Economics of Education Review 93 (2023) 102367

Contents lists available at ScienceDirect

Economics of Education Review

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

Student effort response to shifts in university admission policies☆

Viviana Rodriguez

University of Texas at San Antonio, United States of America

ARTICLE INFO

Keywords:
Student incentives
Student effort
Student engagement
Postsecondary admission policy

ABSTRACT

Policies which change the incentives students face are important levers for policymakers to improve learning. I study a statewide postsecondary admissions policy in North Carolina which introduces minimum admission standards for students wishing to attend college in-state. Regression discontinuity estimates show that high school students respond by increasing GPA and decreasing absences and suspensions, suggesting increased engagement. However, I find that the boost in GPA is driven by students substituting away from demanding coursework. Finally, I document stark heterogeneity of response across demographics, as minority and less affluent students have higher effort allocations but do not engage in strategic course-taking.

1. Introduction

Most US elementary and secondary education policies seek to shift incentives for effort of teachers and schools rather than students. Despite being abundant in K-12 education settings, incentive structures that shape student behavior have received less attention from practitioners and policymakers as a lever to encourage students' investments in school. Thus, researchers seeking to understand the effect of incentives on student effort focus mainly on evaluating the student incentive-effort connection via researcher-designed pay-for-performance incentives schemes (Bettinger, 2010; Fryer, 2011; Levitt et al., 2016).[1] Incentive structures embedded within policies for access to higher education provide an alternative, more realistic, setting in which to examine student effort response to education policies.

This paper investigates how students' effort responds to incentive structures in a salient and high-stakes setting driven by a change in admission requirements across one of the largest university systems in the US. In an attempt to improve graduation rates at all University of North Carolina (UNC) campuses and prevent dropouts by only admitting college-ready students, UNC set minimum GPA and SAT admission standards across the system in 2009, generating a discontinuity in the incentives for effort for high school students across SAT performance. Given that students take the SAT and learn about their admission eligibility before completing their high school coursework, students right below and right above the threshold face different returns to their effort in their last year of high school.

I leverage the discontinuity in the returns to effort generated by UNC's 2009 admission policy to examine students' response to sharp and transparent incentives. Under a regression discontinuity design, I rely on rich administrative data for the universe of high school students in the state of North Carolina to compare the effort allocation of students right above and below the SAT threshold.[2] I find that students just above the SAT threshold have fewer 12th grade absences (−18.5%) and suspensions (−27.4%) than students right below. This suggests that students just above the threshold realize they have the possibility of getting admitted into a UNC campus, and increase effort in the form of attendance and less disruptive behavior. At the same time, students just above the SAT threshold obtain 12th grade GPAs that are 2.8% higher than students right below.

While the interpretation of fewer absences and suspensions is relatively straight forward, the interpretation of higher GPAs is less so. Higher student GPA can reflect either increased effort allocation, substitution away from more demanding courses, or both. To further

☆ I would like to thank Alex Eble, Brooks Bowden, Hugh Macartney, Sarah Cohodes, Judy Scott-Clayton, Daniel Jones, Peter Arcidiacono, Peter Bergman, Jordan Matsudaira, Henry Levin, Rob Garlick, Todd Stinebrickner, Jonathan Moreno, Briana Ballis, Takeshi Yanagiura, Tatiana Velasco, Maya Escueta, Rahul Gupta, participants of the Association for Education Finance and Policy Conference, Teachers College Economics and Education Workshop, Duke University's Public and Labor Economics groups, IZA Workshop on the Economics of Education, Western Economic Association International Conference, and an anonymous reviewer for helpful comments and suggestions. All remaining errors are my own.

E-mail address: viviana.rodriguez@utsa.edu.

[1] A notable exception is Fruehwirth (2013) who exploits a student accountability policy change in the state of North Carolina that required fifth grade students to achieve above a certain level to be automatically promoted to the next grade. Fruehwirth (2013) finds that this policy induced students in danger of scoring below the threshold to exert more effort in order to pass to the next grade.

[2] I follow previous literature and use student absences and suspensions as inverse proxies for effort and student GPA as an outcome that is a function of student effort (Hastings et al., 2012).

https://doi.org/10.1016/j.econedurev.2023.102367
Received 20 May 2022; Received in revised form 17 January 2023; Accepted 24 January 2023
Available online 9 February 2023
0272-7757/

explore the mechanisms driving this effect, I explore differences in course-taking patterns for students above and below the SAT admission requirement. I find that strategic course selection is an important dimension of student effort adjustments as higher GPAs of students right above the SAT cutoff are driven by students' substitution away from more demanding coursework.

Finally, in unpacking these results, I document stark heterogeneity of student response across demographics. I find that minority and less affluent students just above the cutoff have higher effort allocations via decreased absences and suspensions but engage to a lesser extent in gaming practices via course selection. These results are in line with the literature examining student response to incentives in postsecondary settings in two important ways. First, this literature highlights how incentives have heterogeneous effects across student demographics and performance. Second, while incentives generated by postsecondary policy have been found to induce students to increase effort, they have also been found to induce students into less productive behaviors such as strategic course-taking, course withdrawal, and strategic effort allocation relative to the timing of eligibility criteria (Angrist et al., 2009; Casey et al., 2018; Cornwell et al., 2003; Lindo et al., 2010; Scott-Clayton, 2011; Scott-Clayton & Schudde, 2019).[3]

A first order concern of the validity of the research design of this paper is SAT-retaking. Administrative records of NC student SAT performance includes only the last SAT administration. Given student SAT-retaking patterns (Goodman et al., 2020), test retaking can be an important source of endogeneity if students systematically retake the SAT in order to place themselves to the right of the cutoff to gain UNC admission eligibility. To address this concern, I provide several internal validity checks including tests of manipulation of the running variable, for policy relevance of the UNC cutoff, and of covariate smootheness. I further bound the regression discontinuity point estimates following Gerard et al. (2020) to correct for possible manipulation of the running variable (driven by SAT retaking). Taken together, results from these tests suggest that is unlikely that manipulation of the running variable and selection into treatment is such that is completely drives away the results of this paper.

This paper contributes to a recent and growing literature that examines student response to incentives schemes embedded in postsecondary admission policies, in particular, to preferential admission programs (such as percent plan and affirmative action programs). This literature documents important shifts in student effort and learning as a response to the implementation of preferential admission programs (Akhtari et al., 2020; Cortes & Zhang, 2012; Golightly, 2019; Khanna, 2013). Because the UNC system set *minimum* standards for admission, required GPA and test score cutoffs are low. Therefore, this paper adds to this literature by evaluating salient and high-stakes shift in incentives for students at the *lower* end of the test score distribution.

Postsecondary admission policies have also been found to have unintended consequences of student response. For example, percent plan programs have been shown to induce students to sort into lower-achieving high schools (Cullen et al., 2013), decrease student effort for seniors who already know their admission guarantee (Leeds et al., 2017), and decrease effort for students who overestimate their probability of being in the top of their class (Tincani et al., 2020). This paper contributes to this literature by showing that while incentive structures can shape student effort, they can also induce students to adjust on less productive ways such as via strategic course selection.

The remainder of the paper is organized as follows. The next section describes the Minimum Admission Requirement policy implemented by the UNC Board of Governors in 2009 and documents changes in admission practices across campus selectivity. Section 3 describes the data used in the analysis, and provides summary statistics. Section 4 establishes the research design and estimation procedure, as well as the method used for recovering bounds of regression discontinuity estimates which are presented as part of the validity exercises in Section 6. I present regression discontinuity results in Section 5 and conclude in Section 7.

2. Institutional background

University of North Carolina System is one of the largest university systems in the US, enrolling approximately 225,000 students in 16 university campuses across the state of North Carolina.[4] In addition to its national importance, UNC is the main source of college education for North Carolina high school graduates. Upwards of 70% of all Bachelor's degrees earned in the state of North Carolina are issued by the UNC system and around 85% of all UNC first-time freshmen undergraduate students are in-state students (UNC, 2015).

Beginning with the entering freshman class of 2009, in an attempt to improve graduation rates at all UNC campuses and prevent dropouts by only admitting college-ready students, the UNC Board of Governors set minimum admission standards for all campuses in the system.[5] These minimum admission standards were set at a 2.0 weighted high school GPA cutoff and a 700(15) SAT(ACT) score cutoff for all first-time undergraduate students applying for the Fall 2009.[6] The SAT score requirement was based on the sum of the critical reading and mathematics subtests, a sum that has a possible range of 400–1600. For the reminder of this paper, SAT test scores presented will reflect this measure instead of the full SAT score which also includes the writing subtest.

Note that the policy was first implemented in 2009 and affected UNC admission eligibility of students graduating high school in the Spring of 2009 and beginning college in the Fall of 2009. Therefore, the high school graduating cohort of 2009 presents the cleanest natural experiment to examine student response to incentives as, for these students, 12th grade performance was the only potential dimension of adjustment to this policy. Therefore, the focus of this paper will be on this cohort of students.

A key feature of the UNC system is its substantial variation of selectivity across campuses. Given this variation, the minimum admission requirement policy might not have been binding for the most selective institutions within the system. I draw from IPEDS data to explore campus-level trends in selectivity based on SAT test scores.[7] Admission officers at each institution report to IPEDS the 25th percentile SAT performance of their admitted undergraduate student pool each academic

[3] This literature focuses on incentives generated by performance standards, merit aid requirements, and academic probation in college.

[4] Institutions belonging to the UNC system include: Appalachian State University, East Carolina University, Elizabeth City State University, Fayetteville State University, North Carolina A&T State University, North Carolina Central University, North Carolina State University, UNC Asheville, UNC Chapel Hill, UNC Charlotte, UNC Greensboro, UNC Pembroke, UNC School of Arts, UNC Wilmington, Western Carolina University, and Winston-Salem State University.

[5] Before 2009, UNC had no official minimum GPA, SAT, or ACT admission requirements. However, the system had required applicants to have a minimum high school course requirement since 2006. These requirements consisted of 4 credits of math, 4 credits of English, 3 credits of science, 2 credits of social studies, and 2 credits of language. After 2009, minimum GPA, SAT, and ACT requirements were implemented in addition to minimum course requirements. Any given campus could grant exceptions to these new requirements to up to 1% of their admitted pool.

[6] These requirements were further increased in 2011, 2013, and 2021. In 2011, the cutoffs increased to a high school GPA of 2.3 and an SAT(ACT) score of 750(16). In 2013, the cutoffs increased to a high school GPA of 2.5 and an SAT(ACT) score of 800(17) (see Fig. A.1). In 2021 the Minimum Admission Requirements were changed to Minimum Admissibility Requirements which required applicants to have a 2.5 high school GPA *or* a 1010(19) SAT(ACT) test score.

