Affirmative Action and Pre-College Human Capital*

Mitra Akhtari[†] Natalie Bau[‡] Jean-William Laliberté[§]

November 4, 2019

Abstract

Race-based affirmative action policies are widespread in higher education. Despite the prevalence of these policies, there is limited evidence on whether they affect students *before* they reach college. We exploit the 2003 Supreme Court ruling in *Grutter v. Bollinger*, which overturned affirmative action bans in Texas, Louisiana, and Mississippi, but not in other states, to study the effect of affirmative action on high school students' outcomes. We analyze four data sets, including nationwide SAT data and administrative data for the state of Texas. The SAT data allow us to leverage state and time variation in difference-in-differences and synthetic control group analyses. Within Texas, variation in race, time, and ex ante ability further help us to isolate the effects of the policy change on secondary school grades, attendance, and college applications. Across data sets, outcomes, and identification strategies, the results all point toward gains for underrepresented minority students and reductions in the racial achievement gap. These gains were concentrated among students in the top of the ability distribution, who also experienced the largest increases in the returns to pre-college human capital in college admissions due to the policy change. This suggests that students increased their human capital investment in response to increases in the returns to effort.

*We gratefully acknowledge funding from the Lab for Economic Applications and Policy and the Connaught Fund. The Texas Education Agency and a large urban school district provided invaluable administrative data for this project. We are grateful to Josh Angrist, Roland Fryer, Brent Hickman, Caroline Hoxby, Asim Khwaja, Louis-Philippe Morin, Phil Oreopoulos, Sarah Reber, Alex Whalley, Wesley Yin, and seminar and conference participants at the NBER Summer Institute, IZA, Harvard, Brown, UCL, the Ohlstadt workshop, CEA, University of Calgary, and UCLA for their helpful comments. We also thank Graham Beattie for his help with the newslibrary.com database.

[†]Airbnb Data Science. Contact: mitra.akhtari@airbnb.com

[‡]UCLA, CEGA, and CIFAR. Contact: nbau@ucla.edu

[§]University of Calgary. Contact: jeanwilliam.lalibert@ucalgary.ca

1 Introduction

Affirmative action policies that weigh race or ethnicity as one factor in the college admissions process are widespread in higher education in numerous countries, including the United States, Canada, Brazil, and India. In the United States, affirmative action policies in universities have repeatedly been challenged by court cases at the sub-national and national level,[1] and eight states have banned race-based affirmative action at all public universities. Despite the importance of race-based affirmative action policies and the controversy surrounding them, relatively little is known about whether or how these policies affect students *prior* to reaching college. Yet, closing racial achievement gaps in secondary school may be an important tool for reducing inequities in long-term socio-economic outcomes.

Affirmative action policies may affect pre-college human capital investment by incentivizing students, their parents, and their teachers to change their behavior in response to changes in the returns to pre-college human capital. Theoretically, affirmative action policies favoring students from underrepresented minority (URM) groups in the college admissions process have ambiguous average effects on human capital attainment prior to college entry. On the one hand, for very high ability students, affirmative action policies may reduce the returns to pre-college human capital investment by lowering the threshold for college admissions (Coate and Loury, 1993). In this case, affirmative action may disincentivize human capital investments by these students (or their teachers and parents). On the other hand, affirmative action policies may incentivize higher pre-college human capital attainment for URM students – particularly those who are on the margin of being accepted by a selective university – by increasing the probability that increased human capital will translate into college admission (Fryer and Loury, 2005). Even if affirmative action does not directly affect students' perceptions of the likelihood of being admitted, by increasing the observed number of URM students admitted, it may increase aspirations or the perception that selective schools are welcoming to URMs, ultimately increasing pre-college human capital investment. Since the theoretical effects of affirmative action are ambiguous and may also depend on where students are in the ability distribution, we seek to empirically estimate the effects of affirmative action on both the average student and on students in different parts of the test score distribution.

[1] Such cases include: *Regents of the University of California v. Bakke* in 1979, *Hopwood v. Texas* in 1996, *Grutter v. Bollinger* and *Gratz v. Bollinger* in 2003, *Fisher v. University of Texas* in 2013, *Schuette v. Coalition to Defend Affirmative Action* in 2014, *Fisher v. University of Texas* in 2016, and *Students for Fair Admissions v. Harvard* in 2019.

To investigate the effects of affirmative action[2] on human capital investments of high school students, we exploit a natural experiment that induced a policy reversal in Texas, Louisiana, and Mississippi. In 2003, the Supreme Court decision in *Grutter v. Bollinger* ruled that race-conscious admissions processes that do not amount to quota systems are constitutional. This effectively reversed a 1996 lower court ruling in *Hopwood v. Texas* that prohibited any use of race in the admissions process in public universities in these three states. We exploit this exogenous policy change to estimate the effects of affirmative action on secondary school students' outcomes using two main identification strategies in three administrative data sets and one survey data set. In cases where we have administrative data for all (secondary school attendance and college applications, admissions, and graduation) or part (grades) of Texas, we use a difference-in-differences strategy that compares the change in URM (black and Hispanic) and white students' outcomes following the policy. This strategy can be interpreted as estimating the effect of affirmative action on the racial achievement gap. In cases where we have data across multiple states (SAT data), we separately compare the change in URMs' and whites' outcomes in states that were and were not affected by the policy using difference-in-differences and synthetic control group approaches. This second strategy allows us to identify potential spillover effects on whites. Finally, in the SAT data, we use an additional triple-differences strategy and interact cohort, geographic, and racial variation. This strategy estimates the differential change in the racial achievement gap due to the policy and controls for any changes in treated states over time that affected both whites and URMs.

There are four main findings. First, across data sets and identification strategies, we find that URMs respond to affirmative action policies, including by increasing their pre-college human capital investment. URMs increase their number of college applications and applications to selective schools, as well as their SAT scores, grades, and attendance. Our estimates suggest that the re-instatement of affirmative action helped close the racial gap in applications to selective schools by 13%, the gap in math SAT scores by 5%, the gap in secondary school grades by 18%, and the gap in secondary school attendance by 13%. These positive effects are concentrated among high school students in the upper-half of the ability distribution (as measured by middle school test scores), who are likely on the margin of admission to selective Texas public universities. Indeed, we estimate the change in the return to a URM moving up in the test score distribution on college admission after the policy and find that the returns to human capital investment did increase for URMs in the

[2] For simplicity, unless otherwise noted, we use "affirmative action" to refer to race-based affirmative action in the college admissions process, as opposed to "race-blind" affirmative action policies.

top half of the distribution.

Second, we find no evidence that pre-college human capital investment fell for URMs anywhere in the ability distribution. The reintroduction of affirmative action policies in Texas has no detectable disincentivizing effect on effort by URMs. Third, we find that white students also experience small increases in their SAT scores in response to the policy, consistent with either positive spillovers from URMs or whites also responding to increased returns to effort due to increased competition for admissions. Thus, the reductions in the racial achievement gap we observe may be smaller than the aggregate effect of the policy on URMs' pre-college human capital investment. Fourth, we explore the effect of the affirmative action policy on the racial gap in college graduation with the caveat that the net effect conflates effects on pre-college human capital investment, college match, and college quality. Following the policy, the gap in college graduation falls by 15% for students in the top quintile of the ability distribution.

Turning to mechanisms, several pieces of evidence suggest that students increase their effort in response to affirmative action policies. In addition to the fact that the students whose outcomes improve are the same ones who experience increases in the returns to pre-college human capital, attendance – an outcome likely to be indicative of effort – increases. Moreover, in survey data, we find no evidence that parents or guidance counselors respond to the policy, but we do find that URMs spend 10% more time on homework.

Broadly our results contribute to a large literature on the effects of affirmative action policies. This literature has focused primarily on affirmative action policies in higher education and their impact on college application behavior (Card and Krueger, 2005), admissions and campus diversity (Bowen and Bok, 1998; Arcidiacono, 2005; Rothstein and Yoon, 2008; Hinrichs, 2012, 2016), major choice (Arcidiacono et al., 2016), and college graduation (Bowen and Bok, 1998; Hinrichs, 2012). Arcidiacono et al. (2015) reviews this literature.

This paper is most closely related to a smaller literature about the effects of affirmative action on student behavior *prior* to college. In the United States, the evidence on educational outcomes from this literature is mixed.[3] Antonovics and Backes (2014) conclude that SAT scores and high school GPA changed little after California banned affirmative action by public universities. Cotton et al. (2015) simulate affirmative action for younger students

[3]The evidence from abroad is also mixed. Outside of the U.S., Ferman and Assunção (2015) and Estevan et al. (2018) study the effects of race-based and SES-based university admissions quotas in Brazil on high school students, while Khanna (2016) and Cassan (2019) study the effects of affirmative action on pre-college education in India. Ferman and Assunção (2015) find that affirmative action reduced student effort; Estevan et al. (2018) finds little effect on test preparation; and Khanna (2016) and Cassan (2019) finding positive effects on education.

in a math competition between younger and older students and find that their intervention increases younger students' time spent studying and improves their math performance. Bodoh-Creed and Hickman (2018) structurally estimate the U.S. college admissions market. Their structural estimates indicate that moving to a race-blind counterfactual increases pre-college human capital for URMs with a low cost of effort and decreases it for those with a higher cost of effort.

We contribute to this literature in two ways. First, we exploit a policy experiment to directly estimate the effects of the re-instatement of a real affirmative action policy on students' outcomes in the U.S. Thus, we complement Cotton et al. (2015) and Bodoh-Creed and Hickman (2018) by providing model-independent, reduced-form estimates of affirmative action's effects as it is implemented in practice. Second, we exploit large and detailed administrative data sets, allowing us to examine affirmative action's effects on a variety of dimensions and to trace out these effects across the ability distribution.

This study also relates to broader literatures on the incentive effects of college admissions (Cortes and Zhang, 2011; Leeds et al., 2017; Golightly, 2019) and the anticipatory effects of changes in the returns to human capital investment on children and their parents' investment decisions (Jayachandran and Lleras-Muney, 2009; Jensen, 2010, 2012; Oster and Steinberg, 2013; Leeds et al., 2017). While much of the evidence in the latter literature is from low-income countries, our results suggest that students also respond to changes in the returns to human capital investment in the United States.

The remainder of this paper is organized as follows. Section 2 introduces the context in more detail, and Section 3 discusses our different data sources. In Section 4, we report our estimates of the average and distributional effects of affirmative action on student outcomes using both the nation-wide SAT data and Texas administrative data sets. Section 5 provides suggestive evidence that the returns to investment in college admissions increased for the same set of students for whom we observe increases in pre-college human capital. Section 6 uses survey data to test which mechanisms drive the estimated effects, and Section 7 discusses whether alternative educational policies, such as No Child Left Behind, can explain our results. Section 8 concludes.

2 Context & Policy Change

In this section, we describe the Texas context and the policy change that this paper studies. We first sketch out a brief time line of events over the course of our study period (1997-2010) before describing the policy change and its effect on college admissions in more detail. In the last subsection, we examine whether universities' stated commitment to affirmative action translated into real changes in admissions.

Timeline of Events. In 1996, the U.S. Court of Appeals for the Fifth Circuit, which has jurisdiction over Texas, Louisiana, and Mississippi, ruled in *Texas v. Hopwood* that universities may not use race as a factor in deciding which applicants to admit. In the wake of this ruling, the Texas legislature passed the "Top 10% Rule" in 1997, which guaranteed admissions to *any* state-funded university in Texas to students graduating in the top 10% of their class. This law was passed as a means to promote diversity in universities by ensuring college access to high-achieving students from across Texas' somewhat segregated high schools. Then, in June 2003, the Supreme Court ruled in *Grutter v. Bollinger* that a race-conscious admissions process that does not amount to a quota system is constitutional. This Supreme Court decision overturned the previous decision banning the use of race as a factor in the admissions process in Texas, Louisiana, and Mississippi.[4] Thus, public universities in Texas, Louisiana, and Mississippi were unable to legally use race explicitly in the admissions process prior to 2003 and were able to do so again after 2003. We use this 2003 policy reversal to assess the effect of the introduction of race-based affirmative action on high school students' performance.[5]

The Top 10% Rule remained in place from 1997 onward, with the only change occurring at the very end of the study period. In 2009, the Texas legislature passed a law allowing UT Austin to cap the percent of its class admitted through the "Top 10% Rule" at 75%. As a result, following the new law's implementation in 2011, only the top 7% of students were admitted to UT Austin.

Grutter v. Bollinger. The *Grutter v. Bollinger* ruling was a close 5-to-4 ruling, with the deciding vote cast by moderate justice Sandra Day O'Connor. Prior to the ruling, the outcome of the case was viewed as impossible to predict, with *USA Today* writing in 2002, "Both sides think it's their best chance of winning the AA battle...O'Connor is the 5th vote

[4] As the ruling in *Grutter v. Bollinger* only established the constitutionality of affirmative action, states like California, Washington, and Florida, which had banned affirmative action due to ballot measures or executive orders, were unaffected.

[5] We don't focus on the earlier policy change in 1996 for two reasons. First, it combines a ban on race-based affirmative action and the introduction of the Top 10% Rule a year later. Therefore, the 1996 policy change does not provide a clean experiment for estimating the effects of an affirmative action ban on students' outcomes. Second, the scarcity of data from the pre-1996 period make credibly estimating the effect of the ban difficult.

but her moderate history does not indicate her direction." Consistent with this, the Supreme Court majority opinion expressed ambivalence over affirmative action policies, striking down the ban on considering race holistically while upholding a ban on explicitly assigning points for admissions based on race.[6]

The decision was heavily covered by the media. Appendix Figure A1, which plots the number of articles in US newspapers mentioning affirmative action by day, shows the spike in coverage around the ruling. The policy was also heatedly discussed in Texas. On June 29, 2003 (5 days after the ruling), *every* letter to the editor published in the *Austin-American Statesman* was about the case.

Policy Response to *Grutter v. Bollinger*. On the day that the *Grutter v. Bollinger* decision was issued, UT Austin's president, Larry Faulkner, stated that the Texas flagship campus intended to return to considering race in the admissions process. This response was well-publicized, with Faulkner shown making comments to this effect on the NBC nightly news the same day of the ruling. Only the University of Texas Board of Regents could authorize the implementation of such a change and in August 2003, the Board of Regents voted to allow all its campuses to return to considering race.[7] The Texas Tech University Board of Regents also outlined a plan in October 2003 to include race as an element in admitting prospective students. Thus, from the onset of the 2003 Supreme Court ruling, it was clear that the state flagship university, UT Austin, and other public universities in Texas would return to using affirmative action in the admissions process.

Race-based affirmative action co-exists with the Top 10% rule. Texas public universities first admit students who qualify for automatic admission through the 10% rule. Students who are not eligible for automatic admission (i.e. are not in the top decile of their graduating class) are admitted based on a "holistic" review process. Following the policy change, race or ethnicity could again play a role in this admission process. While a portion of public university classes are admitted under the Top 10% Rule, holistic admissions are also important. UT Austin, which has the highest percentage of freshmen admitted under the Top 10% Rule, admitted one-third of its freshman class through this process in 2003 (Office of the President, 2008). As described above, under current rules, UT Austin admits no more than 75% of its

[6]The majority ruling read, "The court takes the Law School at its word that it would like nothing better than to find a race-neutral admissions formula and will terminate its use of racial preferences as soon as practicable. The court expects that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today."

[7]University of Texas campuses consist of Austin, Arlington, Dallas, El Paso, Rio Grande Valley, San Antonio, Tyler, and Permian Basin.

class based on high school ranking cut-offs.

Did Affirmative Action Policies Affect Admissions? To evaluate whether universities' stated commitment to affirmative action translated into different admissions decisions on the ground, we now consider how it affected both university composition and admissions. Appendix Figure A2 uses the IPEDS data to calculate the racial composition of UT Austin's Fall entering class by year. Following Fall 2003, there is a trend-break in the share of blacks and Hispanics, with both rising precipitously. These descriptive statistics are consistent with the findings of Hinrichs (2012), who shows that affirmative action bans decrease the enrollment of URMs at selective universities. In contrast, the upward trend in the share of Asians, who are not considered an underrepresented minority, flattened from 2003 onward.

Similarly, the reversal of the ban appears to have affected UT Austin and other selective Texas universities' admissions behavior. Using administrative data from the Texas Education Agency, Appendix Figure A3 plots event study graphs of URM students' likelihood of being admitted to UT Austin, University of Houston, Texas Tech, and Texas A & M relative to whites by the year in which students attended 9th grade.[8] Students who ended 9th grade in 2001 were the first group whose admissions were affected by the re-instatement of affirmative action, although these students would have had little time to change their pre-college human capital. The likelihood of admissions for URMs following 2003 grew at UT Austin, the University of Houston, and Texas Tech. In contrast, there is no clear positive trend in URM admissions at Texas A & M, consistent with the fact that Texas A & M publicly stated that they would not use race-based affirmative action in admissions at the time of the ruling (Parker, 2018). Altogether, these results suggest that lifting the affirmative action ban did affect URM students' admissions probabilities at selective Texas universities.

3 Data

In this section, we describe our four data sets: (1) the administrative data for all Texas students from the Texas Education Agency (TEA), (2) the administrative data from a large urban school district (LUSD), (3) the panel of race-state-year SAT scores, and (4) the survey data from the Texas Higher Opportunity Project (THEOP).

[8] The estimation procedure for these event study graphs is identical to the one used to produce graphs for our outcome variables from the Texas Education Agency data later in this paper and is described in detail in Section 4.1.

3.1 Texas Education Agency (TEA) Administrative Data

Our first set of administrative data are individual-level records for all Texas elementary, middle, and high school students from the Texas Education Agency. The records include yearly school attendance, test scores on standardized tests, and demographic characteristics (e.g. race/ethnicity, gender, gifted status, socio-economic status). Thus, these data allow us to analyze the effects of affirmative action by ability on college applications, admissions, graduation, and school attendance.

These data have several important advantages. First, one key feature of the TEA data is that the files are linked to (in-state) college administrative data, allowing us to study the impact of *Grutter v. Bollinger* on college applications *and* college completion. Thus, we observe which *Texan* universities a student applies to and whether they graduate from a *Texan* university.[9] Second, since they cover every student in Texas, they allow us to estimate the population average treatment effects of affirmative action. In contrast, data sets like the SAT are restricted to students who take the exam. Data sets like the Integrated Post-Secondary Education Survey only capture information on students who actually enroll in college. Third, the large size of the TEA data set, as well as its panel structure, are important for estimating heterogeneity in the effects of affirmative action by ability. Because we observe 6th grade ability measures, in many cases we observe a student's location in the ability distribution *before* the policy change. The scale of the data also allows us to estimate heterogeneous effects by 6th grade test score with statistical precision.

Since use of the individual-level TEA data is restricted outside of a secure data room in Texas, we constructed a data set of aggregate observations for outside analysis. To examine the heterogeneous effects of affirmative action by academic ability, we collapsed these data at the school district-cohort-race-ability level.[10] Ability is determined by a student's 6th grade standardized test score, and students are classified into quintiles according to their rank in the cohort-specific test score distribution for the entire state of Texas.[11] Cohorts are defined using the academic year students first entered 9th grade. For most of our analysis, we focus on the 1997 to 2010 cohorts.[12] This analytical sample represents close to 3 million students.

