Winners and Losers?
The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes

Sandra E. Black
Columbia University
IZA, NBER, and NHH
sblack@columbia.edu

Jeffrey T. Denning
Brigham Young University
IZA and NBER
jeffdenning@byu.edu

Jesse Rothstein
University of California, Berkeley
IZA and NBER
rothstein@berkeley.edu

June 2021

Abstract

Selective college admissions are fundamentally a question of tradeoffs: Given capacity, admitting one student means rejecting another. Research to date has generally estimated average effects of attending a selective college, and has been unable to distinguish between the effects on students gaining access and on those losing access under alternative admissions policies. We use the introduction of the Top Ten Percent rule and administrative data from the State of Texas to estimate the effect of access to a selective college on student graduation and earnings outcomes. We estimate separate effects on two groups of students. The first--highly ranked students at schools which previously sent few students to the flagship university--gain access due to the policy; the second--students outside the top tier at traditional "feeder" high schools--tend to lose access. We find that students in the first group see increases in college enrollment and graduation with some suggestive evidence of positive earnings gains in some years after college. In contrast, students in the second group attend less selective colleges but do not see declines in overall college enrollment, graduation, or earnings.

The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. This work was partially supported by the Research Council of Norway through its Centres of Excellence Scheme, FAIR project No 262675. We thank Eleanor Golightly, Yonah Meiselman, and Tomas Monarrez for excellent research assistance. We thank Zach Bleemer, Joshua Goodman, Sarena Goodman, and participants in seminars at Brigham Young University, the University of Hawaii, the University of Texas at Austin, Yale, NBER Labor Studies, and the Society for Labor Economists, for their helpful comments. We also thank Rodney Andrews, Scott Imberman, and Mike Lovenheim for sharing their data on the Century/Longhorn Scholars.

1. Introduction

Selective college admissions are fundamentally a question of tradeoffs: Given capacity, admitting one student means rejecting another. Many recent debates, including challenges to the use of affirmative action (e.g., *Fisher v. University of Texas*[1]) or to the consideration of non-academic factors (as in the recent lawsuit over Harvard admissions or in the adoption of "SAT optional" policies) turn explicitly on the fact that admissions rules that benefit one group of students necessarily displace another. Assessing an admissions policy change requires understanding both the effect of attending the selective college on the students admitted under the policy and the effect on the students who are displaced.

There is an extensive literature examining the returns to attending a more selective institution. Several recent studies find significant benefits to students of attending higher quality colleges (Cohodes and Goodman 2014; Hoekstra 2009; Zimmerman 2014; Goodman, Hurwitz, and Smith 2017; Ge, Isaac, and Miller 2018; Bleemer 2020, 2021; Kozakowski 2019). However, these studies generally estimate (local) average effects and so do not answer the policy relevant question of *which* students benefit most from access, or how admissions can be designed so that scarce slots go to the students who will benefit the most.

In this paper, we take advantage of the introduction of the Texas Top Ten Percent rule (hereafter, "TTP")—a policy that guaranteed admission to any Texas public university to anyone in the top ten percent of their high school class in Texas—to identify the effects of access to more selective public universities on student short- and long-run success in a particularly policy-relevant setting. We draw on administrative data covering the entire State of Texas that tracks students from high school through college and into the labor market, allowing us to measure TTP's impact on enrollment, graduation, and post-college labor market outcomes up to 15 years after high school graduation.

Prior to TTP's implementation in 1998, students were admitted separately to each University of Texas campus based on a combination of class rank, test scores, and other factors

[1] 570 U.S. 297 (2013); 758 F. 3d 633 (2016); 579 U.S. ____ (2016), 14-981.

such as the student's personal statement.[2] With TTP, all students whose grades placed them in their school's top decile were guaranteed admission even to the most selective campuses.[3]

We use the introduction of TTP to identify the effect of access to a selective institution on students, many from disadvantaged backgrounds, who were previously unlikely to attend selective institutions. Importantly, and in contrast to other work on college selectivity, we are also able to identify students who lost access to the flagship campus, the University of Texas at Austin (hereafter referred to as "UT Austin" or "UT"), as a result of the policy, and to estimate the impacts on them. Accommodating the new TTP students required tightening admissions standards on other margins, leading some students who would have previously attended a selective Texas university to be denied admission. Students outside the top decile of their high schools faced much stiffer competition; at high schools that previously sent disproportionate shares of students to the flagship, many students who previously would have attended UT Austin were no longer able to do so.

We define identifiable groups of students of each type. The first group, with relatively high performance at schools that had traditionally sent few, if any, students to the University of Texas flagship campus in Austin, became more likely to attend UT Austin.[4] We refer to these students, the nominal target of TTP, as "Pulled In" by the policy. The second group of students, who were ranked outside of the top ten percent at high schools that had previously sent a relatively large share of their students to UT, became less likely to attend. We refer to these students as "Pushed Out" by TTP.

The Pulled In students had higher state test scores, took more AP classes, and had fewer absences than Pushed Out students, but they came from schools with lower average test scores and above average shares of underrepresented minorities and low-income students. In contrast, Pushed Out students came from schools that were more advantaged than average. The Pulled In students were also more racially diverse than either Pushed Out students or the average UT student, though less so than the overall Texas high school population. We view the ability to

[2] Before the entering class of 1997, universities also engaged in race-based affirmative action. This ended with the *Hopwood* court decision in 1996, discussed further below.
[3] Other states, including California and Florida, have since implemented similar policies.
[4] As we discuss below, we measure students' ranking within their high schools primarily based on their test scores. We thus fail to identify students who scored poorly on tests relative to their grades, who might also have benefitted from the TTP admissions rule even if they attended traditional feeder schools. Data limitations prevent us from examining effects on these students.

examine both groups of students as an important contribution; knowing effects on both margins informs policy so that benefits to new enrollees can be weighed against the costs to the displaced students (Kane 1998). Many admissions policy controversies concern efforts to draw in relatively successful students from non-traditional backgrounds, creating tradeoffs like the one we study.

A key challenge for our analysis is that data on class rank were not systematically collected prior to the implementation of TTP.[5] Thus, although the TTP would seem to lend itself naturally to a regression discontinuity strategy (as in several prior studies of college selectivity effects), data limitations make this infeasible. Instead, we use difference-in-differences and event study designs that rely on students unlikely to be admitted to a flagship Texas campus before or after TTP as a control group.

Another consequence of this data limitation is that we cannot precisely identify either the pre-TTP students who would have qualified for TTP admissions had the policy been in place or the post-TTP students who would have been admitted under the pre-TTP statewide standards. We use machine learning methods to implement data-driven strategies for identifying groups of students who are likely to fall into each category. Specifically, we use data on post-TTP students, for whom we observe eligibility for TTP admissions, to train a random forest prediction of top-ten percent status, then use this prediction model to impute the likelihood of being in the top ten percent for students in all years. Using this imputation, we can identify both a group of students who are highly likely to be eligible for TTP admission and a second group of still-high-achieving students for whom this is unlikely.

We then classify high schools by the share of students who they sent to UT Austin prior to TTP. Our "Pulled In" treatment group consists of students with high predicted probabilities of being in their high school's top decile, at high schools where very few students attended UT Austin before the TTP policy was implemented. Our "Pushed Out" group consists of students with high absolute achievement but less impressive relative performance at schools where many non-top-ten percent students attended Austin prior to TTP. These students have low probabilities of qualifying for TTP guarantees and many were crowded out as a result of the policy. We then

[5] For this reason, most past TTP research has eschewed statewide administrative data, instead using data from individual districts where class rank data were available (Daugherty, Martorell, and McFarlin 2014), survey data (Tienda and Niu 2006b), or administrative data from a subset of universities in the state (Kapor 2015).

compare changes in each group's enrollment, graduation, and labor market outcomes following the implementation of the TTP with those of a control group unlikely to have been affected by TTP.

We contribute to the substantial literature on the returns to college quality in higher education. The research to date suggests that college quality has positive effects on student success, although this conclusion is not unanimous. Many of the recent studies use regression discontinuity designs based on admissions or price thresholds. Perhaps most similar to our setting, Bleemer (2021) compares students who just qualify and just miss the threshold for the University of California's "Eligibility in the Local Context" admissions rule (also known as the "four percent plan"), showing that gaining access to more selective campuses benefits students. Cohodes and Goodman (2014) study a Massachusetts merit financial aid program that influenced the enrollment choices of price-sensitive students who were not necessarily near the admissions margin, and also find positive effects of attending a more selective college. Finally, Daugherty, Martorell, and McFarlin (2014) use data from a single urban Texas school district to implement a regression discontinuity research design that compares students with class rank just above and below the TTP cut-off. They find that eligibility for guaranteed admissions increases enrollment at Texas flagship universities as well as the number of semesters enrolled. They also find that the effects are concentrated in schools with high college-sending rates, concluding that these automatic admission plans may have little effect on students in the most disadvantaged schools (see also Cortes and Klasik 2020).[6]

However, the effects of college quality are likely heterogeneous, perhaps different for marginal students than for average students and perhaps even varying across different admission margins. The above studies have limited ability to identify heterogeneous effects, and especially contrasts between effects at different admissions margins. A parallel literature on "mismatch" tests for potentially negative effects of college selectivity on students admitted due to affirmative

[6] Other, similar studies include Hoekstra (2009), Zimmerman (2014), Goodman, Hurwitz, and Smith (2017), Smith, Goodman, and Hurwitz (2019), Kozakowski (2019), and Anelli (2020). Another strategy originates with Dale and Krueger (2002, 2014), who estimate the effect of attending a more selective college based on comparisons of students who made different matriculation decisions within similar choice sets. They find little effect of selectivity on average, though the (small) subsample of black students did appear to benefit. Ge et al. (2018) update Dale and Krueger and, using different sample selection criteria, find benefits for female students. Mountjoy and Hickman (2020) apply this strategy to the Texas ERC data. Similar to Dale and Krueger, they find small effects of attending more selective campuses.

action preferences, with decidedly mixed results.[7] This literature explicitly emphasizes potential heterogeneity: A presumption is that students at the traditional admissions margin – those who would be admitted under race-blind admissions rules – would not suffer from mismatch, though few studies model this heterogeneity explicitly.[8]

Our paper advances these lines of research in several ways. First, our difference-in-differences identification strategy, while relying on a traditional "parallel trends" assumption, allows us to identify effects away from the admissions discontinuity and, in particular, allows us to analyze separately the effects on Pulled In and Pushed Out students. Second, we use data on the population of Texas students rather than a single school district, giving a larger view of the effects of the policy. Third, we follow students beyond college, considering labor market outcomes using administrative earnings records linked to high school and college data.

Consistent with past research, we find evidence that TTP dramatically changed student enrollment patterns (Long, Saenz, and Tienda 2010, Niu and Tienda 2010). Pulled In students became more likely to attend both the flagship UT campus at Austin and the other, less selective four-year campuses as a result of the policy. This was not merely a shift from community colleges. Instead, on net, TTP pulled students into the Texas public higher education system (from not attending college, from private colleges, or from out of state institutions).[9] More distally, we find that TTP increased the share of Pulled In students who graduated with bachelor's degrees within six years after high school. The increases in graduation are similar to what would be expected given average graduation rates at the institutions that students were induced to attend. We also find suggestive evidence that it increased log wages nine to eleven years after high school, though not beyond that point. These are reduced-form effects, so they

[7] See, for example, Rothstein and Yoon (2008); Sander and Taylor (2012); Dillon and Smith (2017); Dillon and Smith (2020); Arcidiacono, Aucejo, and Hotz (2016); and Bleemer (2020). These papers focus on academic mismatch; there is other work that discusses other potential causes for mismatch, such as a type of cultural mismatch where students from disadvantaged backgrounds are less able to successfully navigate the college experience. See, for example, work by Armstrong and Hamilton (2013) and Jack (2016).

[8] Two exceptions are Kapor (2015) and Bleemer (2020). Kapor uses a structural model to estimate the effect of TTP on college enrollment and academic performance, allowing college treatment effects to vary linearly with students' SAT scores. He finds that minority students admitted under TTP achieved higher college GPAs than minority students admitted under a points-based affirmative action policy would have achieved. Bleemer (2020) uses the elimination of affirmative action in California to study effects on Black and Hispanic students, contrasting these to separate estimates for marginal white and Asian applicants.

[9] Similar to our result, Dynarski et al. (2020) find that a clearly communicated policy offering access to a highly selective college, the University of Michigan, draws in students who otherwise would not have attended college at all.

combine effects of increased selectivity with returns to college accruing to those induced to attend college at all by TTP. Our research design does not allow us to isolate the two components. We can establish, however, that that the net effect of TTP on Pulled In students is positive in terms of enrollment and graduation; that there is no indication of a negative effect on earnings; and that the Pulled In students who attend UT Austin as a result of TTP have graduation rates comparable to the average UT Austin student, suggesting that these students were not mismatched.

For Pushed Out students the pattern is different. As expected, TTP reduced Pushed Out students' enrollment at UT Austin. About two-thirds of the displaced students enrolled in less selective public four-year colleges in Texas, while another one-third enrolled in Texas community colleges. The net effect on total enrollment at public colleges and universities in Texas is near zero. Thus, for the Pushed Out group, the policy experiment amounts to a reduction in college selectivity with no change at the extensive margin of enrollment. We find no reduction in Pushed Out students' college graduation probabilities, in part because the colleges that they attended had only slightly lower graduation rates than in the pre-TTP counterfactual and in part because the Pushed Out students previously had below-average graduation rates at UT Austin. We do not find any sign that TTP reduced wages for Pushed Out students. This suggests that the benefits of attending a more selective public institution may be quite small for these students.

Our results pose a puzzle that we cannot fully resolve: Why does gaining access to selective institutions help students whereas losing access does not affect measured student outcomes? We speculate, but cannot prove, that this pattern reflects differences between the two groups of students. Pushed Out students are likely to come from families with more support for college success, so may be less dependent on inputs received from the college itself. This is consistent with evidence, from a variety of settings, that disadvantaged students are more sensitive to school inputs or characteristics (see, e.g., Krueger and Whitmore 2001; Dale and Krueger 2014; and Deming et al. 2014). Another interpretation consistent with our results and the results from the broader literature is that there are positive returns to university attendance (Zimmerman 2014; Goodman, Hurwitz, Smith 2017; Mountjoy and Hickman 2020), which increased in our Pulled In group but did not fall in our Pushed Out group, but that returns to college selectivity within the four year sector are smaller (Dale and Krueger 2002, 2014).[10]

[10] A notable exception is Hoekstra (2009).

Regardless, taken together our results suggest that access to UT Austin improved outcomes for students who would not have attended absent the TTP and did not substantively damage graduation rates or earnings for students who were displaced. Contrary to claims that expanding access to students disadvantaged by more standard admissions policies will induce mismatch, the TTP experience indicates that, if anything, this would improve student outcomes on average.

The paper unfolds as follows. Section 2 provides institutional details on the Texas higher education system and the Top Ten Percent plan. Section 3 describes the data and Section 4 details our empirical strategy. Section 5 presents our results and Section 6 presents a variety of tests to verify the robustness of our results. Section 7 then provides a discussion and concludes.

2. Background and Institutional Detail

Texas has a large public higher education system, with over 30 four-year universities and over 60 two-year colleges. The flagship of the University of Texas System, and its most selective university, is the University of Texas at Austin (known as "UT"); Texas A&M University is also nationally prominent though less selective. Most students who attend college attend public universities, with only 18.8 percent of Texas students (vs. 26.2 percent nationally) attending private colleges, and only 9.8 percent attending out-of-state colleges.[11]

Prior to the entering class of 1997, admission to UT was based primarily on continuous high school class rank and SAT or ACT scores, with affirmative action preferences for students from underrepresented groups. In 1996, the 5th Circuit ruled in *Hopwood v. Texas*[12] that the consideration of race in admissions at UT, and by extension throughout the state, was impermissible. Following *Hopwood*, the UT entering class of 1997 was admitted based on academic preparation, distilled in an "Academic Index" (AI), and a "Personal Achievement Index" (PAI), assigned by readers based on essays, leadership, extracurricular activities, and special circumstances such as family structure, among other factors.

In May 1997, the Texas legislature passed HB 588, creating the TTP. This guaranteed admission to any public university in Texas for students in the top decile of their high school

[11] Authors' calculations from the Integrated Postsecondary Education Data System (IPEDS), distributed by the National Center for Education Statistics.
[12] 78 F.3d 932 (5th Cir. 1996).

class, as calculated by the high school and measured at the end of the junior year.[13] This was primarily binding for students applying to UT; at other campuses, nearly any top decile student would have been admitted even under the pre-TTP rule. For students outside of the top decile, the AI and PAI were used to admit students, though given the large number of TTP students, UT admissions were quite competitive for non-TTP applicants.

Effectively, UT shifted from a single admissions rule that was based to a large extent on a weighted average of SAT scores and high school performance, with preferences for minority students, to a regime where it used two separate rules: One that used only the within-high-school component of high school grades and a second that maintained the old SAT and rank weighting, albeit without racial preferences, but raised the bar considerably.

The TTP was an attempt to maintain diversity among admitted students without explicitly considering race in admission decisions by taking advantage of the substantial racial and economic segregation across Texas public schools (Tienda and Niu 2006a). Admissions from a statewide pool (based, for example, on SAT scores) disproportionately draw students from high-income, primarily white high schools. But if students in the top ten percent of their high school classes are roughly representative of the high schools as a whole, a university admissions pool comprised of them will be roughly representative of the statewide student population.[14]

TTP was not just a change in admissions, but a change in students' perceived admissions chances. The ten percent threshold was more transparent than the earlier rules, so students could assess their prospects more easily. Moreover, the TTP law mandated that every high school post a sign explaining the law and that a letter be sent to every parent of a qualifying student, and the policy change was widely covered in media. Flagship institutions, concerned about declines in minority enrollment following *Hopwood*, also increased efforts to recruit traditionally

[13] University of Texas at Austin Office of Admissions (2008) indicates that UT Austin implemented a top-10 percent admissions rule through an institutional policy for the class entering in 1997, prior to HB 588, though this was not widely known. We treat the 1997 high school class as prior to TTP. As we discuss below, TTP affected enrollments in large part through its impact on student application decisions – students guaranteed admission were more likely to apply than when admission was merely highly likely – and this mechanism did not operate in 1997, when the new rule was not known to be in place. Insofar as a TTP-like admissions rule was in use in 1997, this should attenuate our estimates of the impact of TTP on all dimensions.

[14] In addition to the shift from a statewide competition to a series of school-by-school competitions, TTP also shifted emphasis from SAT scores to high school grades. Minority and low-income students typically do better on the latter than the former metric (e.g., Rothstein 2004). Of course, students at the top of their classes are not likely to be fully representative. See, e.g., Harris and Tienda (2010); Long and Tienda (2008); Long (2004); Niu, Tienda, and Cortes (2006); Long, Saenz, and Tienda (2010).

underrepresented students, including opening outreach centers, visiting high schools that were outside of traditional feeder-school networks, and reorienting scholarships to target non-feeder high schools rather than minority students.[15]

TTP transformed enrollment at the UT flagship, by far the most selective campus. In 1998, 41 percent of freshmen from Texas high schools were admitted under the TTP. By 2003, this share had reached 70 percent (University of Texas at Austin Office of Admissions 2007), though many of these students also ranked highly statewide and would have been admitted in any case under the pre-TTP policy.[16] While the admissions rule did not change during this period, the TTP share grew due to changes in application patterns of students who gradually came to understand that they were now guaranteed admission.[17] As we show below, post-TTP admissions were notably less concentrated at traditional feeder schools than under the pre-TTP regime.

3. Data

We use linked individual-level secondary school, higher education, and workforce administrative data from the Texas Education Research Center (ERC). The data include all students in public secondary schools in Texas and cover enrollment, courses taken, absences, and standardized test scores, in addition to demographic information such as race, gender, and free and reduced lunch status. At the post-secondary level, the data contain enrollment, major, and graduation information for the population of students from all public universities and colleges in Texas. These data are matched to quarterly earnings records from the Texas Unemployment Insurance (UI) system through 2017.

[15] Andrews, Imberman, and Lovenheim (2020) evaluate the Longhorn Opportunity Scholars and Century Scholars programs at UT Austin and Texas A&M respectively. These were targeted outreach programs at low-income high schools.

[16] In 2011, SB 175 allowed UT to limit the number of automatically admitted students to 75 percent of the incoming class by setting the threshold higher than the top ten percent. This was after the period covered by our data, which end with the high school class of 2002.

[17] Cullen et al. (2013) find that some students switched to less competitive high schools in order to qualify for TTP admissions. This group was quantitatively very small – Cullen et al. (2013) estimate only 211 students statewide. As we describe later, we test the sensitivity of our results to this response by limiting our sample to the first two post-TTP cohorts, for which the policy could not have induced mobility because students who arrived at a school after 10th grade were not guaranteed TTP admissions even if they ranked in the top 10 percent of the new school. Golightly (2019) finds that TTP raised high school attendance and graduation rates throughout the ability distribution, suggesting the information aspect of the program may have played an important role.