[7] Similar trends are found when ACT 25th percentile scores are used instead of SAT scores (see Fig. A.2).

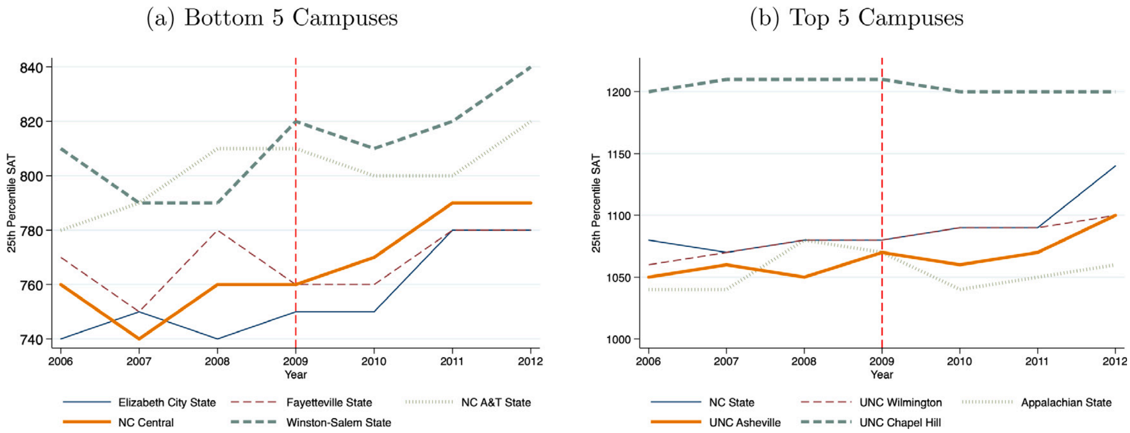

Fig. 1. 25th percentile SAT of admitted students.
Notes: This figure shows the evolution of the 25th percentile SAT score for admitted students of UNC campuses over time. Panel (a) shows the evolution for the bottom five campuses. Panel (b) shows the evolution for the top five campuses. The dash red line represent the year in which the UNC Minimum Admission Requirements were implemented. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Source: IPEDS.

year. Fig. 1 presents the trend over time of the 25th percentile SAT scores of admitted students for the five least selective (panel (a)) and the five most selective institutions (panel (b)) within the UNC system.

Panel (b) of Fig. 1 shows that among the most selective campuses within the UNC system, the new admission policy did not seem to affect their admission decisions. For these campuses only 25% of their admitted pool scored SATs below 1040 before 2009. However, panel (a) of Fig. 1 shows that the UNC admission policy likely changed admission decisions for the least selective institutions. For these institutions, 25% of their admitted applicant pool score below 750 on the SAT before 2009. In fact, on average, least selective UNC campuses increase their 25th percentile SAT scores of their admitted pool after 2009.

3. Data and descriptive statistics

I rely on rich administrative data from North Carolina public school students to form an unbalanced panel of approximately 75,000 public high school students belonging to the graduation cohort of 2009.[8] These data allow me to capture key information on student outcomes that proxy effort and engagement. These proxies include student GPA, credit-taking, absences and suspensions.

GPA and credit-taking patterns are recovered from administrative records of high school student transcripts. These transcripts record information on the courses taken by each student, the subject, the academic level (e.g. honors, AP, etc.) and the final mark obtained. Although high school transcripts are available going back to the 2004–05 school year, coverage of transcripts only becomes universal for the 2008–09 school year. Therefore, I have GPA and course-taking information for all 12th graders in 2009, but not for their 11th, 10th, and 9th grades.

Information on absences are collected by the state of North Carolina for accountability purposes during each testing period (end of Fall and Spring terms). Furthermore, data on student suspensions come from administrative records on reportable incidents. Schools in North Carolina are required to report any incidents that occur within the school that lead to a student suspension, expulsion and/or transfer to an alternative learning program.

In addition to key behavior proxies, the NCERDC data also includes information on student SAT performance in their latest administration from 2009 to 2013 for those students who take the SAT. Therefore, the first year for which data on SAT performance is available is for 12th graders in 2009.[9] As described in Section 2, the UNC minimum admission requirement policy also set minimum admission cutoffs for ACT-takers. Data on ACT performance would be valuable to (a) be used as a running variable among ACT-takers and (b) examine student test choice between ACT and SAT as a response to the policy. Unfortunately, ACT performance is not available in the data until after 2013, the year in which North Carolina adopted the ACT as part of the state's school accountability program. Thus, the remainder of this paper will focus on SAT-takers for the estimation of the models presented in Section 4.

Table 1 provides an overview of the data for high school students belonging to the graduating cohort of 2009. Columns (1) through (3) present descriptive statistics for students who score close to the UNC cutoff on their SATs. These students make up the regression discontinuity sample, which are the data used for the main estimates of this paper. Columns (4) though (6) present descriptive statistics for the full graduating cohort of 2009.

Overall, students close to the UNC cutoff are more likely to be female, economically disadvantaged (EDS), and more likely to be Black compared to the average North Carolina public school student.[10] Furthermore, they are less likely to be classified as academically gifted. By construction, all students in the RD sample take the SAT and score less than the average student who takes the SAT in the state. Finally, in terms of pre-treatment effort measures, students belonging to the RD sample generally obtain lower GPAs, take slightly more coursework and are less absent than their counterparts throughout their high school years.

4. Research design

This section details the empirical framework of this paper. I first describe the regression discontinuity design and estimation used for the main estimates which are presented in Section 5. I then describe the framework used to set identify regression discontinuity estimates in the presence of potential manipulation of the running variable following Gerard et al. (2020). Estimates of these bounds are presented in Section 6.

[8] In partnership with NC Department of Public Instruction, the North Carolina Education Research Data Center (NCERDC) stores and manages data on the state's public schools, school districts, students, and teachers.

[9] Unfortunately, this feature of the data does not allow me to observe trends of student behavior as a function of SAT performance *prior* to the implementation of the UNC admission policy.

[10] EDS is proxied by student eligibility for free or reduced-price lunch.

Table 1
Descriptive statistics.

	RD sample			Full sample		
	(1) Mean	(2) SD	(3) Obs	(4) Mean	(5) SD	(6) Obs
Panel A. student demographics						
Female	0.61	0.49	3,745	0.52	0.50	75,656
Economically disadvantaged	0.58	0.49	3,750	0.38	0.48	75,720
White	0.28	0.45	3,750	0.63	0.48	75,720
Black	0.61	0.49	3,750	0.26	0.44	75,720
Hispanic	0.05	0.21	3,750	0.05	0.23	75,720
Asian	0.03	0.16	3,750	0.02	0.15	75,720
Student with disability	0.09	0.28	3,750	0.09	0.28	75,720
Academically gifted	0.02	0.15	3,750	0.21	0.41	75,720
Took SAT	1.00	0.00	3,750	0.52	0.50	75,720
SAT score	718.08	40.36	3,750	1004.80	190.87	39,222
Panel B. pre-treatment effort and engagement outcomes						
GPA 9th grade	2.63	0.61	3,392	3.02	0.87	67,313
GPA 10th grade	2.53	0.61	3,441	3.00	0.92	68,629
GPA 11th grade	2.56	0.63	3,555	3.07	1.00	71,061
Credits 9th grade	8.28	1.65	3,393	8.07	1.53	67,347
Credits 10th grade	8.25	1.59	3,441	8.06	1.52	68,649
Credits 11th grade	8.17	1.57	3,556	7.97	1.52	71,081
Absences 9th grade	4.91	4.75	3,399	5.48	5.60	67,269
Absences 10th grade	5.64	5.53	3,500	6.39	6.37	69,541
Absences 11th grade	6.51	6.34	3,645	7.48	7.77	72,103
Suspensions 9th grade	0.23	0.84	3,453	0.18	0.78	68,261
Suspensions 10th grade	0.18	0.71	3,531	0.17	0.71	69,979
Suspensions 11th grade	0.18	0.72	3,662	0.17	0.74	72,623

Notes: This table shows descriptive statistics of the data. Panel A. provides summary statistics of student demographics. Panel B. provides summary statistics of student effort and engagement outcomes in pre-treatment years (pre 12th grade). Columns 1 through 3 show descriptive statistics of the sample used for the estimation of the main results shown in Table 2 using the optimal bandwidth estimated for GPA. Columns 4 through 6 show descriptive statistics of the all students belonging to the graduating cohort of 2009.

4.1. Regression discontinuity design

The sharp cutoff set by UNC's new minimum admission requirement, I argue, generated discontinuous incentives for effort for students close to the SAT threshold given the implied jump in the probability of admission to a UNC campus at the cutoff. Students that score just below the SAT requirement have a zero probability of admission to any UNC campus. However, students who score just above, perceive a positive probability of admission. Students who perceive this jump in the probability of admission will have increased incentives to perform well during their remaining academic coursework. Several mechanisms could drive the change in student behavior at the cutoff. Students can choose to exert more effort in order to meet the UNC high school GPA admission requirement once they have fulfilled the SAT requirement. Students can also choose to exert more effort to further increase their probability of admission via increased high school GPAs. Finally, passing the first UNC requirement might increase motivation and effort regardless of student strategic behavior with respect to the admission process.

I use student SAT test score data under a regression discontinuity framework to leverage quasi-random variation from students' SAT scores to compare effort of students who perform on either side of the threshold. Students just above and just below the SAT cutoff can be thought to be similar on several dimensions except for their UNC admission eligibility.

The regression discontinuity specification is given by:

$$y_i = \gamma + \beta D_i + f(SAT_i, D_i) + \epsilon_i \quad (1)$$

where i indexes students. The variable y_i represents the 12th grade outcome of interest, D_i is a treatment indicator variable that takes the value of one when students' SAT score meet the relevant UNC cutoff. SAT_i represents students' performance on their SAT test and is the running variable of this specification. The estimated coefficient, $\hat{\beta}$, should be interpreted as the effect of differential incentives embedded in the UNC admission policy on students' effort during their senior year.

Throughout this paper, estimation of Eq. (1) is carried out via local linear regression (Gelman & Imbens, 2019), using optimal bandwidths and robust confidence intervals proposed by Calonico et al. (2014). Because optimal bandwidths are estimated separately for each student outcome, the number of observations in each estimation may vary. However, estimates are robust to bandwidth selection (see Appendix A.2). Finally, standard errors are clustered at the individual levels of the running variable to correct for misspecification due to discrete running variable (Lee & Card, 2008).