While the TEA data also include data from Texas' state-wide standardized tests, these

[9]In 2004, only 8% of Texan residents enrolled in an institution of higher education were enrolled in an institution outside of Texas (Center for Education Statistics, 2004).

[10]For confidentiality reasons, all cells with less than 5 students are dropped (7% of all students).

[11]The fraction of students with valid 6th grade test scores varies slightly across cohorts and is generally in the 70-75% range.

[12]Years are based on the Spring semester. For example, the 2000 cohort refers to students who were in 9th grade in the 1999-2000 academic year.

tests underwent a substantial version change at roughly the same time as affirmative action was re-instated. In 2003, the standardized exam changed from the TAAS to the TAKS.[13] As a result, we cannot disentangle the effects of affirmative action from the effects of the version change on URMs' test scores. Thus, to examine additional measures of human capital, such as grades and test scores, we turn to a complementary administrative data set from one Texas school district, which we will describe below.

Summary statistics in the top panel of Table 1 provide an overview of the students in the TEA data. These statistics are reported separately for whites and URMs and for cohorts that were and were not affected by *Grutter v. Bollinger*. The fraction of Texas students identified as URMs increases sharply over time, entirely driven by an increase in the Hispanic population. URMs are much more likely to be from poor households than whites (60% vs 12% in 1997-2000) and have lower 6th grade test scores (average decile of 4.4 vs 6.6). Prior to the re-instatement of affirmative action, 17% of URMs applied to any 4-year university (within 4 years after starting high school), while 29% of whites did so. The gap is smaller in the later period, with 26% of URMs applying and 34% of whites doing so. Racial gaps are even larger in terms of applications to selective universities. For example, for the 1997-2000 cohorts of 9th graders, the average number of applications sent to selective institutions by URM students is 0.06, while it is 0.21 for whites.[14] Finally, 11% of all (pre-AA) black and Hispanic students eventually obtain a degree from a public university, while 25% of whites do.[15]

3.2 Large Urban School District (LUSD) Administrative Data

Our second source of administrative data is drawn from a large, urban school district in Texas. These data consist of repeated cross-sections of individual-level data for all 11th graders in the school district between 2001 and 2008.[16] The data contain information on students'

[13]These tests differ meaningfully. First, TAAS was administered to grades 3-8 and grade 10. In contrast, TAKS is administered to grades 3-11, with the higher-stakes exit exam taking place in grade 11 instead of 10. Second, the TAKS high school version includes social studies while TAAS does not (Tutson, 2002).

[14]The selective Texas universities to which we observe applications in our data are UT Austin, University of Houston, Texas A&M and Texas Tech.

[15]We exclude the 2007-2010 cohorts for college completion, since these later cohorts were less likely to have completed college by 2014, the last year of data we have. For instance, the overall college completion rate is less than 6% for the 2007 cohort of 9th graders.

[16]We focused on 11th graders to reduce the substantial administrative burden of constructing the data set for the school district. We believed this group to be most likely to be affected by affirmative action, as they had not yet applied to college but were close enough to the college applications stage to make decisions based on college admissions policies.

demographics (race/ethnicity, gender, age and zip code), courses and course grades, and test scores on the norm-referenced Stanford Achievement Test (hereafter, Stanford), a low-stakes achievement test that the school district has administered since 2000. Prior academic records for the three preceding years (e.g. course grades in 2003, 2004 and 2005 for students enrolled in 11th grade in 2006) also allow us to observe students' grades and test scores in 8th grade, as long as they were enrolled in the same district. We focus on grades as our main outcome in this data since the Stanford test underwent a version change from the Stanford 9 to the Stanford 10 in 2004, our first post-treatment year. While this change was less dramatic than the version change between the TAAKs and TAAS exams, we still view evidence from the Stanford test as suggestive.

The bottom panel of Table 1 shows summary statistics for the sample of 11th graders from this school district. The majority of students in the district are black or Hispanic. In a typical campus, 85% of students are black or Hispanic, and these students have lower achievement than white students along all dimensions. Black and Hispanic students score significantly lower on the Stanford standardized test in terms of national percentile ranking compared to white students, have lower grades in their courses (both in 8th and 11th grade), and have lower attendance rates.

3.3 SAT Data

To analyze the effects of the re-instatement of affirmative action on SAT scores, we collected data on mean math and verbal SAT scores and the number of test-takers at the state-race-year level from 1998 to 2010 from the College Board's publicly available reports. As in our other administrative data sets, we define URMs as Hispanic and black students and use white students as our comparison, non-URM group.

One important benefit of these data is the inclusion of states that were not affected by the policy change. This allows us to separately estimate the effect of *Grutter v. Bollinger* on URMs and whites and to estimate the differential change in URMs' outcomes relative to whites in the treated states. Summary statistics of the SAT panel data are reported in Appendix Table A1. These summary statistics reveal a substantial racial achievement gap, with average math and verbal scores for whites of 530 and 528 respectively and for URMs of 439 and 441 over the 1998-2003 period.

3.4 Texas Higher Education Opportunity Project Data

Our final data set complements our administrative and SAT data with survey data from the Texas Higher Education Opportunity Project (THEOP). THEOP surveyed high school seniors from a random sample of 105 public high schools in Texas in 2002 and in 2004 regarding their demographics, college perceptions, parental involvement, and other activities in high school. The timing of the survey allows us to observe students' responses right before and after affirmative action was re-introduced, with the caveat that the fact we observe only two cross-sections of the data makes it impossible to assess whether pre-trends drive the results. While the two waves of the survey are not identical, the set of questions that are consistent across these waves allow us to compare several outcomes that shed light on what mechanisms may drive affirmative action's effects.

THEOP records time spent on homework outside of school, a student-reported measure of effort. The survey also records whether the student applied to their first choice college, providing additional information on whether college applications behavior changed. We also combine a series of questions about parental behavior into a "parental involvement index," with values ranging from 5 to 20.[17] This index captures whether parents changed their behavior or educational investments in response to affirmative action. Finally, a question about whether the student discussed the college applications process with his/her guidance counselor captures changes in guidance counselor involvement. Appendix Table A2 reports summary statistics for these data.

4 Effects of Affirmative Action on Students' Outcomes

In this section, we empirically investigate the effect of the reinstatement of affirmative action on several measures of students' behavior using our three non-survey data sets. We first report the effect of affirmative action on URMs' college application behavior relative to whites. As this outcome is the most malleable and the most directly connected to affirmative action policies, we view a positive effect of affirmative action on college applications as evidence that students were aware of and responded to the policy change. We then complement these results by estimating the effect of affirmative action on URMs' SAT scores using difference-in-

[17] These questions ask "How often do your parents ... (i) give you special privileges because of good grades, (ii) try to make you work harder if you get bad grades, (iii) know when you are having difficulty in school, (iv) help with your school work, and (v) talk with you about problems in school." Students' responses range from "very rarely" (1) to "almost all the time" (4). We sum across the answers to these questions to construct the "parental involvement index" so that a higher index corresponds to more involvement along these dimensions.

differences, triple-differences, and synthetic control group approaches that compare changes in scores in states that re-instated affirmative action (Texas, Louisiana, Mississippi) to changes in unaffected states. Next, we focus on the single large, urban Texas school district where we observe grades and standardized test scores. Using these data, we estimate the effects of the reinstatement on the within-school-year racial achievement gap in standardized tests and course grades. As an additional measure of secondary school effort, we also use administrative data from the entire state of Texas to estimate the effects of affirmative action on attendance. Finally, we estimate the effects of the reinstatement of affirmative action on college completion, though we caution that this outcome is a function of both pre-college human capital and the college to which a student is matched.

4.1 Impact of Affirmative Action on College Applications

Difference-in-Differences Empirical Strategy. To assess the effects of affirmative action on students' college application behavior, we use Texas-wide administrative data. Our difference-in-differences strategy compares the change in URMs' college application behavior following the reinstatement of affirmative action to the change in whites' behavior. Recalling that an observation in this data is a race-ability quintile-district-cohort cell, we estimate

$$\begin{aligned} y_{dcea} = &\beta_1(URM_e \times PartTreat_c) + \beta_2(URM_e \times FullTreat_c) \\ &+ \mathbf{\Gamma X_{dcea}} + \alpha_{dca} + \alpha_{dea} + \epsilon_{dcea} \end{aligned} \tag{1}$$

where d indexes a school district, c indexes a 9th grade cohort (the year the student entered 9th grade), e indexes an ethnicity, and a indexes an ability quintile in the state standardized test in 6th grade. The variable URM_e is an indicator variable for belonging to a URM group, and $\mathbf{X_{dcea}}$ is a vector of average student characteristics for the observation cell (age, sex, immigrant status, low-income status, gifted, ESL, special education status, and limited English proficiency). The dependent variable, y_{dcea}, is either the fraction of students who applied to *any* 4-year university[18] or the average number of applications sent to selective Texan institutions. Since the impact of affirmative action may not be immediate, we allow the effect to vary across cohorts. In our main parametric specification, we distinguish between partially treated cohorts who were already in high school at the time of the policy change and fully treated cohorts who started high school after the policy change. Thus, $PartTreat_c$ is equal to 1 if a student was in 9th grade between 2000 and 2003, while $FullTreat_c$ is equal

[18]This measure includes non-selective institutions and community colleges.

to 1 if a student was in 9th grade after 2003. α_{dca} denotes district-cohort-ability fixed effects, and α_{dea} denotes district-ethnicity-ability fixed effects.

Our main coefficients of interest, β_1 and β_2 respectively capture the short and medium-run effects of affirmative action on college applications behavior. Later cohorts may have had greater opportunities to adjust their human capital investment in high school in response to the re-instatement of affirmative action. This in turn may have affected their likelihood of being accepted to college and therefore their propensity to apply in the first place, relative to earlier treated cohorts. Thus, we expect $\beta_2 > \beta_1$.

In this difference-in-differences specification, the effect of affirmative action is identified by comparing URM students to non-URM students of the same ability, in the same cohort, and the same school district. The fixed effect α_{dca} accounts for any time trends that may vary across districts or ability levels, as long as they are not differential by race. The fixed effect α_{dea} accounts for any differences across races, districts, or ability levels (or any combination thereof), as long as these differences do not vary over time. To account for correlated outcomes in districts over time, we also cluster the standard errors at the district-level.

One important limitation of this strategy is that whites' outcomes may also be affected by affirmative action. If, for example, whites decrease their college applications in response to the reinstatement of affirmative action, we would estimate positive values for β_1 and β_2 in equation (1), even if URMs' behavior is unchanged. To assess whether this could be driving our results, we also separately graph trends in application behavior by race. Then, we can observe directly if URMs experience a jump or trend break when they are affected by the policy and whether whites are negatively affected.

Event Study Specification. In this difference-in-differences empirical strategy, identification of the causal effect of affirmative action relies on the assumption that the college applications behavior of URM and comparable non-URM students would have evolved the same way in the absence of the ruling. To examine the plausibility of this assumption, we plot the relative effect of being a URM on college applications separately for each cohort. Doing so allows us to establish if trends in college applications for URMs and whites were parallel prior to the re-introduction of affirmative action. Plotting these point estimates also allows us to observe whether the treatment effects of affirmative action accumulate, justifying our decision to separate partially and fully treated cohorts. To do so, we estimate the following model:

$$y_{dcea} = \sum_{t=1997}^{1999} \beta_t (URM_e \times \mathbf{I_{ct}}^{Grade\ 9}) + \sum_{t=2001}^{2010} \beta_t (URM_e \times \mathbf{I_{ct}}^{Grade\ 9}) + \mathbf{\Gamma X_{dcea}}$$
$$+ \alpha_{da} + \alpha_{ca} + \alpha_{ea} + \epsilon_{dcea}, \qquad (2)$$

Equation (2) is more flexible than our parametric difference-in-differences specification in that it includes fewer fixed effects to allow us to estimate these cohort-specific coefficients with more precision. If the parallel trends assumption is valid, for $t < 2000$, we expect that β_t will be indistinguishable from zero. If the effects of affirmative action accumulate over time as students have more time to adjust their behavior, we expect that after 2000, β_t will generally be greater for greater values of t. Additionally, if the effects we estimate in the difference-in-differences strategy are due to affirmative action, we expect to see an increase in the values of β_t soon after 2000.

Difference-in-Differences Results. We report coefficients from equation (1) in column (1) of Table 2. In panel A, the outcome is is the average number of applications to selective institutions, and in panel B, the outcome is the probability of applying to any college. In panel A, on average, fully treated URM students apply to 0.02 more selective Texas colleges,[19] indicating that affirmative action closed the racial gap by 13%. Turning to Panel B, on average, lifting the ban on affirmative action increased URMs' probability of applying to at least one college relative to whites by 0.8 percentage points for cohorts who were in high school at the time of *Grutter v. Bollinger* and by 2.9 percentage points for cohorts who entered high school after the ban was lifted. For fully treated cohorts, the policy closed the pre-affirmative action racial achievement gap in applying to any college by 25%. In both panels, the estimates are precisely estimated and statistically significant at the 1% level.

For both panels, these average effects mask substantial heterogeneity. The remaining columns of the table estimate the effects for students in different ability quintiles. In Panel A, while bottom quintile students are no more likely to apply to selective institutions, top quintile students apply to 0.04 more selective institutions, and the second highest quintile applies to 0.03 more selective institutions. Turning to Panel B, we do find a small positive effect on applying to any college (0.0101) for fully treated students in the bottom quintile of the ability distribution, but the effect is five times larger among the highest ability students (0.0545). In both cases, this heterogeneity accords with where we would expect affirmative

[19]Results are reported separately for black and Hispanic students in Appendix Tables A3 and A4.

action to have the strongest effects on college applications, as affirmative action is most likely to affect admissions for students already on the margin of being admitted. The small, positive effects we estimate for lower quintile students could reflect both noise in the quintile assignment, which is based on 6th grade test scores, and positive spillovers from higher ability students.

Additionally, we examine applying to any of the campuses of the University of Texas. Since the University of Texas Board of Regents promptly allowed its campuses to consider race in admissions, the effects of affirmative action should materialize for these institutions. Appendix Table A5 reports the difference-in-differences results for applications to any UT campus, and the patterns are similar to the results for the other applications outcomes.

Finally, to verify that the positive effects reported in Table 2 are not driven by declining applications by whites, we plot application behavior separately by race in Appendix Figure A4. The unadjusted figures plot the cohort effect on applications (normalized to the cohort in 9th grade in 2000) without controls, while the adjusted figures include the full set of controls from equation (1), as well as race-ability and district-ability fixed effects. Appendix Figure A5 further shows the results by racial group for the top quintile. Taken together, these figures show that there is a trend break in URMs' behavior around the reintroduction of affirmative action, and that the difference-in-differences estimates are not due to reduced applications by whites.

Event Study Results. We now turn to the event study graphs estimated by equation (2) to examine whether pre-trends drive our findings. Figure 1 plots the year-specific coefficients β_t for the number of applications to selective universities. Due to our finding that the effect of affirmative action on selective college applications was concentrated in higher ability cohorts, we examine trends separately for students in the top and bottom quintiles of the ability distribution.[20] Cohorts between the solid and dashed vertical lines are partially treated, while cohorts to the right of the dashed vertical line are fully treated. The point estimates for bottom quintile students are indeed very small and statistically insignificant both before and after the policy change. For top-quintile students, there appears to be a weak negative pre-trend, but these year-specific coefficients are small and not systematically statistically significant, and a strong positive effect emerges directly after the policy change.

Figure 2 plots the point estimates for the probability of applying to any university, and 95% confidence intervals are shown using dashed lines. Again, results are shown separately

[20] Appendix Figure A6 shows average effects for the full sample.

for students in the top and bottom quintiles of the ability distribution. For bottom quintile students, there is a small upward trend in URM college applications relative to whites prior to the policy change, but most year-specific coefficients are close to zero and statistically indistinguishable from the base year. The point estimate for the 2001 cohort indicates that the probability of applying to a university for URM students increases at the time of the policy change, but the jump is considerably more pronounced for top quintile students. Finally, Appendix Figure A7 reports the event study estimates for applying to any university of Texas. The pattern of the point estimates is very similar to Figures 1 and 2.

Across all three graphs, the treatment effects appear to accumulate over time, with affirmative action having a larger effect on fully treated cohorts. Thus, allowing students to have more years to adjust in response to the affirmative action policy appears to strengthen the policy's effect. This could be because students respond to these policies by increasing their pre-college human capital, a hypothesis that we begin to investigate in the next subsection.

Robustness. Before moving on to assess the effects of affirmative action on pre-college human capital, we first ensure that our results are robust to an additional test. One concern is that in later-cohorts our 6th grade ability measure, which we use to estimate the heterogeneous effects of the policy, is observed after the policy change. Thus, observed ability may be endogenously changing as a result of the reintroduction of affirmative action. To ensure this is not affecting our results, in Appendix Table A6, we re-estimate equation (1), dropping cohorts who were in 6th grade following the policy change. The pattern in the results is the same as before: affirmative action has its strongest positive effects on URMs in the upper part of the ability distribution.

4.2 Impact of Affirmative Action on SAT scores

Difference-in-Differences Empirical Strategy. To measure whether affirmative action affected students' human capital, we now examine whether it affected students' SAT scores. To measure the effects of affirmative action, we exploit both time variation in whether students took the SAT after *Grutter v. Bollinger* and geographic variation in whether students lived in a state where *Grutter v. Bollinger* eliminated a previous ban on affirmative action. This difference-in-differences strategy allows us to estimate the effect of affirmative action *separately* for URMs and whites.

To implement this strategy, we use a panel of average math and verbal SAT scores at the state-race-year level. Using this data, for URMs and whites, we separately estimate

$$y_{ket} = \beta(Treated\_State_k \times Post2003_t) + \alpha_k + \alpha_t + \alpha_e + \varepsilon_{ket}. \tag{3}$$

where k indexes a state, e indexes a racial group, and t indexes a year. Then, y_{ket} is either the mean math or verbal test score for group e in state k and year t, $Treated\_State_k$ is an indicator variable equal to 1 if the observation belongs to a state that was treated (Texas, Louisiana, and Mississippi), $Post2003_t$ is an indicator variable equal to 1 if the year is greater than 2003, α_k is a state fixed effect, α_t is a year fixed effect, and α_e is a race fixed effect. Additionally, we weight race-state-year cells by the number of test-takers and cluster our standard errors at the state-level.

Triple-Differences Empirical Strategy. In addition to exploiting time and geographic variation to estimate the effect of affirmative action, we also use a triple-differences strategy. Since we expect URMs to be more affected by affirmative action, we use race as a third difference. This identifies the change in the racial achievement gap due to the policy, in line with our within-Texas results. This approach controls for any time-varying shocks in states affected by the policy but may under or over-estimate the policy's effects on URM's outcomes if the policy also affected whites. To estimate the differential effect of affirmative action on URMs relative to non-URMs, we estimate

$$y_{ket} = \beta_1(Treated\_State_k \times Post2003_t \times URM_e) + \alpha_{ke} + \alpha_{et} + \alpha_{kt} + \varepsilon_{ket}, \tag{4}$$

where URM_e is defined in the same way as before, α_{ke} is a state-race fixed effect, α_{et} is a race-year fixed effect, and α_{kt} is a state-year fixed effect. While the triple-differences strategy requires us to include controls for all three sources of variation and their double interactions, these are subsumed by the fixed effects in this specification.