Our universe consists of students who graduated from Texas public high schools between 1996 and 2002 for whom we have 10th grade standardized test scores.[18] We refer to students by the year that they graduated high school. Thus, the first treated year is the 1998 students. A key limitation is that the data do not include class rank or high school GPA. Beginning with the graduating class of 1999, however, we observe for each student who applied to any Texas public higher education institution a single indicator of whether the student was TTP-eligible (i.e., in the top ten percent of his/her high school class). As we discuss below, we use this measure to impute a probability of being in the top ten percent for every student, before and after TTP, regardless of whether he or she applied to college. Our imputation uses information about students' positions within their high schools' test score distributions, as well as course-taking patterns and absences.

We consider several outcome measures. First, we examine college enrollment in the year after high school graduation. We distinguish several classes of institutions: community colleges, four-year campuses, and the two most selective campuses, Texas A&M and UT Austin. Impacts on enrollment serve as a sort of "first stage" for our analysis, though for reasons discussed below we do not compute two-stage-least-squares estimates of the effects of college selectivity. Second, we examine college completion, measured as whether an individual graduated with a bachelor's degree from a Texas public institution within six years of high school graduation, and attainment of a bachelor's degree in a scientific (STEM) field. In general, bachelor's degrees may be granted by a different institution than the one where the student initially enrolled, though we also separately examine UT graduation rates for initial UT enrollees. Lastly, we examine labor market outcomes, looking at earnings nine to fifteen calendar years after high school graduation. Because students may be absent from the earnings records either because they are not working or because they are working but not in Texas, in some analyses we average earnings only over the years where non-zero earnings are reported, excluding years with no observed earnings. We also

[18] We use the Texas high school exit exam, TAAS, administered in 10th grade. Students who were absent on the day of the test or who failed on the first attempt could re-take the exam at later dates, as it was offered several times each year. We use each student's first recorded score. We limit our sample to high school graduates. Golightly (2019) finds that TTP affected high school graduation. She finds, however, that this occurs throughout the top 80 percent of students and is not driven by students at the top of the test score distribution. In order for the graduation response to affect our results, it would need to be differential for the directly treated relative to those just below TTP eligibility.

analyze an indicator for ever appearing in the earnings data, which captures both long-term non-employment and absence from the state.[19]

Columns 1 and 2 of Table 1 present summary statistics for our full sample of Texas high school graduates. In the pre-TTP 1996 and 1997 cohorts, shown in column 1, 55 percent of students enrolled in college in the year after high school graduation and 18 percent graduated with a bachelor's degree within six years. The sample is 28 percent Hispanic, 12 percent Black, and 52 percent female. Three quarters of students had positive earnings in Texas at some point in the ninth, tenth, and eleventh years after high school graduation, with average annual earnings of $35,487 in the years that they worked or $25,816 across all three years.[20] Thirteen to fifteen years after high school graduation, the average annual earnings were $43,567 in the years they worked or $29,668 across all three years. These statistics are quite similar in the post-TPP data (column 2), covering the 1998-2002 high school cohorts.

Texas is highly segregated based on race and socioeconomic status. The average student in our sample in the pre-TTP 1996 and 1997 cohorts attends a school that is 12 percent Black and 30 percent Hispanic, but the average Black student attends a school that is 39 percent Black and the average Hispanic student attends a school that is 62 percent Hispanic. Similarly, while 23 percent of students receive free or reduced-price lunches, the average free lunch recipient attends a school where this rate is 39 percent. This segregation is closely related to UT Austin attendance: Prior to TTP, the average free or reduced-price lunch student attended a high school where 1.9 percent of students attended UT Austin, while the average non-subsidized-lunch student attended a school where the fraction attending UT Austin was almost double this, at 3.7 percent.

Insofar as the top decile of each high school is demographically representative of the school as a whole, in the presence of segregation across high schools, admission by high school would yield a much more diverse class than statewide admissions. However, schools are

[19]UI records cover employers who pay at least $1500 in gross wages to employees or have at least one employee during twenty different weeks in a calendar year. We winsorize earnings at the 99th percentile. All earnings are in 2012 dollars. A key concern is that we only observe individuals who work in Texas. We discuss this at length below. Andrews, Li, and Lovenheim (2016) show that students who leave the state do not have substantially different wages than students who do not.

[20] We classify earnings as observed if a student has at least one quarter of non-zero earnings during this three-year period, indicating that he or she was likely living in Texas at that time. Our first average earnings measure excludes years with no observed earnings and is missing for those never observed with earnings. Our second measure assigns zeros to any year in which no earnings are observed, so zero average earnings to those never observed with earnings.

internally stratified as well as segregated, with top ranked students whiter and richer than their peers. The impact of TTP on racial diversity has thus been controversial from the start, and the evidence since has been mixed. It appears that there was an initial small dip in minority students' enrollment after the policy was implemented, but that minority enrollment increased subsequently (Tienda and Sullivan 2009; Tienda et al. 2003). More recently, Kapor (2015) finds that TTP did increase minority representation at flagships. Columns 3-4 of Table 1 show that the demographics of UT Austin enrollees did not change dramatically in the early years of TTP, though the post-TTP enrollees are more highly ranked within their high schools' test score distributions.

4. Empirical Strategy

We use a difference-in-differences (DD) strategy for identifying the effect of TTP on students' outcomes. We compare changes in outcomes following the implementation of TTP for students affected by the policy to those for students who were not directly affected.

As mentioned above, a key feature of our study is that we can distinguish between the effects of TTP on two different groups of students for whom we expect the impact could be quite different. The first group includes students with high class ranks from schools that have traditionally sent few students to the flagship Austin campus—these students were "pulled in" to selective campuses by the policy. The second group of students includes students with lower class ranks and test scores from traditional feeder schools—these students were "pushed out." We estimate separate difference-in-differences coefficients for Pulled In and Pushed Out students, comparing each to a control group of students who were above average in achievement but were unlikely to be admitted to UT Austin under either the pre-TTP or the post-TTP regime.

Our key identifying assumption is that outcomes for these three groups of students would have evolved similarly between the 1996 and 2002 cohorts had admissions policies been held stable. While we cannot test this directly, we present suggestive evidence in support of this assumption below. Specifically, we see changes in relative outcomes that coincide with the introduction of TPP, with few changes in the preceding or subsequent years.[21]

[21] Our estimates of TTP effects would be biased if there were a sudden change in 1998 in some other determinant of outcomes that differentially affected the three groups. We are aware of no such change. As noted above, a few years after the TTP plan the flagship institutions implemented programs—the Century Scholars and Longhorn Scholars--

The major challenge in implementing our research design is identifying which students were pulled in and pushed out by the policy. The policy's effect on a student is a function of the student's class rank, but this is not measured in the statewide pre-TTP data and only limited information is available after TTP. This makes it difficult to identify students' counterfactual admissions outcomes under alternative policies.

We address this issue by estimating each student's likelihood of being in the top ten percent given her other observables, most notably her rank within her high school's test score distribution, and by exploiting the fact that high schools differed substantially in the likelihood that their top students attended UT Austin in the pre-TTP period. We identify Pulled In students as those with high probabilities of being at the top of their classes at schools with low pre-TTP sending rates. As we show below, less than four percent of these students attended UT prior to TTP. Similarly, we identify Pushed Out students as those with strong relative performance, but not strong enough to be likely to be in the top decile, at "feeder" schools with high pre-TTP sending rates. Prior to TTP, 17 percent of these students attended UT. This is nearly quadruple the rate for the Pulled In students, despite the latter group having higher academic achievement. Following TTP, the Pulled In group's UT enrollment rate more than doubled, while the Pushed Out group's rate fell by one-quarter.

Figure 1 illustrates the general strategy schematically. We array students on two dimensions, by class rank, r, and a measure of the fraction of students from a particular high school who were attending UT Austin before TTP, s. This latter dimension captures several sources of variation, including socioeconomic composition, distance to Austin, and academic strength. We interpret it as a propensity for a relatively high-achieving student to apply to, be admitted to, and choose to attend UT Austin. Prior to TTP, enrollment can be approximated as depending on the sum of (appropriately scaled versions of) r and s: At schools with low s, no one attended UT Austin; at schools with moderate s, only the highest ranked students did; and at schools with the highest s, Austin dipped deeper into the pool. This admissions rule is represented by the grey triangle at the upper right of the first panel of Figure 1.

After TTP was implemented, any student in the top decile of his or her high school class was guaranteed admission. This had little effect on top decile students at the highest-s schools,

targeting students at selected high schools (Andrews, Imberman, and Lovenheim 2020). As we show later, our results are robust to the exclusion of these schools.

where any such student who wanted to attend was already likely to do so. However, at lower *s* schools it induced a large number of new students to apply and/or be admitted. The new enrollees are represented by the blue bar at the top of the second panel of Figure 1.

Accommodating these students required making space available by reducing admissions offers to students outside the top decile, who necessarily came from high-*s* schools. This is indicated by the smaller grey area in Figure 1; the students who would have attended but were crowded out are indicated by the orange region. In our analyses, the Pulled In students are represented by the blue region, while the Pushed Out students are represented by the orange area. We compare each to a control group of students with high *r* and *s*, but not high enough to attend UT Austin under either regime. These are indicated by the cloud in Figure 1.[22]

Figure 1 is purely schematic. Because we do not observe class rank (*r*), we cannot implement this design exactly. However, we can construct a proxy for *r* that allows us to implement a version of it. We develop our methodology in three steps. First, we describe our calculation of the likelihood that each student is in the top ten percent of his or her high school class. Second, we describe how we use that imputed probability along with *s* to define the three comparison groups. Finally, we present the difference-in-differences estimator, which can be interpreted as an intention-to-treat estimator that captures the reduced-form effect of the policy on access across the range of Texas institutions.[23]

Step One:

Our first step is to construct a measure of the likelihood of qualifying for TTP admissions that is defined consistently throughout our sample. Let $\boldsymbol{X}$ be a vector of student characteristics, let W be an indicator for a student from the 1999-2002 cohorts who applied to at least one public

[22] Note that we are implicitly assuming that the control group is untreated, but there is some chance that they experienced spillover effects from the policy change. As one way to examine whether there is evidence of spillovers to the control group, we have plotted the means of a variety of characteristics to see if there were meaningful changes in the control group in response to TTP. While we do see that the control group had small declines in enrollment in higher education, we also see slight declines in levels for Pushed Out students. Control and Pushed Out students were trending similarly prior to the policy and follow very closely after the policy. As another check, we take advantage of the fact that the degree of the spillover is likely to be smallest early on, when the number of affected students is smallest. When we focus on the first two affected cohorts after TTP, we obtain similar results, suggesting that spillover effects are unlikely to be large.

[23] Bleemer (2021) discusses the interpretation of reduced-form estimates like ours in terms of the effects of specific institutions.

college in Texas, and let $T=\mathbf{1}(r>0.9)$ be an indicator for being in the top ten percent of the high school class. We observe T only for the $W=1$ subsample.

We include in $\boldsymbol{X}$ observable characteristics that are measured consistently throughout our sample period. These include TAAS exit exam scores in reading, writing, and math, measured both in statewide percentiles and as the percentile within the school; 16 indicators for math and science course-taking (e.g., advanced math in 11th grade); the number of foreign language courses taken in high school; the number of courses failed in high school; the number and percentage of school days absent in 12th grade; an indicator for being 18 upon graduation; the school's racial, gender, and socioeconomic (free & reduced-price lunch) composition; and the share of students at the school who are classified as special education. A complete list of variables is reported in Appendix Table 1, along with summary statistics.

To impute the probability of being in the top ten percent of an individual's high school class for the full sample, we assume that the relationship between $\boldsymbol{X}$ and T does not vary with W: $E[T \mid \boldsymbol{X}, W=1] = E[T \mid \boldsymbol{X}]$.[24] Our challenge is thus to estimate $p = E[T \mid \boldsymbol{X}, W=1]$ as flexibly as possible. A straightforward approach is to fit a simple logit. Coefficients and average marginal effects are reported in Appendix Table 1. Not surprisingly, the school-level percentiles are by far the strongest predictors among the test score variables, though course-taking, school composition, and course failures are also strongly predictive.

One can substantially increase the predictive accuracy by adding non-linear and interactive terms to the model; with nearly 200,000 observations, we can easily afford to do this. There is ambiguity in the correct interaction terms to include so we have used a parsimonious specification. Through cross-validation exercises, we have found that it is easy to over-fit the data when adding these terms, even in our large sample. Over-fitting would make our predicted top-ten percent probabilities better proxies for actual top-ten percent status in the post-TTP

[24] W reflects not just cohort but also application decisions; it equals 1 only for those who applied to a Texas public institution in 1999 and thereafter. Differences between applicants and non-applicants in the relationship between covariates and top ten status would violate this assumption. The findings in Long and Tienda (2010) are broadly consistent with this assumption. They find small changes in applicant characteristics at UT Austin and six other colleges after the implementation of the TTP policy. (Note that they examine data for only a subset of colleges in Texas, and do not examine changes in the overall applicant pool even at those colleges.) We exclude non-applicants from the post-policy cohorts when estimating the probability of top-ten status, effectively treating their T as missing at random. We show below that we obtain nearly identical results when we impute $T = 0$ to these students and include them in the estimation.

training sample than in the pre-TTP data, potentially generating spurious changes in enrollment rates and outcomes.

To accommodate nonlinearities while avoiding over-fitting, we use a random forest model to predict top ten percent status. The random forest builds on repeated decision trees, which allow for arbitrary non-linearities and interactions, but reduces over-fitting by averaging across many trees, each generated probabilistically by considering branches based on random subsets of the predictors and of the data.[25] There is no compact description of a random forest model analogous to the coefficients of a logit regression. In Appendix Table 1, we summarize the model by regressing the predicted values from the random forest on the predictor variables in an OLS regression. The resulting coefficients approximate the highly nonlinear prediction function in much the same way as do average marginal effects for logit models. They are generally similar to the logit marginal effects, though some variables seem to play larger or smaller roles in the random forest model. We use the random forest predictions in our main analysis, though our results are similar when we use the logit model instead.[26]

Having fit the prediction model, we generate for each student in both the pre- and post-TTP subsamples a predicted probability of being in the top ten percent of their high school. We label this $\hat{p}$ to emphasize both that it is a probability, not an estimate of the continuous rank, and that it is estimated with some error. Figure 2 shows the distribution of $\hat{p}$ in the pre- and post-periods. Unlike a logit, a random forest model can generate predicted probabilities of exactly zero. Indeed, we find that fully 23 percent of students have zero predicted probability of being in the top ten percent of their high school. The second panel of Figure 2 shows the distribution among those with $\hat{p} > 0.1$. We see only small differences in the distribution of $\hat{p}$ across years, with about 1 percent more students having high $\hat{p}$ values in the post-TTP period than prior to TTP.

Our strategy relies on $\hat{p}$ being an equally accurate prediction of the student's true class rank in each year. Threats to this assumption include the possibilities that the measurement of

[25] See James et al. (2013) for a thorough discussion of random forests. We allow trees to branch until each leaf has 50 observations and take the average prediction from 1000 trees. Random forests can still over-fit; we show, however, that predictions are very highly correlated across forests fit to different subsets of the data, and discuss other validation strategies below.

[26] One way to quantify the improvement in predictive power is to note the correlation between the $\hat{p}$ and actual top 10 percent status. We find that our logit predictions are correlated at .59 whereas our random forest predictions are correlated at .72. Another, related summary is the mean squared prediction error, which is 30% lower in the random forest model than in a logit (0.076 vs. 0.107).

some of our predictors may have changed over time, that the relationship between class rank and other characteristics changed, or that our model is over-fit to the years that we use to estimate it. Another issue is selection into application to college, as we use only applicants to train our prediction model.

We take several approaches to assess these threats. First, we show that our estimates of $\hat{p}$ are highly robust (Appendix Table 2). Our logit and random forest models generate $\hat{p}$s that are correlated 0.94 with each other,[27] and varying the smoothing in the random forest model (e.g., by allowing larger or smaller final "leaves") does not generate meaningful differences in $\hat{p}$. Second, we show that our estimates are invariant to the specific data used to fit the models in Appendix Table 3. A random forest model fit to 1999-2000 data is correlated 0.96 with a model fit to 2001-2002 data, strongly suggesting that there are no meaningful changes in the measurement of $\boldsymbol{X}$ or its relationship with class rank over this period. As we discuss below, the treatment effects of TTP on Pulled In and Pushed Out students are not sensitive to the specific $\hat{p}$ measure used.

To address threats coming from selection into application we first note that the relevant issue is selection into application to *any* public four-year college or university in Texas rather than selection into applying to UT Austin. The assumption that this is ignorable seems relatively mild, because the policy did not affect students' admissibility to the less selective campuses. In any event, when we re-estimate our prediction model including non-applicants and imputing T=0 for them (and thereby modeling P(T & W) rather than P(T | W), we obtain nearly identical predictions, correlated 0.98 with our preferred estimates. This suggests that selection into application is not a major factor for our index.

It is convenient for the rest of our analysis to discretize the $\hat{p}$ distribution. We create categories that each correspond to two percent of our sample, combining the 23 percent with $\hat{p} = 0$ into a single category.[28] We let $\hat{q}$ represent the category number, labeling the first bin 1 and the remaining bins 12 through 50. Figure 3 shows $E[\hat{p} \mid \hat{q}]$. Because most individuals have low predicted probabilities of being in the top ten percent, $E[\hat{p} \mid \hat{q}]$ is low for most $\hat{q}$. However, starting around the 44th or 45th group (the 88th or 90th percentile of $\hat{p}$), we see that the predicted probability increases dramatically. In the very top group ($\hat{q}$=50), the probability of being in the

[27] The standard deviation of $\hat{p}$ is 0.20 using the random forest model and 0.19 with the logit model.

[28] Our $\hat{p}$ is an average across 1,000 trees, each of which predicts that each individual will either be in the top ten percent or not. It therefore has precision 0.001. There are smaller mass points at $\hat{p} = 0.001$, 0.002, 0.003, and 0.004, each of which leads us to skip values in the $\hat{q}$ sequence.

top decile of the class is 85 percent. The curves are essentially identical in the pre- and post-TTP periods.

Step Two:

The second dimension of Figure 1 is the high school's propensity to send students to the selective university. We measure this as the share of students from each high school in the 1996 and 1997 (pre-TTP) cohorts who enroll at UT Austin. We divide high schools into deciles, denoted by *s*. Figure 4 shows the share of students from schools in each decile who attend UT Austin, in both the pre-TTP and post-TTP cohorts.[29] Schools are highly skewed prior to TTP: The top decile of schools sends 13 percent of students to UT Austin, while the bottom five deciles each send less than two percent of students. In the post-TTP data, this skew is still evident but a bit reduced: The share of students from the top-decile schools who attend Austin falls to 11 percent, while the schools that previously sent few students to Austin send slightly more.[30]

We use our estimated top-ten-percent probability categories $\hat{q}$ and school sending rate deciles *s* to define the Pulled In, Pushed Out, and control groups. Figure 5 shows the empirical analog to the schematic in Figure 1. We array students by school type *s,* on the horizontal axis, and rank group $\hat{q}$, on the vertical axis

Panel A shows the change in UT shares between 1996-7 and 1998-2002. The expected pattern emerges: For high $\hat{q}$*,* low *s* students, who are likely to have benefitted from the TTP rule, the share attending UT Austin rose substantially, while for students with high *s* and moderate $\hat{q}$ – who are unlikely to be in the top decile, as indicated by Figure 3 – the share fell. For example, consider two cells, first the $\hat{q}$=50, s=2 cell, at the upper left, and then the $\hat{q}$=21, s=10 cell, in the middle of the right column. The first cell represents high-ranking students from non-traditional schools that TTP was meant to identify. 3.2 percent of students from this cell attended Austin prior to TTP, but after TTP over 10 percent did, exactly as intended. By contrast, the second cell contains students who are very unlikely to be at the top of their classes, from schools that sent many students to Austin prior to TTP. In this cell, the pre-TTP share attending UT Austin was

[29] Appendix Table 4 shows average student characteristics by *s* group. In Figure 4, the *s*=1 and *s*=2 groups are suppressed by ERC's disclosure avoidance rules because the share attending UT in the pre period is so low.
[30] We have examined contrasts across other periods to verify that this is not simply reversion to the mean.

around 8 percent, but it fell to 5.5 percent following TTP. Overall, we interpret this graph as showing that our prediction of *T* captures the probability of being in the top ten percent and the corresponding access to UT Austin fairly well.

We use a simple rule for defining three groups—Pulled In, Pushed Out, and control--though we explore alternative methods as well. Any cell where the share attending UT Austin rose (respectively, fell) by more than 3 percentage points is included in the Pulled In (Pushed Out) group. In order to avoid giving too much weight to idiosyncratic and/or sampling variation, cells with changes larger than 1.75 percentage points in absolute magnitude that are bracketed (above and below, or on the left and right) by included cells are also included. We add an ad hoc rule to exclude the $\hat{q}$=50, s≥9 cells from the Pulled In group, as they would otherwise be included but appear distinct from the rest of this group.[31] For the control group, we use two blocks of cells that are close to the treated groups but show small changes in the probability of enrollment at UT Austin: $\hat{q}$ between 25 and 40 and *s* in the 6th through 8th deciles, and $\hat{q}$ between 40 and 45 and *s* in the 3rd through 5th deciles, except where they are otherwise included in our treatment groups. The groups are marked in Panel B of Figure 5. Note that, despite our interpolation, the Pushed Out group is far from contiguous.