4.2. Regression discontinuity bounds under manipulated running variable

A first order concern of the estimation presented in Eq. (1) is the fact that I do not observe all student SAT administrations. Instead, I only observe the last one. Student SAT retaking can be an important source of endogeneity if students systematically retake the SAT in order to place themselves to the right of the cutoff to gain UNC admission eligibility. This selection into the right side of the cutoff can be an important threat to the internal validity of the regression discontinuity design presented in Eq. (1).[11] I address this concern by bounding the regression discontinuity point estimates following Gerard et al. (2020).

Gerard et al. (2020) develop a framework to estimate sharp bounds of regression discontinuity treatment effects under a potentially manipulated running variable. For the purposes of this paper, consider two types of students: manipulators ($M_i = 1$) and non-manipulators ($M_i = 0$). Manipulators have perfect control of the *side* of the cutoff they are in and manipulate in only one direction. Therefore, manipulators can perfectly retake the SAT until they are on the right side of the cutoff and they will never retake to be on the left side of the cutoff. Under

[11] In the past, studies that use SAT scores as the running variable in regression discontinuity designs, as for example, Goodman et al. (2017), find no differences in estimated treatment effects when using students' first and last SAT administration scores.

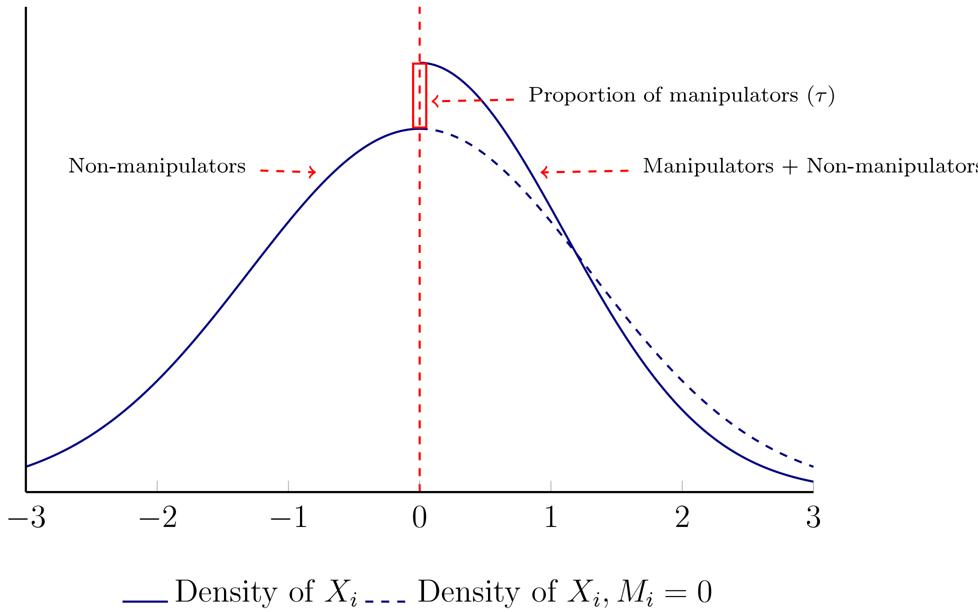

Fig. 2. Identification of manipulators.
Notes: This figure presents an illustration of how the proportion of manipulators can be recovered from the data.
Source: Gerard et al. (2020).

this setting, the parameter of interest is the treatment on the treated for non-manipulators[12]:

$$\begin{aligned}\Gamma_0 &= E(Y_i(1) - Y_i(0)|X_i = c, M_i = 0) \\ &= E(Y_i|X_i = c^+, M_i = 0) - E(Y_i|X_i = c^-, M_i = 0)\end{aligned} \tag{2}$$

Note that, because manipulation is assumed to have one direction (students retake the SAT to be on the right side of the cutoff), we know that $E(Y_i|X_i = c^-, M_i = 0) = E(Y_i|X_i = c^-)$. However, we cannot identify the term $E(Y_i|X_i = c^+, M_i = 0)$ of Eq. (2), because on the right side of the cutoff we cannot disentangle manipulators from non-manipulators. Therefore, we cannot point identify Γ_0. However, we can bound it using the estimated proportion of manipulators (τ) that is identified from the data:

$$\tau \equiv Pr(M_i = 1|X_i = c^+) = \frac{f_x(c^+) - f_x(c^-)}{f_c(c^+)} \tag{3}$$

To understand why the proportion of manipulators (τ) is identified from the data, Fig. 2 presents an illustration of the density of a manipulated running variable. All students to the left of the cutoff are non-manipulators by definition. However, students on the right side of the cutoff are both manipulators and non-manipulators. Since manipulators are able to control the side on the cutoff they are in, this generates a discontinuity of the density of the running variable at the cutoff. This discontinuity represents the proportion of students who manipulate the running variable.

Using the proportion of manipulators, lower bounds can be derived by trimming quantile $(1 - \tau)$ of the outcome distribution, and therefore, excluding units with the highest outcomes from the estimation. Analogously, upper bounds can be derived by trimming quantile τ of distribution, and excluding the units with the lowest outcomes. Then, sharp lower and upper bounds on Γ_0 are given by:

$$\begin{aligned}\Gamma_0^L &= E(Y_i|X_i = c^+, Y_i \le Q_{Y|X=c^+}(1-\tau)) - E(Y_i|X_i = c^-) \\ \Gamma_0^U &= E(Y_i|X_i = c^+, Y_i \ge Q_{Y|X=c^+}(\tau)) - E(Y_i|X_i = c^-)\end{aligned} \tag{4}$$

where $Q_{Y|X}$ is the conditional quantile function of the outcome variable. Estimation of both the density function and the conditional quantile are carried out via a local linear polynomial approximations. For a more detailed explanation of estimation and inference see Gerard et al. (2020).

5. Results

This section presents regression discontinuity estimates of Eq. (1) for NC's 2009 cohort of high school students. First, I present core results on student effort and engagement proxies in Section 5.1. Then I present mechanisms in Section 5.2, by exploring effects on student course-taking patterns. Finally, in Section 5.3, I present heterogeneity of the results across student demographics.

5.1. Core results

In this section, I present Eq. (1) estimates for student 12th grade GPA, absences, times in suspension, and credit-taking. Panels (a) through (d) of Fig. 3 plot average 12th grade outcomes for each SAT score bin with a global linear fit on either side of the 700 cutoff.[13] Visual evidence provided by these plots suggest a possible student response on 12th grade GPA, absences, and times in suspension.

Table 2 presents regression discontinuity estimates of Eq. (1) for GPA, absences, suspensions, and credit-taking in columns (1) to (4), respectively. These estimates indicate that students who score just above the UNC SAT minimum admission threshold obtain GPAs that are, on average, 0.08 points higher compared to students who score just below. This represents an increase of 2.8% from the control group mean and a 12th grade GPA that is 0.11 standard deviations higher than students with SATs just below the cutoff. In addition, threshold crossing

[12] The treatment effect for manipulators cannot be recovered as, by assumption, all manipulators on the right side of the cutoff and are treated (Gerard et al., 2020).

[13] The UNC minimum admission requirements sets a GPA threshold that is based on students' high school weighted GPA. Weighted GPA assigns additional grade points for Honors and AP/IB courses. I use weighted GPA instead of unweighted GPA as an outcome variable is to reflect this policy. However, Appendix A.3 shows that results are robust to using unweighted GPA as the outcome.

Fig. 3. Global linear fit.
Notes: This figure shows discontinuities in senior-year weighted GPA, absences, times in suspension, and course-taking (Panels (a), (b), (c), and (d). All panels are created by plotting the average outcomes across each SAT score bin and fitting the data using a linear regression without controls on either side of the cutoff.

Table 2
Effects on student effort and engagement outcomes.

	(1) GPA	(2) Absences	(3) Suspensions	(4) Credits
RD Estimate	0.078**	−1.479***	−0.074*	−0.084
	(0.026)	(0.172)	(0.030)	(0.060)
Bandwidth	79.40	65.40	82.34	90.29
Observations	3,748	2,836	4,298	4,858
Control Mean	2.75	8.00	0.27	7.50

Notes: this table shows regression discontinuity estimates of senior-year gpa, course-taking, absences and days in suspension. Optimal bandwidths are selected following Calonico et al. (2014). Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

also reduces absences by 1.5 days (0.19 SD) and suspensions by 0.07 times (0.09 SD). Finally, column (4) shows that there is no statistically significant effect on the number of courses students take in 12th grade. Overall, these results suggests that changes in incentives brought by the design of UNC's new admission policy induced an increase in high school student engagement.

5.2. Mechanisms

I interpret decreased absences and suspensions shown in Table 2 as a reflection of increased engagement in school work due to differential incentives generated by the implementation of UNC's new admission policy. Nevertheless, interpretation of increased student GPA is less straightforward. Results presented in Table 2 suggest that students who score just above the UNC SAT minimum admission threshold obtain higher 12th grade GPAs than their counterparts. While, this effect can be driven by students choosing to exert more effort in their coursework, it can also be drive by students choosing to substitute away from more demanding coursework.

I explore this mechanism by splitting the effects on student GPA and credit-taking into core and non-core subject effects. I define core courses as courses in the subjects of English, Math, Science, and Social Studies. Non-core courses are all other course subjects. Presumably, courses belonging to core subjects are, on average, more demanding. Table 3 presents regression discontinuity estimates for core and non-core GPA and course-taking in 12th grade.

Estimates presented in panel A of Table 3 suggest that the effects of student GPA presented in Table 2 are driven by higher non-core-specific GPA. Additionally, estimates for core and non-core credit-taking indicate that the null effects for credit-taking presented in Table 2 mask important dimensions of student response. Students on the right side of the threshold take, on average, 0.225 (−6.4%) less courses belonging to core subjects than students who score just below the SAT requirement, while also taking 0.221 (+5.5%) more credits in non-core subjects. I further disaggregate these results and present credit-taking effects across core subjects in panel B of Table 3. Threshold crossing induces students to take less courses in math, social studies, and ELA. However, effects are larger and statistically significant for diminished credit-taking in social studies and ELA.

While panels A and B of Table 3 provide evidence that students are substituting away from more demanding coursework as they meet

Table 3
Effects for credit-taking.

	(1)	(2)	(3)	(4)
Panel A. outcomes across core and non-core subjects				
	Weighted GPA		*Credits*	
	Core	Non-Core	Core	Non-Core
RD Estimate	0.022	0.087*	−0.225***	0.221***
	(0.018)	(0.039)	(0.049)	(0.028)
Panel B. credit-taking across core subjects				
	Math	Science	Social	ELA
RD Estimate	−0.014	−0.005	−0.084*	−0.109***
	(0.046)	(0.040)	(0.034)	(0.008)

Notes: This table shows regression discontinuity estimates of senior-year GPA, course-taking of core and non-core courses. I define core courses as courses in the subjects of English, Math, Science, and Social Studies. Non-core courses are all other course subjects. Optimal bandwidths are selected following Calonico et al. (2014). Estimated bandwidths for core GPA, non-core GPA, core credits, and non-core credits are 77.31, 100.7, 59.4, and 51.76, respectively. Estimated bandwidths for math, science, social studies, and ELA course-taking are 91.10, 80.85, 58.13, and 64.61, respectively. Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

Table 4
Credit-taking effects across course difficulty.