This strategy controls for all the same potential sources of bias as the difference-in-differences strategy. Both strategies use fixed effects to account for level differences in SAT scores between states and over time. In addition, the triple-differences strategy includes the fixed effect α_{kt}, which controls for any state-specific differences over time. Thus, this triple-differences strategy is valid even if Texas, Louisiana, and Mississippi have different time trends from other states, as long as those time trends also don't vary by race.

Event Study. As with college applications, we also use event study graphs to assess whether the parallel trends assumption of our difference-in-differences strategy is violated. To do so, we estimate the following equation separately for whites and URMs

$$\begin{aligned} y_{ket} = &\sum_{l=1998}^{2002} \beta_l (Treated\_State_k \times \mathbf{I_{tl}}) + \sum_{l=2004}^{2010} \beta_l (Treated\_State_k \times \mathbf{I_{tl}}) \\ &+ \alpha_k + \alpha_e + \alpha_t + \varepsilon_{ket}, \end{aligned} \tag{5}$$

where $\mathbf{I_{tl}}$ is an indicator variable equal to 1 if $t = l$. The omitted year is 2003, the year before the policy change. This event study specification estimates the differential effect of a test-taker being in a treated state for each year, β_l. If pre-trends between treated and non-treated states are parallel, we expect that β_l should be small and insignificant prior to 2003.

Synthetic Control Group Strategy. While event study graphs help us to assess the appropriateness of the parallel trends assumption, synthetic control group methods provide us with an alternative way of verifying that our results are robust to accounting for differential time trends. Based on these methods, developed by Abadie and Gardeazabal (2003) and Abadie et al. (2010), we construct a synthetic control group of states by matching those states' pre-trends in test scores to the pre-trends of the treated unit (the weighted average of Texas, Mississippi, and Louisiana).[21] We match the pre-treatment values of the number of URM and white test takers, the math SAT scores of URM and white students, and the verbal SAT scores of URM and white students. Our estimated effect of the reinstatement of affirmative action is then the difference between the change in test scores in the weighted average of the treated states and the synthetic control.

To assess the significance of our estimates, we use permutation tests. For all possible combinations of three untreated states, we apply the synthetic control method and calculate the post/pre-treatment ratio of root mean squared prediction errors (RMSPE).[22] We then plot the distribution of these ratios and examine the rank of the real treatment unit in that distribution.

[21]When generating the synthetic control groups, we exclude South Dakota, North Dakota, Wyoming, and Washington DC from the pool of potential controls because SAT scores are missing for some ethnic groups in some years in these states due to small samples. We follow the standard practice of minimizing the mean squared prediction error of our outcome variable over the entire pre-treatment period. In Appendix B, we show that our results are robust to using fewer pre-treatment years to construct the synthetic control group.

[22]Since the donor pool contains 44 control units, the number of possible combinations of three states is 13,244.

Difference-in-Differences and Triple-Differences Results. Table 3 reports the coefficients from equation (3) (panels A and B) and equation (4) (panel C) for SAT scores. Column (1) shows that math scores for both URMs and whites improved in treated states following 2003, though URMs' test scores improved by twice as much (8 points relative to 4 points). The student-level national standard deviation in math SAT scores was 115 points in 2003, so the math effects are equivalent to 0.07sd and 0.035sd. There is no effect on verbal scores for either group.

Whites' human capital appears to be positively affected by the abrogation of the ban. This could occur if whites increased their effort in response to intensifying competition. That is, if the distribution of the cost of pre-college human capital investment is different for whites than URMs, the introduction of affirmative action could increase the returns to pre-college human capital investment for both the average white and URM student. This mechanism is consistent with both the theoretical model and empirical findings of Cotton et al. (2015), who show that students who do not benefit from a simulated affirmative action policy may also be incentivized to increase their effort. Alternatively, increased human capital investment by URMs could have had positive spillovers on whites. Altogether, the findings in Table 3 indicate that, since the policy increased both whites' and URMs' pre-college human capital investment, estimates of the effect of the policy on the racial achievement gap are likely to under-estimate its aggregate effects on pre-college human capital.

In the last panel of Table 3, we report the results of the triple-differences specification. We find that URMs' math SAT scores improved relative to whites' in treated states by a statistically significant 4 points, equivalent to 0.035sd or a 5% reduction in the racial achievement gap. The triple-differences results further confirm that the positive difference-in-differences estimates are not merely due to differential time trends in states that were not affected by *Grutter v. Bollinger*.

In the last two columns, we evaluate whether the policy change affected test-taking. In column (3), the outcome is the raw number of SAT test-takers, and cells are weighted by the average number of test-takers in the first three years of the panel, 1998-2000. In column (4), we generate a measure of the probability of taking the SAT by dividing the number of test-takers by the number of 17-19 year-olds in each cell (using yearly ACS population counts) to account for fluctuations in the number of test takers coming from changes in population size. Both metrics suggest there was no significant change in the probability of taking the SAT in treated states relative to untreated states.

One potential concern is that our estimates might be contaminated by other affirmative action policy changes that occurred between 1998 and 2010 in the states that were not affected by *Grutter v. Bollinger*. For instance, four states banned affirmative action in college admissions during these years. In Appendix A, we verify that the SAT results are robust to controlling for these bans and discuss their effects. Additionally, in this appendix, motivated by Hinrichs (2012), who argues that Louisiana and Mississippi were less bound by the original *Hopwood v. Texas* ruling than Texas, we show that the results are unchanged if we drop Lousiana and Mississippi from the analysis.

Event Study Results. Figure 3 reports test-taking year-specific coefficients from equation (5) separately for URM and non-URM students. The plot shows a negative pre-trend in math SAT scores for students in treated states relative to those in non-treated states. That is, prior to *Grutter v. Bollinger*, students in treated states were falling behind the rest of the country on the math SAT.[23] Following the reinstatement of affirmative action, there is a reversal of fortunes, and the negative trend turns positive right after 2004. Importantly, the post-treatment positive trend for math scores appears to be considerably steeper for URM students than for whites. Consistent with the point estimates in Table 3, there is no clear change in verbal scores over time for either URMs or whites. Altogether, Figure 3 provides further evidence that our estimates are not driven by differential time trends between treated and untreated states.

Synthetic Control Results. The top panel of Figure 4 shows the evolution of SAT math scores over time for our treatment unit and the associated synthetic control group separately for white and URM students. Interestingly, for both URMs and whites, California, which banned affirmative action in public universities prior to our study period, contributes to the control group. In both cases, the synthetic control group closely tracks the treatment unit prior to the re-instatement of affirmative action, and the two trends diverge considerably from 2004 onward. This is true both for both groups, but the divergence is greater in magnitude for URMs. The implied treatment effects are larger than our baseline difference-in-differences estimates. Whites' math scores increase by 6.3 points, and URMs' increase by 10.8 points. The placebo tests suggest that these results are not due to chance. The treatment unit's post-pre ratio of RMSPE is at the 99.2th percentile of the distribution for whites, and at the 96.6th percentile for URMs. Appendix B provides further robustness tests of the synthetic control estimates, including dropping states that enacted AA bans during the study period from the donor pool.

[23] Appendix Figure A8 shows that a similar pattern holds for Asians.

Appendix Figure A9 is a synthetic control plot of the differences between treated and untreated units separately for white and URM students. For both racial groups, the differences are close to zero prior to treatment and then exhibit large increases following *Grutter v. Bollinger*. Again, the effect is greater for URMs than whites. The implied triple-differences estimate from differencing out the whites' difference from the URMs' is 4.5 points, which is similar to our conventional triple-differences estimate.

Having found evidence that students respond to affirmative action by improving their SAT scores, we next investigate whether students also increase other dimensions of their human capital. SAT scores may just reflect better SAT-specific test-taking skills. Thus, examining other outcomes allows us to evaluate if affirmative action affects human capital more broadly.

4.3 Impact of Affirmative Action on Grades

Empirical Strategy. In this subsection, we turn to our data from the large, urban Texas school district (LUSD) to examine the effect of affirmative action on students' grades in 11th grade. Our econometric specification is similar to equation (1), with some alterations to accommodate the different structure of the school district's administrative data. In particular, unlike our Texas-wide regressions, which use aggregate district-year-race-ability data, for the LUSD, an observation is an individual. The specification is

$$y_{isec} = \beta\left(URM_i \times Post2003_i\right) + \mathbf{\Gamma X_i} + \alpha_{sc} + \alpha_e + \epsilon_{isec} \tag{6}$$

where i denotes an individual, s denotes a school, e denotes a racial group, and c denotes a cohort.[24] The treatment variable $Post2003_i$ is an indicator variable equal to 1 if the outcome is realized after the policy change, so a student is observed in 11th grade after 2003. α_{sc} denotes a school-cohort fixed effect, and α_e is a race-specific fixed effect. We include α_{sc} to account for the fact that grades may not be comparable across schools or across years.[25] Thus, the effect of affirmative action in this regression is identified by comparing URM and white students in the same school in the same year. The basic controls $\mathbf{X_i}$ consists of controls for age, sex, and home zip code fixed effects. Additionally, in a more conservative, "value-added" specification, we control for a lagged measure of ability (8th grade grades).[26] This

[24]Since the LUSD data consists of repeated cross-sections of 11th graders, in this data set, a cohort refers to the year students attended 11th grade.

[25]For example, this would be the case if course offerings or grading standards are changing over time.

[26]The fact that we use 6th grade test scores as our ability measure in the TEA data and 8th grade test scores as our ability measure in the LUSD simply reflects differences in the availability of lagged scores across

control accounts for any changes in the ability distribution of URMs over time that might otherwise be attributed to affirmative action (such as changes due to cohort composition or migration). As before, the coefficient of interest, β, represents the effect of affirmative action on URM students relative to non-URM students.

In addition to using this difference-in-differences approach to estimate the effect of affirmative action, we also estimate cohort-specific coefficients and plot them in an event study graph. To do so, we simply alter equation (6) to estimate a different coefficient on the variable URM_i for every cohort. As in our previous analyses, the event study graph sheds light on whether the results we observe are driven by pre-trends.

Difference-in-Differences Results. The difference-in-differences estimates from equation (6) are reported in Table 4. The point estimates confirm that affirmative action had a positive effect on school grades in 11th grade. Our baseline estimates of equation (6) in column (1) indicate that grades increased by 0.9 points (on a 0-100 scale), equivalent to 0.1sd, following the reinstatement of affirmative action, closing the pre-AA racial gap by 18%. In column (2), we estimate a value-added specification, where we control for school grades in middle school (8th grade). The difference-in-differences coefficient is almost identical and remains strongly statistically significant.

In column (3), we re-arrange the data set into a panel that includes two entries per student (one for 11th grade and one for 8th grade) and estimate a specification with student fixed effects. In this model, our main explanatory variable becomes a triple-difference interaction term ($URM_e \times Treat_c \times I_g^{11th\,Grade}$), where $I_g^{11th\,Grade}$ is an indicator variable equal to 1 when a student is enrolled in 11th grade. Here, the effect of affirmative action is identified from within-student changes in outcomes between 8th and 11th grade for students who were in 11th grade after the policy change and in 8th grade before the change. This alternative specification further accounts for any unobserved changes in URM students' characteristics across cohorts that might otherwise bias our estimate of the effect of affirmative action. Again, the results of this alternative specification are nearly identical to our previous results.

In columns (4) to (6), we examine whether the effects are heterogeneous by prior ability. To do so, we calculate school-by-cohort specific terciles of the distribution of grades in 8th grade within school-years. In this data, we focus on terciles instead of quintiles, as we did in the TEA data, because of the much smaller sample size. We then re-estimate equation (6) separately for students in the bottom, middle, and top terciles. While the point estimates for

the two data sets.

the effect of affirmative action are positive for all three ability categories, they are particularly large for top-ability students (an effect of 1.4 percentage points or 0.2 sd). This is what one would expect if these students are most likely to apply to selective colleges and therefore to benefit from the policy change.[27]

Event Study Results. The top panel of Figure 5 reports year-specific coefficients on the URM_i indicator variable when the outcome is mean student grades. There are no significant pre-trends, with the racial gap in school grades remaining constant over the 2001-2003 period. School grades for URM students improve relative to their non-URM peers upon the re-instatement of affirmative action and remain at this higher level through 2008. The bottom panel of Figure 5 reports the year-specific coefficients under the value-added specification, which controls for 8th grade test scores. The results across the two specifications are very similar.

4.4 Impact of Affirmative Action on the Stanford Exam

The data from the large, urban school district also allows us to estimate the effects of affirmative action on the standardized Stanford test, a low-stakes exam that the school district itself administered. To estimate the effects on the Stanford exam, we follow the exact same difference-in-differences strategy as we did for grades in Section 4.3. The only difference is that the outcome variable is now a student's mean percentile on the Stanford exam, where percentiles are based on the national distribution. Appendix Table A9 reports the estimates. On average, Stanford test scores increase by 4.78 percentiles for URMs relative to whites (equivalent to 0.2sd). The effect is again the largest for the top tercile, which experiences gains of 7.47 percentiles (0.3sd). Appendix Figure A10 plots the event study graph for the Stanford exam. We again see little evidence of pre-trends and the immediate positive effect of affirmative action on URMs' test scores at the time of the policy change.

4.5 Impact of Affirmative Action on Attendance

Having shown that grades and test scores increase as a result of affirmative action, we now consider a more direct measure of student effort. Returning to the Texas-wide administrative TEA data, we test whether affirmative action affects URM students' attendance. Our empir-

[27] Appendix Tables A7 and A8 re-estimate the specifications in Table 4 separately for math and English grades.

ical strategy for examining attendance in the TEA data follows our strategy for estimating effects on college applications (see equation (1)).

Table 5 reports the regression results for 10th and 11th grade attendance. Difference-in-differences estimates indicate a positive average effect on the fraction of days present of 0.0036 percentage points in 10th grade (panel A) and of 0.0024 percentage points in 11th grade (panel B). The latter effect is equivalent to 13% of the pre-AA racial gap in attendance rates.[28] While the effects on attendance occur throughout the distribution in grade 10, for grade 11, they are concentrated again in the top part of the distribution.

Figure 6 reports the event study plots for attendance in grades 10 and 11. For these outcomes, because our data is organized in cohort-time, the first treated cohort for 10th grade attendance is the 2003 cohort, and the first treated cohort for 11th grade attendance is the 2002 cohort. Reassuringly, the timing of increases in attendance rates is consistent with a positive treatment effect at the time affirmative action was re-instituted rather than simple differences in attendance rates across cohorts. Attendance rates for the 2002 cohort of 9th graders are greater than for the 2001 cohort in 11th grade but not in 10th grade (both cohorts were in 10th grade before *Grutter v. Bollinger*, but only the 2002 cohort was in 11th grade post-treatment). Overall, the plots show no discernible pre-trend, and they suggest that there was a positive effect on attendance in high school.

4.6 Affirmative Action and College Completion

Thusfar, our analyses have documented the positive effects of affirmative action in undergraduate college admissions on URMs' college applications and human capital prior to reaching college. In this subsection, we estimate the effect of affirmative action on the probability of completing a college degree using administrative data from the TEA.

In Section 4.1, we showed that more URM students applied to college as a result of the reinstatement of affirmative action. However, this need not result in an increase in the fraction of URM students who obtain a post-secondary degree. On the one hand, if marginal students are now matched to colleges for which they are not prepared, they may be less likely to complete their degrees. This is essentially the mismatch argument of Sander (2004). Then, affirmative action might reduce the fraction of degree holders. On the other hand, if increased effort in high school contributes to the accumulation of human capital, the probability of completing a college degree may increase. Additionally, if students are matched to better schools that have higher returns to education, incentivizing students to

[28] Results are quantitatively similar, but considerably less precise, in the LUSD.

graduate, or that are more able to ensure students graduate, graduation rates may increase. To measure the direction of the effect of affirmative action on college graduation, we employ the same empirical strategy that we used in the TEA data to measure college application behavior (see equation (1)).

Table 6 reports the difference-in-differences estimates. Pooling all students together (column (1)), we find no effect of affirmative action on students who had little opportunity to adjust their level of effort in high school (the partially treated cohorts). For fully treated cohorts, the probability of graduating increases by 0.46 percentage points (3%). As in our other analyses, in columns (2) to (6) we estimate the effect separately for each quintile of the ability distribution. We find no significant evidence of gains for partially treated cohorts for any of the quintiles, though the estimate is positive for top ability students. For fully treated cohorts, the point estimates are positive throughout the ability distribution, but are much larger for the top ability quintile. The estimates indicate that students in the top quintile of the ability distribution who started high school post-policy experienced a 1.4 percentage point increase (4%) in the probability of completing college.

Figure 7 is an event study plot of the effect of the reinstatement of affirmative action on the probability of completing a 4-year college degree, with separate estimates for students in the bottom and top quintiles of the ability distribution. The probability of completing a college degree is roughly flat for low-ability students throughout the study period, as one might expect. For high ability students, the relative probability of graduating college appears to increase post-policy change. Graduation rates vary noisily around zero for cohorts that were never treated (i.e. who would have started college prior to the court ruling), appear to start increasing with cohorts that were partially treated (i.e. who were in 9^{th} grade between 2001 and 2003), and stabilize at higher values for cohorts who started high school post 2003. This pattern is suggestive evidence in favor of the human capital accumulation channel. The cohorts who had the most time to adjust their human capital investment in secondary school also experience the largest increase in college graduation.

Taking all our results together, higher ability URM students increased their effort in high school as measured by attendance, increased their pre-college human capital, increased the number of applications they sent to selective institutions, and became more likely to graduate from college. For college graduation, any decrease in match-quality in parts of the distribution that may have resulted from the reinstatement of affirmative action was more than made up for by positive effects on effort, application rates, and college quality.

5 Changes in Returns to Pre-College Human Capital

The previous section provided evidence that the effects of affirmative action on URM students' pre-college human capital are concentrated in the top half of the ability distribution. If students are indeed responding to changes in the returns to human capital investment caused by the policy, then these should also be the students for whom the returns increased. In this section, we provide suggestive evidence that this is the case.

To do so, we return to the TEA data on university admissions. Taking advantage of our 6th grade ability measure, we estimate the change in the marginal effect of moving up an ability *decile* on university admissions. The estimating equation is

$$\begin{aligned} y_{dcea} = &\sum_k \beta_{1,k}(URM_e \times PartTreat_c \times \mathbf{I_a}^{a\geq k}) + \sum_k \beta_{2,k}(URM_e \times FullTreat_c \times \mathbf{I_a}^{a\geq k}) \\ &+ \mathbf{\Gamma X_{dcea}} + \alpha_{dca} + \alpha_{dea} + \epsilon_{dcea} \end{aligned} \tag{7}$$

where a denotes a decile, y_{dcea} is a college admissions outcome, and $\mathbf{I_a}^{a\geq k}$ is an indicator variable if a student's ability decile a is greater than or equal to k. The controls are the same as in equation (1) except that they are now fully interacted with ability decile fixed effects. Thus, $\beta_{1,k}$ and $\beta_{2,k}$ capture the change in the marginal effect of moving from decile $k-1$ to k due to the policy for those who are partially and fully treated. Since increased human capital investment can allow a student to move up in the distribution relative to her peers, we interpret these coefficients as a proxy for the change in the returns to investment in university admissions, with the caveat that part of the increased admissions may be due to changes in students' applications behavior due to the policy.