Table 2 displays summary statistics for each group, in both the pre- and post-TTP subsamples. Columns 1-3 show the Pulled In students, 4-6 the Pushed Out students, and 7-9 the control group. It is notable that the racial minority share of Pulled In students is lower than for high school graduates as a whole, but higher than among all UT students. Pulled In students do come from schools that have above average fractions of underrepresented minorities. They also have high state test scores, with average test score percentiles (in the statewide distribution) ranging from 85-89, and an average predicted probability of being in the top ten percent of around 70 percent. Pulled In students have relatively high college enrollment rates even before TTP, at 65 percent, with 49 percent attending four-year institutions but only 4 percent attending UT Austin. We observe employment nine to eleven years after high school graduation for 78 percent of the students, and average yearly earnings in this period, when working, are $46,682. When we compare the pre-TTP and post-TTP means, we see that for this group, college

[31] The increase in the UT attendance rates in these cells are consistent with other research that TTP drew some top students into UT who would otherwise have gone to private colleges (Daugherty et al. 2014). Our results are unchanged when we include these cells in the Pulled In group; see Appendix Table 5.

enrollment increases, UT enrollment increases, and college graduation increases following TTP. These preview our DD results.

Pushed Out students are substantially less likely to be underrepresented racial minorities than the Pulled In students, but are more likely to be Asian. As expected, Pushed Out students come from schools that have fewer underrepresented minorities than the sample as a whole. Their average test scores range from the 75-80th percentile (statewide) depending on the subject – notably, this is lower than for the Pulled In group.[32] Prior to TTP, the Pushed Out students also enroll in college at relatively high rates, with 61 percent attending any college and 47 percent attending 4-year colleges. While these are similar rates to those seen in the Pulled In group, the share attending UT Austin, 17 percent, is not. Among all Pushed Out students, 39 percent graduate with a BA within six years, lower than the rate among Pulled In students.

These summary statistics yield insights into the types of students affected by TTP. Marginally admitted students were more academically prepared, as measured by the state tests; came from more diverse high schools; and were more likely to be racial minorities, relative to marginally pushed out students. The share of students in the Pulled In group who attended UT Austin more than doubled with the policy, a large effect. However, this change represents only about 200 new Pulled In students at UT in each cohort, not enough to substantially shift the composition of the class (Cortes and Klasik 2020).

These summary statistics highlight what we think is a new fact about the TTP: Pulled In students have higher qualifications than displaced students as measured by the metrics available to us such as statewide test scores, the number of AP courses taken, and total absences. It is worth noting, however, that our identification strategy allows us to identify only subsets of students who gained and lost access to UT under the policy. With different data, we could use different dimensions to predict who is affected by the policy and generate different groups of affected and unaffected students. For example, our Pulled In group likely excludes students at feeder schools with high grades but low test scores, who may have benefitted from the policy but who we cannot identify due to our use of test scores to predict class rank. Nevertheless, while we can identify only a subset of pulled in students, the dimension that we identify is a natural one

[32] This is perhaps due in part to our use of scores rather than grades to measure class rank. There may be another group of students with low test scores but high grades who are also pulled in by TTP, but we are unable to identify them.

given the intent of the policy: students at the top of their classes at formerly underrepresented high schools. In addition, it seems like we are identifying the large fraction of Pulled In students in the bottom *s* groups; based on our calculations, the students we identify as Pulled In represent approximately 20% of all top ten students, and we identify 60-65 percent of all students in the top ten in the lowest two s groups (1-2).

We note that our procedure is designed to identify students in the top ten percent, not those close to but outside the top ten. However, there are several reasons to think that our procedure does successfully identify pushed out students. First, in a supplementary analysis (not reported) we find that many students in our Pushed Out group are indeed ranked between the 11th and 25th percentiles of their high school classes, which is also reported in our data. Second, Table 2 indicates that the UT Austin enrollment share fell by 5 percentage points for our Pushed Out group, consistent with the view that this group indeed saw reduced access.[33]

Our control group is somewhat less academically successful than either treatment group, but the differences are not large. Students have average test scores that range from the 64th to 66th percentile in the statewide distribution, depending on the subject. The underrepresented minority share lies in between those of the Pulled In and Pushed Out groups, while the fraction receiving free or reduced price lunch is slightly higher than that of the Pulled In group (19 percent). College enrollment rates are somewhat lower than those of the Pulled In and Pushed Out groups, and control students had lower bachelor's degree attainment rates. Note that our identification strategy does not require that the treatment and control groups be identical, just that they would have trended in a similar way in the absence of TTP. We will provide a number of tests to validate this assumption.

Given that our rule for defining the three groups is somewhat arbitrary and may over-fit to the change in attendance rates, we have extensively tested the sensitivity of our results to this choice. As an alternative, we used an automated machine learning approach to pick the Pulled In and Pushed Out groups. Conceptually, in our baseline specification, we choose the Pulled In and Pushed Out groups based on the observed mean change in enrollment at UT in the s-$\hat{q}$ cell between the pre and post TPP periods. Figure 5 indicates that there is substantial idiosyncratic noise in these cell means, and as a result the groups – especially the Pushed Out group – are quite

[33] Nevertheless, even after this decline, 12 percent of Pushed Out students enrolled at UT Austin, indicating that not all lost access (just as some students in the Pulled In group were admitted before TTP).

discontinuous. To reduce the influence of noise, in our alternative approach we smooth the cell means using a LASSO estimator, described in the Appendix. Appendix Figures 1 and 2 show the Pulled In, Pushed Out, and control groups based on our LASSO estimation. As we show below, our results are qualitatively similar using this strategy, even though the treatment groups are much more inclusive in LASSO. Our results are also robust to defining treatment in a continuous way rather than discretely, as the (smoothed) change in the share of students in the cell attending UT.

Step 3:

Once the treatment and control groups are defined, the final step is to estimate the effect of TTP on the outcomes of students in the Pulled In and Pushed Out groups. Using the students in these two treatment groups and the chosen control group, we estimate the following equation:

$$Y_{it} = \beta_0 + \beta_1 PulledIn_{it} + \beta_2 PushedOut_{it} + \beta_3 PulledIn_{it} * Post_{it} + \beta_4 PushedOut_{it} * Post_{it} + \boldsymbol{Z_{it}\theta} + \boldsymbol{\delta_t} + \epsilon_{it}$$

Here, *i* indexes students; *t* indexes cohorts; Y_{it} is a student outcome such as graduation; $PulledIn_{it}$ and $PushedOut_{it}$ are indicators for the two treatment groups; $Post_{it}$ is an indicator for the cohorts affected by TPP, from 1998 onward; δ_t represents year indicators; and $\boldsymbol{Z_{it}}$ is a vector of individual characteristics such as gender, race, ethnicity, and free and reduced lunch status. We also include in $\boldsymbol{Z_{it}}$ indicators for the ten *s* deciles, a cubic in $\hat{p}$, and a linear interaction of *s* and $\hat{p}$, to absorb changes in the distribution of student characteristics within the three groups over time that might otherwise confound our estimates. ϵ_{it} is an idiosyncratic error term.

For our main results, we present standard errors clustered at the school district level. However, these do not reflect uncertainty in the estimate of $\hat{p}$ or in the definitions of the three groups. We also estimate bootstrapped standard errors, re-estimating the top ten percent imputation model and allowing the group definitions to vary freely on each bootstrap draw. As we show, standard errors tend to be very similar using the two methods. Because the bootstrapped standard errors are quite computationally intensive, our main results use the analytic standard errors.

The interpretation of the difference-in-differences estimates bears some discussion. Our groups are not defined by actual admission to UT Austin or even by the true change in

admissibility to UT Austin brought about by TTP. Thus, our estimates cannot be seen as the causal effect of enrollment at or admission to UT Austin. Rather, our Pulled In and Pushed Out indicators are imperfect proxies for the groups affected by TTP. Each group includes some students who are eligible for TTP guaranteed admission and some who are not – with more of the former in the Pulled In group and more of the latter in the Pushed Out group. Each group also includes some students who would have been admitted to UT under both the pre-TTP and post-TTP admissions rules. Thus, our estimates can be seen as intention-to-treat (ITT) estimates of the effect of changes in access to selective colleges. The "treatment" is multi-valued, reflecting each of the different colleges that students might attend, and TTP might affect it in complex ways. For some students, the counterfactual in which they are not admitted to UT Austin sees them attending A&M, while for others it is a less selective UT campus, a community college, or no college at all. Moreover, even for students who do not attend UT Austin under either admissions policy, TTP may affect their choices among other alternatives. For these reasons, we do not attempt to construct treatment-on-the-treated estimates, focusing instead on the ITT for the TTP policy.[34]

5. Results

We first preview our results visually in an event study framework. In our difference-in-differences equation, we replace indicators for $Post_{it}$ with indicators for each cohort in the sample and replace the $Post_{it}$-treatment group interactions with interactions between the two treatment groups and the full set of cohort indicators. The 1997 cohort, immediately prior to the policy, is the omitted category. Results are shown in Figures 6, 7, 8, 9, and 10.

The event study specifications are useful for assessing our assumption that the treated and control groups would have moved together absent the change in admissions policy. This

[34] See Bleemer (2021) for further discussion in a similar context. Our ITT will likely substantially understate the impact on students actually affected by TTP. Students we identify as Pulled In have an average probability of being in the top 10 of 0.49. Many of these students would have been admitted to UT under pre-TTP rules, while the remaining 51 percent of students who are considered "pulled in" by our measure were not actually in the top of their classes and did not have access to UT due to TTP. As a result, the net change in UT access in our Pulled In group is no larger than 0.49, and likely much smaller than that. Our estimates of the change in outcomes for Pulled In students are attenuated relative to the change for students who actually gained access by at least a factor of two. Similar logic applies to the Pushed Out group. In this group, 37 percent of students are in fact in the top 10 percent of their high schools, so are not actually displaced but gain an admissions guarantee with TTP. Thus, a lower bound for the effect of losing access to UT can be obtained by dividing our Pushed Out effects by 0.63, or multiplying by 1.5.

assumption would be less credible if they were not moving together *prior* to the policy. While our analysis is limited to the two cohorts prior to the implementation of the policy, we see in each figure that the outcomes of the treatment and control groups do appear to be trending similarly prior to the policy change. This suggests that our underlying identification assumption of similar outcomes in the absence of treatment is reasonable. Following TTP, we see that Pulled In students were more likely to enroll at UT Austin, more likely to enroll in a four-year college overall, more likely to earn four-year degrees, and have higher earnings, in each case relative to the control group. Pushed Out students were less likely to enroll at UT Austin, with no change in their likelihood of enrolling in a four-year college overall or of earning degrees, or in their post-college earnings.

Enrollment

Table 3 presents the DD results (equation 1) in tabular form for a range of outcomes. Column 1 presents the *Pulled In*Post* coefficients. We can see that Pulled In students are no more likely to enroll in community college following TTP, but are 6.6 percentage points more likely to attend a public four-year college in Texas. When we look more narrowly, we see that Pulled In students are 5.3 percentage points more likely to attend UT Austin as a result of the policy change (from a base of 2.8 percentage points). They are 1.0 percentage points *less* likely on net to attend Texas A&M, the second most selective UT campus, although this is not statistically significant.

Unfortunately, as noted above, it is impossible to know the counterfactual enrollment behavior for specific students; we can only recover the net enrollment effects. For example, some students may switch from no college to community college while some may switch from community college to UT Austin. Moreover, some of the increase in college enrollment may have come through students switching from private or out-of-state universities to Texas public institutions, though we think this unlikely to be common for the relatively disadvantaged Pulled In students. To investigate this, we construct outcome measures capturing whether students are present at all in the Texas administrative data, via either a college enrollment or an earnings record, in each year after high school graduation. This would indicate that the student was present in Texas, so we would expect impacts on this measure if the policy led to changes in out-of-state enrollment. Estimates for these outcomes are shown in the first panel of Table 4. We find

no evidence that TTP affected the probability that we observe Pulled In students, consistent with no impact on out of state migration.

We can also examine the overall quality of the schools these individuals attended. We consider several summary measures of college quality: the college graduation rate of students who attended the college; the average math percentile (in the statewide distribution) of the students who attended; the average log earnings in years 9-11 after high school graduation of students who attended the college; and average instructional expenditures per student. We measure all of these using data from the pre-TTP 1996 and 1997 cohorts, and for all but instructional expenditures we consider two variants, one conditioning on enrollment (and so set to missing for those who do not go to college) and the other treating non-enrollment as an "institution," assigning non-enrolled students the average for all of their non-college peers.[35] We find (Table 3) that the policy led Pulled In students to attend colleges with average graduation rates that are 1.9 percentage points higher than the colleges attended by their pre- TPP counterparts. They also attend colleges with better peers than before TTP, by 1.5 percentiles on average. These numbers are conditional on enrollment, which likely leads to an underestimate of the effect on college quality due to the substantial increase in college attendance. If we include students who don't enroll in college by treating them as attendees of an institution with the average college completion rate and math score of non-college students, we see a larger increase in college quality—Pulled In students attended colleges with graduation rates and math test scores that were higher by 3.8 percentage points and 1.9 percentiles in the state distribution, respectively, relative to the pre-period. Overall, the Pulled In students were more likely to attend college, more likely to attend UT Austin, and attended higher quality institutions overall after the implementation of the TTP plan. We also find that following TTP Pulled In students attended schools with higher instructional expenditures and higher earning students.

Column 2 presents the results for Pushed Out students. These students were 3.6 percentage points less likely to enroll at UT Austin after the implementation of the TTP plan. This decrease was offset by increases in enrollment at other four-year schools (2.2 percentage points) and at community colleges (1.1 percentage points). The net effect on total four-year enrollment is a small, statistically insignificant -0.6 percentage points, while the effect on college enrollment overall is zero to three decimal places. For Pushed Out students, then, there was no

[35] Instructional expenditures are measured in 2014, and are set to missing for non-college students.

net change in college enrollment rates, but these students shifted away from UT Austin and toward less selective institutions. This is confirmed when one looks at the peer quality in the institutions they attended; we see meaningful declines in per-pupil expenditures and peer math scores. We see small declines in institution graduation rates and no decline in peer earnings, however – each foreshadowing what we find for the affected students themselves. Interestingly, the null effect on any in-state enrollment suggests that there was very little shifting towards out-of-state schools in response to the policy, as this would have to have been offset by an increase in college-going overall among displaced students, which seems unlikely. This is consistent with work by Tienda and Niu (2006b), which shows that second decile students at feeder schools have preferences for and enroll at out of state institutions at similar rates to first decile students, and with our results on whether or not we observe a student in our data (either working or in public post-secondary school in Texas) in Table 4.

Graduation and Labor Market Outcomes

When looking at graduation outcomes, we see that Pulled In students were 3.9 percentage points more likely to graduate with a BA from UT Austin within 6 years after high school graduation and 3.7 percentage points more likely to graduate from any four-year college in the state after the implementation of TTP, relative to control students. When we look more closely at time-to-degree (Appendix Table 6), we find that graduation effects are present for Pulled In students starting 4 years after leaving high school and persist until we last observe them, 8 years after leaving high school.

The UT Austin effect implies a graduation rate of 74 percent (3.9 / 5.3) for the marginal UT Austin enrollees brought in by TTP. This is quite similar to the *average* UT Austin graduation rate of 75 percent in the post period, and is higher than the pre-period graduation rate of 70 percent, suggesting that marginal students brought in by the TTP do not struggle more than inframarginal students.[36]

[36] This divides all UT graduates by all initial UT enrollees. These groups are not fully comparable, as some UT graduates might have transferred there after first enrolling elsewhere. We have also examined effects on the joint outcome of initially enrolling at UT and then obtaining a BA from there. The difference-in-differences coefficient for this outcome is 0.038 (SE 0.003). This implies that the marginal student who initially enrolled at UT due to TTP has a graduation rate of 72 percent. Note that the UT Austin graduation rate is not confounded by any shifts in enrollment from private/out-of-state schools, and implies that the marginal enrollee had similar or higher graduation rates than the average enrollee, regardless of their counterfactual enrollment.

Given that, on average, the Pulled In students go to colleges with 3.8 p.p. higher graduation rates, and their graduation rates go up by 3.7 percentage point, the implied effect of attending a school with 1 percentage point higher graduation rates is a 1.0 percentage point increase in the likelihood of graduating.[37] This matches Zimmerman's (2014) estimate, also 1.0, and is somewhat lower than Bleemer's (2021) estimate (computed from his Table 4) of 1.1 or Cohodes and Goodman's (2014) estimate of 1.6. However, it is important to note that much of the increase that we find in the institutional graduation rate comes from the extensive margin – students shifted into going to college at all – so neither Bleemer (2021) or Cohodes and Goodman (2014) are not strictly comparable.

Arcidiacono, Aucejo, and Hotz (2016) find that marginally admitted students in California are less likely to graduate in science (STEM) majors. We find that Pulled In students are 0.8 percentage points less likely to graduate with a degree in a STEM major despite the overall increase in bachelor's degree receipt, but this effect is not statistically significant. This means that the increase in graduation we observe is happening despite no net gain in STEM degrees. This is consistent with the Pulled In students not graduating in STEM degrees (on net).

We next consider the labor market effects of the policy. We first show that the probability of being employed, so having observed earnings, does not change as a result of the policy for either the Pulled In or the Pushed Out students. This, like our results in Table 4 discussed below, suggests that selection into observing earnings--as might occur if TTP affected the likelihood that a student remains in Texas--is unlikely to be an issue.

We consider several measures of earnings. Our preferred measure averages only across years when earnings are observed, and excludes students for whom earnings are never observed. Among Pulled In students, earnings nine to eleven years after graduation increase by $692 after TTP with a large standard error; when we include students and years with zero earnings, the estimate falls to $359 and remains insignificant. When we consider the log of average earnings (including zeros in the average and conditioning on observing positive earnings in at least one quarter), however, we find an increase in log earnings of 5.5 percentage points which is statistically significant. When we look at earnings thirteen to fifteen years after graduation, this

[37] Prior to rounding, the effect on institutional graduation rate is 3.77 percentage points, and the effect on the likelihood of graduating is 3.74. The ratio of these is 0.992.

earnings effect disappears. Overall, these results are imprecise but suggest slightly positive effects on earnings.

In Table 4, we explore the sensitivity of the log earnings results. First, we estimate models for log earnings in every year from 9 to 15 after high school graduation, in each year excluding those with zero earnings in that year. Second, we recompute log average earnings, this time winsorizing outliers more aggressively. Our result of positive effects for Pulled In students' log earnings in years 9-11 after high school is robust to this alternative calculation, and the year-by-year estimates indicate somewhat larger effects in years 10 and 11 than in year 9. There is no indication of positive effects in any year after 11. The positive effect for Pushed Out students in years 9-11 is not robust. It falls by more than half in our alternative calculation, and the year-by-year estimates are all smaller than in our original specification and statistically insignificant. There is no indication of a negative effect here, but also not strong evidence of a positive one.

Our earnings effects are imprecise, and do not provide strong evidence that effects change over time. All of our estimates are consistent with an unchanging effect of 0.03 in each year. Nevertheless, they are clearly smaller than the earnings effects implied by Bleemer's (2021) results, which in our setting would imply a log wage effect around 0.1.[38] On the other hand, they offer no evidence for the hypothesis that Pulled In students are substantially mismatched at the schools they are pulled into and therefore worse off than they would have been without TTP. Overall, we find that following TTP Pulled In students were more likely to attend UT Austin and other four-year universities, were more likely to graduate from college, and experienced no negative impact on earnings. This is consistent with estimates from Goodman et al (2017), Hoekstra (2009), Bleemer (2021), and Zimmerman (2014) which indicate that access to more selective colleges improves student graduation and earnings outcomes, though magnitudes vary.

When we examine the outcomes for students Pushed Out of UT Austin by TTP, we find little evidence that they were harmed. While they were significantly less likely to graduate with a

[38] Bleemer (2021, Table 4) finds that students whose enrollment was affected by Eligibility in the Local Context, California's version of the TTP, saw increases of 26.8 percentage points in institutional graduation rates and obtained increases in log wages 7 to 9 years after high school graduation of 0.76. The ratio of these is 0.76/26.8 = 0.028, so a program the size of TTP, which raised Pulled In students' institutional graduation rates by 3.8 percentage points, would be expected to increase log wages by 0.028*3.8 ≈ 0.1. Note, however, that Bleemer's estimates reflect only intensive-margin changes in college quality, where ours combine intensive-margin changes with extensive-margin changes in enrollment rates.

BA from UT Austin (2.1 percentage points), this was fully offset by increases in BA attainment at other public schools in Texas. The net change in BA attainment within 6 years after high school graduation for the Pushed Out group as a whole was -0.1 percentage points with a standard error of 0.6 percentage points. This is consistent with the fact that Pushed Out students attended schools with similar graduation rates after the policy. When we explore alternative time-to-degree thresholds, coefficients are positive, though always close to zero. Pushed Out students are slightly less likely to obtain BAs in STEM majors (0.2 percentage points), but this is again not statistically significant. One contributing factor to the absence of negative graduation effects is that the Pushed Out students were not particularly successful before TTP: The UT graduation rate for the marginal UT enrollees pushed out by TTP that is implied by our estimated effect on UT degrees is only 57 percent.