	Average letter courses			Deviation from average	
	(1) A-Courses	(2) B-Courses	(3) C-Courses	(4) Harder	(5) Easier
RD Estimate	0.119	−0.047	−0.043	−0.197**	0.027
	(0.073)	(0.051)	(0.061)	(0.065)	(0.115)
Observations	3115	3142	2230	3678	4278

Notes: This table shows regression discontinuity estimates of senior-year course-taking across courses for which students on average obtain an A (Column (1)), a B (Column (2)), and a C (Column (3)). Optimal bandwidths are selected following Calonico et al. (2014). Estimated bandwidths for A, B, and C courses are 88.67, 90.29, 69.07, and 57.57, respectively. Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

the SAT requirement, course subject categories are broad and might not be a precise measure of course difficulty. To address this concern, I provide two alternative measures of course difficulty. In the first, I recover school-year-course-level grade averages. With these averages, I classify a course as an "A-course", if on average, students obtain an A in that course. I do the analogous classification for "B-courses" and "C-courses". I then recover the number of courses students take for each category. The second measure of course difficulty is based on students' deviation on their performance in a course from the average performance on all the courses they take that year. Therefore, a course is "Harder" if, on average, students who take the course perform worse on the course than they do on other courses, and "Easier" if otherwise. See Appendix A.4 for a more detailed explanation of this measure.[14]

Table 4 presents regression discontinuity estimates for the number of courses students take in 12th grade across several difficulty measures. Even though most of the estimates are not statistically significant, the direction of the coefficients follow the same underlying mechanism shown in Table 3. Students seem to be taking more courses where, on average, they obtain an A, and less courses where, on average, they obtain a B or a C (columns (1) through (3)). At the same time, students are taking less courses that are classified as "Harder" (columns(4) and (5)).

[14] I would like to thank and anonymous referee for suggesting this alternative measure of course difficulty.

Overall, estimates presented in Tables 3 and 4 suggest that in response to the incentives generated by UNC's admission policy, students who meet the SAT requirement substitute away from more demanding coursework during their last year of high school. Thus, effects on 12th grade GPA presented in Table 2 might not fully reflect increased effort but rather strategic course choices that boost students' 12th grade GPAs.

5.3. Heterogeneity

In this section I explore heterogeneity of the effects presented in Tables 2 and 3 across different student populations. Heterogeneity in student response is likely to arise in this context for two main reasons. First, the literature examining student response to incentives in postsecondary settings highlights how incentives have heterogeneous effects across student demographics and performance (Angrist et al., 2009; Casey et al., 2018; Cornwell et al., 2003; Lindo et al., 2010; Scott-Clayton & Schudde, 2019). While this literature focuses on incentives generated by post-secondary policies such as performance standards, merit aid requirements, and academic probation in college, one can think of 12th grade high school students responding in similar ways as college-going youth. Second, as discussed in Section 2, UNC's new admission policy likely affected admission decisions of the least selective campuses within the UNC system. These campuses predominately serve minority and less affluent students. As such, one can expect minority and less affluent high school students to respond more strongly to this policy because they foresee changes in admission practices of the UNC campuses they intend to attend.

Table 5 presents estimates of Eq. (1) for GPA, absences, and times in suspension in panel A. Panel B presents estimate for core and non-core credit-taking. All estimates are presented across student race, SES, gender, and 9th grade GPA. Columns (2) and (3) present regression discontinuity estimates for white and non-white students, respectively. Differences across these two groups are statistically significant for GPA, suspensions, and core and non-core credit-taking. Estimates presented in these columns suggest that minority students respond to the incentives embedded by the UNC admission policy by obtaining higher GPA's while at the same time engaging in less strategic course-taking compared to the response of their counterparts.

Columns (4) and (5) presents results for students who are economically disadvantaged and for students who are not, respectively. Statistically significant differences across these two groups include absences, suspensions, core and non-core course-taking. These results suggest that low income students that make the SAT cutoff have higher effort and engage in less strategic course-taking than their counterparts. However, this response does not translate into higher GPAs for less affluent students.

Recent work from Goodman et al. (2020) highlight how low SES students are substantially less likely to retake the SAT.[15] Table 5 shows that low SES student respond to the policy to a greater extent than their peers in some dimensions of student engagement, suggesting a true response of students to the UNC admission policy.

Columns (6) and (7) present estimate for female and male students, respectively. Statistically significant differences across these two groups include GPA, absences, suspensions, non-core course-taking. These results suggest that male students engage in more gaming practices via course selection and obtain higher GPAs than female students without increased effort.

Finally, Columns (8) and (9) present estimates for students with 9th grade GPA below and above the median, respectively. Presumably, students with higher pre-treatment GPAs are more likely to be considering

[15] Internal validity checks of manipulation of the running variable and pre-treatment covariate smoothness at the cutoff suggest little scope of selection into the right side of the cutoff for this subgroup of students (see Appendix A.5).

Table 5
Effects across student demographics.

		Race		SES		Gender		9th GPA	
	(1) Full	(2) White	(3) Non-White	(4) Low	(5) High	(6) Female	(7) Male	(8) Below	(9) Above
Panel A. effort and engagement outcomes									
GPA	0.078**	0.045	0.086***	0.071	0.055	0.043	0.175***	−0.023***	0.095***
	(0.026)	(0.077)	(0.021)	(0.037)	(0.126)	(0.032)	(0.052)	(0.041)	(0.046)
Absences	−1.479***	−0.915	−1.523***	−1.802***	−0.706	−1.935***	−0.247	−1.063	−1.352
	(0.172)	(1.145)	(0.427)	(0.435)	(0.703)	(0.495)	(0.432)	(0.530)	(0.890)
Suspensions	−0.074*	0.023	−0.095*	−0.152*	0.090	−0.080***	−0.072	−0.048	−0.101***
	(0.030)	(0.045)	(0.039)	(0.049)	(0.068)	(0.026)	(0.064)	(0.044)	(0.047)
Panel B. student credit-taking									
Core	−0.225***	−0.449***	−0.119	−0.102	−0.358**	−0.220***	−0.193	−0.060	−0.537***
	(0.049)	(0.103)	(0.073)	(0.081)	(0.130)	(0.028)	(0.130)	(0.069)	(0.075)
Non-Core	0.221***	0.566***	0.122***	0.156	0.201	0.095*	0.374***	0.312**	0.152
	(0.028)	(0.061)	(0.034)	(0.103)	(0.119)	(0.043)	(0.051)	(0.063)	(0.052)

Notes: This table shows regression discontinuity estimates of senior-year GPA across student demographics. Optimal bandwidths are selected following Calonico et al. (2014). Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

a college education, and therefore, more likely to adjust their behavior as a response to new admission requirements. Statistically significant differences across these two groups include GPA, suspensions, core and, non-core course-taking. These results suggest that students with higher pre-treatment GPAs are responding to a greater extent than their peers in both their GPAs and by substituting away from more challenging coursework.

Estimates presented in Table 5 document stark heterogeneity of student response across demographics. Minority and less affluent students just above the cutoff have higher effort allocations via decreased absences and suspensions but engage to a lesser extent in gaming practices via course selection. At the same time, students with higher pre-treatment GPAs are more likely to respond to the policy than their peers.

6. Internal validity checks

My results show that the discontinuity in the probability of admission to a UNC campus, generated by the 2009 admission policy shift, led students to the right of the cutoff to have higher effort levels, as proxied by absences and suspensions, but also led students to substitute away from more demanding coursework in order to boost their GPAs. The ability of my research design to produce causal estimates of these effects rests on the assumption that validity holds. An important threat to the validity of this design is student SAT retaking. Because I am only able to observe students' latest SAT administration, if students retake the SAT until they meet the UNC cutoff, then my results might be reflecting selection into the right side of the SAT cutoff rather than student response to the incentives generated by the design of UNC's new admission policy. In this section, I provide four types of evidence based on observables to address this validity concern.

6.1. Manipulation of the running variable

First, I explore whether there is evidence of manipulation of the running variable. I implement a local polynomial density test following Cattaneo et al. (2019). This test estimates a local polynomial density on either side of the cutoff and tests for the null hypothesis that the limit of both functions when they approach the cutoff from either side are equal. Fig. 4 presents the fit of the local polynomial density functions and the histogram of the SAT test scores. With a p-value of 0.376, I am not able to reject the null hypothesis that the density functions are equal at the cutoff. Therefore, this test provides evidence that there is no systematic manipulation of the running variable.

6.2. Policy relevance

The second type of evidence examines the relevance of the policy cutoff. The idea behind this check is the following: if students are responding to the UNC admission policy alone, there should not be any discontinuous response of student outcomes on values of the running variable other than the relevant cutoff of 700 set by the UNC Board of Governors in 2009. To check whether this is the case, I run regression discontinuity estimates for 12th grade GPA, absences, and suspensions, for all cutoffs close to UNC's 700 cutoff. Table 6 presents estimates of these specifications for all cutoffs between 670 and 730. Estimates presented in Table 6 provide evidence that in fact, discontinuity of student outcomes are only found in the relevant UNC admission policy cutoff.

6.3. Covariate smoothness

The third type of evidence for validity examines covariate smoothness across the SAT cutoff. Figs. 5 and 6 provide scatter plots analogous to the ones shown in Fig. 3 for student covariates and pre-treatment (9th and 10th grade) outcomes, respectively.

Fig. 7 provides regression discontinuity estimates of Eq. (1) using standardized student covariates and pre-treatment (9th and 10th grade) outcomes presented in Figs. 5 and 6. Across most estimates, I find no evidence of a discontinuity except for two covariates: 8th grade math test scores and 9th grade GPA.

Discontinuities of these variables are concerning because they might indicate that higher achieving students systematically retake the SAT in order to be on the right side of the cutoff. To address this concern I run three additional tests. First, I recover bounds for regression discontinuity estimates under possible manipulation of the running variable in Section 6.4. Second, I explore whether pre-treatment GPA is discontinuous at the cutoff for students above and below the 9th grade GPA median. Presumably, students with higher pre-treatment GPAs are more likely to drive selection into the right side of the cutoff because they are more likely to be considering a college education. Table A.2 presents evidence that suggests that this is not the case.