Figure 8 reports $\beta_{2,k}$ for admission to any college and for the number of selective Texas universities to which a student is admitted. Consistent with the fact that affirmative action increases effort and human capital in the top half of the ability distribution, we see that the returns mainly rise in the top half. For instance, the returns to moving from the 9th to the 10th ability decile, in terms of number of admissions to selective universities per student, increased by 0.026 for URMs relative to whites among fully treated cohorts. The fact that there are strong increases in the returns for the top decile isn't inconsistent with the existence of the Top 10% Rule. This is because the deciles do not accord with the cut-offs used by the rule: they are based on performance in 6th grade rather than at the end of high school and are across-school deciles rather than within-school deciles.

Appendix Figure A11 further reports the effects on the "returns to human capital" for admissions to four specific schools: UT Austin, University of Houston, Texas Tech, and Texas A&M. For the first three schools, which were free to practice affirmative action, there are increases in the "returns to human capital" in the top half of the ability distribution (or in the top 30% in the case of Houston). Reassuringly, for Texas A&M, which does not practice affirmative action, there is no systematic effect on the returns to human capital. Given that human capital investment responded to affirmative action exactly among the students with the greatest increase in its returns, the results provide evidence that students increased their effort in response to the change in the returns to pre-college human capital.

6 Suggestive Evidence on Mechanisms

So far, we have provided evidence that affirmative action narrowed the achievement gap between whites and URMs for an array of outcomes. A natural next question is what channels led to these effects. One possibility that is consistent with both the evidence presented in the previous section and the effects on attendance is that high school students changed their behavior in direct response to perceived changes in their likelihood of college admission. Still, teachers may have also become more lenient toward URMs after the policy change or may have focused more on improving URMs' outcomes. While the relative improvement in standardized test scores cannot be explained by teachers grading URMs more leniently, this does not rule out the possibility that they focused more attention on improving URMs' learning.[29] Another alternative explanation is that the change in affirmative action policy changed parents' or guidance counselors' perception of a student's returns to human capital investment and led them to become more involved with the students. To provide suggestive evidence on the drivers of URM students' improved outcomes, we analyze students' responses from the THEOP survey.

As mentioned previously, the THEOP survey asked two cross-sections of high school seniors across Texas about their demographics, college applications behavior, and high school activities in 2002 (pre-affirmative action) and then again in 2004 (post-affirmative action). While the two waves of the survey are not identical, the questions that are consistent across waves allow us to measure student effort in terms of time spent on homework, as well as parental involvement, and guidance counselor involvement. For each outcome, we run the following regression, which closely mirrors our difference-in-differences strategies in the TEA

[29] However, our findings for the SAT suggest that if this is the case, it did not have negative spillovers for whites' outcomes.

and LUSD data:[30]

$$y_{iet} = \beta_1 Post2003_i + \beta_2 \left(URM_i \times Post2003_t\right) + \alpha_e + \varepsilon_{iet}, \tag{8}$$

where i denotes an individual, e denotes an ethnicity, and t denotes a survey round. $Post2003_i$ is an indicator variable equal to 1 for seniors surveyed in 2004, while α_e is an ethnicity fixed effect. This regression compares the change in outcomes between URM and white seniors from 2002 to 2004.

Table 7 reports the results. As column (1) shows, after the implementation of affirmative action, URM high school seniors spend 5 minutes more on homework a day, equivalent to 10% more time on homework outside of school relative to white students. URM students are also 5 percentage points more likely to apply to their first choice college after the policy change compared to whites, consistent with our findings in the TEA data. However, we do not see any changes in an index of variables measuring parental involvement or the likelihood of discussing college applications with guidance counselors after affirmative action is put in place. Overall, Table 7 provides suggestive evidence that students directly responded to the change in affirmative action policies by changing their behavior.

7 Alternative Policies

This section discusses alternative educational policies that were enacted in the early 2000s and discusses whether they can explain our findings. The first subsection discusses No Child Left Behind (NCLB). The second subsection discusses charter school expansions.

No Child Left Behind

One threat to the validity of our findings is that a major national educational policy, No Child Left Behind (NCLB), was signed into law in 2002. NCLB may have also differentially affected URM students' outcomes, confounding our estimates. We believe that this is unlikely to be the case for several reasons. First, as documented by Dee and Jacob (2011) and Deming et al. (2016), Texas has had high-stakes school accountability policies since 1993. These policies, which were adopted under Governor George Bush, served as the later basis for the NCLB policies enacted when George Bush was president (Deming et al., 2016). Second, our

[30] In this analysis, we cannot include campus fixed effects because we do not observe the campus to which the student belongs.

SAT results exploit geographic variation in the reinstatement of affirmative action policies. Since NCLB was a national law, we do not expect it to differentially positively affect URMs specifically in Texas, Louisiana, and Mississippi.[31] Third, we find that affirmative action had the largest effects on high-achieving students who would have been on the margin of college admissions. In contrast, NCLB incentivized schools to ensure students passed relatively low proficiency cut-offs. Consistent with this, Neal and Schanzenbach (2010) show that NCLB and similar policies increased test scores among the middle of the test score distribution. Thus, the distribution of effects we estimate is inconsistent with NCLB's incentive system and with past estimates of the effects of the NCLB program.

Charter School Expansion

In the early-2000s, charter schools expanded rapidly in Texas. Since these charter schools typically serve disadvantaged populations, they may have also differentially affected URM students' outcomes. However, we believe this expansion is unlikely to drive our results since, at the time of the policy change (2003-2004), charter school enrollment made up only 1% of total enrollment in Texas (Texas Education Agency, 2004).[32] Nonetheless, as a robustness test, we also omit Houston and Dallas, the two areas with by far the largest number of students enrolled in charter schools today (Texas Charter Schools Association, 2016), from the analysis and re-estimate our college applications regressions. These results are reported in Appendix Table A10 and are again very similar to the main estimates.

8 Conclusion

In this paper, we study the effects of a 2003 U.S. Supreme Court ruling that effectively reinstated race-based affirmative action policies in public universities in Texas, Louisiana, and Mississippi. We find that the policy increased applications to selective colleges, high school attendance, and college graduation by URMs in Texas. The policy also reduced the racial achievement gap in math SAT scores in the affected states by 5%. In the context of a large, urban school district in Texas, the policy reduced the racial achievement gap in grades by 18%.

[31]If anything, given that Texas should be *less* affected by NCLB due to its pre-existing policies, we should expect our estimates of the change in SAT scores for URMs in Texas, Louisiana and Mississippi will be under-estimates due to NCLB.

[32]In 2003-2004, there were 60,833 students in charter schools in Texas and 4,328,028 enrolled overall (Texas Education Agency, 2004).

The effects we observe are concentrated among higher ability students, who are likely on the margin of admission to a selective public school. We verify that these students experience the greatest change in the effects of moving up an ability decile on admissions to selective universities following the policy. Thus, the students whose returns to human capital investment increase the most are also those whose pre-college human capital investment is most affected by the policy.

Altogether, given the positive effects on attendance, the distribution of the treatment effects, and the evidence from the survey data, our findings suggest that URM students respond to the affirmative action policy by changing their college aspirations and adjust their effort accordingly. We speculate that these results are consistent with work by Hoxby and Avery (2012) and Hoxby and Turner (2013), which shows that qualified, disadvantaged students are less likely to apply to highly selective four-year institutions. If affirmative action leads URM students to perceive admission to a selective school as more attainable or colleges as more welcoming, it may change both their application behavior and their pre-college human capital investment.

Finally, the meaningful effect sizes we estimate on a variety of dimensions suggest that policy debates that ignore the pre-college incentive effects of affirmative action policies ignore a significant effect of these policies. Specifically, we show that affirmative action policies do not merely affect admissions among a given pool of applicants, they affect the composition of the pool of university applicants itself. Given the importance of the racial achievement gap for determining gaps in long-term outcomes, reductions in the achievement gap may translate into substantial changes in welfare later in life.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program," *Journal of the American statistical Association*, 2010, *105* (490), 493–505.

_ **and Javier Gardeazabal**, "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 2003, *93* (1), 113–132.

Antonovics, Kate and Ben Backes, "The effect of banning affirmative action on human capital accumulation prior to college entry," *IZA Journal of Labor Economics*, 2014, *3* (1), 5.

Arcidiacono, Peter, "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524.

_ **, Esteban M Aucejo, and V Joseph Hotz**, "University differences in the graduation of minorities in STEM fields: Evidence from California," *American Economic Review*, 2016, *106* (3), 525–62.

_ **, Michael Lovenheim, and Maria Zhu**, "Affirmative Action in Undergraduate Education," *Annual Review of Economics*, 2015, *7* (1), 487–518.

Bodoh-Creed, Aaron and Brent R Hickman, "Pre-college human capital investment and affirmative action: a structural policy analysis of US college admissions," *Working Paper*, 2018.

Bowen, William G and Derek Bok, *The Shape of the River. Long-Term Consequences of Considering Race in College and University Admissions.*, ERIC, 1998.

Card, David and Alan B Krueger, "Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas," *Industrial & Labor Relations Review*, 2005, *58* (3), 416–434.

Cassan, Guilhem, "Affirmative action, education and gender: Evidence from India," *Journal of Development Economics*, 2019, *136*, 51–70.

Center for Education Statistics, "Enrollment in Postsecondary Institutions, Fall 2004; Graduation Rates, 1998 2001 Cohorts; and Financial Statistics, Fiscal Year 2004," *U.S. Department of Education*, 2004.

Coate, Stephen and Glenn C Loury, "Will affirmative-action policies eliminate negative stereotypes?," *The American Economic Review*, 1993, pp. 1220–1240.

Cortes, Kalena E and Lei Zhang, "The incentive effects of the top 10% plan," *Working Paper*, 2011.

Cotton, Christopher, Brent R Hickman, and Joseph P Price, "Affirmative action and human capital investment: Theory and evidence from a randomized field experiment," *Working Paper*, 2015.

Dee, Thomas S and Brian Jacob, "The impact of No Child Left Behind on student achievement," *Journal of Policy Analysis and management*, 2011, *30* (3), 418–446.

Deming, David J, Sarah Cohodes, Jennifer Jennings, and Christopher Jencks, "School accountability, postsecondary attainment, and earnings," *Review of Economics and Statistics*, 2016, *98* (5), 848–862.

Estevan, Fernanda, Thomas Gall, and Louis-Philippe Morin, "Redistribution without Distortion: Evidence from An Affirmative Action Programme At a Large Brazilian University," *The Economic Journal*, 2018.

Ferman, Bruno and Juliano Assunção, "Does Affirmative Action Enhance or Undercut Investment Incentives? Evidence from Quotas in Brazilian Public Universities," *Working Paper*, 2015.

Fryer, Roland G and Glenn C Loury, "Affirmative Action and Its Mythology," *Journal of Economic Perspectives*, 2005, *19* (3), 147–162.

Golightly, Eleanor, "Does College Access Increase High School Effort? Evaluating the impact of the Texas Top 10% Rule," 2019.

Hinrichs, Peter, "The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities," *Review of Economics and Statistics*, 2012, *94* (3), 712–722.

_ , "Affirmative Action and Racial Segregation," 2016.

Hoxby, Caroline and Sarah Turner, "Expanding college opportunities for high-achieving, low income students," *Stanford Institute for Economic Policy Research Discussion Paper*, 2013, (12-014).

Hoxby, Caroline M and Christopher Avery, "The missing "one-offs": The hidden supply of high-achieving, low income students," *NBER Working Paper*, 2012.

Jayachandran, Seema and Adriana Lleras-Muney, "Life expectancy and human capital investments: Evidence from maternal mortality declines," *The Quarterly Journal of Economics*, 2009, *124* (1), 349–397.

Jensen, Robert, "The (perceived) returns to education and the demand for schooling," *The Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

_ , "Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India," *The Quarterly Journal of Economics*, 2012, *127* (2), 753–792.

Khanna, Gaurav, "Does Affirmative Action Incentivize Schooling? Evidence From India," *Working Paper*, 2016.

Leeds, Daniel M, Isaac McFarlin, and Lindsay Daugherty, "Does student effort respond to incentives? Evidence from a guaranteed college admissions program," *Research in Higher Education*, 2017, *58* (3), 231–243.

Neal, Derek and Diane Whitmore Schanzenbach, "Left behind by design: Proficiency counts and test-based accountability," *The Review of Economics and Statistics*, 2010, *92* (2), 263–283.

Office of the President, "A Report on the Top Ten Percent Law," Technical Report, The University of Texas at Austin 2008.

Oster, Emily and Bryce Millett Steinberg, "Do IT service centers promote school enrollment? Evidence from India," *Journal of Development Economics*, 2013, *104*, 123–135.

Parker, Claire, "UT-Austin has no plans to drop affirmative action policy, despite new Trump administration guidelines," *Texas Tribune*, 2018.

Rothstein, Jesse and Albert H Yoon, "Affirmative Action in Law School Admissions: What Do Racial Preferences Do?," Technical Report, NBER 2008.

Sander, Richard H, "A systemic analysis of affirmative action in American law schools," *Stanford Law Review*, 2004, pp. 367–483.

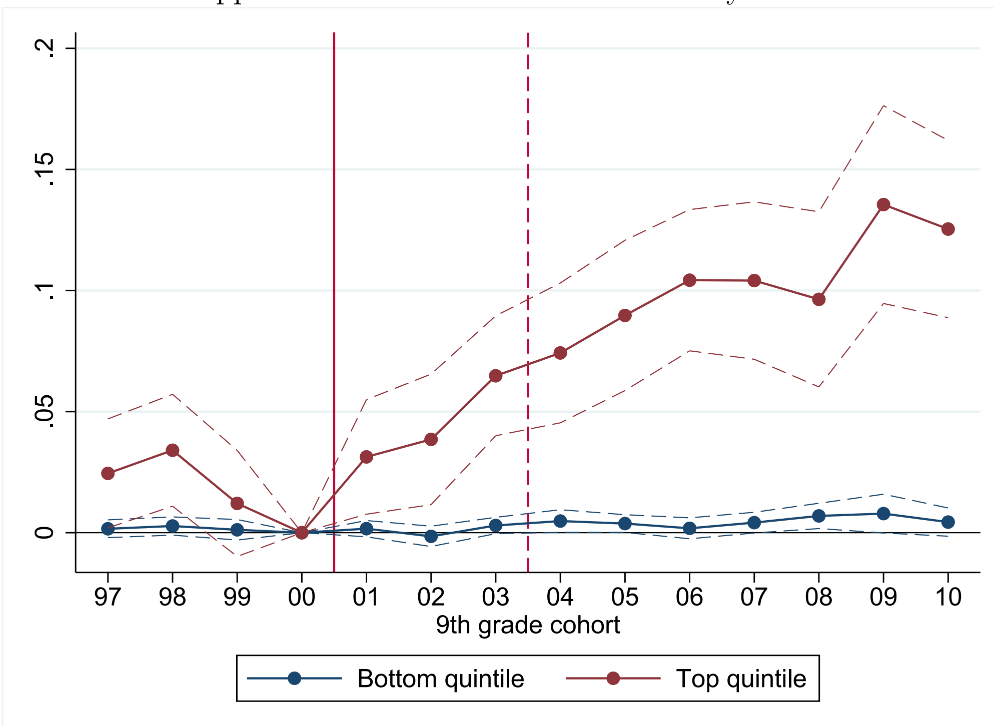
Texas Charter Schools Association, "Texas Public Charter Schools 20 Years Later: A State of the Sector Report," 2016.

Texas Education Agency, "Enrollment in Texas Public Schools 2003-2004," 2004.

Tutson, Teddy, "Students now taking TAKS instead of TAAS tests," *Houston Chronicle*, 2002.

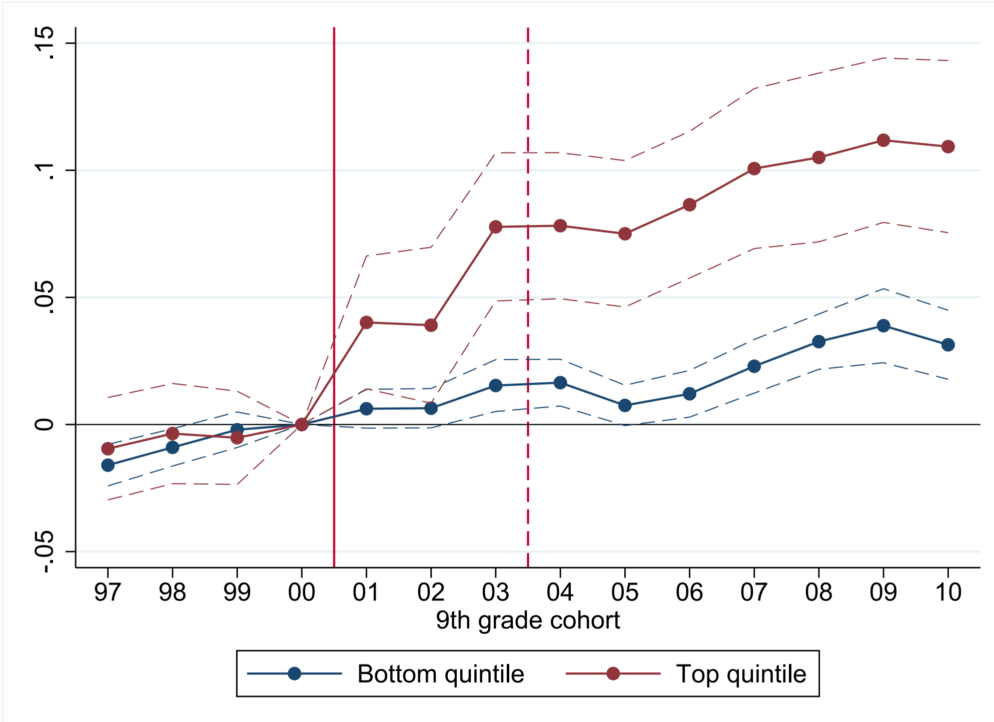
Figures

Figure 1: Number of Applications to Selective Universities by URMs Relative to Whites



Notes: The outcome is the average number of applications sent to selective universities by students within 4 years of starting 9^{th} grade. Dots indicate coefficients from a regression of the outcome on year dummies interacted with URM status separately for students in the bottom and top quintiles of the ability distribution, where ability is given by quintiles of cohort-specific distribution of 6^{th} grade standard test scores. All regressions condition on cohort, race and district fixed effects, as well as means of individual characteristics at the district-cohort-ethnicity-ability quintile level. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