When we turn to labor market outcomes for the Pushed Out students, we see that these students are slightly more likely to be observed with positive earnings in Texas following TTP, though this effect is not statistically significant. This again counters the hypothesis that many Pushed Out students might have left Texas for out-of-state institutions and stayed elsewhere after graduation. We see small gains in earnings in years 9-11, similar to Pulled In students. When we look 13-15 years after graduation, we see no statistically significant effects. Importantly, while our estimates are imprecise, there is no evidence that being displaced from the most selective schools harmed Pushed Out students' labor market outcomes.

Table 5 presents estimates that vary the covariates in the DD specification. The first set of estimates are for a sparse specification that includes just indicators for the two treatment groups, year indicators, and the treatment times post TTP interaction. The second set adds individual controls – race, gender, and free lunch and immigrant status. The third set adds school level racial composition, free lunch share, English language learner share, and special education share. The fourth replaces the school characteristics with indicators for the ten *s* deciles and a linear control for $\hat{p}$. The fifth adds a linear interaction between the school's pre-TTP UT Austin share and $\hat{p}$, while the sixth adds square and cubic terms in $\hat{p}$, as in our specifications in Table 3. None of these yields appreciably different results.

Overall, Pushed Out students were not any less likely to enroll in or graduate from college in Texas. There is also no evidence of negative earnings effects. Taken together, our results suggest that the Top Ten Percent policy helped Pulled In students via increased

graduation rates, with a small increase or at least no reduction in earnings, but did not harm displaced students' graduation rates or earnings. This suggests important heterogeneity in the returns to attending a selective college, which may help to explain the mixed results in the college quality literature. An intriguing, speculative hypothesis is that the heterogeneity may reflect differences in student disadvantage, with college quality mattering more for disadvantaged students.

Heterogeneity

The effects of the TTP policy may vary based on individual characteristics within the treated groups. For example, more disadvantaged students, who may have less knowledge about or exposure to role models with college degrees, may be more affected. We next examine whether the policy had heterogeneous effects based on income (proxied by an indicator of whether the student was eligible for free/reduced price lunch in high school), race, and gender. The results, from separate estimates of equation (1) for each group of students, are presented in Table 6.

We begin with racial and ethnic heterogeneity. The introduction of TTP overlapped with the elimination of affirmative action preferences in Texas, so we might expect smaller enrollment effects for students from underrepresented minority groups. We do not see this. In particular, the increase in UT enrollment among Pulled In Black and Hispanic students is larger than that for white students, while Pushed Out students – where we might expect to find non-white students harmed by the loss of affirmative action preferences– show similar declines in UT enrollment across races.[39]

Across other outcomes, there are few large differences by race/ethnicity. Pulled In Black students see somewhat larger boosts in graduation rates (6.5 percentage points), while Hispanic students see the same boost as whites (3.6 vs. 3.5 percentage points) despite their larger UT Austin enrollment effect. There is no evidence of harm to any Pushed Out students' graduation rates, with a statistically insignificant, positive estimate for Black students. Similarly, earnings effects are, if anything, larger for Pulled In Black students nine to eleven years after graduation, though point estimates are positive for both Pulled In and Pushed Out students from all

[39] Of course, the removal of affirmative action preferences might have had other effects on enrollment that do not align with our TTP treatment groups.

ethnicities. Again, we find no effects on earnings thirteen to fifteen years after graduation for any students.

The second set of columns show differences by family socioeconomic status, contrasting the effects for students who receive free and reduced-price lunch with the effects for those who do not. Estimates are generally similar across groups, with a pattern of somewhat better impacts on free lunch students in the Pulled In group than on non-free-lunch students.

The final set of columns shows heterogeneity by gender. Here, we see a bit more evidence for differences in the effect of TTP. While men and women have similar changes in enrollment patterns, Pulled In men do not see a statistically significant increase in graduation despite increases in college attendance, both overall and at UT Austin, while women's graduation does increase substantially. The UT Austin effects are not significantly different between men and women, but the overall graduation effects are. The earnings effects nine to eleven years after graduation are also very large for Pulled In women, who see a statistically significant 7.9 percentage point increase in earnings, while Pulled In men see an earnings boost one-third as large (though we cannot reject equality across groups).[40] However, we see no effect on earnings for either men or women thirteen to fifteen years after graduation. Gender differences are much smaller for Pushed Out students – the (positive) male earnings effect 9-11 years after graduation is statistically significant while the female effect is not, but point estimates are not so far apart. Again, we see no effect on earnings for either Pushed Out men or women 13-15 years after graduation.

Overall, while we find some evidence of heterogeneity of effects, the general picture is quite consistent across race, gender, and free and reduced-price lunch status.

6. Robustness Checks

We conduct a number of checks to verify the robustness of our conclusions.

Longhorn Opportunity/Century Scholars

Our key assumption is that there were no concurrent policy changes that differentially affected our treated students. A possible violation of this assumption is the initiation of programs

[40] This larger return for women is consistent with Ge et al. (2018).

at UT Austin and Texas A&M to actively recruit students from some high schools. These programs--the Longhorn Opportunity Scholars and Century Scholars programs for UT Austin and Texas A&M respectively--gave a suite of financial and social support to students from targeted high schools to encourage them to attend the University of Texas Austin or Texas A&M (Andrews, Imberman, and Lovenheim 2020). Appendix Table 7 presents results that exclude the high schools targeted under these programs. Our results are unchanged.

Choice of Prediction Method

As discussed above, a key limitation of our analysis is that we do not observe top ten status prior to the implementation of TTP. As a result, we must predict this status for the pre-period data. As noted earlier, we tried a variety of methods to generate predictions of whether a student was in the top 10 percent of his or her high school class, and these predictions were highly correlated (Appendix Table 2).

Because we only observe top ten status in the post-TTP period, one might worry that we are better fitting T in the post period and thus generating spurious relationships between our $\hat{q}$ by *s* cells and enrollment changes. This could also occur if the relationship between the predictors and top ten status changed in the post period; if so, our $\hat{q}$ will better proxy actual top ten percent class rank in the years that we used to fit the prediction model than in earlier years. While the adequacy of our T prediction model for years when T is not observed is inherently untestable, we can test whether results are sensitive to the specific years used among those where T is available. We assess this by dividing our training sample into two subsets, one consisting of the years 1999 and 2000 and the other of 2001 and 2002. We train the prediction of $\hat{p}$ separately using the two subsets, then re-construct the Pulled In and Pushed Out treatment groups and re-estimate the DD model using the full sample. Appendix Table 2 shows that the two alternative $\hat{p}$ predictions are correlated 0.96 with each other and 0.98-0.99 with one constructed using the full sample. Appendix Table 3 presents DD estimates using the alternative estimates.

Columns 1 and 4 present results when we use the 1999-2000 data to generate predictions, and Columns 2 and 5 present results when we use the 2001-2002 data. The results are generally consistent, suggesting that the relationships between covariates and top ten status are not changing over time. This allays some concerns about overfitting. We have also constructed full split sample estimates, using the 1999-2000-trained $\hat{p}$ measures to estimate DD

models using only the non-overlapping 1997, 2001, and 2002 sample, and the 2001-2002-trained $\hat{p}$ measures to estimate the DD model using the 1996, 1999, and 2000 sample. This ensures that there is no mechanical relationship between sampling error in our $\hat{p}$ and the DD estimates, and yields quite similar results. All of these results indicate that the relationship between Xs and top ten status is not meaningfully changing across the years of the post period, lending some support to the assumption that the relationship between the Xs and top 10 status is not changing in the earlier parts of the sample.

Finally, in our main specifications, we used a random forest method of generating our prediction of the top ten status. As a further robustness check, we also predict the top ten status by estimating a simple logit that includes the extensive set of demographic, course, and high school characteristics described earlier, without polynomial or interaction terms. From this, we generate new predicted probabilities of being in the top ten percent of the class. As mentioned earlier and shown in Appendix Table 2, the correlation between the random forest prediction and the logit prediction is quite high—approximately 0.94, which is reassuring. We then re-define our treatment groups and re-estimate the DD specifications using the logit estimates of $\hat{p}$. The results are presented in Columns 3 and 6 of Appendix Table 3. Results are again very consistent, highlighting that our conclusions are insensitive to the choice of prediction strategy.

Choice of Treatment Groups

As discussed earlier, our definition of the Pulled In and Pushed Out groups is somewhat arbitrary. To test the sensitivity of our results to this definition, we use an alternative method-the LASSO technique described in Section 4 and in the Appendix—to choose our treatment and control groups. This method yields much larger treatment groups, as shown in Appendix Figure 2. We present DD results using these groups in Table 7. Because the LASSO treatment groups include a number of cells in which the change in UT Austin enrollment is quite small, the estimated TTP effect on UT Austin enrollment, effectively the first stage for our analysis, is much reduced for both the Pulled In and Pushed Out groups. Not surprisingly, all of the other coefficients are also attenuated. Importantly, the ratios of the effects on earnings or graduation rates to the effects on enrollment are quite similar to the baseline specification.

Another robustness check takes advantage of variation in the intensity of treatment by defining Pulled In and Pushed Out groups more continuously. To do this, rather than using

binary classifications, we use the continuous variation in the change in UT Austin enrollment shares from the LASSO specification. This then identifies the policy effect not just from contrasts between Pulled In/Pushed Out groups and the control group, but also from variation among cells with larger and smaller shares of Pulled In (Pushed Out) students within the Pulled In (Pushed Out) group. To make the resulting estimates comparable to earlier results, we rescale the LASSO-smoothed change in the UT Austin share by dividing by the change seen in the cell with the largest change, separately for cells with positive and negative changes. This ensures that treatment dosages vary between 0 and 1, so that the coefficients represent the effect on the most heavily treated cells relative to those with no change in UT Austin enrollment. We allow the dose-response relationship to differ for cells with increases and reductions in UT Austin enrollment, to permit separate Pulled In and Pushed Out effects.

Table 8 presents these results. In general, the conclusions are very similar to our main specification. Pulled In students are more likely to graduate as a result of TTP and Pushed Out students are no less likely to graduate. We find similar patterns in earnings, with no evidence of negative effects for either Pulled In or Pushed Out students. One slight difference from our main specification is that Pushed Out students are more likely to have observed earnings. As our primary concern was that Pushed Out students would *leave* Texas, creating a sample selection problem for us, this does not much concern us. It does not change our overall conclusion of no negative effects of the TTP on Pushed Out students' earnings.

Inference

Our results thus far present standard errors from the difference-in-differences regression (1), allowing for clustering at the school district level but not accounting for error deriving from our estimation of the $\hat{p}$ model. In Appendix Table 8, we present results from bootstrapping our entire procedure, from the estimation of our top ten percent prediction model through the choice of control groups and from the difference-in-differences regression. We use a clustered bootstrap, with high schools as clusters. Results are nearly identical to the much less computationally intensive analytic standard errors reported elsewhere.

Mobility

As a final specification test, we explore sensitivity to student mobility. Cullen, Long, and Reback (2013; see also Estevan et al., 2017) point out that TTP created incentives for parents to move their children to different high schools where they were more likely to make it into the top decile. They find empirical evidence for this, though the magnitudes are small – Cullen et al. (2013) estimate that only 211 students per cohort, statewide, moved from more competitive schools to take advantage of TTP, while results in Estevan et al. (2017) suggest a somewhat larger response.

To address potential mobility, we take advantage of the fact that, when the policy was implemented, the first two affected cohorts would not have had much opportunity to move in response, as they were already enrolled in the 10^{th} grade and the class rank used for automatic admission is calculated at the end of junior year. (In addition, high schools were permitted to exclude students from their top ten percent calculations who had enrolled after 10^{th} grade, though we do not have data on how often this occurred.) We thus limit attention to the pre-TTP cohorts and the first two post-TTP cohorts. Students in these four cohorts enrolled in 10^{th} grade before HB 588. When we conduct our analysis on this restricted sample in Appendix Table 9 our results are quite similar to those from the full sample, as expected given the consistency of effects across cohorts seen in the event study plots. This exercise also accounts for other changes that may have arisen over the long term, such as rising selectivity at UT Austin.

Discussion/Conclusion

Our results show that the Top Ten Percent rule increased college access and completion in Texas. Pulled In students gained access to more selective institutions, with increased enrollment at the flagship campus. This was not just a reallocation of students across campuses; many Pulled In students would not have attended any college absent the policy. This shift substantially increased the share of students who earn BAs, with no indication that these students suffered from attending more selective colleges. In contrast, Pushed Out students lost access to UT Austin but offset this with higher enrollment rates at less selective campuses, with no change in overall college enrollment. Despite the decline in the quality of the initial college attended by Pushed Out students, we find no evidence of negative effects on graduation rates or earnings for this group.

Our results have interesting implications for the returns to college quality. We show meaningful improvements in graduation rates and suggestive evidence of earnings gains for students who gain access to a selective institution. However, we do not find reductions in graduation rates or earnings for students who lose access to selective institutions. Our finding that college quality does not matter for Pushed Out students contrasts with much of the existing work on college quality, which finds that college quality increases student achievement (Zimmerman 2014, Hoekstra 2009, Goodman et al. 2017, Cohodes and Goodman 2014, Bleemer 2021, Kozakowski 2019).

The primary difference between the Pulled In and Pushed Out groups is the high schools they attended. This yields some insights into possible mechanisms for the different effects of access to selective institutions on these two groups. Our results are consistent with college selectivity mattering for students from disadvantaged schools but not mattering for students from more advantaged schools. Interestingly, our results based on student-level measures of disadvantage suggest that the story is most consistent with school or community levels of disadvantage, rather than individual levels of disadvantage within schools, affecting the returns to college selectivity (Chetty et al. 2014). These different effects may be driven by peers, mentors, or parents who can help insulate students displaced from selective institutions. However, our results are inconsistent with academic mismatch for marginally admitted students under the Top Ten Percent policy.

The TTP reduced the discretion of administrators by placing constraints on who must be accepted. Our results suggest that this rules-based approach does no worse than allowing discretion in college admissions. In this setting, a rule seems to have done as well at identifying students who would succeed while increasing representation of students from disadvantaged schools.

Our results suggest that the benefits of access to selective institutions are not zero sum. In our setting, some students seem to benefit more from access than others, and TTP seems to have redirected the allocation of scarce spots toward students who could benefit more from them than did those who were displaced. College admissions decisions, especially at public colleges, could account for these differences by offering spots to students most likely to benefit. Future research should carefully consider not only if college quality matters, but *when* college quality matters.

References

Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, *74*, 101923.

Andrews, Rodney J., Jing Li, and Michael F. Lovenheim. 2016. "Quantile Treatment Effects of College Quality on Earnings." *Journal of Human Resources* 51 (1): 200-238.

Anelli, Massimo. 2020. "The Returns to Elite University Education: a Quasi-Experimental Analysis." Journal of the European Economic Association. https://doi.org/10.1093/jeea/jvz070

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economic Review* 106 (3): 525-62.

Armstrong, E. A., & Hamilton, L. T. 2013. Paying for the Party. Harvard University Press.

Bleemer, Zachary. 2020. "Affirmative Action, Mismatch, and Economic Mobility After California's Proposition 209." Working paper, November.

Bleemer, Zachary. 2021. "Top Percent Policies and the Return to Postsecondary Selectivity." Working paper, January.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129 (4): 1553-1623.

Cohodes, Sarah R., and Joshua S. Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy." *American Economic Journal: Applied Economics* 6 (4): 251-85.

Cortes, K., & Klasik, D. (2020). *Uniform Admissions, Unequal Access: Did the Top 10% Plan Increase Access to Selective Flagship Institutions?* (No. w28280). National Bureau of Economic Research.

Cullen, Julie B., Mark C. Long and Randall Reback. 2013. "Jockeying for Position: Strategic High School Choice under Texas' Top Ten Percent Plan." *Journal of Public Economics* 97: 32-48.

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

Dale, Stacy B., and Alan B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources* 49 (2): 323-358.

Daugherty, Lindsay, Paco Martorell, and Isaac McFarlin Jr. 2014. "Percent Plans, Automatic Admissions, and College Outcomes." *IZA Journal of Labor Economics* 3 (10).

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2014. "School Choice, School Quality, and Postsecondary Attainment." *American Economic Review* 104 (3): 991-1013.

Dillon, Eleanor W., and Jeffrey A. Smith. 2017. "Determinants of the Match Between Student Ability and College Quality." *Journal of Labor Economics* 35 (1): 45-66.

Dillon, Eleanor W., and Jeffrey A. Smith. 2020. "The Consequences of Academic Match Between Students and Colleges." *Journal of Human Resources*.

Dynarski, S., Libassi, C. J., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721-56.

Estevan, Fernanda, Thomas Gall, Patrick Legros, and Andrew F. Newman. 2017. "The Top-Ten Way to Integrate High Schools." CEPR Discussion Paper DP11910.

Ge, Suqin, Elliott Isaac, and Amalia Miller. 2018. "Elite Schools and Opting-In: Effects of College Selectivity on Career and Family Outcomes." NBER Working Paper 25315.

Golightly, Eleanor. 2019. "Does College Access Increase High School Effort? Evaluating the Impact of the Texas Top 10% Rule." Working paper, University of Texas at Austin, August.

Goodman, Joshua, Michael Hurwitz, and Jonathan Smith. 2017. "Access to 4-year Public Colleges and Degree Completion." *Journal of Labor Economics* 35 (3): 829-867.

Harris, Angel and Marta Tienda. 2010. "Minority Higher Education Pipeline: Consequences of Changes in College Admissions Policy in Texas." *The ANNALS of the American Academy of Political and Social Science* 627: 60-81.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *The Review of Economics and Statistics* 91 (4): 717-724.

Jack, A. A. 2016. "(No) Harm in Asking: Class, Acquired Cultural Capital, and Academic Engagement at an Elite University." *Sociology of Education*, *89*(1), 1-19.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R.* Springer Science+Business Media: New York.

Kane, Thomas. 1998. "Racial and Ethnic Preferences in College Admissions." In *The Black-White Test Score Gap* (Christopher Jencks and Meredith Philips, eds.). Brookings Institution Press.

Kapor, Adam. 2015. "Distributional Effects of Race-Blind Affirmative Action." Yale University Working Paper.

Kozakowski, Whitney. 2019. "Are Four-Year Public Colleges Engines for Mobility? Evidence from Statewide Admissions Thresholds." Unpublished manuscript, Nov 5, 2019.

Krueger, Alan B., and Diane M. Whitmore. 2001. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." *The Economic Journal* 111 (468): 1-28.

Long, Mark C. 2004. "Race and College Admissions: An Alternative to Affirmative Action?" *The Review of Economics and Statistics* 86 (4): 1020-1033.

Long, Mark C., and Marta Tienda. 2008. "Winners and Losers: Changes in Texas University Admissions Post-*Hopwood*." *Educational Evaluation and Policy Analysis* 30 (3): 255-280.

Long, Mark C., and Marta Tienda. 2010. "Changes in Texas Universities' Applicant Pools After the *Hopwood* Decision." *Social Science Research* 39 (1): 48-66.

Long, Mark C., Victor Saenz, and Marta Tienda. 2010. "Policy Transparency and College Enrollment: Did the Texas Top Ten Percent Law Broaden Access to the Public Flagships?" *The Annals of the American Academy of Political and Social Science*, 627 (1): 82-105.

Mountjoy, Jack, and Brent R. Hickman. 2020. "The Returns to College(s): Estimating Value-Added and Match Effects in Higher Education." Unpublished manuscript.

Niu, Sunny X., and Marta Tienda. 2010. "The Impact of the Texas Top Ten Percent Law on College Enrollment: A Regression Discontinuity Approach." *Journal of Policy Analysis and Management* 29 (1): 84-110.

Niu, Sunny X., Marta Tienda, and Kalena Cortes. 2006. "College Selectivity and the Texas Top 10% Law." *Economics of Education Review* 25: 259-272.

Rothstein, Jesse. 2004. "College Performance Predictions and the SAT." *Journal of Econometrics* 121 (1-2): 297-317.

Rothstein, Jesse, and Albert Yoon. 2008. "Affirmative Action in Law School Admissions: What Do Racial Preferences Do?" *University of Chicago Law Review* 75 (2): 649-714.

Sander, Richard, and Stuart Taylor Jr. 2012. *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit it*. Basic Books.

Smith, Jonathan, Joshua Goodman, Michael Hurwitz. 2019. "The Economic Impact of Access to Public Four-Year Colleges"

Tienda, Marta, Kevin T. Leicht, Teresa Sullivan, Michael Maltese, and Kim Lloyd. 2003. "Closing the Gap? : Admissions and Enrollments at the Texas Public Flagships before and after Affirmative Action." Office of Population Research, Princeton University Working Paper 303

Tienda, Marta, and Sunny X. Niu. 2006a. "Capitalizing on Segregation, Pretending Neutrality: College Admissions and the Texas Top 10% Law." *American Law and Economics Review* 8 (2): 312-346.

Tienda, Marta, and Sunny X. Niu. 2006b. "Flagships, Feeders, and the Texas Top 10% Law: A Test of the "Brain Drain" Hypothesis." *The Journal of Higher Education* 77 (4): 712-739.

Tienda, Marta, and Teresa Sullivan. 2009. "The Promise and Peril of the Texas Uniform Admission Law." In The Next Twenty Five Years, Affirmative Action and Higher Education in the United States and South Africa, 155-174.

University of Texas at Austin Office of Admissions. 2007. "Implementation and Results of the Texas Automatic Admissions Law (HB 588) at The University of Texas at Austin."