Finally, given the high dimensionality of the set of pre-treatment variables, I assess covariate smoothness at the cutoff by summarizing the variation of these variables in two different ways. First, I summarize the variation of the data into an index. Second, I recover the fitted values of the outcomes using a regression of the outcome on pre-treatment covariates. Both the index and the fitted values of the outcomes represent the variation that is explained by pre-existing student differences and thus, should not be discontinuous at the cutoff. Regression discontinuity estimates of this exercise (Table A.1) do not provide evidence of discontinuity of pre-treatment variables at the cutoff (see Appendix A.6 for more details).

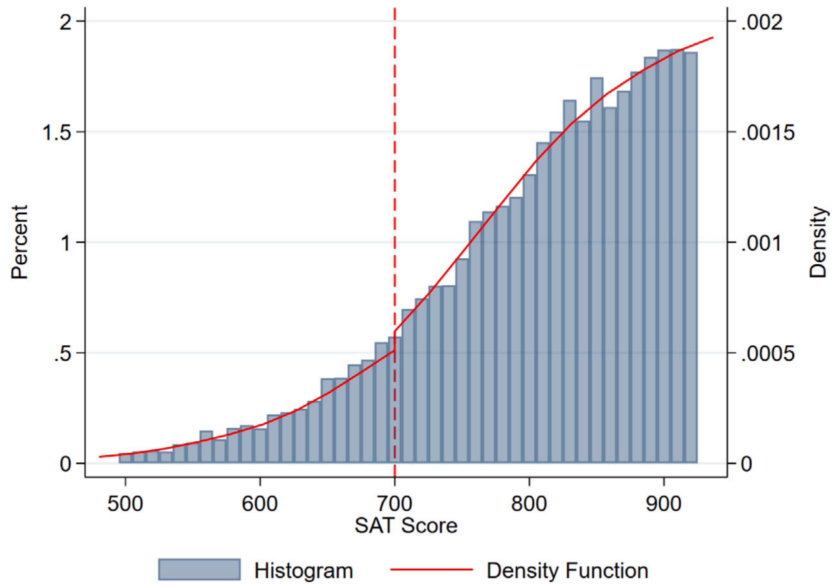

Fig. 4. Local polynomial density test.
Notes: This figure shows the histogram of the SAT test scores in blue bars and the local polynomial density estimation on either side of the cutoff in red solid lines estimated following Cattaneo et al. (2019). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Regression discontinuity bounds of effects on student outcomes.

	(1) 670	(2) 680	(3) 690	(4) 700	(5) 710	(6) 720	(7) 730
GPA	0.057	−0.030	−0.025	0.078**	0.053	0.026	0.015
	(0.033)	(0.039)	(0.029)	(0.026)	(0.035)	(0.019)	(0.015)
Absences	0.665*	−0.147	−0.717	−1.479***	−0.485	0.628*	0.644
	(0.266)	(0.332)	(0.446)	(0.172)	(0.514)	(0.263)	(0.359)
Suspensions	−0.124	−0.015	−0.019	−0.074*	−0.048	0.067*	0.033
	(0.064)	(0.073)	(0.045)	(0.030)	(0.032)	(0.031)	(0.038)

Notes: This table shows regression discontinuity estimates of senior-year GPA, absences and times in suspension for different SAT cutoffs. Optimal bandwidths are selected following Calonico et al. (2014). Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

6.4. Regression discontinuity bounds

Despite evidence of no systematic manipulation of the running variable given by the local polynomial density test implemented in Fig. 4, discontinuities of 8th grade math test scores and 9th grade GPA found in Fig. 7 are of concern as they could indicate selection into treatment. In order to address this concern, I present a fourth type of evidence for validity that recovers bounds for the regression discontinuity estimates found in Table 2 under the framework presented in Section 4.2. These bounds are presented in Table 7. Overall, the estimated bounds of the regression discontinuity results are fairly tight. The reason why the bounds are fairly tight is due to the fact that the actual jump in density at the cutoff is estimated to be small.

As explained in Section 4.2, regression discontinuity estimates are bounded using the estimated proportion of manipulators, τ, to trim upper and lower tails of the outcome distribution. This proportion is identified by the discontinuity of the density function at the cutoff via Eq. (3). In order to assess how sensitive the estimated regression discontinuity bounds presented in Table 7 are to the estimated density discontinuity (τ), Fig. 8 presents how these bounds change as the estimated τ changes.

As expected, larger values of τ lead to larger the bounds. However, as seen in Fig. 8, the estimated regression discontinuity bounds are robust to the estimated proportion of manipulators for all student outcomes. The estimated bounds for student 12th grade GPA and student 12th grade absences start to include zero after τ is set to be 9 times and 6 times the actual estimated value, respectively. Therefore, actual manipulation of the running variable would need to be 6 to 9 times as large as the one observed in the data in order for my results to go away. Furthermore, the estimated bounds for student 12th grade times in suspension never include zero even if the estimated proportion of manipulators, τ, is set to be 10 times the actual estimated value. Thus, it is unlikely that manipulation of the running variable and selection into treatment is such that it completely drives away the results presented in Section 5.

7. Conclusion

Changing the incentives students face is one lever for educators and policymakers to improve student learning in the presence of student disengagement. However, education economists and policymakers lack an understanding of the ways in which these incentive structures can shape student responses. Generally, education accountability policies seek to shift incentives for effort of teachers and schools rather than students. Thus, there have not been many opportunities to evaluate the student incentive-effort connection outside of researcher designed pay-for-performance incentives schemes.

Incentive structures embedded within policies for access to higher education provide an alternative, more realistic, setting to examine student effort. In this paper, I leverage exogenous variation from a

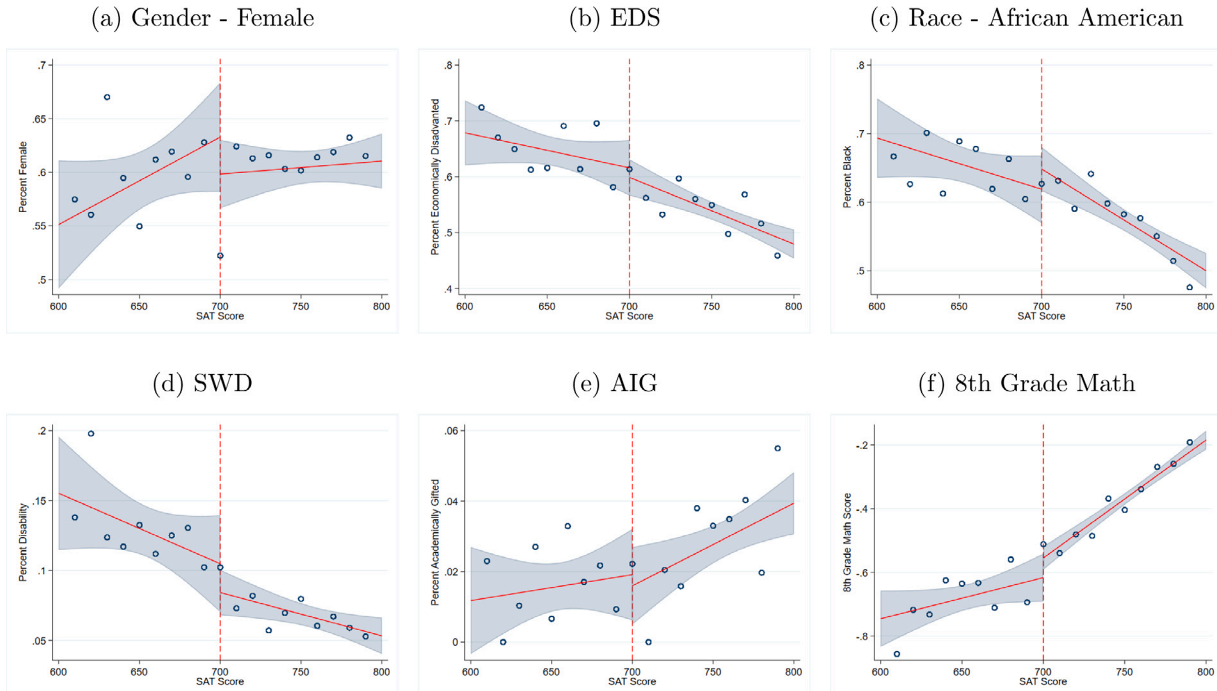

Fig. 5. Smoothness of student pre-treatment covariates.
Notes: This figure shows discontinuities in student pre-treatment covariates. All panels are created by plotting the average outcomes across each SAT score bin and fitting the data using a linear regression without controls on either side of the cutoff. 95% confidence intervals are displayed in the shaded blue areas.

Table 7
Regression discontinuity bounds of effects on student outcomes.

	(1) GPA	(2) Course-Taking	(3) Absences	(4) Suspensions
RD Estimate	0.078**	−0.084	−1.479***	−0.074*
	(0.026)	(0.060)	(0.172)	(0.030)
Bounds	[0.069, 0.094]	[−0.135, −0.053]	[−2.351, −1.190]	[−0.108, −0.074]
Density Jump	0.007	0.006	0.037	0.007

Notes: This table shows regression discontinuity estimates of senior-year GPA, absences and times in suspension for students all students in Column (1), for students who want to attend a UNC campus in Column (2), and for students who do not want to attend a UNC campus in Column (3). Optimal bandwidths are selected following Calonico et al. (2014). Local linear regressions are estimated on either side of the cutoff. Standard errors are clustered at the individual levels of the running variable to account for misspecification due to discrete running variable (Lee & Card, 2008). * p<0.1, ** p<0.05, *** p<0.01.

change in admission requirements in one of the largest university systems in the US that provided a salient and high-stakes setting to examine how student effort responds when incentive structures shift.

Using rich administrative data from the North Carolina public school system, I evaluate whether students' investments in schoolwork shifted as a results of this exogenous shift in incentives for effort. I find that students respond by increasing GPA and decreasing absences and suspensions. While these effects suggest an increase in student engagement as a result of the admission policy, further exploration of course-taking suggests that student increases in GPA are driven by changes in course composition as students substitute away from more demanding coursework. These unintended consequences of admission policies on student course-taking decisions can lead students to miss important learning opportunities in high school, possibly generating detrimental effects on student postsecondary success.

Finally, in line with the literature examining student response to incentives in postsecondary settings, I document stark heterogeneity of student response across demographics. Minority and less affluent students just above the cutoff have higher effort allocations via decreased absences and suspensions but engage to a lesser extent in gaming practices via course selection. At the same time, students with higher pre-treatment GPAs are more likely to respond to the policy than their peers.

Data availability

The authors do not have permission to share data.

Appendix

A.1. ACT admission trends for UNC campuses

See Figs. A.1 and A.2.

A.2. Robustness of regression discontinuity estimates to bandwidth selection

See Fig. A.3.