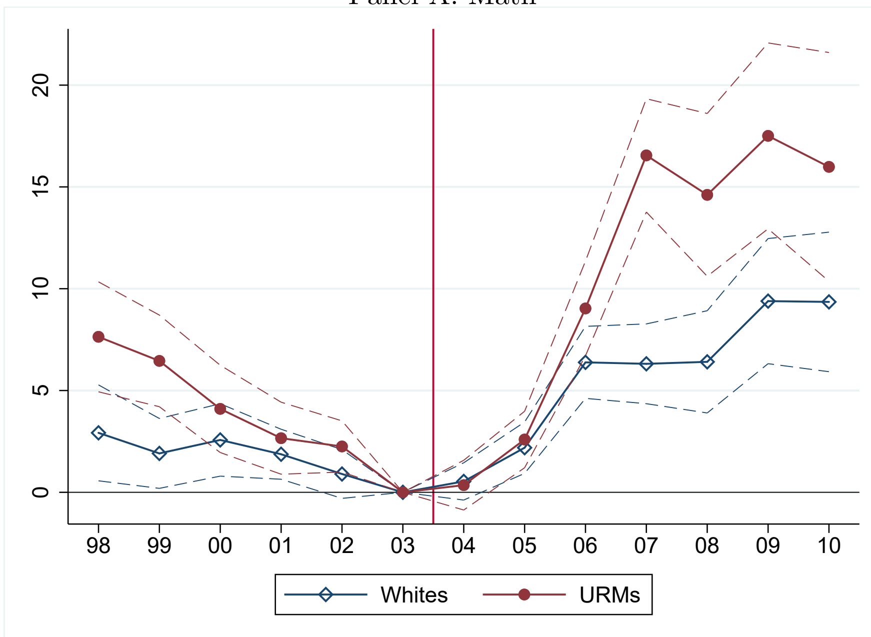
Figure 2: Probability of Applying to Any University by URMs Relative to Whites



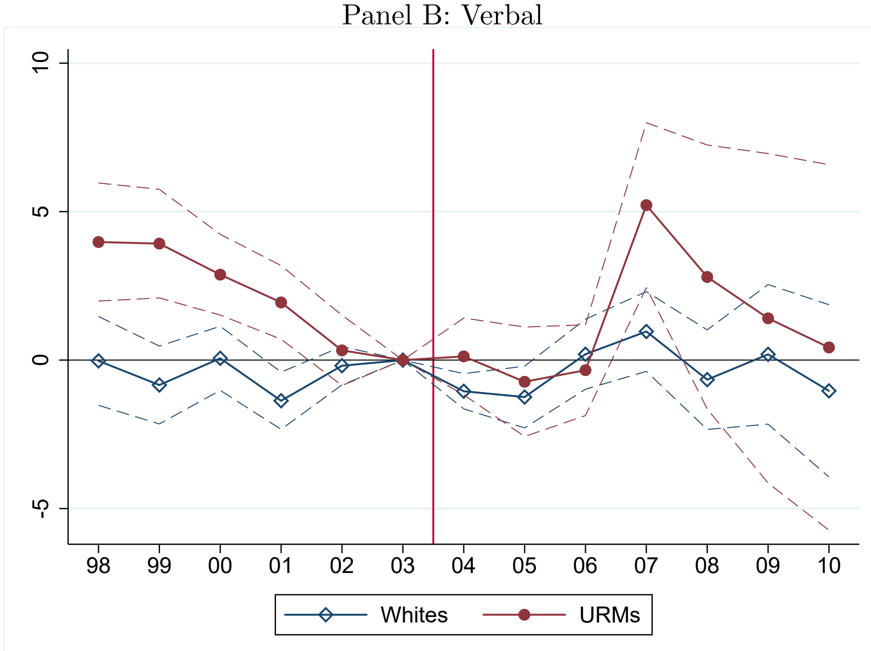
Notes: The outcome is the probability of applying to any university within 4 years of starting 9^{th} grade. Dots indicate coefficients from a regression of the outcome on year dummies interacted with URM status separately for students in the bottom and top quintiles of the ability distribution, where ability is given by quintiles of the cohort-specific distribution of 6^{th} grade standard test scores. All regressions condition on cohort, race and district fixed effects, as well as means of individual characteristics at the district-cohort-ethnicity-ability quintile level. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Figure 3: Effect of AA on SAT Scores for URMs and Whites

Panel A: Math

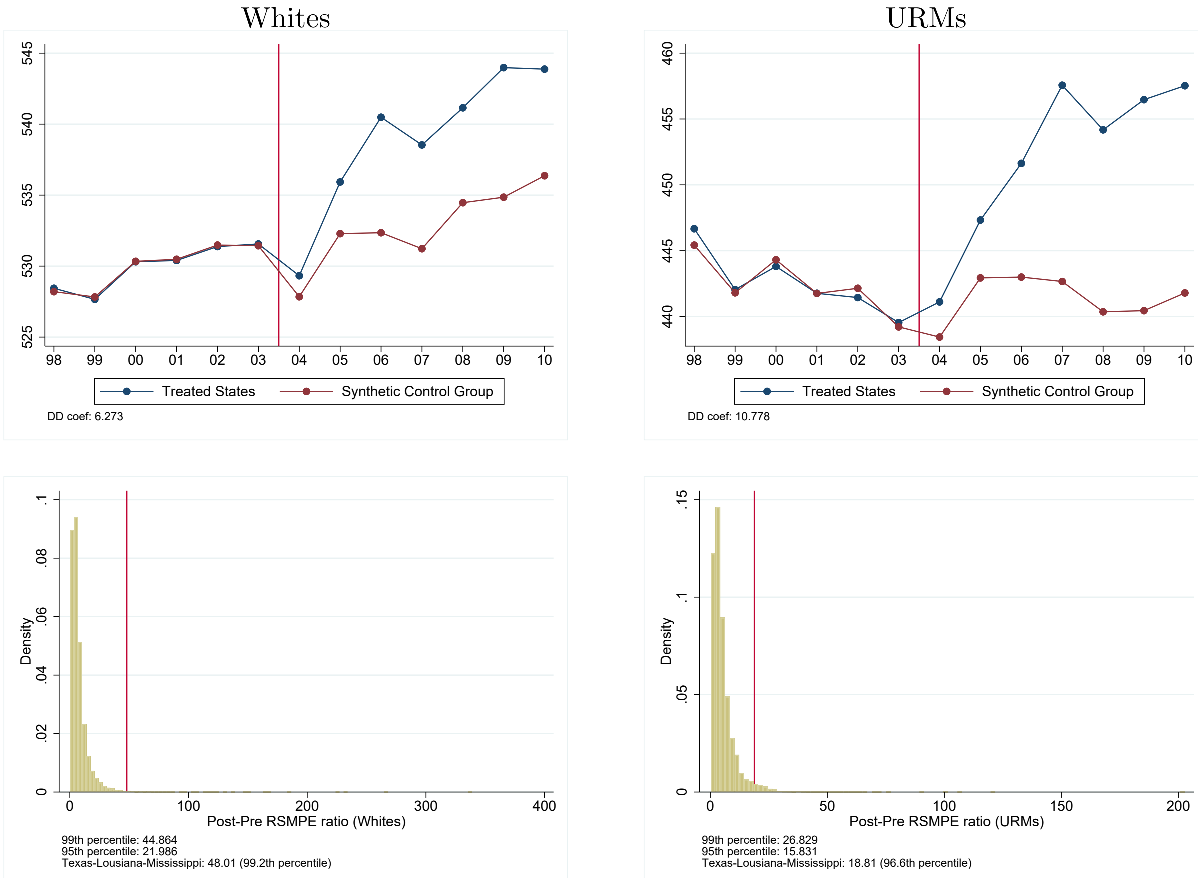


Panel B: Verbal



Notes: The outcome is average SAT scores at the state-year level. Dots indicate coefficients of regressions of the outcome on year dummies interacted with an indicator variable for the three treated states, estimated separately for white and URM students. Cells are weighted by the number of SAT test takers. Dashed lines show 95% confidence intervals for standard errors clustered at the state level.

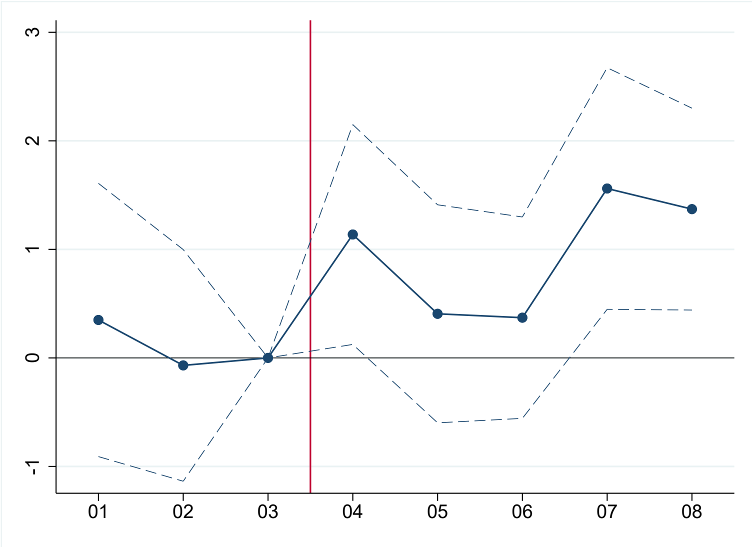
Figure 4: Synthetic Control Estimates of the Effect of AA on SAT Math Scores



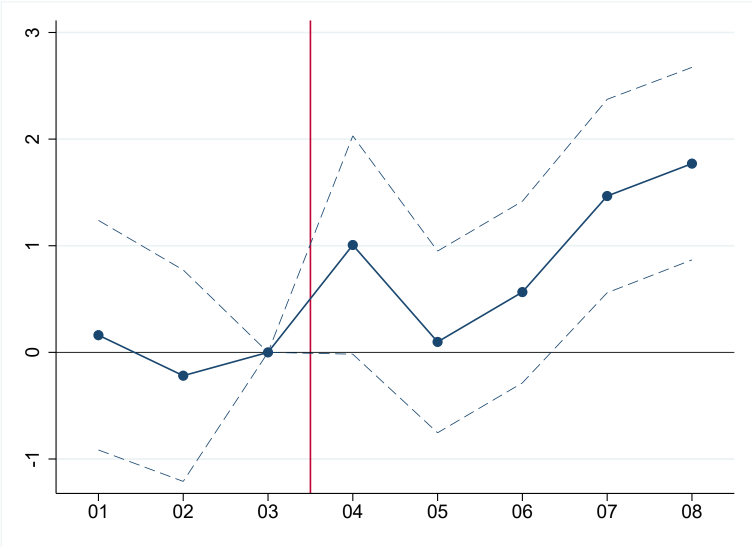
Notes: This figure reports synthetic control analyses separately for whites and URMs. The top panel shows SAT math scores for the treated states (Texas, Mississippi and Louisiana) and for the associated synthetic control group. The bottom panel shows the distribution of post/pre RMSPE ratio for placebo estimates. The vertical red line in the bottom panels indicates the post/pre RMSPE ratio for the treated states. For whites, weights on control units are 42.5% (California), 40.8% (Florida), 8.3% (Pennsylvania), 6.2% (New York), and 2.2% (Indiana). All other states have a weight of zero. For URMs, weights on control units are 33.2% (Oregon), 28.4% (New Jersey), 20.6% (California), and 17.8% (Pennsylvania).

Figure 5: Raw and Value-Added Estimates of AA's Effect on Mean Grades for URMs Relative to Whites

Panel A: Raw Grades

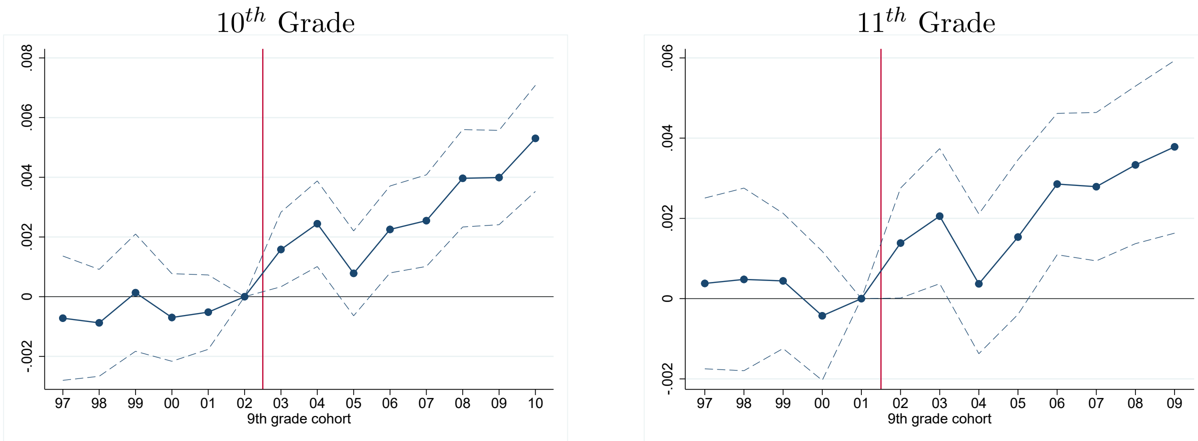


Panel B: Value-Added Specification



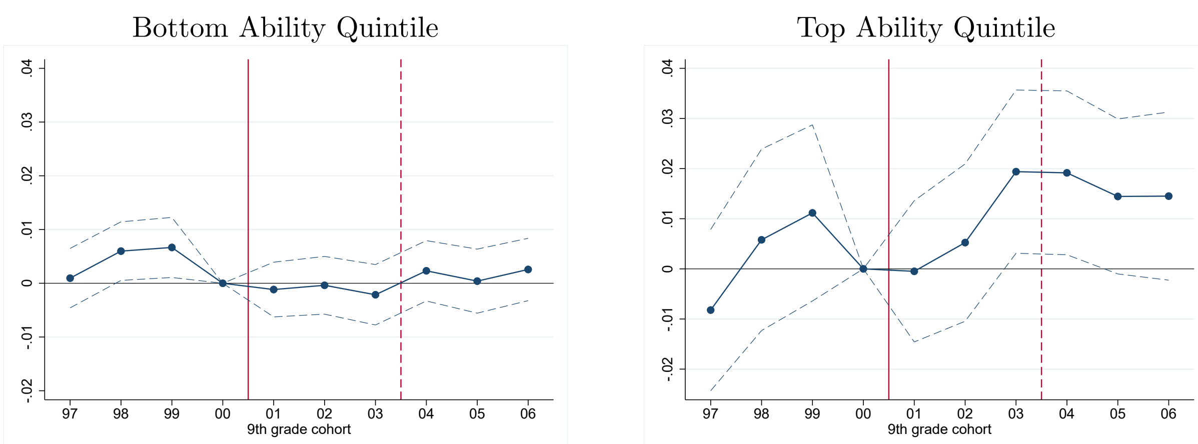
Notes: The outcome in both figures is mean grades across subjects in 11^{th} grade. Dots indicate the coefficients from a regression of the outcome on year dummies interacted with an indicator variable for URM status. The regression also includes school-cohort, race, and ZIP code fixed effects, as well as controls for age and gender. The bottom figure additionally includes controls for 8th grade grades, so the coefficients are in value-added terms. Dashed lines show 95% confidence intervals for standard errors clustered at the school-cohort level.

Figure 6: Effect of AA on Attendance for URMs Relative to Whites



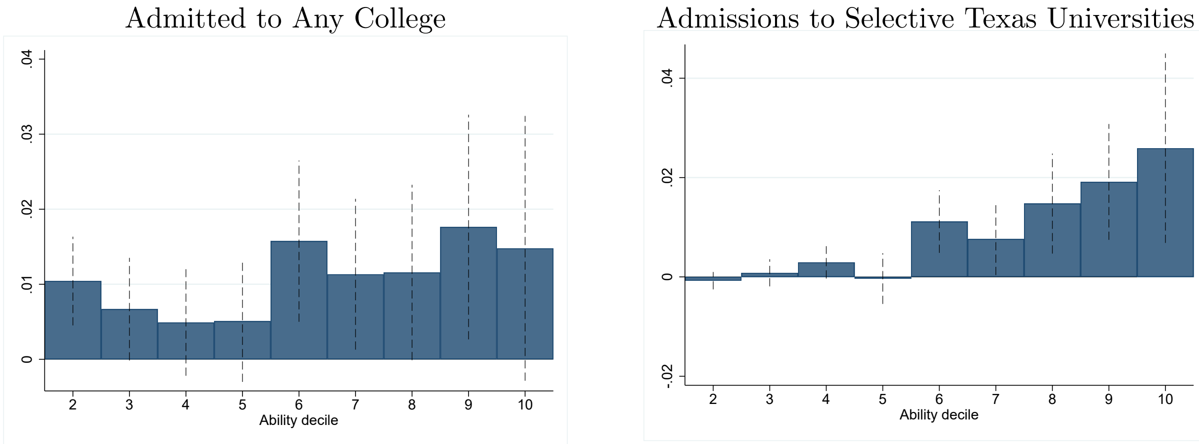
Notes: The outcomes are attendance rates in grades 10 and 11. Dots indicate coefficients from a regression of the outcome on year dummies interacted with URM status. All regressions condition on cohort-ability, race-ability and district-ability fixed effects, where ability is given by deciles of cohort-specific distribution of 6^{th} grade standard test scores. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Figure 7: Effect of AA On College Graduation for URMs Relative to Whites



Notes: The outcome is college graduation. Dots indicate coefficients from a regression of the outcome on year dummies interacted with URM status, separately for students in the bottom and top quintiles of the ability distribution. All regressions condition on cohort, race and district fixed effects, as well as means of individual characteristics. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Figure 8: Relative Change in Returns to Moving Up an Ability Decile in College Admissions



Notes: The outcome in the left graph is admission to any Texas college. The outcome in the right graph is the number of selective Texas schools to which a student is admitted. Bars are the coefficients from equation (7), which capture the change in the marginal effect of moving up an ability decile on college admissions. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Tables

Table 1: Summary Statistics

	URMs		Whites	
TEA Administrative Data				
Cohorts (grade 9)	1997-2000	2001-2010	1997-2000	2001-2010
Age (grade 9)	14.2684	14.1983	14.1648	14.1412
Limited English Proficiency (LEP)	0.0671	0.0477	0.0004	0.0004
Special Ed status	0.0772	0.0521	0.0784	0.0590
English as a Second Language (ESL)	0.042	0.0379	0.0001	0.0002
Gifted	0.0771	0.0852	0.1583	0.1599
Immigrant	0.0051	0.0014	0.0010	0.0004
Poor	0.6022	0.6628	0.1246	0.1570
Female	0.508	0.5079	0.4988	0.4963
Ability (decile)	4.3648	4.5541	6.6283	6.6263
Attendance rate (grade 10)	0.9343	0.9405	0.9541	0.9554
Attendance rate (grade 11)	0.9305	0.9335	0.9493	0.9494
University application rate (within 4 years)	0.1734	0.2583	0.2900	0.3350
Applications to selective universities (within 4 years)	0.0603	0.1012	0.2098	0.2384
College graduation rate	0.1126	0.0969	0.2488	0.2265
District-cohort-ability cells	12,492	36,462	17,414	41,614
Number of students	357,973	1,176,595	405,005	971,850
Number of districts	522	680	803	844
LUSD Administrative Data				
Cohorts (grade 11)	2001-2003	2004-2008	2001-2003	2004-2008
Age (grade 11)	16.3936	16.4087	16.2100	16.2234
Female	0.5377	0.5346	0.5057	0.5202
Mean school grades (grade 11)	77.3440	78.1689	82.2364	83.4534
Mean school grades (grade 8)	82.4995	81.9075	86.6246	86.8627
Attendance rate (grade 11)	0.9286	0.9274	0.9431	0.9482
Stanford test percentile rank (grade 11)	36.1245	49.7647	69.2039	77.8087
Number of students	17,620	34,107	3,623	5,779
Number of schools	42	49	36	42

Notes: This table reports summary statistics from the Texas Education Agency (TEA) administrative data and the administrative data from a large, urban school district (LUSD). An observation in the TEA data is a district-ability-cohort cell. The LUSD data consists of repeated cross-sections of 11th graders, and an observation is a student.