University of Texas at Austin Office of Admissions. 2008. "Implementation and Results of the Texas Automatic Admissions Law (HB 588) at The University of Texas at Austin."

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics* 32 (4): 711-754.

Figures

Figure 1. Schematic description of Top Ten Percent plan effect on enrollment

A. Enrollment prior to TTP **B. Changes with TTP**

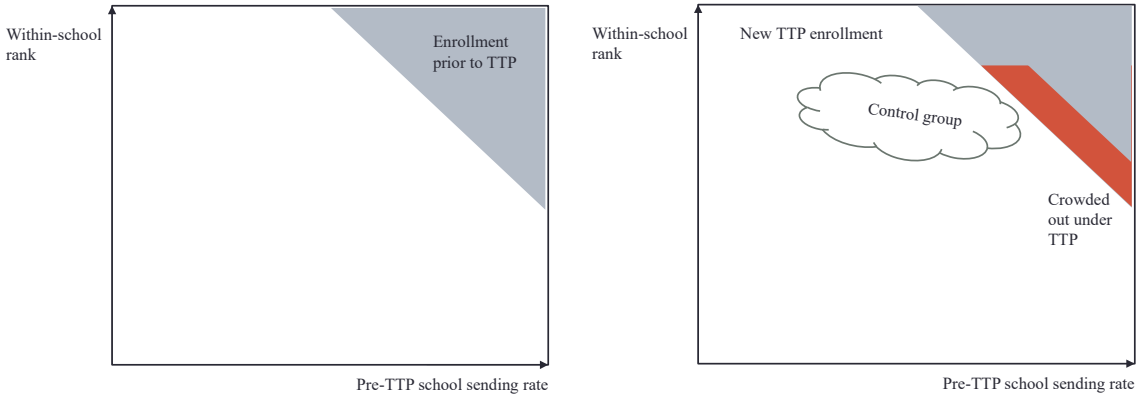
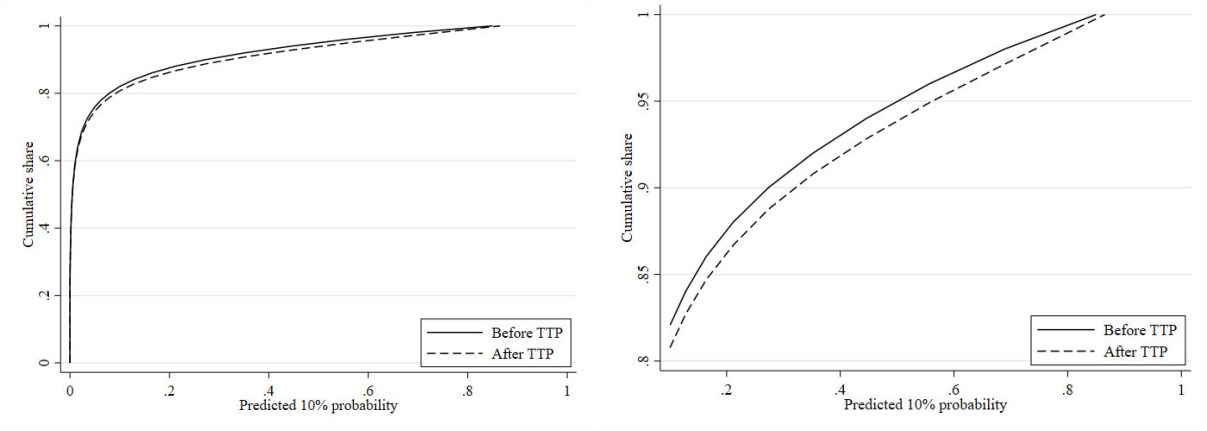


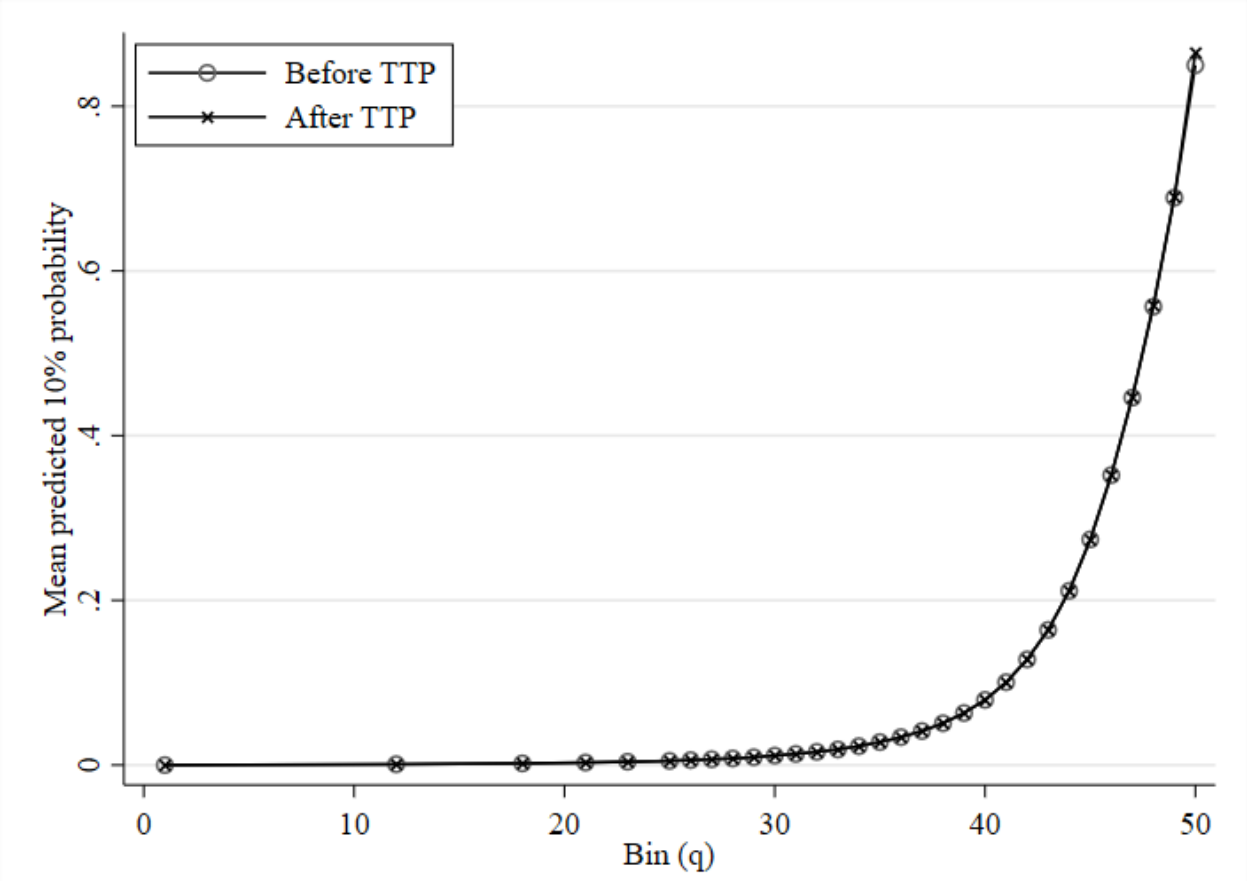
Figure 2. Distribution of predicted top ten probability ($\hat{p}$) before and after TTP

A. Full distribution **B. Excluding $\hat{p}$<0.1**



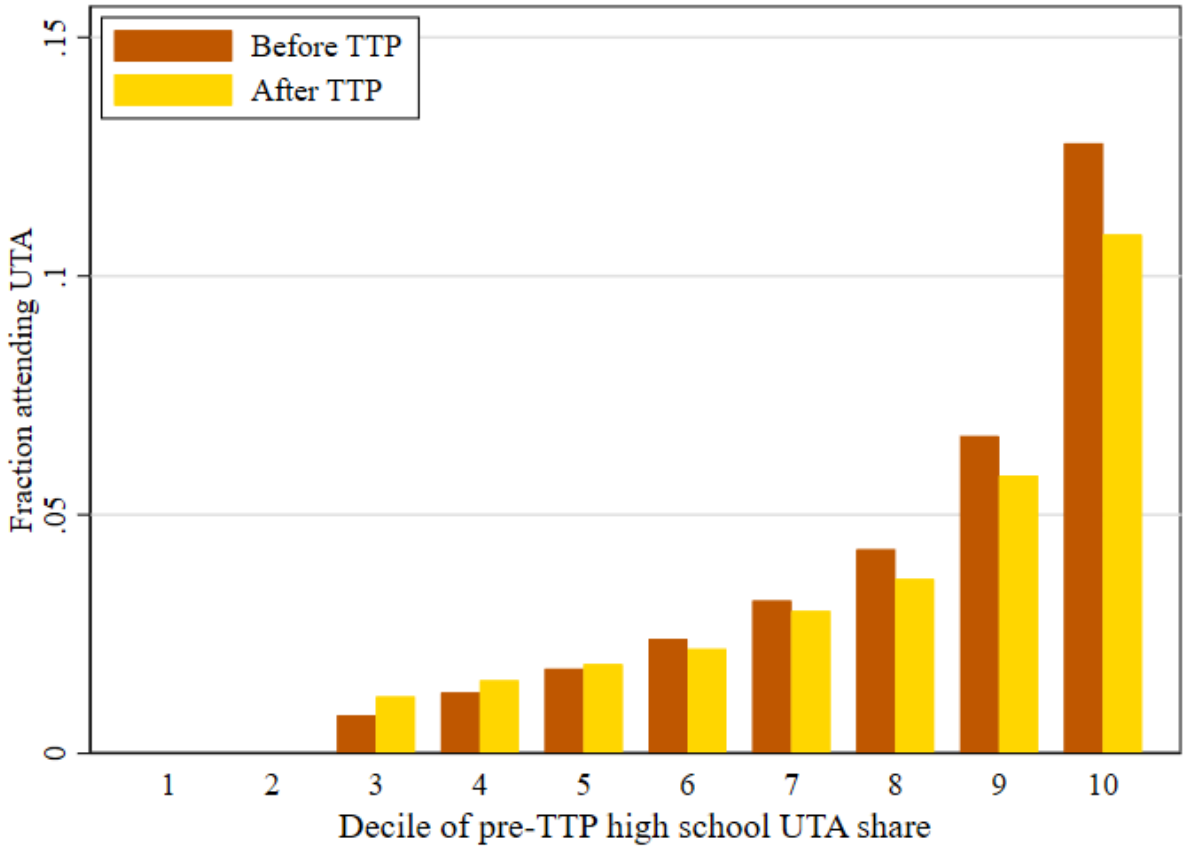
Notes: Figures show the CDF of estimated top-ten-percent probability ($\hat{p}$) across all students in the sample, separately for 1996-1997 (Before TTP) and 1998-2002 (After TTP). Predicted probabilities of being in the top ten percent are derived from a random forest model fit to 1999-2002 data. Right panel shows the portion of the CDF for $\hat{p} > 0.1$.

Figure 3. Mean predicted top ten probability ($\hat{p}$) by bin ($\hat{q}$)



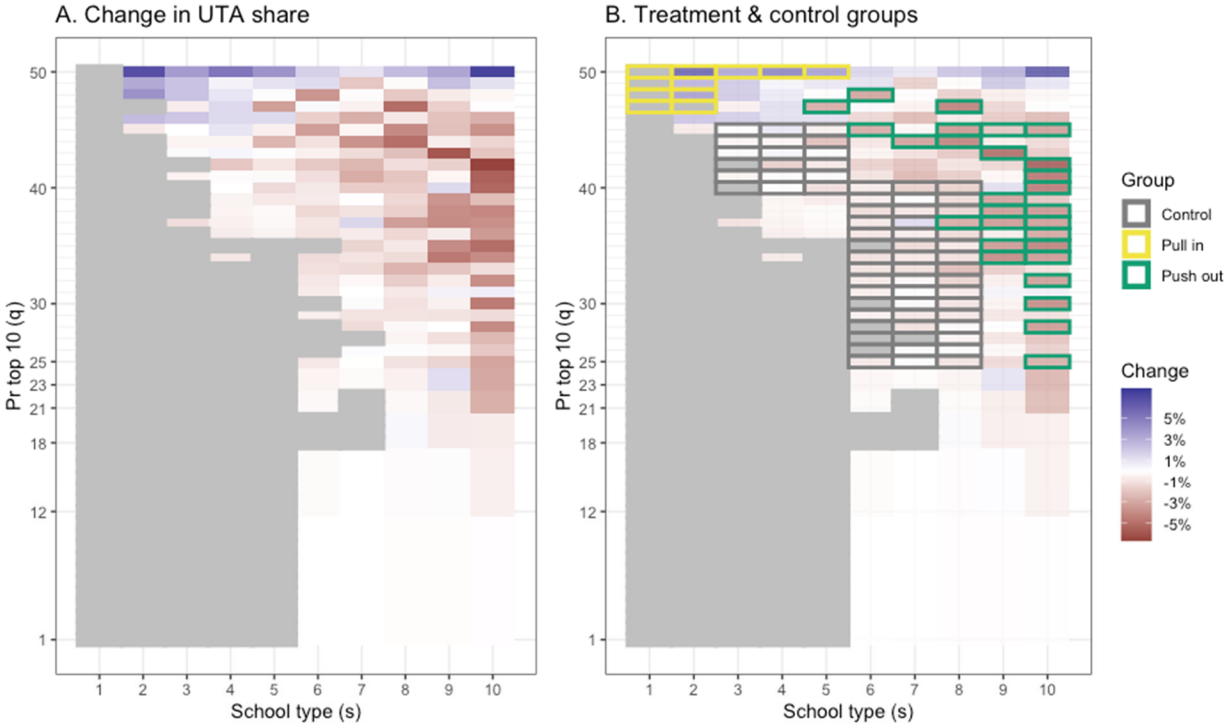
Notes: The x-axis represents $\hat{q}$ bins, ordered from 1-50, with bins defined by dividing the distribution of $\hat{p}$ (the predicted top ten percent probability) into 50 cells. There are fewer than 50 bins because of mass points in the $\hat{p}$ distribution – for example, bin 1 contains all observations with $\hat{p}$=0. The y-axis shows the mean of $\hat{p}$ within each bin.

Figure 4. Share of students from high school attending UT Austin, by school pre-TTP decile (s)



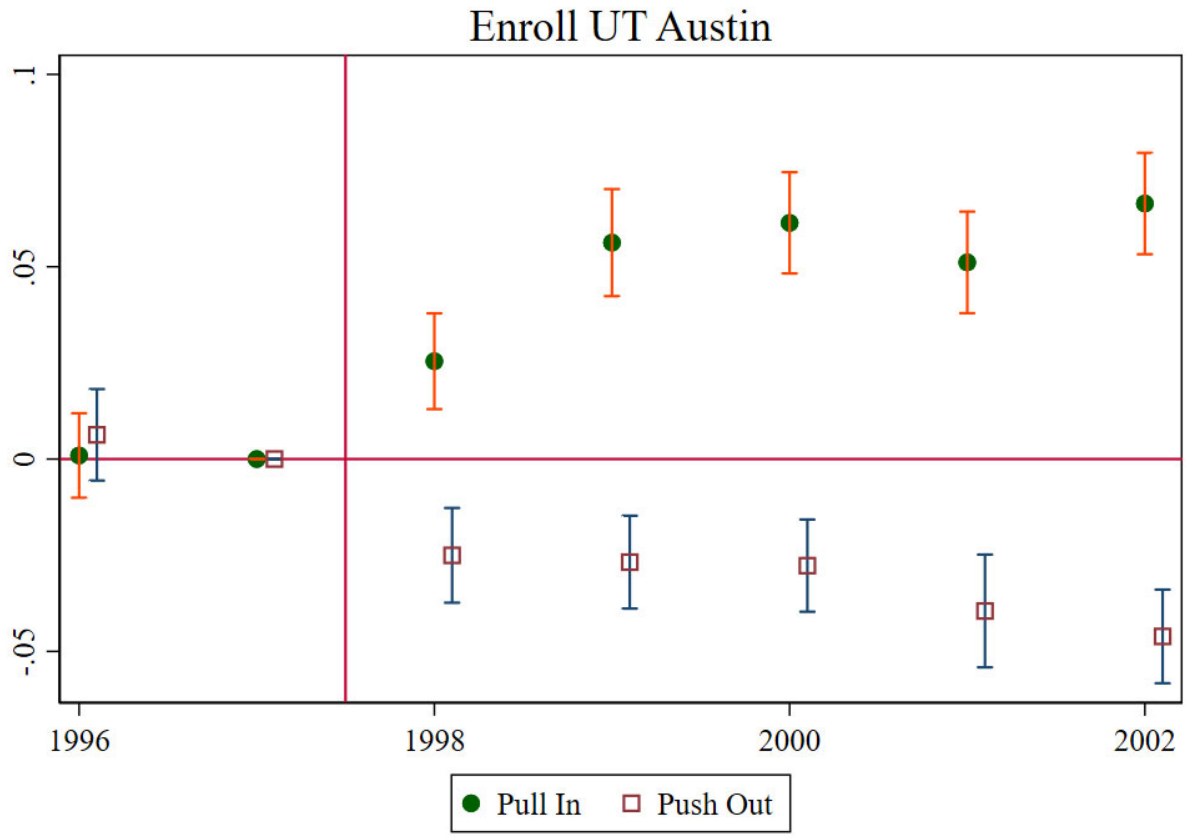
Notes: High schools are divided into deciles based on the share of students who attended UT Austin in the 1996 and 1997 graduating classes. Deciles 1 and 2 have very low UT Austin sending shares, and are suppressed for disclosure avoidance.

Figure 5. Share of students attending UT Austin by school sending group (s) and predicted top ten probability bin ($\hat{q}$)



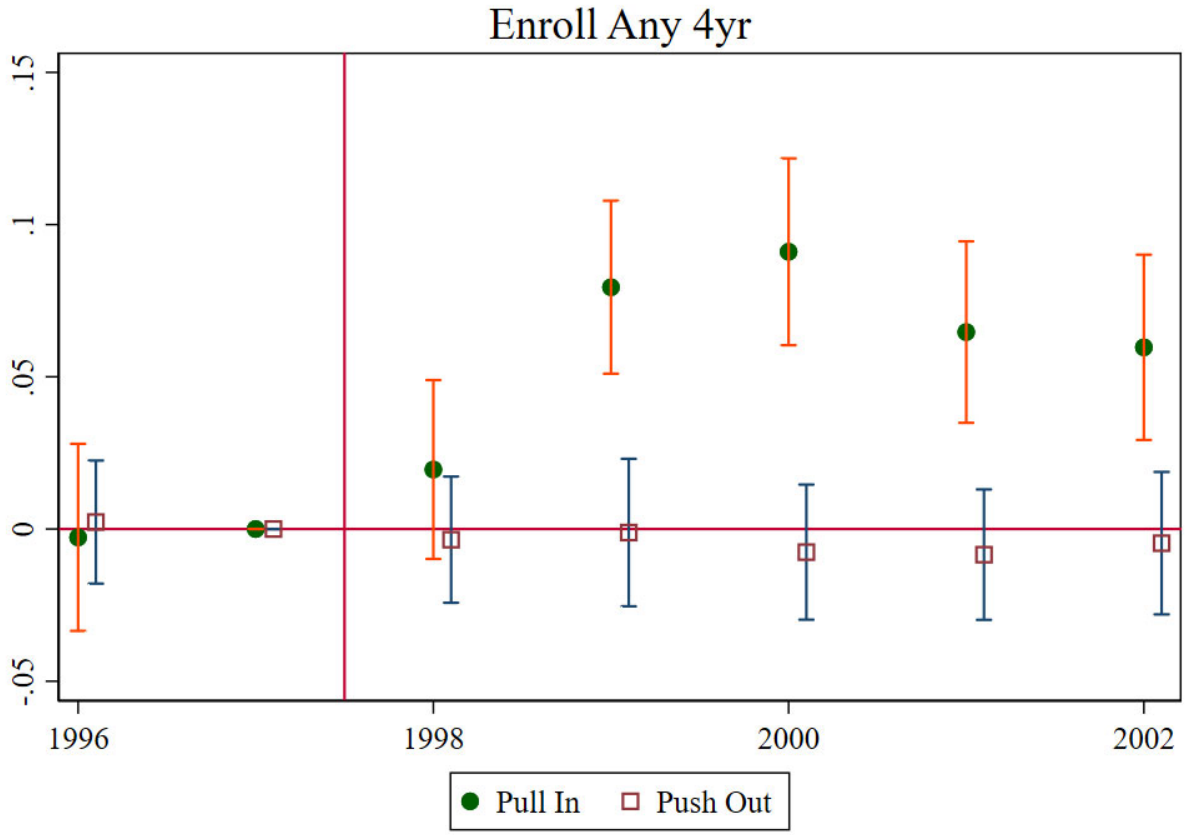
Notes: X-axis represents the high school pre-TTP UT Austin share decile, as illustrated in Figure 4. Y-axis represents the $\hat{q}$ bins of the student's predicted top-ten-percent probability, based on a random forest model fit to 1999-2002 data. Panels A and B show the difference between the two periods; Panel B overlays the definitions of treatment and control groups using the algorithm described in the text. Cells shaded grey are suppressed for disclosure avoidance, due to too few students attending UT. Treatment and control groups illustrated in Panel B are defined based on the unsuppressed values.