A.3. Results for raw GPA v. Weighted GPA

See Fig. A.4.

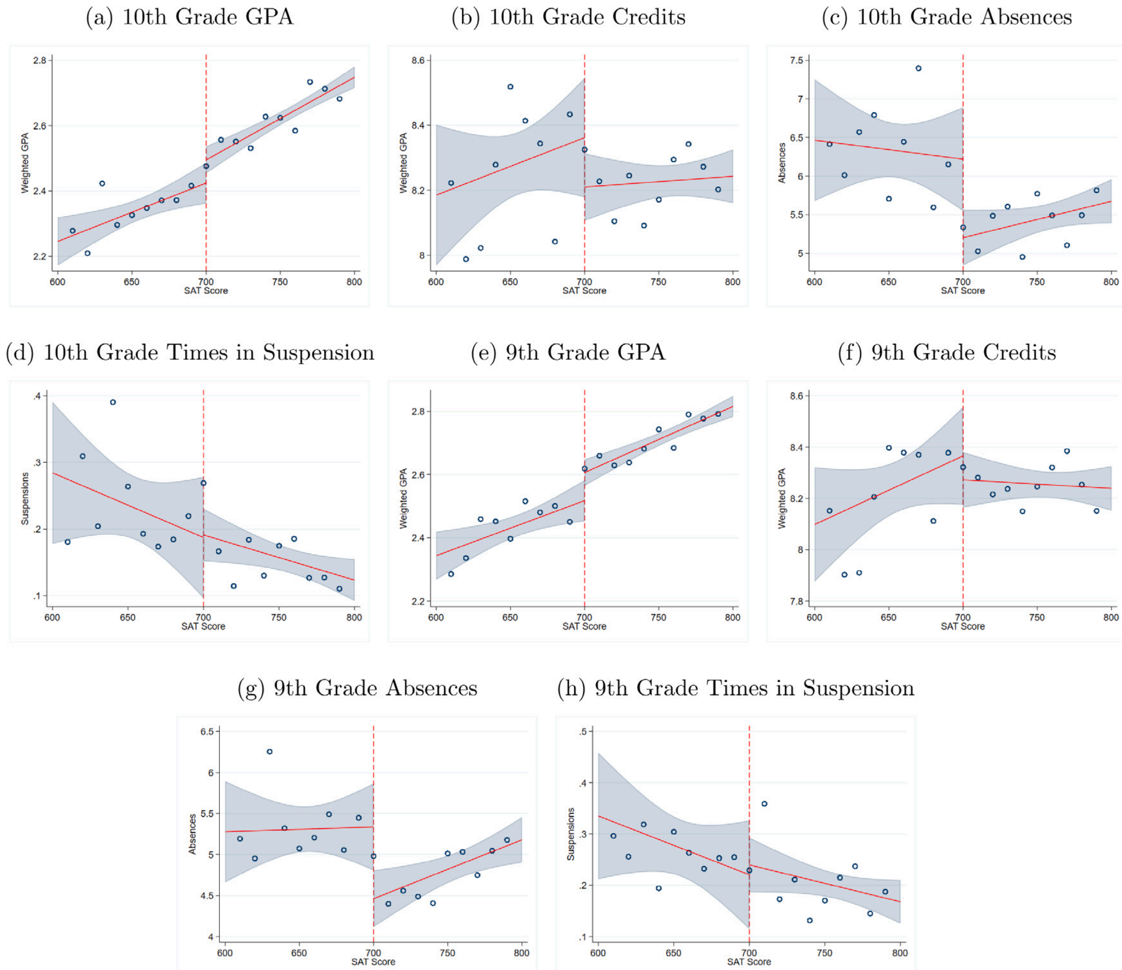

Fig. 6. Smoothness of student pre-treatment outcomes.
Notes: This figure shows discontinuities in student pre-treatment outcomes. All panels are created by plotting the average outcomes across each SAT score bin and fitting the data using a linear regression without controls on either side of the cutoff. 95% confidence intervals are displayed in the shaded blue areas.

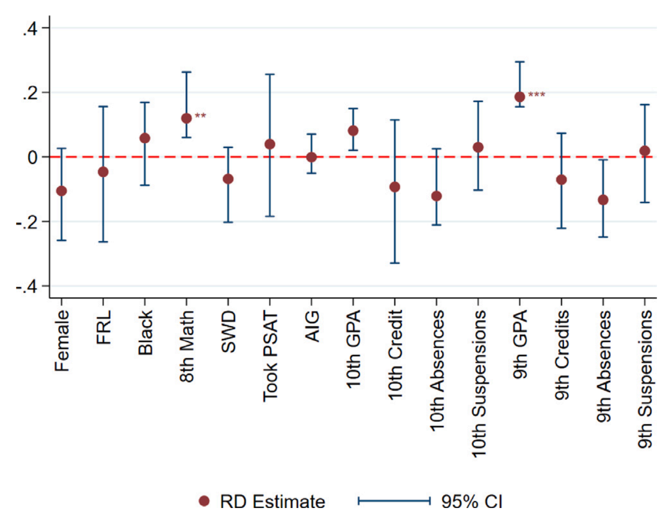

Fig. 7. Discontinuity in student demographics and pre-treatment covariates.
Notes: This figure shows regression discontinuity estimates of Eq. (1) using standardized pre-treatment student covariates and outcomes as the dependent variable. Robust 95% confidence intervals are estimated following Calonico et al. (2014). I apply the Bonferroni correction to adjust significance for multiple comparisons. * p<0.1, ** p<0.05, *** p<0.01.

A.4. Alternative measure of course difficulty

This section provides a detailed overview of the procedure I followed to recovered the course difficulty measure used in columns (4) and (5) of Table 4. For every student, i, taking a course, c, at year t, let g_{ict} represent the grade the student obtains in that course. Let d_{ict} represent the deviation of the student's performance in the course with respect to that student's overall GPA that year, such that:

$$d_{ict} = g_{ict} - GPA_{it} \tag{5}$$

Given this measure, I can recover the average deviation, d_{ct}, for students in every course:

$$d_{ct} = \frac{1}{N_c} \sum_{N} d_{ct} \tag{6}$$

With this measure at hand, I classify courses as "Harder" if $d_{ct} >= 0$. Therefore, a course is difficult if, on average, students who take the course perform worse on the course than they do on other courses they take that year. I classify a course as "Easier" if $d_{ct} < 0$.

A.5. Histograms and covariate smoothness for low SES students

See Fig. A.5.

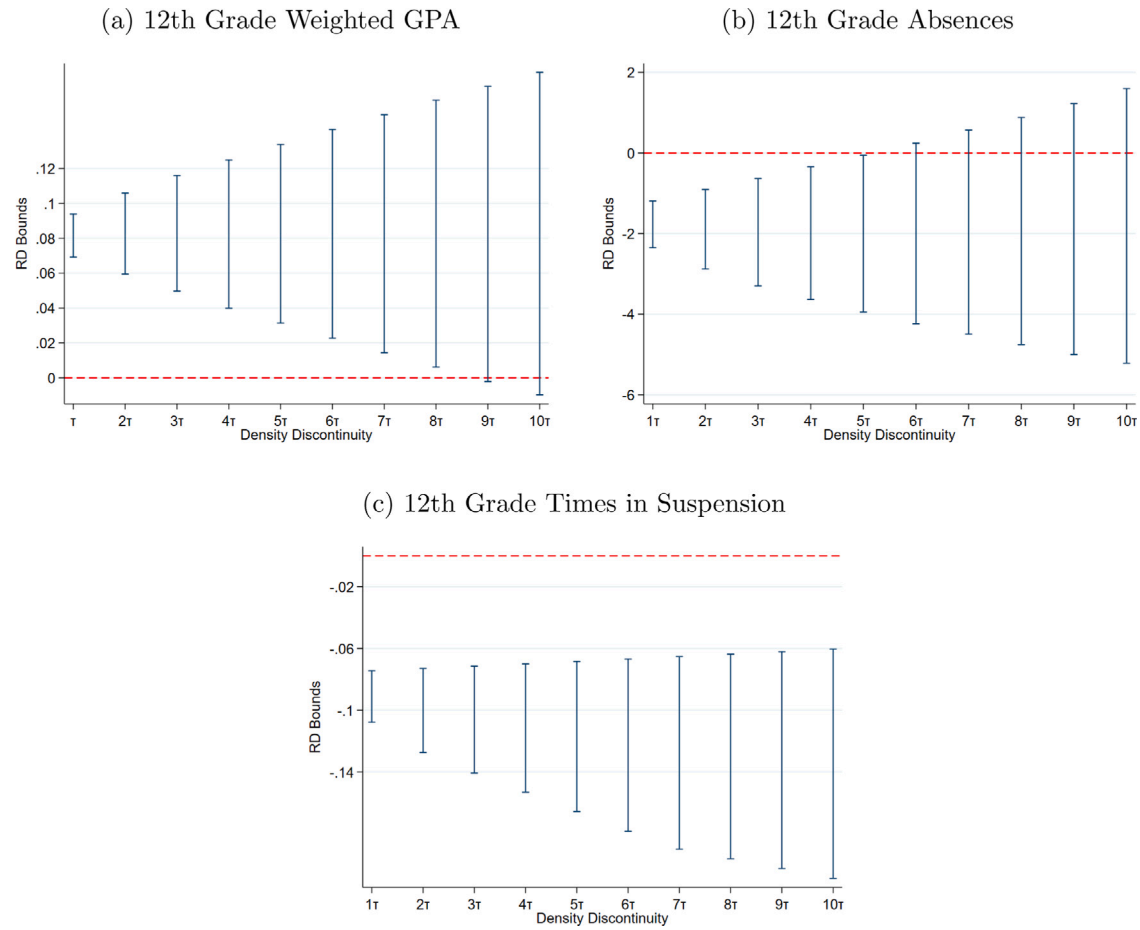

Fig. 8. Sensitivity of bounds to estimated density discontinuities.
Notes: This figure shows discontinuities in senior-year weighted GPA, absences and times in suspension (Panels (a), (b), and (c). All panels are created by plotting the average outcomes across each SAT score bin and fitting the data using a linear regression without controls on either side of the cutoff.

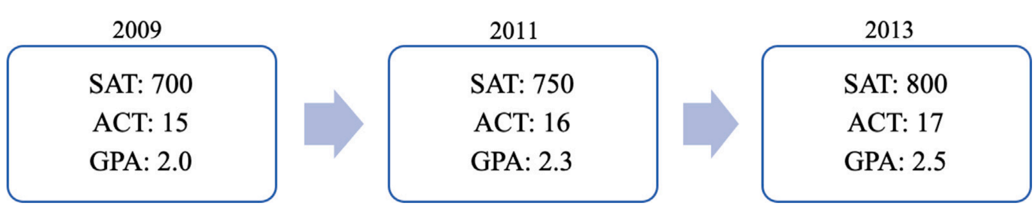

Fig. A.1. UNC system minimum admission requirements.