Table 2: Effect of AA on College Applications Behavior for URMs Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Applications to selective colleges					
Partial treatment	0.0095***	0.0017	0.0020	0.0022	0.0145**	0.0276***
	(0.0027)	(0.0019)	(0.0025)	(0.0034)	(0.0066)	(0.0086)
Full treatment	0.0190***	0.0016	0.0044*	0.0145***	0.0344***	0.0429***
	(0.0033)	(0.0014)	(0.0025)	(0.0040)	(0.0057)	(0.0099)
Observations (cells)	97121	18380	20681	20974	19960	17126
R^2	0.913	0.492	0.646	0.738	0.798	0.838
Mean dependent variable	0.1584	0.0100	0.0376	0.0941	0.2120	0.4426
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]			0.0032 [8.7649]			
Full treatment: p-value [F-stat]			0.0000 [18.2058]			
	Panel B: Application to any college					
Partial treatment	0.0078***	0.0086***	0.0046	0.0011	0.0086*	0.0222***
	(0.0026)	(0.0027)	(0.0035)	(0.0044)	(0.0052)	(0.0075)
Full treatment	0.0286***	0.0101***	0.0132***	0.0263***	0.0432***	0.0545***
	(0.0035)	(0.0027)	(0.0035)	(0.0051)	(0.0054)	(0.0086)
Observations (cells)	97121	18380	20681	20974	19960	17126
R^2	0.915	0.798	0.824	0.814	0.803	0.781
Mean dependent variable	0.2785	0.0789	0.1595	0.2505	0.3708	0.5330
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]			0.0902 [2.8770]			
Full treatment: p-value [F-stat]			0.0000 [28.8031]			
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' college applications behavior. The regressions use the TEA data, an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. Partial treatment is the coefficient on the interaction between an indicator for being a URM and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and being a URM. The outcome variable in panel A is the average number of selective colleges to which students applied. The outcome variable in Panel B is the fraction of students in a cell that applied to any college. Standard errors are clustered at the district level.

Table 3: Effect of AA on SAT Scores for URMs and Whites

	Math	Verbal	# Test takers	% Test takers
	(1)	(2)	(3)	(4)
	Panel A: URMs			
DD coefficient	7.998***	-0.779	531.8	0.0026
	(1.497)	(1.755)	(1162.0)	(0.0053)
Observations (cells)	1904	1901	1904	1116
R^2	0.844	0.795	0.802	0.877
State, year and ethnicity FE	X	X	X	X
	Panel B: Whites			
DD coefficient	4.145***	0.0247	1546.0	0.0052
	(0.994)	(0.878)	(1268.7)	(0.0045)
Observations (cells)	663	663	663	561
R^2	0.968	0.971	0.987	0.978
State, year and ethnicity FE	X	X	X	X
	Panel C: Triple-Differences			
DDD coefficient	3.975***	1.083	-379.4	-0.0021
	(0.872)	(0.825)	(1071.8)	(0.0025)
Observations (cells)	2555	2552	2555	1677
R^2	0.998	0.998	0.999	0.993
State-year FE	X	X	X	X
State-ethnicity FE	X	X	X	X
Ethnicity-year FE	X	X	X	X

Notes: This table reports differences-in-difference and triple-differences estimates of the effect of affirmative action on SAT scores. Each observation is a state-race-year group. In columns (1) and (2), cells are weighted by the number of test-takers in a group. In column (3), cells are weighted by the average number of test-takers in the pre-treatment years 1998-2000. In column (4), cells are weighted by the number of 17-19 year olds in the population group (from ACS), and the dependent variable is (# of test-takers)/(# of 17-19 years old). In Panels A and B, the DD coefficient reports the interaction of an indicator variable for belonging to a treated state (Texas, Louisiana, Mississippi) and being tested after *Grutter v. Bollinger* (post 2003). In Panel C, the coefficient is on the interaction between being a URM, being tested post 2003, and belonging to a treated state. Standard errors are clustered at the state level.

Table 4: Effect of AA on School Grades for URMs Relative to Whites

	All students			Ability distribution		
				Bottom tercile	Middle tercile	Top tercile
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.8770***	1.0024***	0.9552***	0.8816*	0.3996	1.3859***
	(0.3086)	(0.2979)	(0.3114)	(0.5102)	(0.3906)	(0.4207)
Lagged dep. var. (grade 8)		0.5552***				
		(0.0092)				
Observations	61089	46346	92847	15874	15621	14776
R^2	0.226	0.345	0.784	0.189	0.224	0.208
Mean dependent variable	78.67	79.48	81.11	75.79	79.49	83.46
S.D. dependent variable	8.67	7.80	7.37	7.43	6.99	6.97
Test: Bottom tercile = Top tercile						
p-value [F-stat]			0.4412 [0.5948]			
School-year FE	X	X	X	X	X	X
Ethnicity FE	X	X		X	X	X
Demographic controls	X	X		X	X	X
Student FE			X			
Grade-year FE			X			
Grade-ethnicity FE			X			

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on grades in a large urban school district. An observation is a student, and the sample consists of repeated cross-sections of 11th graders. "Treated" is the coefficient on the interaction between being a URM and being observed post 2003. Ability terciles are assigned based on 8th grade average school grades. Standard errors are clustered at the school level.

Table 5: Effect of AA on School Attendance for URMs Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Attendance in grade 10					
Treated	0.0036***	0.0045***	0.0024***	0.0039***	0.0036***	0.0035***
	(0.0005)	(0.0012)	(0.0008)	(0.0008)	(0.0005)	(0.0006)
Observations (cells)	97071	18340	20677	20970	19958	17126
R^2	0.757	0.629	0.617	0.597	0.604	0.634
Mean dependent variable	0.9464	0.9238	0.9386	0.9479	0.9561	0.9653
Test: Bottom quintile = Top quintile						
p-value [F-stat]				0.4208 [0.6488]		
	Panel B: Attendance in grade 11					
Treated	0.0024***	0.0019	0.0012	0.0028***	0.0024***	0.0038***
	(0.0006)	(0.0014)	(0.0009)	(0.0009)	(0.0007)	(0.0006)
Observations (cells)	89849	16910	19120	19438	18532	15849
R^2	0.713	0.577	0.585	0.589	0.607	0.647
Mean dependent variable	0.9405	0.9199	0.9322	0.9409	0.9494	0.9596
Test: Bottom quintile = Top quintile						
p-value [F-stat]				0.1569 [2.0076]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' school attendance. The regressions use the TEA data, an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. The reported coefficient is the coefficient on the interaction between an indicator for being a URM and an indicator variable for being observed after 2003. The outcome variables in Panels A and B are the average percent of days students in a cell attended school in 10th and 11th grade respectively. Standard errors are clustered at the district level.

Table 6: Effect of AA on College Completion for URMs Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
Partial treatment	-0.0009	-0.0011	-0.0011	-0.0055	-0.0022	0.0098
	(0.0022)	(0.0018)	(0.0030)	(0.0036)	(0.0037)	(0.0063)
Full treatment	0.0046*	0.0006	0.0023	0.0033	0.0054	0.0141**
	(0.0025)	(0.0023)	(0.0031)	(0.0041)	(0.0049)	(0.0071)
Observations (cells)	68509	12933	14515	14809	14145	12107
R^2	0.890	0.556	0.640	0.690	0.708	0.707
Mean dependent variable	0.1688	0.0202	0.0695	0.1415	0.2398	0.3714
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.0955 [2.7856]		
Full treatment: p-value [F-stat]				0.0592 [3.5714]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' college graduation. The regressions use the TEA data, and an observation is at the district-cohort-race-ability quintile level. The ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. Partial treatment is the coefficient on the interaction between an indicator for being a URM and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and being a URM. The outcome variable is the fraction of students in a cell who completed college. Standard errors are clustered at the district level.

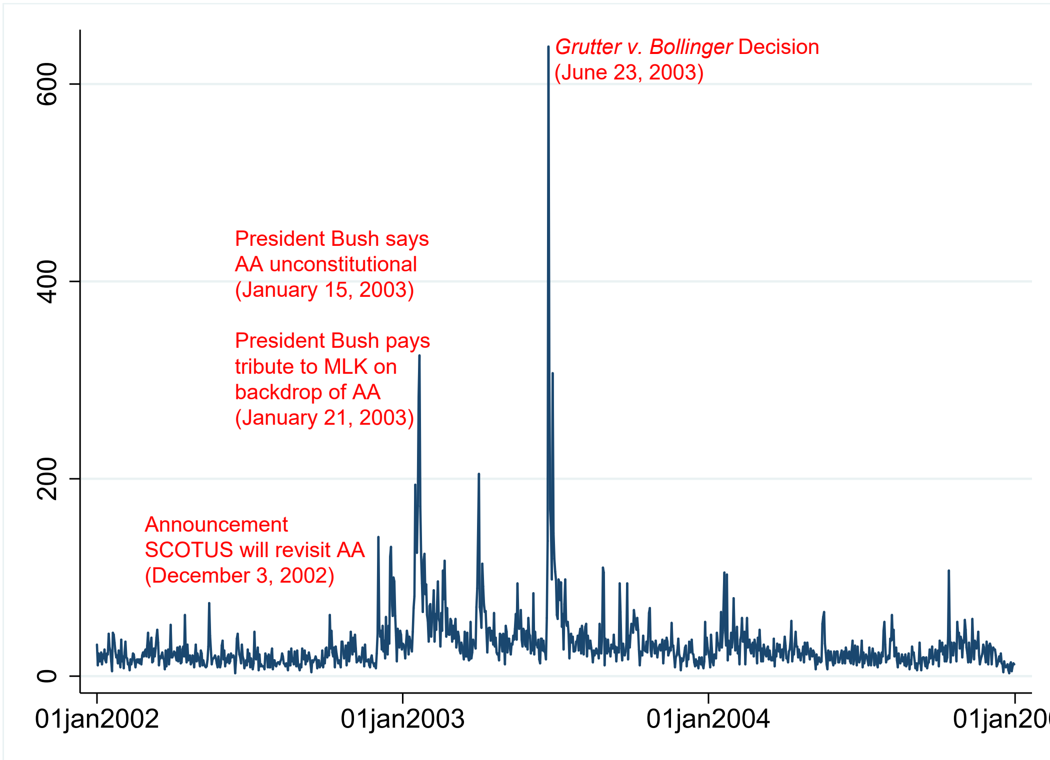
Table 7: Student and Parent Behavior and Affirmative Action

	(1) Time on Homework	(2) Applied to First Choice College	(3) Parental Involvement	(4) Guidance From Counselor
$URM \times Post2003$	5.452**	0.048**	0.176	-0.025
	(2.496)	(0.023)	(0.166)	(0.018)
Mean Whites Pre-2003	51.585	0.732	10.635	0.614
N	13,452	9,993	13,558	13,699
Adjusted R^2	0.061	0.024	0.038	0.026
Race Fixed Effects	X	X	X	X
Year control	X	X	X	X

Notes: This table reports differences-in-differences analyses using survey data from two cohorts, both in their senior year, of the Texas Higher Education Opportunity Project (THEOP). The earlier cohort was surveyed in 2002, and the later cohort was surveyed in 2004. For the measure of how many minutes per day students spend on homework, students were asked how many hours per day they spent on their homework and were given the options zero hours, less than 1 hour, 1 to 2 hours, 3 to 4 hours, and 5+ hours. We convert these to minutes so that 0 hours is 0 minutes, less than 1 hour is 30 minutes, 1 to 2 hours is 90 minutes, and so on. The parental involvement index is constructed using questions that ask "How often do your parents ... (i) give you special privileges because of good grades, (ii) try to make you work harder if you get bad grades, (iii) know when you are having difficulty in school, (iv) help with your school work, and (v) talk with you about problems in school." Students' responses range from "very rarely" (1) to "almost all the time" (4). We sum across the answers to these questions to construct the "parental involvement index" in a way that a higher index corresponds to more involvement along these dimensions. Standard errors are heteroskedasticity robust.

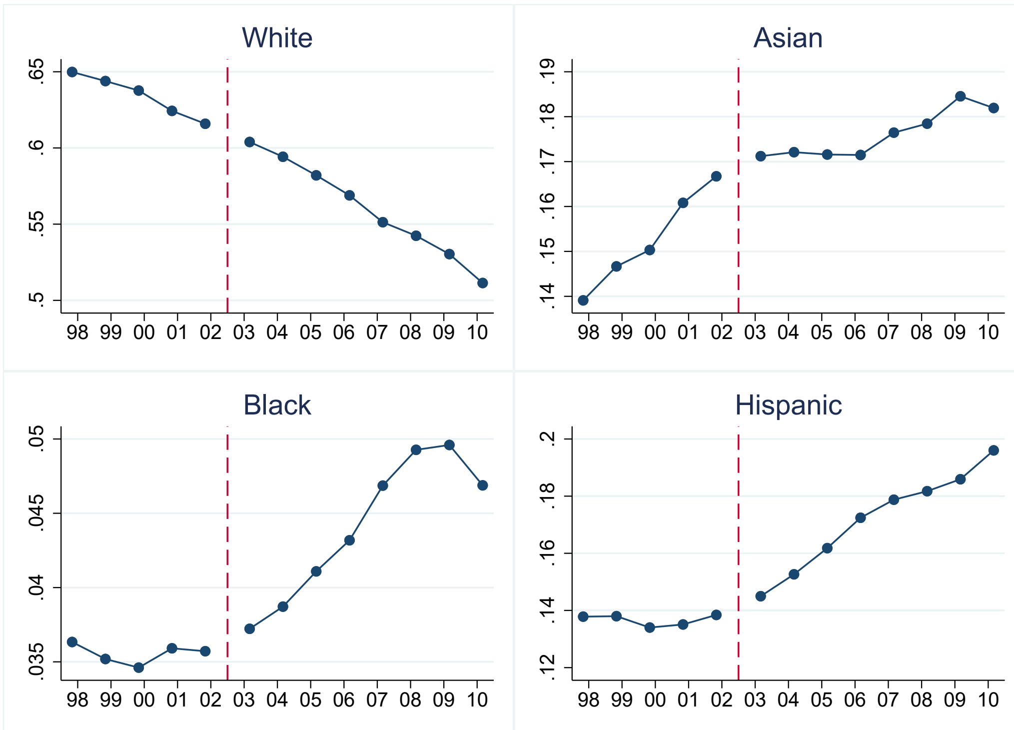
Appendix Figures

Figure A1: Number of Articles Mentioning Affirmative Action by Day, 2002-2004



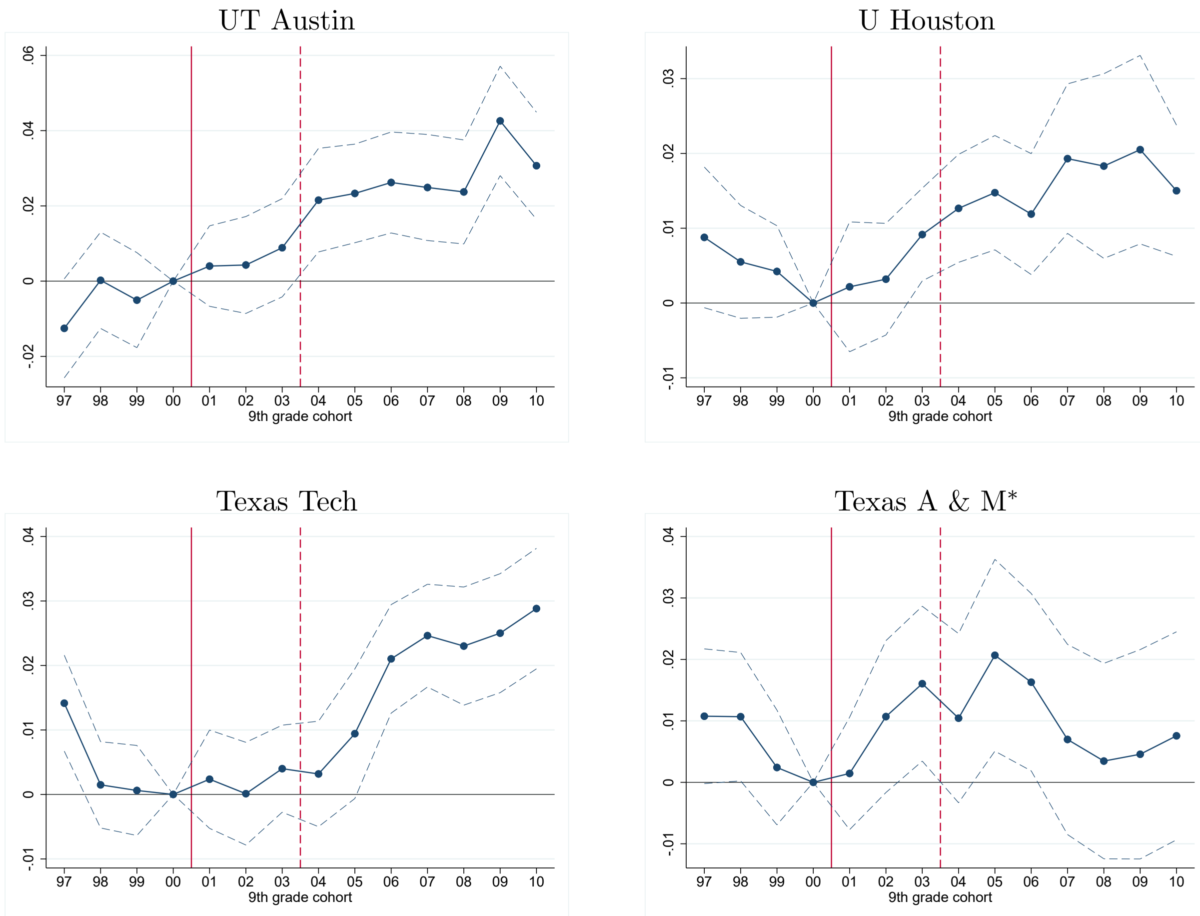
Notes: This figure reports the number of US newspaper articles by day that contained the phrase "affirmative action" on newslibrary.com.

Figure A2: Racial Composition of UT Austin by Year



Notes: This figure reports the racial composition of UT Austin's fall enrollment by year using data from the Integrated Postsecondary Education Data System (IPEDS).