Figure 6. Event study of effect of TTP on UT Austin enrollment for students in Pulled In and Pushed Out groups



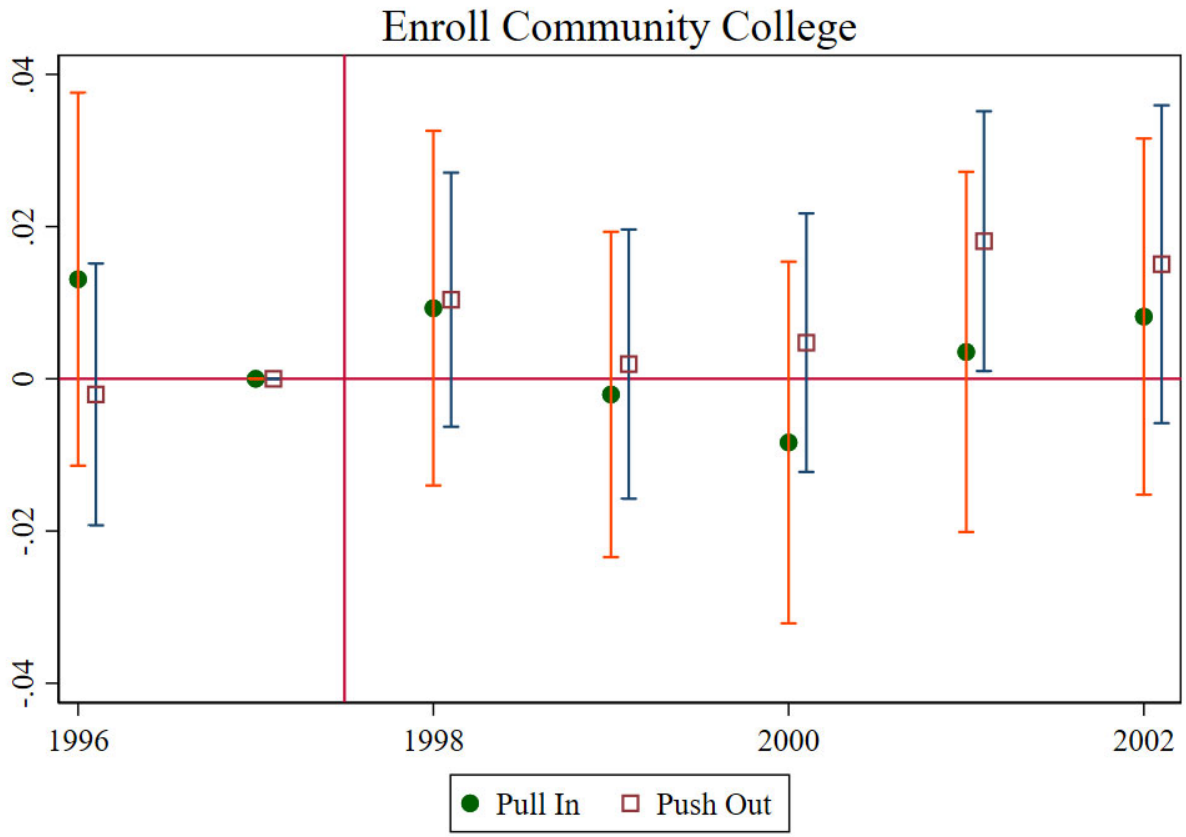
Notes: Figures shows point estimates and 95% confidence intervals for $PulledIn_{it} * Year_t$ and $PushedOut_{it} * Year_t$ interactions in an event study version of Equation 1 that replaces the $Post_{it}$ indicator with a set of $Year_t$ indicators. 1997 is the excluded category. Dependent variable is an indicator for enrollment at UT Austin following high school. Standard errors are clustered at the school district level. The Pushed Out coefficients have been shifted slightly horizontally to aid it in readability.

Figure 7. Event study of effect of TTP on 4-year college enrollment for students in Pulled In and Pushed Out groups



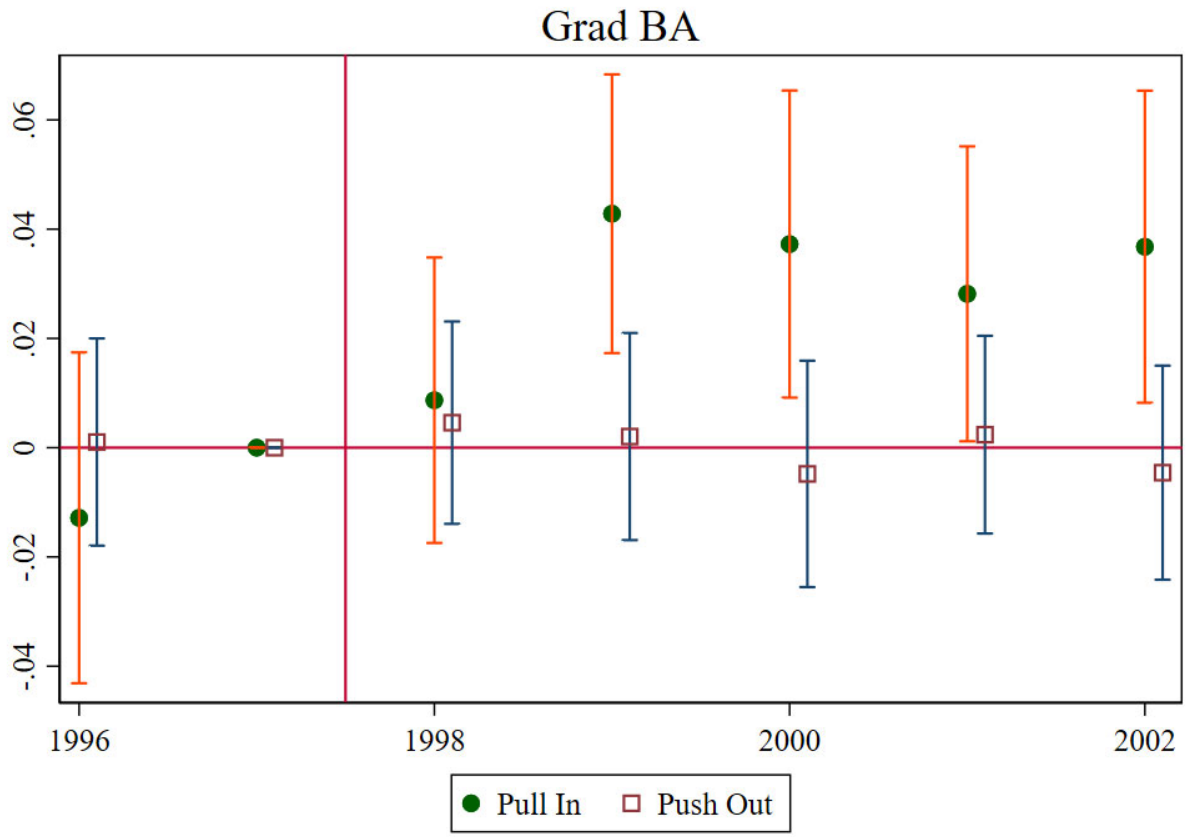
Notes: See notes to Figure 6. Dependent variable is an indicator for enrollment at any Texas public four-year college following high school.

Figure 8. Event study of effect of TTP on 2-year college enrollment for students in Pulled In and Pushed Out groups



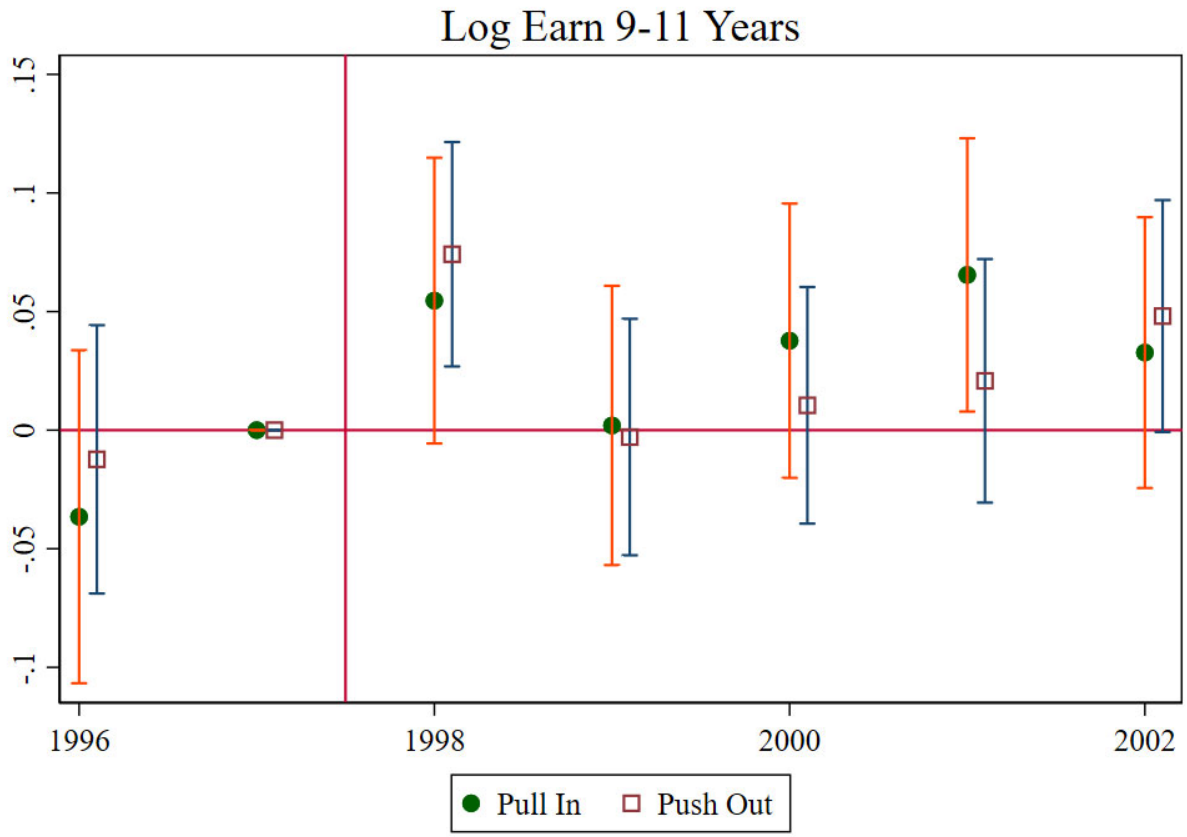
Notes: See notes to Figure 6. Dependent variable is an indicator for enrollment at any Texas public community college following high school.

Figure 9. Event study of effect of TTP on college completion (BA attainment) for students in Pulled In and Pushed Out groups



Notes: See notes to Figure 6. Dependent variable is an indicator for attainment of a bachelor's degree from any Texas public institution within six years following high school.

Figure 10. Event study of effect of TTP on earnings 9-11 years after high school completion for students in Pulled In and Pushed Out groups



Notes: See notes to Figure 6. Dependent variable is the log of average earnings in the 9th, 10th, and 11th calendar years following high school. Years in which no earnings are recorded are excluded from the average; students with no recorded earnings in any year are excluded from the analysis.

Table 1. Summary Statistics

	Full sample		UT Enrollees	
	1996-1997	1998-2002	1996-1997	1998-2002
	(1)	(2)	(3)	(4)
N	258,240	766,962	8,633	24,805
Demographics & high school characteristics				
Hispanic	28%	30%	13%	14%
Black	12%	12%	3%	3%
Asian	3%	3%	14%	16%
Female	52%	52%	50%	52%
Free/Reduced Lunch	23%	22%	5%	7%
Math score (statewide %ile)	50.3	50.7	81.5	81.2
Reading score (statewide %ile)	50.3	50.6	77.9	78.5
Writing score (statewide %ile)	50.4	50.6	77.6	78.1
Math score (%ile within school)	44.4	45.8	67.5	70.7
Reading score (%tile within school	44.2	45.7	64.2	67.6
Writing score (%ile within school)	43.9	45.6	63.8	67.5
Number of AP courses taken	0.78	1.36	3.24	5.16
Days absent	8.4	8.4	5.8	5.5
School: Fr. Black	12%	12%	10%	10%
School: Fr. Hispanic	30%	29%	21%	22%
School: Fr. FRL	22%	22%	13%	14%
School: Attend UT Austin	3.3%	3.4%	7.8%	7.3%
P & S variables				
Actually Top 10 (applicants only)		23.0%		53.5%
Applied & top 10 (all)		6.1%		43.3%
Fitted pr(Top 10)	8.4%	9.4%	28.5%	36.0%
School pre-policy UT rate				
Bottom two deciles	20.0%	19.5%	0.8%	5.0%
Top two deciles	20.2%	20.9%	58.4%	53.7%
Enrollment				
UT Austin	3.3%	3.2%	100%	100%
Any 4-year	25%	24%	100%	100%
Community College	32%	31%	4%	4%
Any college	55%	53%	100%	100%
Six-year Graduation				
UT Austin	2.9%	2.9%	70%	75%
Any BA	18%	18%	74%	79%
Any degree	23%	24%	74%	79%
Employment outcomes				
9-11 Years After HS Graduation				
Employment (0/1)	76.0%	74.9%	76.7%	76.3%
Average wage (with 0s)	$25,816	$24,708	$35,376	$35,687
Average wage (excluding 0s)	$35,487	$34,547	$49,212	$49,971
13-15 Years After HS Graduation				
Employment (0/1)	71.1%	71.4%	69.5%	70.7%
Average wage (with 0s)	$29,688	$30,210	$43,257	$46,405
Average wage (excluding 0s)	$43,567	$44,075	$66,598	$69,820

Notes: Years at top of columns refer to the high school graduating class.

Table 2. Summary Statistics for Treatment and Control Groups

	Pulled in		Pushed out		Control	
	1996-97	1998-02	1996-97	1998-02	1996-97	1998-02
	(1)	(2)	(3)	(4)	(5)	(6)
N	5,004	19,809	13,917	41,065	31,137	90,235
Demographics & high school characteristics						
Hispanic	20%	24%	13%	15%	26%	29%
Black	9%	8%	5%	6%	7%	9%
Asian	3%	3%	8%	9%	3%	3%
Female	53%	57%	51%	52%	52%	52%
Free/Reduced Lunch	18%	20%	7%	8%	19%	19%
Math score (statewide %ile)	88.8	88.3	79.6	77.7	66.2	65.6
Reading score (statewide %ile)	84.9	83.8	75.1	74.3	64.6	63.2
Writing score (statewide %ile)	86.7	85.4	75.1	74.3	64.2	63.5
Math score (%ile within school)	81.2	81.4	64.3	66.6	59.2	60.0
Reading score (%tile within school)	77.5	78.1	60.4	62.7	57.3	57.6
Writing score (%ile within school)	78.0	78.5	60.4	63.2	56.4	57.5
Number of AP courses taken	2.40	4.39	2.09	3.16	0.86	1.53
Days absent	3.6	3.6	6.7	7.3	7.7	7.8
School: Fr. Black	15%	14%	10%	10%	11%	11%
School: Fr. Hispanic	32%	33%	19%	20%	31%	31%
School: Fr. FRL	30%	30%	11%	11%	22%	22%
School: Attend UT Austin	0.6%	0.6%	8.8%	8.6%	2.6%	2.6%
P & S variables						
Actually Top 10 (applicants only)		85.9%		23.1%		13.4%
Applied & top 10 (all)		47.7%		10.5%		4.0%
Fitted pr(Top 10)	69.6%	72.5%	13.9%	14.6%	6.5%	6.6%
School pre-policy UT rate						
Bottom two deciles	65.3%	62.8%	0.0%	0.0%	0.0%	0.0%
Top two deciles	0.0%	0.0%	73.4%	71.5%	0.0%	0.0%
Enrollment						
UT Austin	4%	9%	17%	12%	2.7%	2.1%
Any 4-year	49%	53%	47%	44%	33%	30%
Community College	18%	18%	16%	19%	32%	33%
Any college	65%	69%	61%	60%	63%	61%
Six-year Graduation						
UT Austin	3.6%	7.4%	13.7%	10.9%	2.3%	1.9%
Any BA	44%	47%	39%	37%	25%	23%
Any degree	48%	52%	41%	40%	31%	31%
Employment outcomes						
9-11 Years After HS Graduation						
Employment (0/1)	77.7%	76.3%	70.3%	69.9%	76.5%	75.3%
Average wage (with 0s)	$34,587	$33,573	$29,857	$28,647	$27,992	$26,484
Average wage (excluding 0s)	$46,682	$46,083	$44,951	$43,333	$38,189	$36,735
13-15 Years After HS Graduation						
Employment (0/1)	73.1%	72.3%	63.9%	65.3%	71.4%	71.5%
Average wage (with 0s)	$42,715	$41,907	$35,495	$36,197	$32,222	$32,280
Average wage (excluding 0s)	$61,142	$60,588	$58,608	$58,425	$46,925	$46,922

Notes: Years at top of columns refer to the high school graduating class. Groups are defined as explained in text, using the baseline algorithm and random forest predictions.

Table 3. Baseline Difference-in-Difference Analysis

	DD estimates		Pre-policy means [SDs]		
	Pulled In	Pushed Out	Control	Pulled In	Pushed Out
	(1)	(2)	(3)	(4)	(5)
Enrollment outcomes					
UT Austin	0.053	-0.036	0.03	0.04	0.17
	(0.005)	(0.004)	[0.16]	[0.19]	[0.37]
Texas A&M	-0.010	0.008	0.04	0.14	0.09
	(0.007)	(0.004)	[0.19]	[0.34]	[0.28]
Any college	0.052	0.000	0.63	0.65	0.61
	(0.011)	(0.007)	[0.48]	[0.48]	[0.49]
Any 4-year	0.066	-0.006	0.33	0.49	0.47
	(0.011)	(0.008)	[0.47]	[0.50]	[0.50]
Any community college	-0.004	0.011	0.32	0.18	0.16
	(0.008)	(0.005)	[0.47]	[0.38]	[0.37]
Any other 4-year	0.024	0.022	0.26	0.31	0.22
	(0.009)	(0.006)	[0.44]	[0.46]	[0.41]
Characteristics of institution attended (fixed at pre-policy levels)					
Graduation rate (conditional on enrollment)	0.019	-0.006	0.41	0.53	0.55
	(0.005)	(0.004)	[0.18]	[0.20]	[0.22]
Graduation rate (with non-enrollment as institution)	0.038	-0.006	0.27	0.36	0.36
	(0.006)	(0.005)	[0.23]	[0.28]	[0.30]
Math state %ile (conditional on enrollment)	1.47	-1.14	57.1	64.2	68.0
	(0.33)	(0.26)	[12.2]	[13.1]	[13.7]
Math state %ile (with non-enrollment as institution)	1.87	-0.84	52.4	57.3	58.9
	(0.32)	(0.24)	[11.4]	[14.1]	[15.7]
Instructional expenditures per student (conditional on enrollment)	656	-406	7,270	8,899	10,774
	(112)	(83)	[3,523]	[4,234]	[4,921]
Average log earnings in years 9-11 (conditional on enrollment)	0.011	0.000	10.13	10.21	10.24
	(0.005)	(0.003)	[0.17]	[0.20]	[0.16]
Average log earnings in years 9-11 (with non-enrollment as an institution)	0.026	-0.001	10.01	10.07	10.08
	(0.005)	(0.004)	[0.20]	[0.24]	[0.24]
Degree attainment within 6 years					
Bachelors from UT Austin	0.039	-0.021	0.02	0.04	0.14
	(0.004)	(0.003)	[0.15]	[0.19]	[0.34]
Bachelors from any institution	0.037	-0.001	0.25	0.44	0.39
	(0.010)	(0.006)	[0.43]	[0.50]	[0.49]
Associates or better	0.032	-0.006	0.31	0.48	0.41
	(0.010)	(0.007)	[0.46]	[0.50]	[0.49]
Bachelors with STEM major	-0.007	-0.001	0.03	0.13	0.07
	(0.006)	(0.003)	[0.16]	[0.33]	[0.26]
Labor market outcomes 9-11 years after HS graduation					
Employment (0/1)	-0.000	0.008	0.77	0.78	0.70
	(0.009)	(0.006)	[0.42]	[0.42]	[0.46]
Average annual earnings (excluding 0s)	692	-122	38,189	46,682	44,951
	(525)	(479)	[24,104]	[27,594]	[28,867]
Average annual earnings (including 0s)	359	305	27,992	34,587	29,857
	(596)	(384)	[26,329]	[30,645]	[30,979]
Log average annual earnings (excluding 0s)	0.055	0.036	10.24	10.47	10.38
	(0.019)	(0.017)	[1.01]	[0.99]	[1.08]
Labor market outcomes 13-15 years after HS graduation					
Employment (0/1)	-0.007	0.012	0.71	0.73	0.64
	(0.010)	(0.006)	[0.45]	[0.44]	[0.48]
Average annual earnings (excluding 0s)	-977	-68	46,925	61,142	58,608
	(836)	(675)	[31,818]	[39,690]	[40,995]
Average annual earnings (including 0s)	-1130	719	32,222	42,715	35,495
	(878)	(506)	[33,550]	[41,888]	[41,346]
Log average annual earnings (excluding 0s)	-0.004	-0.004	10.43	10.73	10.64
	(0.020)	(0.017)	[1.04]	[0.98]	[1.05]

Notes: Each row represents a separate difference-in-differences regression. Standard errors, clustered at the school district, in parentheses; standard deviations in square brackets. N=201,167 for DD specifications, with smaller samples when outcome is not available for all observations (e.g., log earnings).