A.6. Alternative covariate smoothness test

Given the high dimensionality of the set of pre-treatment variables, I run two alternative ways to assess covariate smoothness at the cutoff. First, I summarize the variation of pre-treatment variables into one index. I do so following Anderson (2008) and create two indices. The first uses pre-treatment outcomes (9th and 10th grade GPA, absences, suspensions, and credit-taking) and the second uses pre-treatment covariates and outcomes. I then recover regression discontinuity estimates using these indices as outcome variables.

The use of summary indices has three advantages over testing on individual outcomes (see Anderson (2008) for further details). First, it is more robust to over-testing by reducing the number of carried tests. Second, it provides a statistical test for whether the pre-treatment covariates jump at the cutoff in 'general'. Finally, it is potentially more powerful than individual tests as outcomes that approach marginal significance can aggregate into an index that does attain statistical significance. Columns (1) and (2) of Table A.1 present these regression discontinuity estimates and panels (a) and (b) of Fig. A.6 provide the corresponding scatter plots. Overall, this exercise provides evidence of no discontinuity of pre-treatment variables at the cutoff.

The second way to use covariate variation to assess smoothness at the cutoff is to regress the outcome of interest on pre-treatment covariates and outcomes and partition the outcomes variation into: (a) a part that is explained by pre-existing differences across students (the fitted values), and (b) a part that cannot be explained by pre-existing differences (the residuals). If pre-existing student differences do not drive the results, then the fitted outcome variables should not

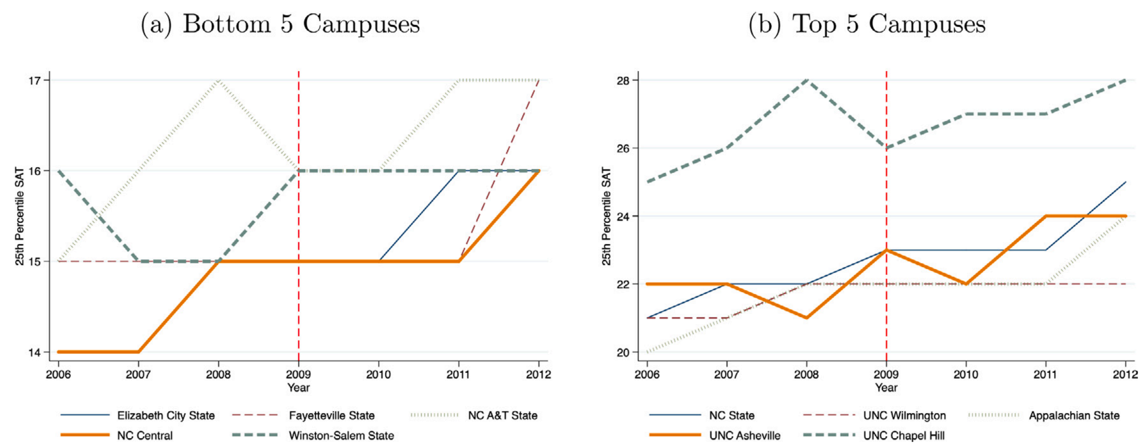

Fig. A.2. 25th percentile ACT among admitted students.
Notes: This figure shows the evolution of the average 25th percentile ACT score for admitted students of UNC campuses over time. Panel (a) shows the evolution for the bottom five campuses. Panel (b) shows the evolution for the top five campuses. The dash red line represent the year in which the UNC Minimum Admission Requirements were implemented. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Source: IPEDS.

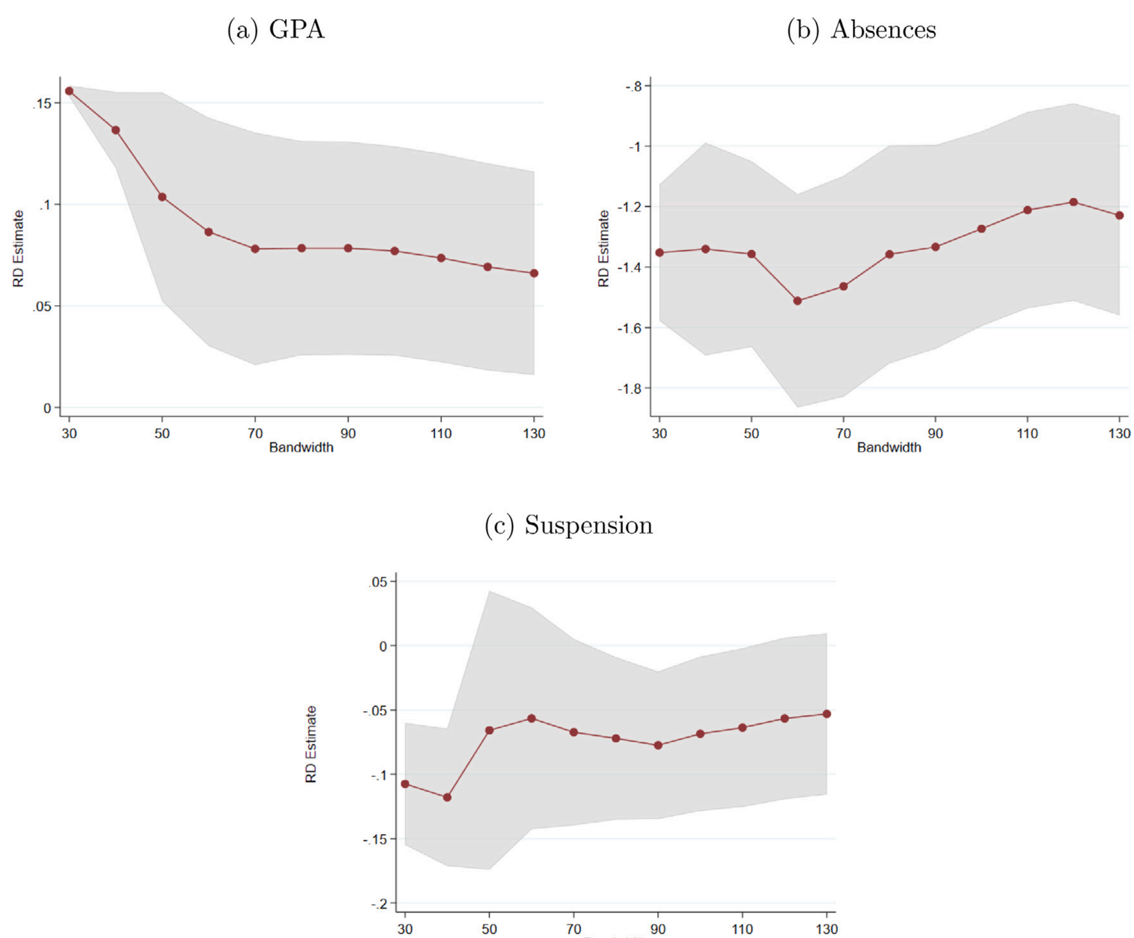

Fig. A.3. Regression discontinuity estimates with different bandwidths.
Notes: This figure shows regression discontinuity estimates for student senior-year outcomes as the bandwidth used for estimation increases from 30 to 120. Panel (a) presents estimates for student GPA, panel (b) presents estimates for student absences, and panel (c) presents estimates for student suspensions.

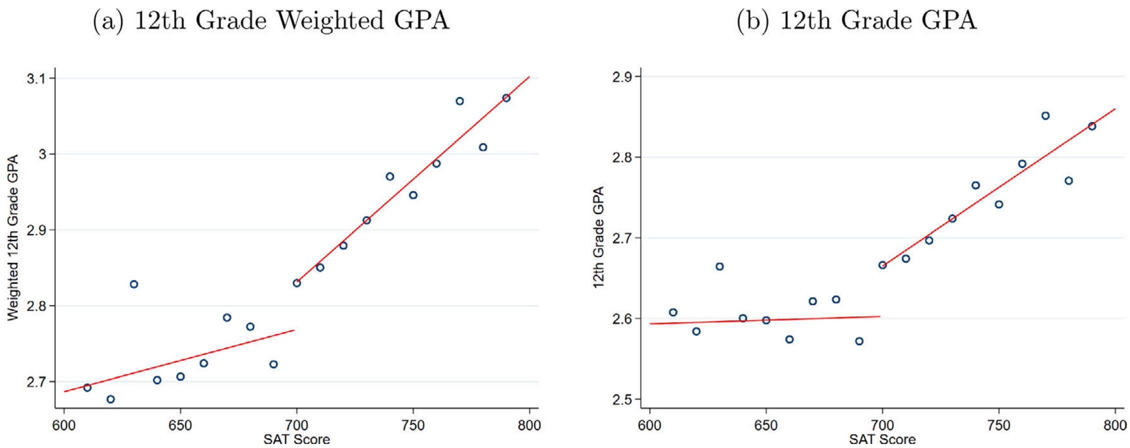

Fig. A.4. Global linear fit.
Notes: This figure shows discontinuities in senior-year weighted and unweighted GPA. All panels are created by plotting the average outcomes across each SAT score bin and fitting the data using a linear regression without controls on either side of the cutoff.

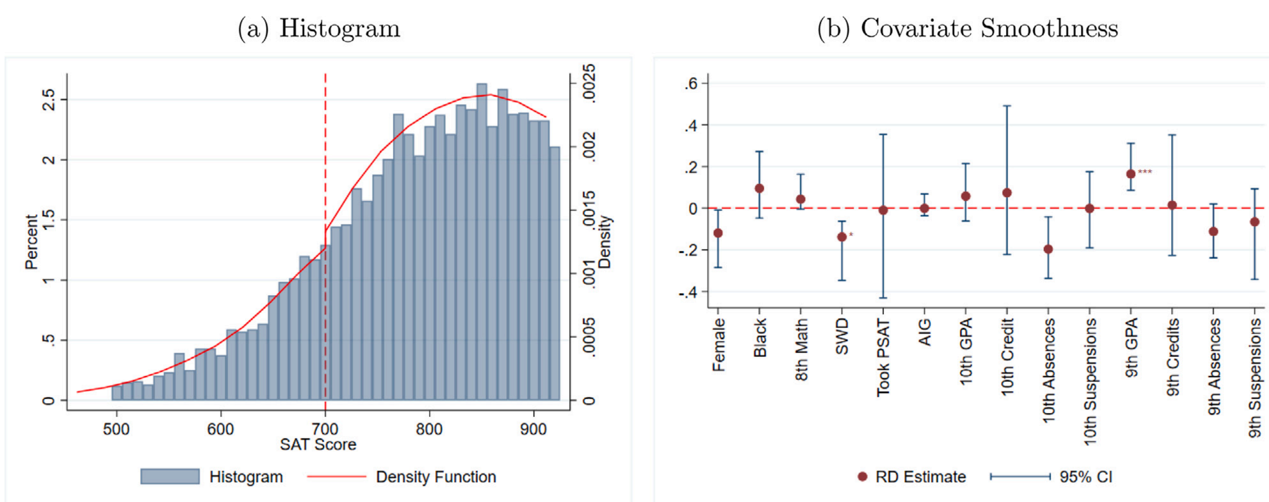

Fig. A.5. Histogram and covariate smoothness for eds students.