Figure A3: Effect of AA on Admissions to Selective Institutions for URMs Relative to Whites



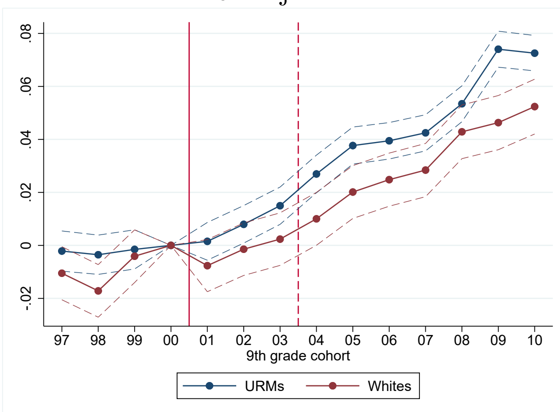
Notes: This figure reports event study graphs for the probability of a URM student receiving admission to each institution relative to a white student by students' 9th grade cohort. The regressions use the TEA data. Dotted lines report 95% confidence intervals with standard errors clustered at the district level.

*Texas A & M publicly announced that it would *not* use race (Parker, 2018).

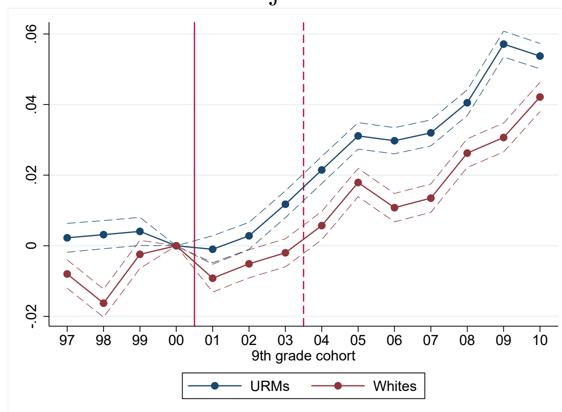
Figure A4: Trends in College Applications by Race

Panel A: Number of Applications to Selective Universities

Unadjusted

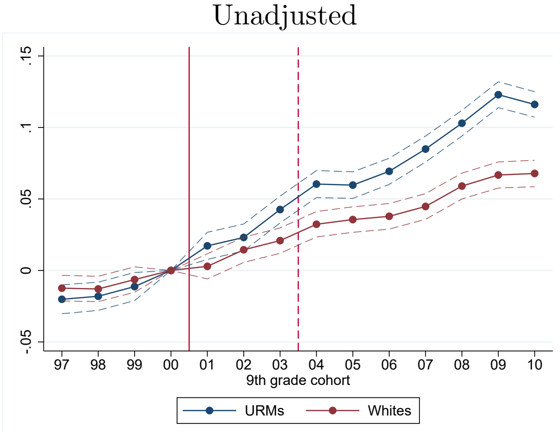


Adjusted

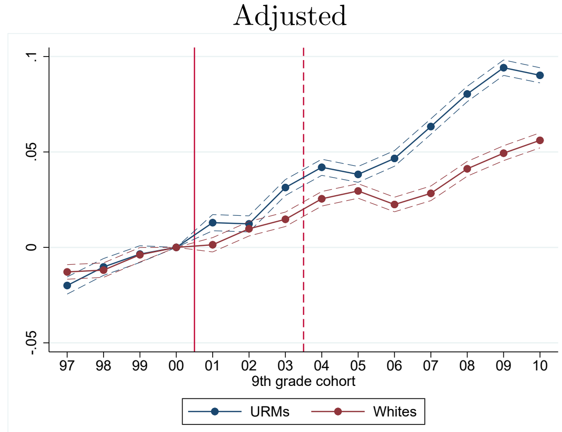


Panel B: Probability of Applying to Any University

Unadjusted



Adjusted

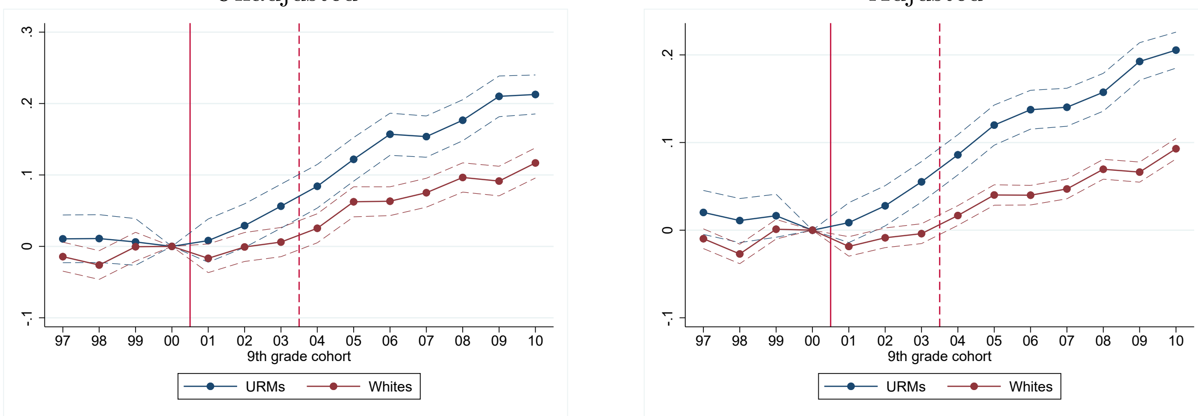


Notes: This figure reports trends in college applications behavior in our analytical sample. Time series are normalized relative to the cohort that entered 9th grade in 2000. Unadjusted figures directly plot raw averages. Adjusted figures are residuals from regressions on mean individual characteristics, race-ability fixed effects, and district-ability fixed effects.

Figure A5: Trends in College Applications in the Top Ability Quintile by Race

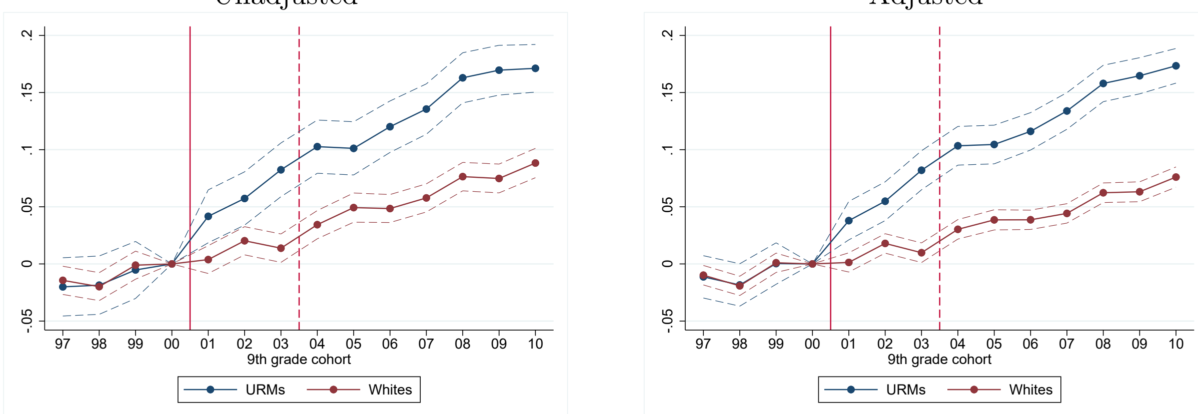
Panel A: Number of Applications to Selective Universities

Unadjusted　　　　Adjusted



Panel B: Probability of Applying to Any University

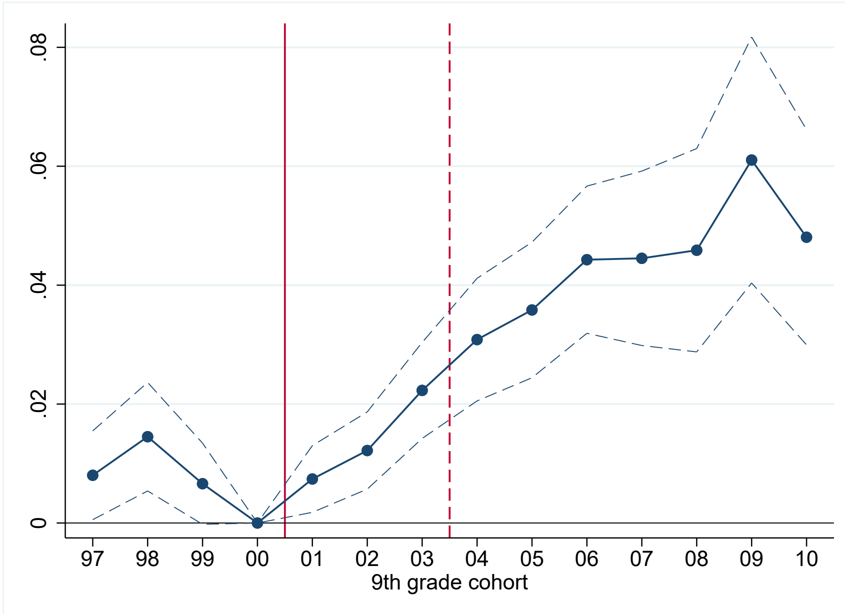
Unadjusted　　　　Adjusted



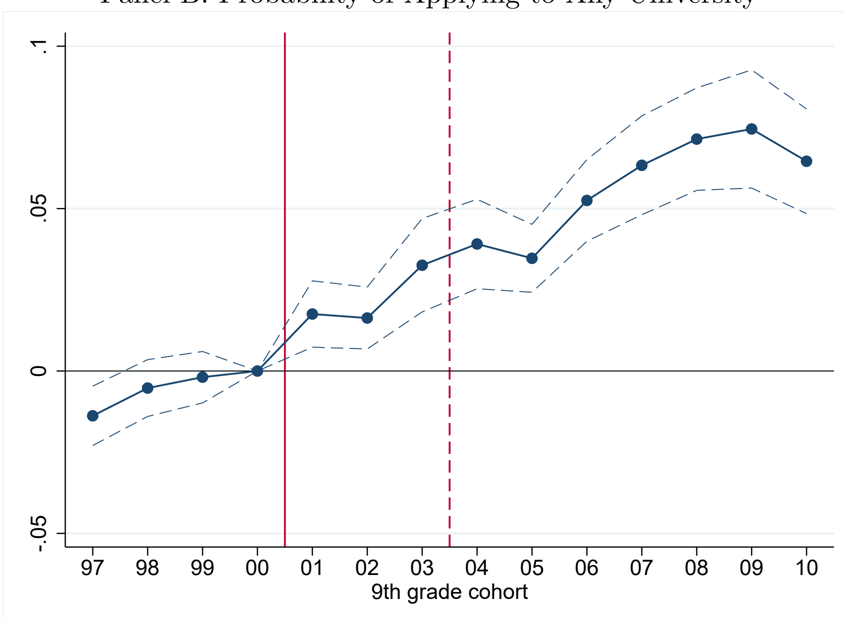
Notes: This figure reports trends in college applications behavior in our analytical sample. Time series are normalized relative to the cohort that entered 9th grade in 2000. Unadjusted figures directly plot raw averages. Adjusted figures are residuals from regressions on individual characteristics, race-ability fixed effects, and district-ability fixed effects.

Figure A6: Average Effect of AA on College Applications for URMs Relative to Whites

Panel A: Number of Applications to Selective Universities

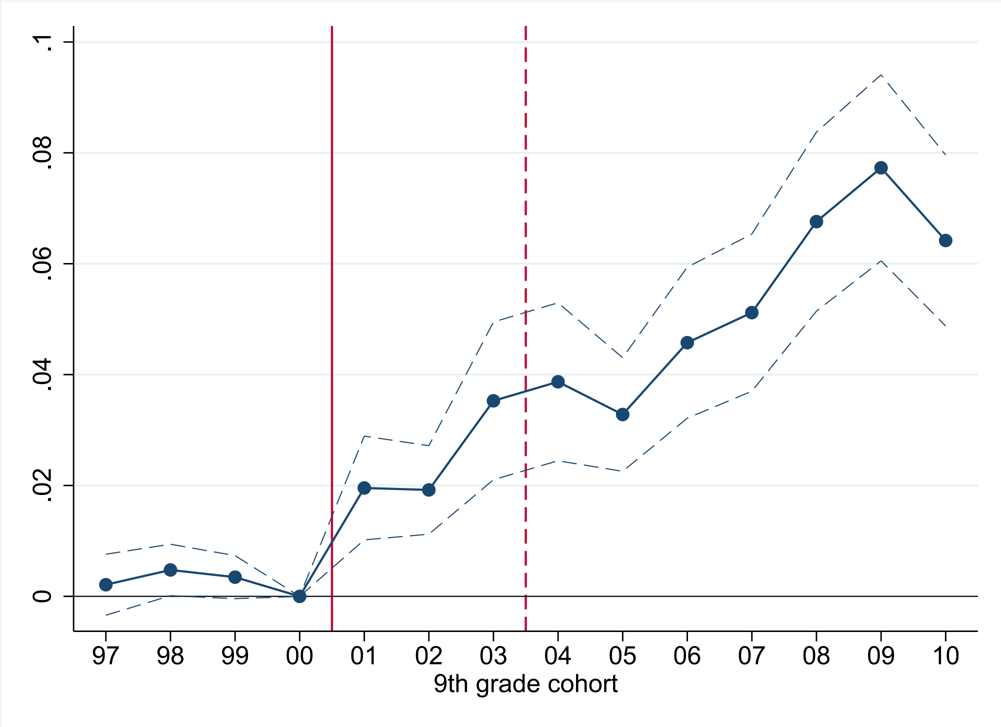


Panel B: Probability of Applying to Any University



Notes: The outcome is the average number of applications sent to selective universities by students in Panel A, and the probability of applying to any university in Panel B. Dots are coefficients from a regression of the outcome on year dummies interacted with URM status. All regressions condition on cohort-ability, race-ability and district-ability fixed effects, as well as means of individual characteristics, where ability is given by deciles of cohort-specific distribution of 6^{th} grade standard test scores. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

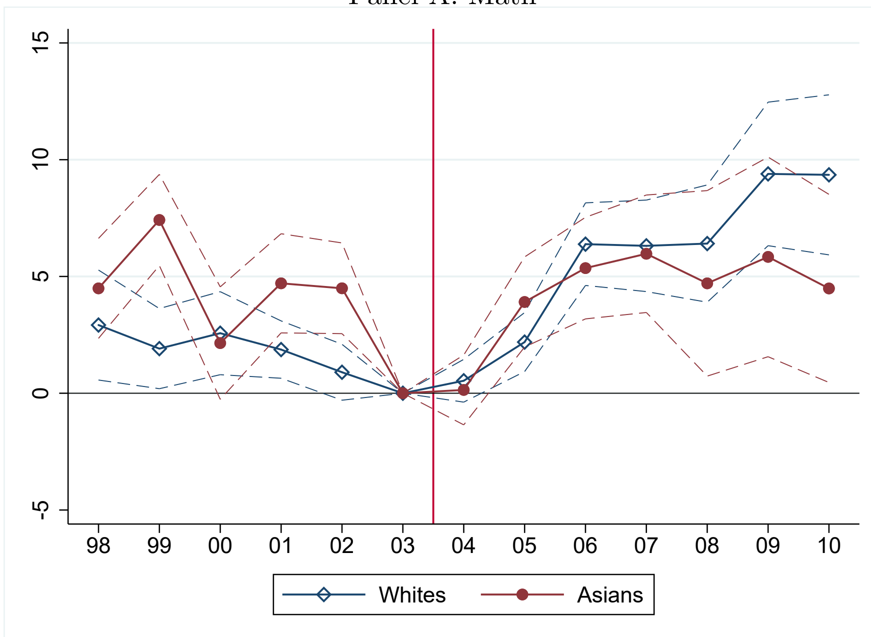
Figure A7: Effect of AA on Probability of Applying to Any UT Institution for URMs Relative to Whites



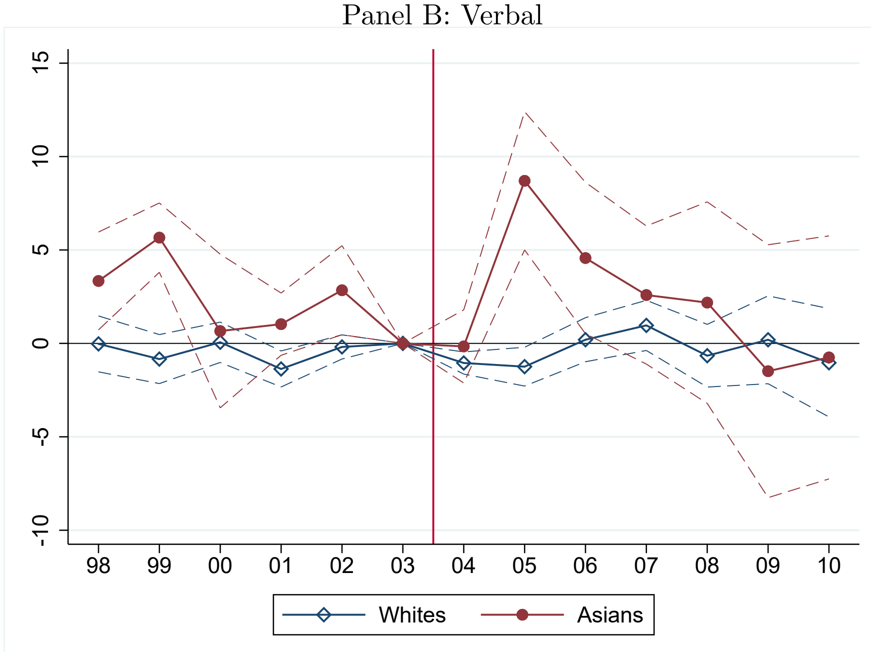
Notes: The outcome is the probability of applying to any university in the University of Texas System within 4 years of starting 9^{th} grade. Dots are coefficients from a regression of the outcome on year dummies interacted with URM status. All regressions condition on cohort-ability, race-ability and district-ability fixed effects, where ability is given by deciles of cohort-specific distribution of 6^{th} grade standard test scores. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Figure A8: Effect of AA on SAT Scores for Asians and Whites