Table 4. Baseline difference-in-differences analysis of alternative outcomes

	DD estimates		Pre-policy means [SDs]		
	Pulled In	Pushed Out	Control	Pulled In	Pushed Out
	(1)	(2)	(3)	(4)	(5)
Any activity in data					
Year 2	-0.006	-0.003	0.85	0.89	0.84
	(0.010)	(0.007)	[0.35]	[0.31]	[0.36]
Year 3	-0.005	-0.005	0.83	0.86	0.81
	(0.010)	(0.007)	[0.38]	[0.35]	[0.39]
Year 4	0.008	-0.004	0.81	0.84	0.79
	(0.010)	(0.007)	[0.39]	[0.37]	[0.41]
Year 5	0.001	0.003	0.81	0.82	0.77
	(0.010)	(0.007)	[0.40]	[0.38]	[0.42]
Year 6	0.005	0.006	0.78	0.78	0.73
	(0.010)	(0.007)	[0.41]	[0.41]	[0.45]
Year 7	0.011	-0.002	0.76	0.75	0.70
	(0.010)	(0.006)	[0.43]	[0.43]	[0.46]
Year 8	0.006	-0.002	0.74	0.74	0.68
	(0.010)	(0.006)	[0.44]	[0.44]	[0.47]
Year 9	0.004	0.007	0.73	0.74	0.66
	(0.009)	(0.007)	[0.44]	[0.44]	[0.47]
Log annual earnings, winsorized (excluding years with zeros)					
Year 9	0.019	-0.006	10.19	10.43	10.35
	(0.020)	(0.021)	[1.07]	[0.97]	[1.06]
Year 10	0.064	0.022	10.29	10.51	10.44
	(0.021)	(0.019)	[1.06]	[1.03]	[1.10]
Year 11	0.039	0.010	10.36	10.60	10.53
	(0.020)	(0.022)	[1.04]	[0.98]	[1.07]
Year 12	-0.010	0.010	10.42	10.70	10.60
	(0.019)	(0.020)	[1.02]	[0.92]	[1.05]
Year 13	-0.000	-0.003	10.45	10.74	10.65
	(0.022)	(0.020)	[1.05]	[0.98]	[1.03]
Year 14	-0.013	-0.004	10.48	10.77	10.70
	(0.021)	(0.019)	[1.05]	[0.98]	[1.05]
Year 15	-0.007	-0.007	10.53	10.81	10.74
	(0.020)	(0.015)	[1.02]	[0.97]	[1.05]
Log average earnings, excluding years with zeros & winsorizing at 5th and 95th percentiles					
Log(9-11 average)	0.040	0.022	10.28	10.49	10.41
	(0.015)	(0.014)	[0.80]	[0.77]	[0.84]
Log(13-15 average)	-0.004	0.002	10.47	10.73	10.65
	(0.016)	(0.013)	[0.82]	[0.78]	[0.83]
Log(9-15 average)	0.017	0.020	10.37	10.61	10.51
	(0.014)	(0.013)	[0.77]	[0.74]	[0.80]
Log annual earnings, including years with zeros					
9-11 years	0.065	0.045	10.11	10.34	10.22
	(0.022)	(0.020)	[1.21]	[1.18]	[1.30]
13-15 years	0.002	-0.000	10.31	10.62	10.50
	(0.024)	(0.019)	[1.23]	[1.14]	[1.23]
9-15 years	0.002	0.008	10.37	10.64	10.54
	(0.015)	(0.015)	[0.84]	[0.78]	[0.88]

Notes: Each row represents a separate difference-in-differences regression. Standard errors, clustered at the school district, in parentheses; standard deviations in square brackets. N=201,167 for DD specifications, with smaller samples when outcome is not available for all observations. "Any activity" in a year is measured by presence of the individual in college enrollment records in either the fall or spring semester, or in earnings records in any quarter of the academic year (Q3 to Q2).

Table 5. Sensitivity of Difference-in-Difference Estimates to Controls

	Pulled In						Pushed Out					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Controls												
Individual demographics		y	y	y	y	y		y	y	y	y	y
School demographics			y						y			
S group dummies				y	y	y				y	y	y
Predicted top-ten probability				y	y	y (cubic)				y	y	y (cubic)
s-by-p dummies					y	y					y	y
Enrollment outcomes												
UT Austin	0.057	0.057	0.057	0.050	0.054	0.053	-0.038	-0.038	-0.036	-0.036	-0.036	-0.036
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Any college	0.054	0.052	0.052	0.053	0.051	0.052	0.001	0.001	-0.001	-0.000	-0.000	0.000
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Any 4-year	0.074	0.075	0.074	0.065	0.066	0.066	-0.006	-0.006	-0.005	-0.007	-0.007	-0.006
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Degree attainment within 6 years												
Bachelors from UT Austin	0.043	0.043	0.043	0.037	0.041	0.039	-0.023	-0.023	-0.021	-0.021	-0.021	-0.021
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bachelors from any	0.051	0.048	0.048	0.037	0.038	0.037	0.001	-0.000	0.002	-0.001	-0.001	-0.001
institution	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Associates or better	0.044	0.040	0.039	0.031	0.032	0.032	-0.004	-0.005	-0.005	-0.006	-0.006	-0.006
	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Labor market outcomes 9-11 years after HS graduation												
Employment (0/1)	-0.001	-0.001	-0.001	0.002	0.000	-0.000	0.009	0.009	0.006	0.008	0.008	0.008
	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Log average annual	0.060	0.065	0.064	0.051	0.053	0.055	0.035	0.036	0.039	0.035	0.035	0.036
earnings (excluding 0s)	(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.017)
Labor market outcomes 13-15 years after HS graduation												
Employment (0/1)	-0.009	-0.009	-0.008	-0.006	-0.007	-0.007	0.012	0.013	0.010	0.012	0.012	0.012
	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Log average annual	0.002	0.011	0.012	-0.006	-0.002	-0.004	-0.005	-0.003	0.001	-0.004	-0.005	-0.004
earnings (excluding 0s)	(0.022)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)

Notes: Each row presents estimates from six separate difference-in-differences regressions: One model is presented in columns 1 and 7, another in 2 and 8, and so on. Standard errors, clustered at the school district, in parentheses. All specifications include calendar year and group (Pulled In, Pushed Out) indicators. Individual demographics are indicators for race (Hispanic, Black, Asian, other), gender, free or reduced price lunch, and immigrant status. School demographics are the school fraction Black, Hispanic, Asian, free or reduced price lunch, English language learner, or special education.

Table 6. Heterogeneity

		By race				By free lunch status			By gender		
		Black	Hisp.	White/ Asian	p	No	Yes	p	Fem.	Male	p
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enrollment outcomes											
UT Austin	Pulled In	0.07	0.06	0.04	0.03	0.05	0.07	0.02	0.06	0.05	0.16
		(0.01)	(0.01)	(0.01)		(0.00)	(0.01)		(0.01)	(0.01)	
	Pushed Out	-0.03	-0.05	-0.03	0.43	-0.04	-0.04	0.71	-0.03	-0.04	0.17
		(0.02)	(0.01)	(0.00)		(0.00)	(0.01)		(0.01)	(0.01)	
Any college	Pulled In	0.07	0.08	0.04	0.20	0.05	0.05	0.77	0.06	0.05	0.49
		(0.03)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.04	-0.03	0.00	0.06	0.00	-0.01	0.37	-0.00	0.00	0.59
		(0.02)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Any 4-year	Pulled In	0.05	0.07	0.07	0.92	0.07	0.04	0.15	0.08	0.05	0.13
		(0.03)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.03	-0.03	-0.00	0.11	-0.00	-0.03	0.19	-0.01	-0.00	0.68
		(0.03)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Degree attainment within 6 years											
BA from UT Austin	Pulled In	0.05	0.05	0.03	0.06	0.04	0.04	0.33	0.04	0.03	0.20
		(0.01)	(0.01)	(0.01)		(0.00)	(0.01)		(0.01)	(0.01)	
	Pushed Out	-0.02	-0.02	-0.02	0.87	-0.02	-0.02	1.00	-0.02	-0.02	0.53
		(0.01)	(0.01)	(0.00)		(0.00)	(0.01)		(0.00)	(0.00)	
BA from any institution	Pulled In	0.06	0.04	0.04	0.39	0.03	0.06	0.08	0.06	0.02	0.02
		(0.02)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.02	-0.01	0.00	0.27	0.00	0.01	0.37	-0.00	0.00	0.46
		(0.02)	(0.01)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Associates or better	Pulled In	0.06	0.04	0.03	0.44	0.02	0.07	0.02	0.05	0.01	0.03
		(0.02)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.02	-0.02	-0.00	0.37	-0.01	0.01	0.25	-0.01	-0.00	0.41
		(0.02)	(0.01)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Labor market outcomes 9-11 years after HS graduation											
Employment (0/1)	Pulled In	-0.04	0.02	-0.00	0.06	-0.01	0.04	0.02	0.00	-0.01	0.42
		(0.02)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.01	0.01	0.01	0.96	0.01	-0.01	0.45	0.00	0.01	0.47
		(0.02)	(0.01)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Log average annual earnings	Pulled In	0.10	0.04	0.05	0.78	0.04	0.10	0.16	0.08	0.03	0.17
		(0.07)	(0.04)	(0.02)		(0.02)	(0.04)		(0.03)	(0.03)	
	Pushed Out	0.02	0.05	0.03	0.91	0.03	0.09	0.29	0.03	0.04	0.67
		(0.06)	(0.04)	(0.02)		(0.02)	(0.05)		(0.02)	(0.02)	
Labor market outcomes 13-15 years after HS graduation											
Employment (0/1)	Pulled In	-0.02	0.02	-0.01	0.14	-0.02	0.03	0.01	-0.00	-0.01	0.75
		(0.03)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
	Pushed Out	0.01	0.01	0.01	0.92	0.01	0.01	0.94	0.01	0.02	0.63
		(0.02)	(0.02)	(0.01)		(0.01)	(0.02)		(0.01)	(0.01)	
Log average annual earnings	Pulled In	-0.06	0.00	-0.00	0.63	-0.02	0.04	0.26	0.02	-0.02	0.33
		(0.06)	(0.05)	(0.03)		(0.02)	(0.04)		(0.03)	(0.03)	
	Pushed Out	-0.03	0.00	-0.01	0.93	-0.00	-0.01	0.88	0.01	-0.01	0.55
		(0.07)	(0.04)	(0.02)		(0.02)	(0.05)		(0.02)	(0.02)	

Notes: Each stacked pair of entries (Pulled In and Pushed Out) represents a separate difference-in-differences regression, fit to the indicated subgroup. Specifications are as in Table 3, but for the sample changes. Columns 4, 7, and 10 present p-values for the hypothesis that the coefficients are identical across the preceding groups (which are mutually exclusive).

Table 7. Alternative Definitions of Treatment Groups

	Baseline		LASSO	
	Pulled In	Pushed Out	Pulled In	Pushed Out
	(1)	(2)	(3)	(4)
Enrollment outcomes				
UT Austin	0.053	-0.036	0.016	-0.016
	(0.005)	(0.004)	(0.002)	(0.002)
Any college	0.052	0.000	0.030	0.010
	(0.011)	(0.007)	(0.005)	(0.005)
Any 4-year	0.066	-0.006	0.017	-0.013
	(0.011)	(0.008)	(0.005)	(0.005)
Degree attainment within 6 years				
Bachelors from UT Austin	0.039	-0.021	0.011	-0.010
	(0.004)	(0.003)	(0.002)	(0.002)
Bachelors from any institution	0.037	-0.001	0.015	-0.004
	(0.010)	(0.006)	(0.004)	(0.004)
Associates or better	0.032	-0.006	0.016	-0.002
	(0.010)	(0.007)	(0.005)	(0.004)
Labor market outcomes 9-11 years after HS graduation				
Employment (0/1)	-0.000	0.008	0.002	0.011
	(0.009)	(0.006)	(0.004)	(0.004)
Log average annual earnings (excluding 0s)	0.055	0.036	0.016	-0.008
	(0.019)	(0.017)	(0.009)	(0.009)
Labor market outcomes 13-15 years after HS graduation				
Employment (0/1)	-0.007	0.012	-0.004	0.009
	(0.010)	(0.006)	(0.004)	(0.004)
Log average annual earnings (excluding 0s)	-0.004	-0.004	0.014	0.009
	(0.020)	(0.017)	(0.011)	(0.009)

Notes: Columns 1-2 repeat specifications from Table 3. Columns 3-4 present identical specifications, this time defininig treatment and control groups based on LASSO estimates described in text. Standard errors, clustered at the school district, in parentheses.

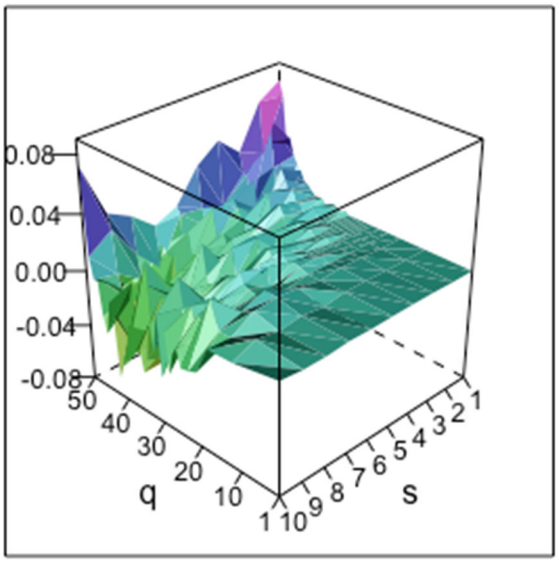
Table 8. Continuous measures of pulled in/pushed out

	Random forest		Logit	
	Increase in Pr(UT Austin) - pulled in	Reduction in Pr(UT Austin) - pushed out	Increase in Pr(UT Austin) - pulled in	Reduction in Pr(UT Austin) - pushed out
	(1)	(2)	(3)	(4)
Enrollment outcomes				
UT Austin	0.078	-0.040	0.072	-0.068
	(0.007)	(0.005)	(0.006)	(0.006)
Any college	0.106	0.009	0.060	0.008
	(0.015)	(0.009)	(0.015)	(0.012)
Any 4-year	0.103	-0.018	0.061	-0.025
	(0.016)	(0.009)	(0.016)	(0.011)
Degree attainment within 6 years				
Bachelors from any institution	0.062	-0.008	0.038	-0.010
	(0.014)	(0.006)	(0.012)	(0.008)
Bachelors with STEM major	-0.014	-0.003	-0.013	-0.010
	(0.009)	(0.003)	(0.009)	(0.004)
Labor market outcomes 9-11 years after HS graduation				
Employment (0/1)	0.014	0.019	0.004	0.029
	(0.012)	(0.008)	(0.011)	(0.009)
Log average annual earnings (excluding 0s)	0.073	0.009	0.046	0.001
	(0.028)	(0.018)	(0.026)	(0.021)
Labor market outcomes 13-15 years after HS graduation				
Employment (0/1)	-0.001	0.022	-0.012	0.027
	(0.012)	(0.008)	(0.012)	(0.010)
Log average annual earnings (excluding 0s)	0.018	0.010	0.018	0.005
	(0.029)	(0.016)	(0.028)	(0.019)

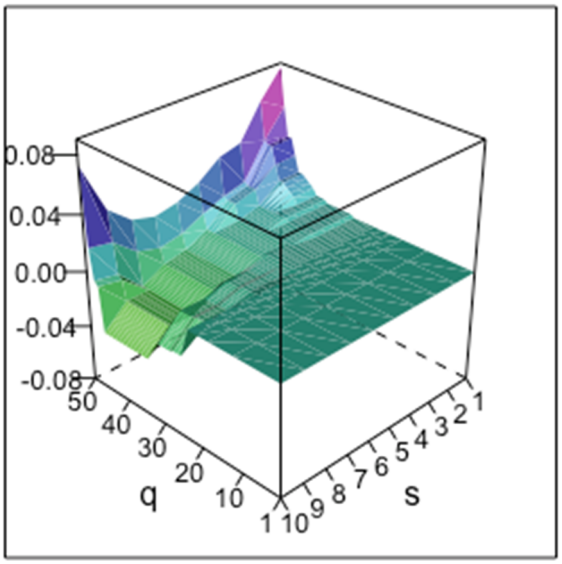
Note: Table reports estimates from the same difference-in-differences specification as in Table 3, except that the Pulled In and Pushed Out variables (and their interactions) are defined to range continuously from 0 to 1 rather than being binary. In columns 3-4, the predicted top-10% probability is obtained from a logit rather than a random forest model.

Appendix Figure 1. LASSO-smoothed changes in UT Austin enrollment shares

A. Raw

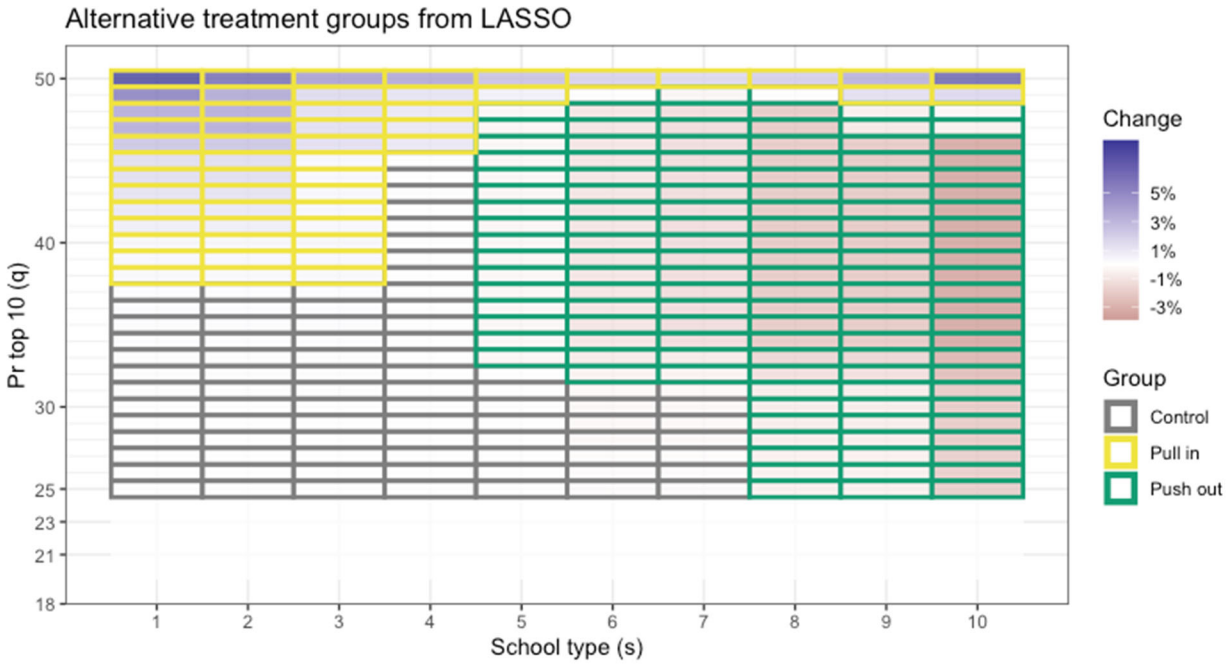


B. With LASSO regularization



Notes: Panel A shows the same data that is displayed in Figure 5A; the vertical axis is the change in the share of students in the cell who attended UT Austin between 1996-1997 and 1998-2002. Panel B presents estimates obtained by fitting a LASSO model to the data in Panel A, as described in the Appendix, then fitting OLS using only the explanatory variables whose estimated LASSO coefficients are not zero.

Appendix Figure 2. Pulled in, pushed out, and control groups using LASSO



Notes: Axes are defined as in Figure 5, though only the upper half ($\hat{q} \geq 18$)is shown. Outlines indicate treatment and control groups selected using the LASSO method described in the text.

