.80)	(0.80)	(0.10)	(0.11)	(0.11)
Graduated: top 5	0.085	0.031	4.11	0.72	0.0045	0.51	0.21	0.11
	(0.0035)	(0.0027)	(1.00)	(0.99)	(0.99)	(0.13)	(0.13)	(0.13)
B. Parents at most high school								
Graduated: any	1.06	0.47	104.5	77.4	66.1	17.4	13.4	11.7
	(0.014)	(0.012)	(2.44)	(2.44)	(2.43)	(0.39)	(0.39)	(0.39)
Graduated: top 10	0.27	0.077	31.4	22.5	20.6	2.87	1.92	1.64
	(0.0046)	(0.0039)	(1.14)	(1.14)	(1.14)	(0.15)	(0.16)	(0.15)
Graduated: top 5	0.097	0.039	6.85	3.75	2.83	0.76	0.44	0.30
	(0.0061)	(0.0049)	(1.58)	(1.56)	(1.55)	(0.21)	(0.21)	(0.21)
C. Parents at least college								
Graduated: any	0.69	0.29	106.7	84.1	78.8	18.1	15.6	14.9
	(0.059)	(0.051)	(14.6)	(14.4)	(14.3)	(2.11)	(2.08)	(2.07)
Graduated: top 10	0.29	0.040	45.9	32.6	31.8	4.97	3.74	3.65
	(0.0081)	(0.0062)	(2.44)	(2.48)	(2.48)	(0.31)	(0.32)	(0.32)
Graduated: top 5	0.097	0.033	1.29	-2.87	-3.49	0.19	-0.18	-0.26
	(0.0074)	(0.0055)	(2.40)	(2.38)	(2.38)	(0.30)	(0.30)	(0.30)
Baseline controls	✓	✓	✓	✓	✓	✓	✓	✓
Admission opportunities		✓		✓	✓		✓	✓
Compulsory school GPA					✓			✓

Notes: Earnings are measured in 2013 (top-coded at 99th percentile; mean: 287). All individuals born 1972-1983 are included. The number of observations is 1,075,747. Baseline controls include female, cohort by graduation year interactions, immigrant, and parental education dummies. Admission opportunities are controlled for via a full set of HSGPA-ventile-by-SAT-score interactions.

Table A10: The effect of a higher SAT score on repeating and the maximum life-time score

	(1) Pr(repeat)	(2) Maximum life-time score
A. All (parental education observed)		
Above cutoff	-0.071	0.57
	(0.012)	(0.028)
Mean of dep. var.	0.44	16.3
Obs. (total—person-exam—person)	27,222—20,445—20,445	
B. Parents at most high-school educated		
Above cutoff	-0.074	0.67
	(0.018)	(0.040)
Mean of dep. var.	0.37	16.0
Obs. (total—person-exam—person)	11,232—8,398—8,398	
C. Both parents hold college degree		
Above cutoff	-0.087	0.51
	(0.028)	(0.081)
Mean of dep. var.	0.57	16.8
Obs. (total—person-exam—person)	4,404—3,303—3,303	

Notes: Results from estimating (2) are reported. The unit of observation is person-test-threshold. Samples are restricted to individuals whose GPA is below the local median. Only first tests are included. Regressions include indicators for female, foreign born, the test date, and a full set of age-at-test dummies. Standard errors, clustered by individual, in parentheses.

Table A11: The effects of receiving the top SAT score on enrollment and graduation

	Two years after test		Twelve years after test		
	(1) Enrolled	(2) Cumulative credits	(3) Cumulative credits	(4) Sufficient credits	(5) Degree
A. All (parental education observed)					
Above cutoff	9.44	0.17	0.090	-1.50	6.67
	(2.70)	(0.13)	(0.31)	(3.27)	(3.65)
Mean of dep. var.	68.6	3.0	8.0	67.7	56.5
Obs. (total—person-exam—person)	5,011—5,011—4,505		3,366—3,366—3,105		
B. Parents at most high school educated					
Above cutoff	11.5	-0.021	0.97	10.3	13.4
	(5.48)	(0.27)	(0.58)	(6.33)	(6.75)
Mean of dep. var.	63.3	3.0	7.3	62.2	50.5
Obs. (total—person-exam—person)	1,357—1,357—1,256		980—980—921		
C. Both parents hold college degree					
Above cutoff	7.65	0.36	-0.81	-11.5	-2.47
	(4.84)	(0.25)	(0.61)	(5.62)	(6.68)
Mean of dep. var.	72.7	3.2	8.9	74.8	65.2
Obs. (total—person-exam—person)	1,445—1,445—1,268		930—930—837		
Difference between *B.* & *C.*	3.81	-0.38	1.78	21.85	15.85
p-value of difference	0.60	0.30	0.03	0.01	0.09

Notes: Results from the same specifications as in Table 2 are shown, except that only the highest threshold is included in the sample. See also the notes to Table 2.