Table A.1
Pre-treatment covariate index.

	Index		Fitted outcomes		
	(1) Outcome	(2) All Covs.	(3) GPA	(4) Absences	(5) Suspensions
RD Estimate	−0.004	0.008	0.032	−0.315	0.006
	(0.029)	(0.017)	(0.024)	(0.259)	(0.016)
Bandwidth	94.13	72.82	79.40	65.40	82.34
Observations	3975	3064	3062	2345	3522

Table A.2
Heterogeneity in 9th grade GPA discontinuity.

	Below			Above		
	(1) 9th	(2) 12th	(3) Δ	(4) 9th	(5) 12th	(6) Δ
RD Estimate	0.052	−0.023	−0.074	0.018	0.095*	0.077
	(0.039)	(0.041)	(0.067)	(0.012)	(0.046)	(0.048)
Bandwidth	79.40	79.40	79.40	79.40	79.40	79.40
Observations	1698	1697	1697	1694	1693	1693
Control Mean	2.09	2.56	0.47	3.07	3.05	−0.02

be discontinuous at the cutoff.[16] Columns (3) through (5) in Table A.1 provide regression discontinuity estimates for fitted outcome variables (panels (c) through (e) of Fig. A.6 provide corresponding scatter plots). Again, this exercises provides evidence to suggest that selection into the right side of the cutoff is not the main driver of the results presented in this paper.

Finally, to address concerns around 9th grade GPA discontinuity being the main driver of 12th grade GPA discontinuity, I run a heterogeneity exercise for students above and below the 9th grade GPA median. The idea behind this exercise is to test whether 9th grade GPA discontinuity is driven by students who are likely responding to the change in UNC admission criteria. Presumably, students wishing to pursue a college degree after high school, who also have higher pre-treatment GPAs, are more likely to respond to incentives generated

[16] I thank an anonymous reviewer for suggesting the exercises to check covariate smoothness at the cutoff.

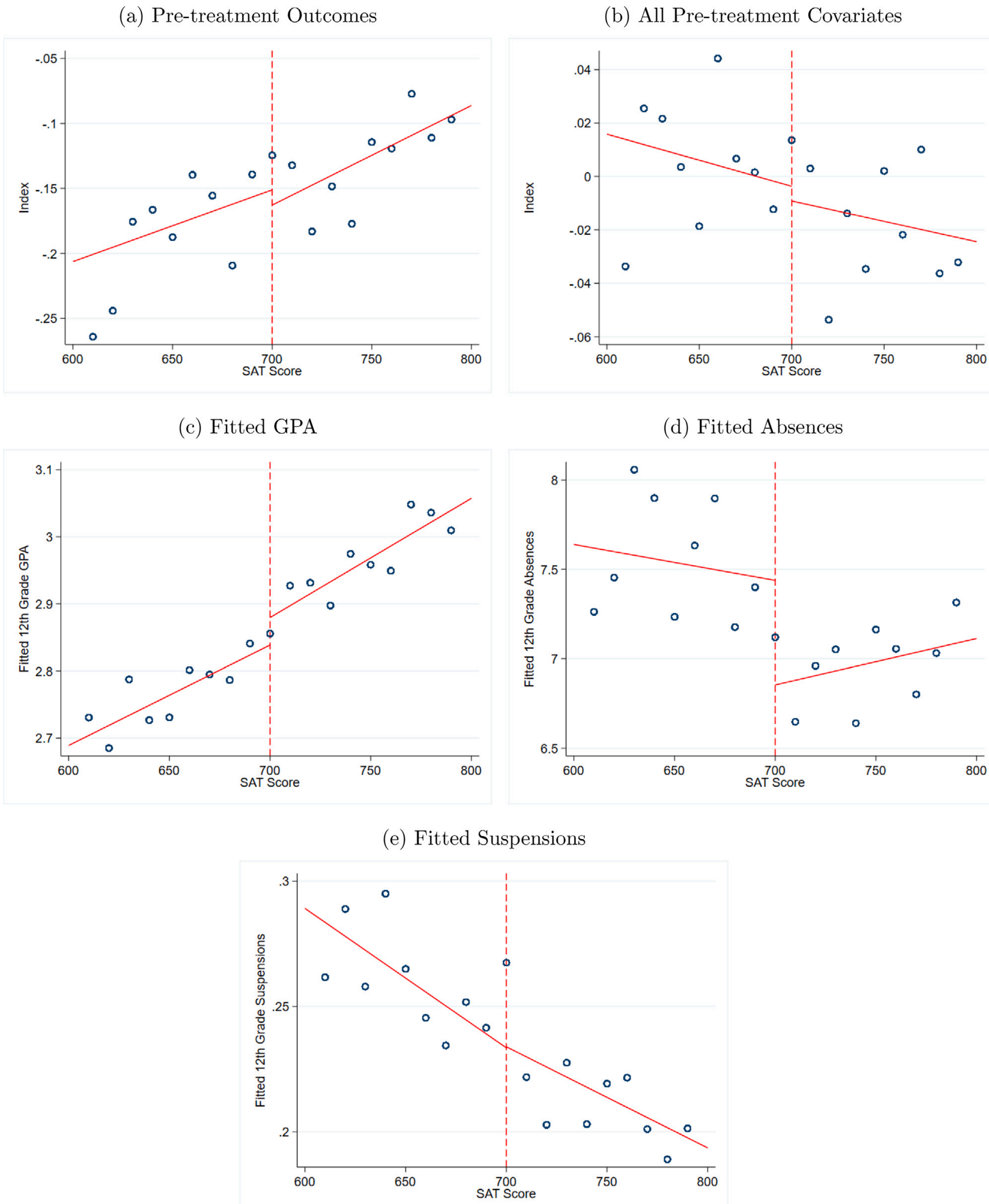

Fig. A.6. Pre-treatment covariate index plots.

by this policy. Table A.2 present regression discontinuity estimates for these two groups of students. Overall, these estimates suggest that the observed discontinuity across the SAT threshold in 9th grade GPA is mainly driven by students below the 9th grade median GPA. Furthermore, these students do not seem to respond to the SAT threshold, as their 12th GPA is comparable across the threshold. Students above the median GPA in 9th grade do see discontinuous increases of 12th grade GPA at the cutoff increases and even in this reduced sample. Finally, even after controlling for 9th grade GPA performance, the estimated magnitude for the jump for students above the 9th grade median GPA is large (0.077), whereas I find an almost opposite effect for those below the 9th grade median GPA (−0.074).

References

Akhtari, M., Bau, N., & Laliberté, J.-W. (2020). *Affirmative action and pre-college human capital*: *Working paper series 27779*, National Bureau of Economic Research.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495.

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *1*(1), 136–163.

Bettinger, E. P. (2010). *Paying to learns: The effect of financial incentives on elementary school test scores*: *Working paper series 16333*, National Bureau of Economic Research.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *6*(82), 2295–2326.

Casey, M. D., Cline, J., Ost, B., & Qureshi, J. A. (2018). Academic probation, student performance, and strategic course-taking. *Economic Inquiry*, *56*(3), 1646–1677.

Cattaneo, M. D., Jansson, M., & Ma, X. (2019). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 1–7.

Cornwell, C., Lee, K. H., & Mustard, D. B. (2003). *The effects of merit-based financial aid on course enrollment, withdrawal and completion in college*: *Working paper 820, Discussion paper series*, IZA Institute of Labor Economics.

Cortes, K. E., & Zhang, L. (2012). *The incentive effects of the top 10% plan*: *Working paper, unpublished manuscript, June*.

Cullen, J. B., Long, M. C., & Reback, R. (2013). Jockeying for position: Strategic high school choice under Texas' top ten percent plan. *Journal of Public Economics*, *97*, 32–48.

Fruehwirth, J. C. (2013). Identifying peer achievement spillovers: Implications for desegregation and the achievement gap. *Quantitative Economics*, *4*(1), 85–124.

Fryer, R. G. (2011). Financial incentives and student achievement: Evidence from randomized trials. *Quarterly Journal of Economics*, *126*(4), 1755–1798.

Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*.

Gerard, F., Rokkanen, M., & Rothe, C. (2020). Bounds on treatment effects in regression discontinuity designs with a manipulated running variable. *Quantitative Economics*, *11*(3), 839–870.

Golightly, E. (2019). *Does college access increase high school effort? Evaluating the impact of the Texas top 10% rule*: *Working paper, unpublished manuscript University of Texas at Austin, August*.

Goodman, J., Gurantz, O., & Smith, J. (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, *12*(2), 115–158.

Goodman, J., Hurwitz, M., & Smith, J. (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics*, *35*(3), 829–867.

Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2012). *The effect of school choice on intrinsic motivation and academic outcomes*: *Working paper series 18324*, National Bureau of Economic Research.

Khanna, G. (2013). Does affirmative action incentivize schooling? Evidence from India. *The Review of Economics and Statistics*, *102*(2), 219–233.

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *2*(142), 665–674.

Leeds, D. M., McFarlin, I., & Daugherty, L. (2017). Does student effort respond to incentives? Evidence from a guaranteed college admissions program. *Research in Higher Education*, *58*(3), 231–243.

Levitt, S. D., List, J. A., Neckermann, S., & Sadoff, S. (2016). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy*, *8*(4), 183–219.

Lindo, J. M., Sanders, N. J., & Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics*, *2*(2), 95–117.

Scott-Clayton, J. (2011). On money and motivation a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources*, *46*(3), 614–646.

Scott-Clayton, J., & Schudde, L. (2019). The consequences of performance standards in need-based aid: Evidence from community colleges. *Journal of Human Resources*, *55*(4), 1105–1136.

Tincani, M. M., Kosse, F., & Miglino, E. (2020). *Student beliefs and the (perverse) incentives of preferential college admissions*: *Working paper, unpublished manuscript, July*.

UNC (2015). *Statistical abstract of higher education in North Carolina 2014–15*: *Technical report*, North Carolina: The University of North Carolina Chapel Hill.