Appendix Table 1. Predicting the Top 10%

	Summary statistics		Logit			Random forest predictions (OLS)	
	Mean	SD	Coeff.	SE	Marg. effect	Coeff.	SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Math score (statewide percentile/100)	0.51	[0.29]	0.35	(0.23)	0.025	0.067	(0.027)
Reading score (statewide percentile/100)	0.51	[0.29]	-0.02	(0.27)	-0.001	0.023	(0.020)
Writing score (statewide percentile/100)	0.51	[0.29]	0.08	(0.28)	0.005	-0.043	(0.016)
Math score (percentile within school/100)	0.46	[0.26]	3.05	(0.23)	0.220	0.222	(0.025)
Reading score (percentile within school/100)	0.46	[0.26]	2.32	(0.28)	0.167	0.271	(0.015)
Writing score (percentile within school)/100	0.46	[0.26]	2.09	(0.26)	0.151	0.162	(0.020)
Indicator for taking regular math in 9th grade	0.35	[0.48]	-0.47	(0.07)	-0.032	-0.056	(0.008)
Indicator for taking regular math in 10th grade	0.51	[0.50]	0.08	(0.04)	0.006	0.012	(0.005)
Indicator for taking regular math in 11th grade	0.52	[0.50]	0.06	(0.05)	0.004	-0.006	(0.004)
Indicator for taking regular math in 12th grade	0.27	[0.44]	0.39	(0.05)	0.029	0.025	(0.005)
Indicator for taking advanced math in 9th grade	0.14	[0.34]	-0.10	(0.07)	-0.007	-0.023	(0.011)
Indicator for taking advanced math in 10th grade	0.15	[0.36]	0.32	(0.06)	0.025	0.023	(0.007)
Indicator for taking advanced math in 11th grade	0.14	[0.34]	-0.13	(0.07)	-0.009	-0.013	(0.006)
Indicator for taking advanced math in 12th grade	0.07	[0.26]	0.84	(0.07)	0.078	0.266	(0.008)
Indicator for taking any math in 9th grade	0.51	[0.50]	-0.47	(0.07)	-0.036	-0.037	(0.009)
Indicator for taking any math in 10th grade	0.60	[0.49]	-0.48	(0.05)	-0.038	-0.036	(0.005)
Indicator for taking any math in 11th grade	0.61	[0.49]	0.13	(0.05)	0.009	0.015	(0.003)
Indicator for taking any math in 12th grade	0.35	[0.48]	0.19	(0.06)	0.014	0.017	(0.003)
Indicator for taking biology	0.86	[0.35]	0.08	(0.09)	0.006	0.025	(0.004)
Indicator for taking chemistry	0.57	[0.50]	0.13	(0.05)	0.009	-0.001	(0.003)
Indicator for taking physics	0.24	[0.43]	0.07	(0.04)	0.005	-0.014	(0.002)
Indicator for taking remedial algebra	0.17	[0.38]	-1.12	(0.11)	-0.056	-0.037	(0.005)
Number of foreign language courses taken	3.40	[1.99]	0.084	(0.009)	0.006	0.003	(0.000)
Number of days absent	8.35	[8.75]	-0.009	(0.018)	-0.001	-0.003	(0.000)
Percent of schools days absent (/100)	0.05	[0.05]	-8.90	(3.18)	-0.642	-0.260	(0.058)
Percent of school Black (/100)	0.12	[0.17]	1.13	(0.23)	0.082	0.074	(0.015)
Percent of school Hispanic (/100)	0.29	[0.30]	0.11	(0.24)	0.008	-0.005	(0.011)
Percent of school Asian (/100)	0.03	[0.05]	-3.47	(0.69)	-0.250	-0.263	(0.063)
Percent of school other minority race (/100)	0.00	[0.00]	12.00	(4.59)	0.866	0.944	(0.340)
Percent of school on Free/Reduced Lunch (/100)	0.22	[0.21]	0.79	(0.34)	0.057	0.075	(0.019)
Percent of school female (/100)	0.52	[0.03]	1.69	(1.40)	0.122	0.059	(0.090)
Percent of school English Lang. Learner (/100)	0.04	[0.07]	0.93	(0.78)	0.067	0.036	(0.041)
Percent of school Special Education (/100)	0.04	[0.02]	0.99	(0.98)	0.072	0.061	(0.077)
Indicator for being 18 years old	0.16	[0.37]	-0.10	(0.04)	-0.007	0.014	(0.002)
Number of courses failed	0.95	[1.70]	-0.59	(0.09)	-0.042	-0.006	(0.001)

Notes: N=195,710. Columns 6 and 7 present coefficients of an linear (OLS) regression of the predicted values from the nonlinear random forest model on the covariates.

Appendix Table 2. Stability of predicted top 10% probability across specifications

	RF	Logit	RF 1999-2000	RF 2001-2002	RF 25	RF 100
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Full sample						
Random Forest	1					
Logit	0.94	1				
Random Forest (fit to 1999-2000 data)	0.98	0.93	1			
Random Forest (fit to 2001-2002 data)	0.99	0.95	0.96	1		
Random Forest (Leaf=25 obs.)	0.99	0.94	0.97	0.97	1	
Random Forest (Leaf=100 obs.)	0.99	0.93	0.98	0.98	0.97	1
Panel B. 1996-1997 (pre TTP)						
Random Forest	1					
Logit	0.95	1				
Random Forest (fit to 1999-2000 data)	0.99	0.94	1			
Random Forest (fit to 2001-2002 data)	0.99	0.95	0.97	1		
Random Forest (Leaf=25 obs.)	0.99	0.96	0.98	0.98	1	
Random Forest (Leaf=100 obs.)	0.99	0.93	0.98	0.99	0.97	1
Panel C. 1998-2002 (post TTP)						
Random Forest	1					
Logit	0.94	1				
Random Forest (fit to 1999-2000 data)	0.98	0.93	1			
Random Forest (fit to 2001-2002 data)	0.99	0.95	0.96	1		
Random Forest (Leaf=25 obs.)	0.99	0.93	0.97	0.97	1	
Random Forest (Leaf=100 obs.)	0.99	0.92	0.98	0.98	0.97	1

Notes: The baseline random forest model discussed in the text, fit to 1999-2002 data and using a "leaf" size of 50 observations, is labeled "random forest" here. Table shows correlations of predicted values across alternative models. Correlations are estimated at the individual level.

Appendix Table 3. Sensitivity to prediction model

Group:	Pulled In			Pushed Out		
Prediction model:	RF	RF	Logit	RF	RF	Logit
	1999-00	2001-02		1999-00	2001-02	
	(1)	(2)	(3)	(4)	(5)	(6)
Enrollment outcomes						
UT Austin	0.030	0.047	0.049	-0.027	-0.034	-0.036
	(0.007)	(0.006)	(0.004)	(0.008)	(0.006)	(0.003)
Any college	0.017	0.042	0.040	0.010	-0.006	0.007
	(0.012)	(0.014)	(0.010)	(0.011)	(0.009)	(0.007)
Any 4-year	0.018	0.065	0.049	0.005	-0.009	-0.006
	(0.012)	(0.014)	(0.011)	(0.013)	(0.011)	(0.008)
Degree attainment within 6 years						
Bachelors from UT Austin	0.024	0.035	0.034	-0.013	-0.019	-0.023
	(0.006)	(0.006)	(0.003)	(0.007)	(0.007)	(0.003)
Bachelors from any institution	0.012	0.044	0.036	0.007	0.002	0.002
	(0.011)	(0.013)	(0.009)	(0.011)	(0.010)	(0.006)
Bachelors with STEM major	-0.011	-0.003	-0.004	0.010	-0.007	-0.006
	(0.006)	(0.008)	(0.006)	(0.006)	(0.004)	(0.002)
Labor market outcomes 9-11 years after HS graduation						
Employment (0/1)	-0.011	0.004	0.003	0.024	0.007	0.016
	(0.011)	(0.012)	(0.007)	(0.010)	(0.008)	(0.005)
Log average annual earnings (excluding 0s)	0.051	0.045	0.050	0.050	0.040	0.030
	(0.024)	(0.027)	(0.018)	(0.021)	(0.025)	(0.013)
Labor market outcomes 13-15 years after HS graduation						
Employment (0/1)	-0.013	-0.008	-0.006	0.024	0.019	0.016
	(0.011)	(0.013)	(0.008)	(0.009)	(0.011)	(0.006)
Log average annual earnings (excluding 0s)	0.003	0.004	0.020	0.022	-0.006	0.006
	(0.023)	(0.029)	(0.020)	(0.026)	(0.024)	(0.014)

Notes: In columns 1 and 4, predictions of top ten percent status are from a random forest model fit to 1999-2000 data only, while only 1997, 2001, and 2002 data are used to fit the difference-in-differences model. In columns 2 and 5, the random forest model is fit to 2001-2002 data, and 1996, 1999, and 2000 data are used for the difference-in-differences. In columns 3 and 6, predictions are based on a logit model using 1999-2002 data, and all years are used in the difference-in-differences model.

Appendix Table 4. Summary statistics by decile of high school pre-TTP UT sending rate (s)

	Demographics				Mean statewide percentile			Send to UT	
s	Black	Hispanic	Free lunch	ELL	Math	Reading	Writing	Pre	Post
1	18%	28%	31%	3%	46.0	45.7	46.3	NR	0.7%
2	19%	44%	38%	9%	43.7	42.5	44.2	NR	1.1%
3	10%	38%	29%	4%	48.1	47.4	48.2	0.8%	1.2%
4	9%	37%	28%	5%	48.1	47.7	48.3	1.3%	1.5%
5	14%	39%	29%	4%	48.0	48.1	48.8	1.8%	1.9%
6	11%	27%	21%	3%	50.6	50.4	50.3	2.4%	2.2%
7	11%	30%	18%	3%	51.0	51.5	51.3	3.2%	3.0%
8	10%	21%	12%	2%	53.4	54.1	53.5	4.3%	3.7%
9	10%	24%	13%	3%	55.4	55.5	54.4	6.6%	5.8%
10	8%	12%	6%	2%	60.0	60.5	59.1	12.8%	10.9%

Notes: All columns present student-weighted means of school average characteristics, based (except in final column) on pre-TTP data. S=1 represents the decile of high schools with the lowest pre-TTP UT sending rates, and s=10 the decile with the highest rates. "ELL" represents English Language Learners.

Appendix Table 5. Difference-in-Differences Analysis, Including P=50, S=9 and P=50, S=10 Cells in Pulled-In Group

	Pulled In	Pushed Out
	(1)	(2)
Enrollment outcomes		
UT Austin	0.050	-0.036
	(0.004)	(0.004)
Texas A&M	-0.006	0.008
	(0.006)	(0.004)
Any college	0.056	0.000
	(0.010)	(0.007)
Any 4-year	0.069	-0.006
	(0.010)	(0.008)
Any Community College	-0.003	0.011
	(0.008)	(0.005)
Any other 4-year	0.026	0.022
	(0.008)	(0.006)
Characteristics of institution attended (fixed at pre-policy levels)		
Graduation rate (conditional on enrollment)	0.019	-0.006
	(0.005)	(0.004)
Graduation rate (with non-enrollment as institution)	0.040	-0.006
	(0.006)	(0.005)
Math state %ile (conditional on enrollment)	1.29	-1.15
	(0.31)	(0.26)
Math state %ile (with non-enrollment as institution)	1.91	-0.84
	(0.30)	(0.24)
Instructional expenditures per student (conditional on enrollment)	584	-409
	(105)	(82)
Average log earnings in years 9-11 (conditional on enrollment)	0.028	0.000
	(0.005)	(0.003)
Average log earnings in years 9-11 (with non-enrollment as an institution)	0.039	-0.001
	(0.004)	(0.004)
Degree attainment within 6 years		
Bachelors from UT Austin	0.039	-0.021
	(0.004)	(0.003)
Bachelors from any institution	0.043	-0.001
	(0.009)	(0.006)
Associates or better	0.036	-0.006
	(0.009)	(0.007)
Bachelors with STEM major	-0.003	-0.001
	(0.006)	(0.003)
Labor market outcomes 9-11 years after HS graduation		
Employment (0/1)	-0.001	0.008
	(0.009)	(0.006)
Average annual earnings (excluding 0s)	1004	-119
	(503)	(478)
Average annual earnings (including 0s)	507	305
	(567)	(384)
Log average annual earnings (excluding 0s)	0.063	0.036
	(0.018)	(0.017)
Labor market outcomes 13-15 years after HS graduation		
Employment (0/1)	-0.006	0.012
	(0.009)	(0.006)
Average annual earnings (excluding 0s)	-557	-58
	(798)	(675)
Average annual earnings (including 0s)	-758	721
	(814)	(506)
Log average annual earnings (excluding 0s)	0.000	-0.004
	(0.019)	(0.017)

Notes: N=205,693.

Appendix Table 6. Baseline difference-in-differences analysis of alternative outcomes

Graduation outcomes, by year

	DD estimates		Pre-policy means [SDs]		
	Pulled In	Pushed Out	Control	Pulled In	Pushed Out
	(1)	(2)	(3)	(4)	(5)
UT, year 4	0.019	-0.005	0.01	0.02	0.06
	(0.003)	(0.003)	[0.10]	[0.14]	[0.24]
UT, year 5	0.036	-0.017	0.02	0.03	0.12
	(0.004)	(0.003)	[0.14]	[0.17]	[0.32]
UT, year 6	0.039	-0.021	0.02	0.04	0.14
	(0.004)	(0.003)	[0.15]	[0.19]	[0.34]
UT, year 7	0.038	-0.021	0.03	0.04	0.15
	(0.004)	(0.004)	[0.16]	[0.19]	[0.35]
UT, year 8	0.037	-0.019	0.03	0.04	0.15
	(0.004)	(0.004)	[0.16]	[0.20]	[0.36]
University, year 4	0.022	0.006	0.07	0.20	0.15
	(0.007)	(0.005)	[0.26]	[0.40]	[0.36]
University, year 5	0.031	0.003	0.19	0.37	0.32
	(0.009)	(0.006)	[0.39]	[0.48]	[0.47]
University, year 6	0.037	-0.001	0.25	0.44	0.39
	(0.010)	(0.006)	[0.43]	[0.50]	[0.49]
University, year 7	0.035	0.001	0.28	0.47	0.42
	(0.010)	(0.007)	[0.45]	[0.50]	[0.49]
University, year 8	0.038	0.001	0.30	0.49	0.44
	(0.011)	(0.007)	[0.46]	[0.50]	[0.50]
University, STEM, year 4	0.002	0.004	0.01	0.05	0.02
	(0.004)	(0.001)	[0.09]	[0.22]	[0.15]
University, STEM, year 5	-0.006	0.002	0.02	0.11	0.06
	(0.006)	(0.002)	[0.14]	[0.31]	[0.24]
University, STEM, year 6	-0.007	-0.001	0.03	0.13	0.07
	(0.006)	(0.003)	[0.16]	[0.33]	[0.26]
University, STEM, year 7	-0.005	-0.002	0.03	0.13	0.08
	(0.006)	(0.003)	[0.17]	[0.34]	[0.27]
University, STEM, year 8	-0.002	-0.003	0.03	0.13	0.08
	(0.007)	(0.003)	[0.17]	[0.34]	[0.27]

Notes: Each row represents a separate difference-in-differences regression. Standard errors, clustered at the school district, in parentheses; standard deviations in square brackets. N=201,167 for DD specifications, with smaller samples for years 7 and 8, where we do not have graduation outcomes for all cohorts).

Appendix Table 7. Difference-in-Differences Analysis, Excluding High Schools That Participate in Longhorn Opportunity and Century Scholar Programs

	Pulled In	Pushed Out
	(1)	(2)
Enrollment outcomes		
UT Austin	0.047	-0.035
	(0.005)	(0.004)
Texas A&M	-0.011	0.008
	(0.008)	(0.004)
Any college	0.043	0.003
	(0.011)	(0.008)
Any 4-year	0.062	-0.003
	(0.012)	(0.008)
Any Community College	-0.008	0.011
	(0.009)	(0.006)
Any other 4-year	0.026	0.024
	(0.010)	(0.006)
Characteristics of institution attended (fixed at pre-policy levels)		
Graduation rate (conditional on enrollment)	0.017	-0.006
	(0.005)	(0.004)
Graduation rate (with non-enrollment as institution)	0.034	-0.005
	(0.007)	(0.005)
Math state %ile (conditional on enrollment)	1.42	-1.18
	(0.36)	(0.28)
Math state %ile (with non-enrollment as institution)	1.75	-0.81
	(0.35)	(0.26)
Instructional expenditures per student (conditional on enrollment)	583	-399
	(119)	(86)
Average log earnings in years 9-11 (conditional on enrollment)	0.024	0.000
	(0.006)	(0.003)
Average log earnings in years 9-11 (with non-enrollment as an institution)	0.036	-0.000
	(0.005)	(0.004)
Degree attainment within 6 years		
Bachelors from UT Austin	0.036	-0.020
	(0.005)	(0.003)
Bachelors from any institution	0.029	0.000
	(0.012)	(0.006)
Associates or better	0.026	-0.005
	(0.012)	(0.007)
Bachelors with STEM major	-0.013	-0.001
	(0.007)	(0.002)
Labor market outcomes 9-11 years after HS graduation		
Employment (0/1)	-0.001	0.008
	(0.011)	(0.007)
Average annual earnings (excluding 0s)	669	-262
	(605)	(491)
Average annual earnings (including 0s)	311	225
	(699)	(403)
Log average annual earnings (excluding 0s)	0.044	0.032
	(0.021)	(0.018)
Labor market outcomes 13-15 years after HS graduation		
Employment (0/1)	-0.007	0.012
	(0.011)	(0.007)
Average annual earnings (excluding 0s)	-923	-313
	(948)	(665)
Average annual earnings (including 0s)	-1161	613
	(998)	(538)
Log average annual earnings (excluding 0s)	-0.009	-0.007
	(0.023)	(0.018)

Notes: N=180,101.

Appendix Table 8. Bootstrap vs. clustered standard errors

	Pulled In				Pushed Out			
	Baseline		Lasso		Baseline		Lasso	
	Cluster	Bootstrap	Cluster	Bootstrap	Cluster	Bootstrap	Cluster	Bootstrap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Enrollment outcomes								
UT Austin	0.053		0.016		-0.036		-0.016	
	(0.005)	(0.004)	(0.002)	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Any college	0.052		0.030		0.000		0.010	
	(0.011)	(0.011)	(0.005)	(0.008)	(0.007)	(0.008)	(0.005)	(0.005)
Any 4-year	0.066		0.017		-0.006		-0.013	
	(0.011)	(0.011)	(0.005)	(0.008)	(0.008)	(0.008)	(0.005)	(0.006)
Degree attainment within 6 years								
Bachelors from UT Austin	0.039		0.011		-0.021		-0.010	
	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
Bachelors from any institution	0.037		0.015		-0.001		-0.004	
	(0.010)	(0.010)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.004)
Associates or better	0.032		0.016		-0.006		-0.002	
	(0.010)	(0.010)	(0.005)	(0.005)	(0.007)	(0.007)	(0.004)	(0.004)
Labor market outcomes 9-11 years after HS graduation								
Employment (0/1)	0.000		0.002		0.008		0.011	
	(0.009)	(0.008)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.004)
Log average annual earnings (excluding 0s)	0.055		0.016		0.036		-0.008	
	(0.019)	(0.019)	(0.009)	(0.012)	(0.017)	(0.016)	(0.009)	(0.011)
Labor market outcomes 13-15 years after HS graduation								
Employment (0/1)	-0.007		-0.004		0.012		0.009	
	(0.010)	(0.009)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)
Log average annual earnings (excluding 0s)	-0.004		0.014		-0.004		0.009	
	(0.020)	(0.021)	(0.011)	(0.015)	(0.017)	(0.018)	(0.009)	(0.010)

Notes: All specifications use predicted probabilities of being in the top 10% from the random forest model. In odd numbered columns, standard errors of the DD regresson are clustered at the school district level. In even numbered columns, standard errors are computed from bootstrap replications of the entire estimation procedure, from estimation of predicted probabilities to formation of treatment groups and estimation of the DD regression. Bootstrap samples are clustered at the school level; standard errors are the standard deviations across 250 bootstrap replications.

Appendix Table 9. Estimates based on 1996-1999 cohorts only

	Pulled In		Pushed Out	
	Using 1999 to train predictions (1)	Using 1999-2002 to train predictions (2)	Using 1999 to train predictions (3)	Using 1999-2002 to train predictions (4)
Enrollment outcomes				
UT Austin	0.046	0.042	-0.037	-0.029
	(0.006)	(0.005)	(0.005)	(0.004)
Any college	0.046	0.041	-0.017	0.000
	(0.011)	(0.011)	(0.008)	(0.008)
Any 4-year	0.057	0.054	-0.021	-0.003
	(0.011)	(0.011)	(0.008)	(0.009)
Any Community College	-0.002	-0.003	0.008	0.007
	(0.010)	(0.009)	(0.006)	(0.006)
Any other 4-year	0.018	0.026	0.004	0.009
	(0.010)	(0.010)	(0.006)	(0.006)
Degree attainment within 6 years				
Bachelors from UT Austin	0.030	0.033	-0.026	-0.018
	(0.005)	(0.005)	(0.004)	(0.004)
Bachelors from any institution	0.036	0.033	-0.013	0.003
	(0.011)	(0.010)	(0.006)	(0.006)
Bachelors with STEM major	-0.001	-0.003	-0.008	-0.000
	(0.007)	(0.007)	(0.003)	(0.003)
Labor market outcomes 9-11 years after HS graduation				
Employment (0/1)	0.002	-0.004	-0.007	0.000
	(0.009)	(0.009)	(0.008)	(0.006)
Log average annual earnings (excluding 0s)	0.018	0.042	0.011	0.041
	(0.023)	(0.022)	(0.017)	(0.017)
Labor market outcomes 13-15 years after HS graduation				
Employment (0/1)	-0.006	-0.014	-0.002	0.004
	(0.010)	(0.010)	(0.007)	(0.007)
Log average annual earnings (excluding 0s)	0.014	-0.007	0.025	0.006
	(0.026)	(0.024)	(0.019)	(0.020)

Notes: Each row represents a separate difference-in-differences regression. Standard errors, clustered at the school district, in parentheses; standard deviations in square brackets. N=201,167 for DD specifications, with smaller samples for years 7 and 8, where we do not have graduation outcomes for all cohorts).