Panel A: Math

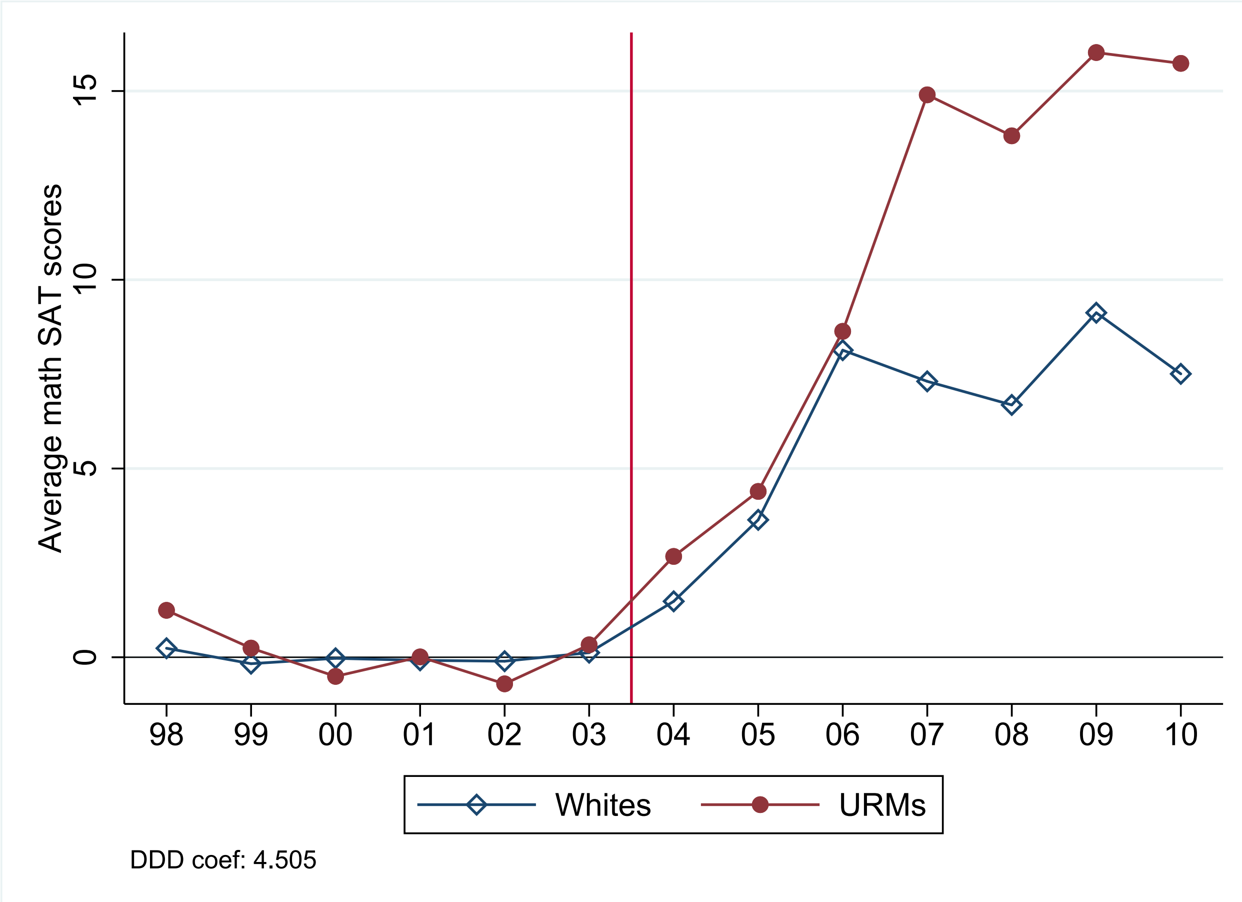


Panel B: Verbal



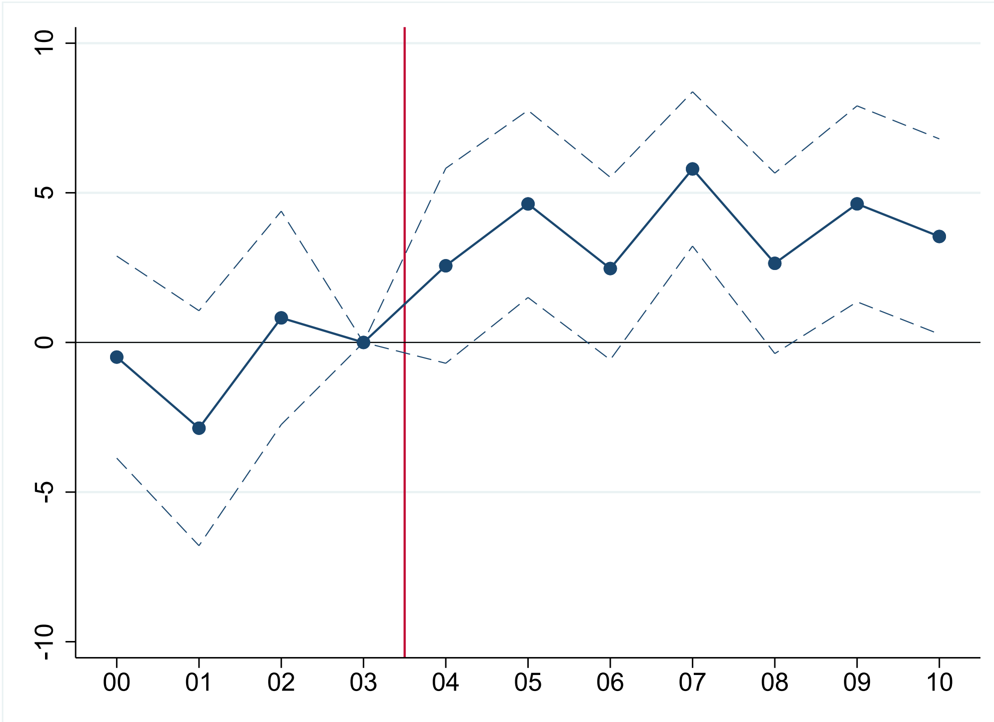
Notes: The outcome is state-year average SAT math and verbal scores. Dots indicate coefficients from a regression of the outcome on year indicator variables interacted with an indicator variable for the three treated states, estimated separately for Asian and white students. Cells are weighted by the number of SAT test takers. Dashed lines show 95% confidence intervals for standard errors clustered at the state level.

Figure A9: Difference Between Treated and Control States in the Synthetic Control Estimates



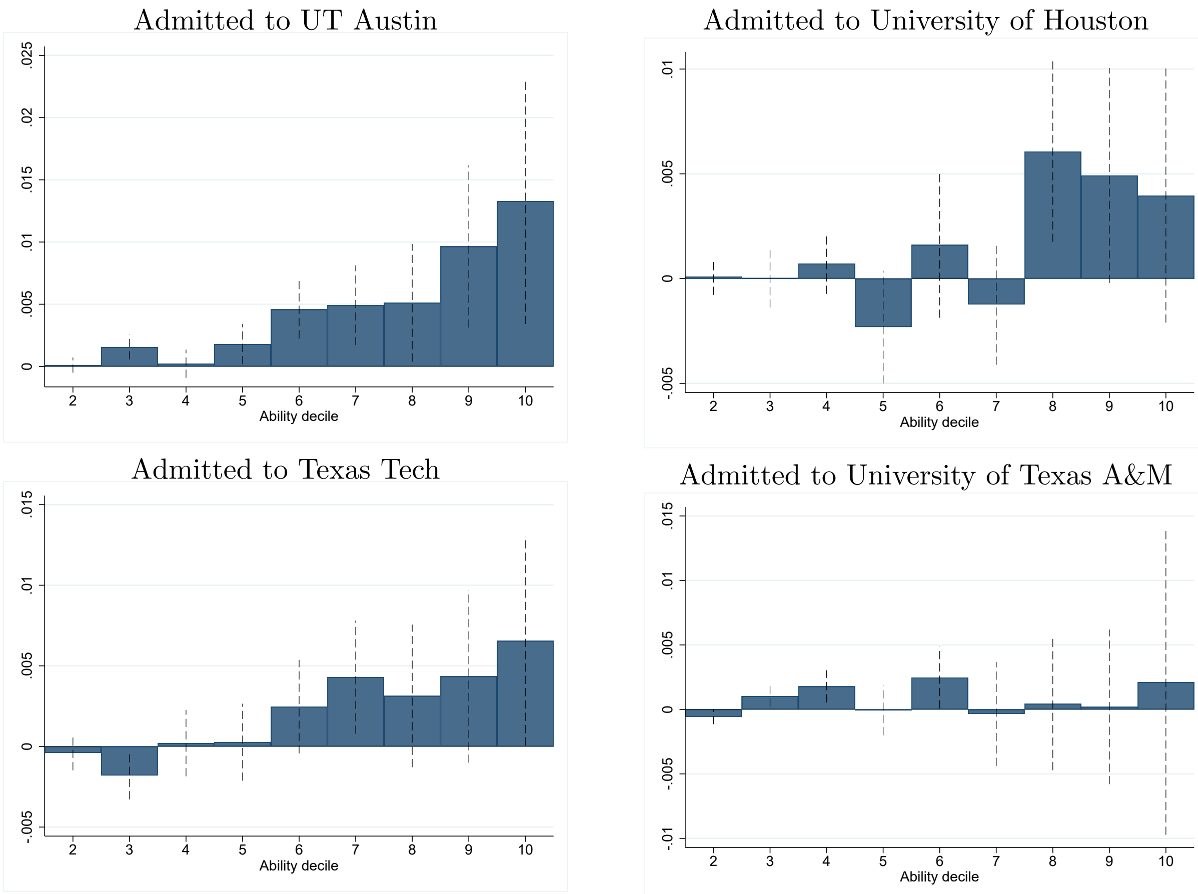
Notes: This figure reports differences in SAT math scores between treated states and synthetic control groups, separately for URMs and white students.

Figure A10: Effect of AA on Mean Stanford Scores for URMs Relative to Whites



Notes: The outcome is the mean percentile rank on the Stanford test in 11^{th} grade. Dots indicate coefficients from a regression of the outcome on year dummies interacted with an indicator variable for URM status. The regression also includes school-cohort, race, and ZIP code fixed effects, as well as controls for age and gender. Dashed lines show 95% confidence intervals for standard errors clustered at the school-cohort level.

Figure A11: Change in Relative Returns to Moving Up an Ability Decile on Admissions to Specific Texas Universities



Notes: The outcomes are admission to each of 4 selective Texas universities. Bars indicate the coefficients from equation (7), which capture the change in the marginal effect of moving up an ability decile on college admissions for URMs relative to whites. Dashed lines show 95% confidence intervals for standard errors clustered at the district level.

Appendix Tables

Table A1: Summary Statistics for SAT Data

	URMs		Whites	
Years	1998-2003	2004-2010	1998-2003	2004-2010
Verbal scores (mean)	440.9	441.7	527.7	528.4
Verbal scores (standard deviation)	21.5	21.7	18.7	19.8
Math scores (mean)	438.7	443.4	530.1	534.7
Math scores (standard deviation)	23.9	23.7	20.2	19.0
Number of cells	878	1,026	306	357
Number of SAT takers	1,194,067	2,159,747	4,136,869	5,634,200

Notes: This table reports summary statistics for the SAT data. An observation is a race-year-state cell. The reported standard deviation is calculated across cells and so mostly reflects variation across states.

Table A2: Summary Statistics for THEOP Survey Data

Panel A: Summary Statistics	Full Sample		Whites		URMs	
	Mean	SD	Mean	SD	Mean	SD
Time (Minutes) Spent on Homework	64.54	56.69	56.06	53.60	70.56	56.26
Applied to First Choice College	0.65	0.48	0.70	0.46	0.60	0.49
Parental Involvement Index (0-15)	5.98	3.87	5.94	3.78	6.18	3.96
Discussed College App. w. Counselor	0.67	0.47	0.65	0.48	0.70	0.46

Panel B: Total Numbers	N
Total Students	13,938
Whites	6,406
URMs	7,532
Students in 2002	11,098
Students in 2004	2,840

Notes: This table reports summary statistics for the Texas Higher Education Opportunity Project (THEOP) survey data for two cohorts of seniors, one in 2002 and one in 2004. For the measure of how many minutes per day students spend on homework, students were asked how many hours per day they spent on their homework and were given the options zero hours, less than 1 hour, 1 to 2 hours, 3 to 4 hours, and 5+ hours. We convert these to minutes so that 0 hours is 0 minutes, less than 1 hour is 30 minutes, 1 to 2 hours is 90 minutes, and so on. The parental involvement index is also constructed using several questions that ask "How often do your parents ... (i) give you special privileges because of good grades, (ii) try to make you work harder if you get bad grades, (iii) know when you are having difficulty in school, (iv) help with your school work, and (v) talk with you about problems in school." Students' responses range from "very rarely" (1) to "almost all the time" (4). We sum across the answers to these questions to construct the "parental involvement index" in a way that a higher index corresponds to more involvement along these dimensions, and renormalize the measure by subtracting 5 so that the minimum score is 0 rather than 5.

Table A3: Effect of AA on College Applications for Black Students Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Applications to selective colleges					
Partial treatment	0.0181***	0.0040	0.0093**	0.0100	0.0216**	0.0679***
	(0.0041)	(0.0030)	(0.0045)	(0.0070)	(0.0104)	(0.0145)
Full treatment	0.0270***	0.0062***	0.0121***	0.0254***	0.0515***	0.0654***
	(0.0052)	(0.0022)	(0.0038)	(0.0059)	(0.0107)	(0.0197)
Observations (cells)	64017	10546	12868	13882	13963	12758
R^2	0.922	0.522	0.664	0.754	0.819	0.860
Mean dependent variable	0.2070	0.0153	0.0482	0.1112	0.2348	0.4681
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]	0.0000 [18.6942]					
Full treatment: p-value [F-stat]	0.0019 [9.6814]					
	Panel B: Application to any college					
Partial treatment	0.0240***	0.0184***	0.0276***	0.0129	0.0225***	0.0581***
	(0.0039)	(0.0040)	(0.0054)	(0.0085)	(0.0084)	(0.0116)
Full treatment	0.0414***	0.0220***	0.0359***	0.0397***	0.0583***	0.0745***
	(0.0044)	(0.0039)	(0.0059)	(0.0068)	(0.0090)	(0.0139)
Observations (cells)	64017	10546	12868	13882	13963	12758
R^2	0.910	0.716	0.771	0.781	0.789	0.791
Mean dependent variable	0.3152	0.0857	0.1640	0.2534	0.3731	0.5374
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]	0.0012 [10.5951]					
Full treatment: p-value [F-stat]	0.0002 [14.4416]					
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on blacks' college applications behavior relative to whites. The regressions use the TEA data, an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. Partial treatment is the coefficient on the interaction between an indicator for black students and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and an indicator for black students. The outcome variable in Panel A is the average number of selective colleges to which students applied, and in Panel B, it is the fraction of students in a cell that applied to any college. Standard errors are clustered at the district level.

Table A4: Effect of AA on College Applications for Hispanic Students Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Applications to selective colleges					
Partial treatment	0.0045	-0.0003	-0.0018	-0.0022	0.0102	0.0128
	(0.0029)	(0.0019)	(0.0024)	(0.0033)	(0.0074)	(0.0101)
Full treatment	0.0141***	-0.0019	0.0007	0.0102**	0.0272***	0.0349***
	(0.0036)	(0.0016)	(0.0024)	(0.0045)	(0.0059)	(0.0107)
Observations (cells)	81946	14155	16963	17864	17493	15471
R^2	0.925	0.538	0.687	0.778	0.826	0.857
Mean dependent variable	0.1662	0.0077	0.0327	0.0888	0.2095	0.4441
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.1878 [1.7374]		
Full treatment: p-value [F-stat]				0.0005 [12.3546]		
	Panel B: Application to any college					
Partial treatment	0.0000	0.0037	-0.0072**	-0.0041	0.0013	0.0095
	(0.0030)	(0.0028)	(0.0035)	(0.0042)	(0.0060)	(0.0087)
Full treatment	0.0217***	0.0010	0.0021	0.0206***	0.0377***	0.0482***
	(0.0041)	(0.0028)	(0.0031)	(0.0059)	(0.0061)	(0.0103)
Observations (cells)	81946	14155	16963	17864	17493	15471
R^2	0.927	0.847	0.845	0.830	0.821	0.802
Mean dependent variable	0.2783	0.0668	0.1408	0.2339	0.3609	0.5309
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.5222 [0.4100]		
Full treatment: p-value [F-stat]				0.0000 [22.1933]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on Hispanics' college applications behavior relative to whites. The regressions use the TEA data, and an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. Partial treatment is the coefficient on the interaction between an indicator for Hispanic students and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and an indicator for Hispanic students. The outcome variable in Panel A is the average number of selective colleges to which students applied. In Panel B, it is the fraction of students in a cell that applied to any college. Standard errors are clustered at the district level.

Table A5: Effect of AA on Any Application to the University of Texas System for URMs Relative to Whites

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Application to any UT					
Partial treatment	0.0029	0.0030**	-0.0025	0.0020	-0.0004	0.0145**
	(0.0019)	(0.0014)	(0.0023)	(0.0030)	(0.0043)	(0.0057)
Full treatment	0.0152***	0.0040**	0.0083***	0.0161***	0.0191***	0.0291***
	(0.0028)	(0.0016)	(0.0017)	(0.0038)	(0.0048)	(0.0065)
Observations (cells)	97121	18380	20681	20974	19960	17126
R^2	0.911	0.857	0.886	0.890	0.875	0.855
Mean dependent variable	0.1191	0.0280	0.0613	0.0991	0.1534	0.2549
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.0587 [3.5847]		
Full treatment: p-value [F-stat]				0.0001 [15.6335]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' college applications behavior. The regressions use the TEA data, an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. Partial treatment is the coefficient on the interaction between an indicator for being a URM and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and being a URM. The outcome variable is the fraction of students in a cell that applied to any UT institution (UT Arlington, UT Austin, UT Dallas, UT El Paso, UT Permian Basin, UT Rio Grande, UT San Antonio, UT Tyler). Standard errors are clustered at the district level.

Table A6: Effect of AA on College Applications – Exogenous Ability Sample

		Percentile of grade 6 test score distribution				
	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Applications to selective colleges					
Partial treatment	0.0097***	0.0018	0.0022	0.0024	0.0145**	0.0297***
	(0.0027)	(0.0019)	(0.0025)	(0.0035)	(0.0067)	(0.0085)
Full treatment	0.0187***	0.0019	0.0046	0.0151***	0.0304***	0.0449***
	(0.0038)	(0.0016)	(0.0030)	(0.0049)	(0.0073)	(0.0105)
Observations (cells)	68509	12933	14515	14809	14145	12107
R^2	0.913	0.469	0.630	0.738	0.800	0.837
Mean dependent variable	0.1484	0.0079	0.0331	0.0877	0.1994	0.4158
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.0013 [10.4874]		
Full treatment: p-value [F-stat]				0.0000 [17.9979]		
	Panel B: Application to any college					
Partial treatment	0.0082***	0.0085***	0.0047	0.0016	0.0090*	0.0238***
	(0.0026)	(0.0027)	(0.0036)	(0.0044)	(0.0052)	(0.0073)
Full treatment	0.0169***	0.0031	0.0048	0.0125**	0.0254***	0.0438***
	(0.0035)	(0.0030)	(0.0040)	(0.0058)	(0.0059)	(0.0085)
Observations (cells)	68509	12933	14515	14809	14145	12107
R^2	0.915	0.788	0.815	0.810	0.802	0.781
Mean dependent variable	0.2603	0.0659	0.1414	0.2312	0.3499	0.5107
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.0552 [3.6870]		
Full treatment: p-value [F-stat]				0.0000 [23.2199]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Note: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' college applications behavior. The regressions use the TEA data, and an observation is at the district-cohort-race-ability quintile level, where the ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. The sample excludes the 2007-2010 cohorts. Partial treatment is the coefficient on the interaction between an indicator for being a URM and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and being a URM. The outcome variable in Panel A is the average number of selective colleges to which students applied. In Panel B, it is the fraction of students in a cell that applied to any college. Standard errors are clustered at the district level.

Table A7: Effect of AA on Math Grades for URMs Relative to Whites

	All students			Ability distribution		
				Bottom tercile	Middle tercile	Top tercile
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.7389*	0.7274*	0.2932	0.2845	0.4446	1.7302***
	(0.4263)	(0.4293)	(0.4272)	(0.6580)	(0.5309)	(0.6590)
Lagged dep. var. (grade 8)		0.4538***				
		(0.0112)				
Observations	55595	41724	83590	14314	14641	13947
R^2	0.148	0.228	0.729	0.136	0.156	0.162
Mean dependent variable	76.12	76.68	79.07	72.67	76.52	81.19
S.D. dependent variable	10.79	10.11	9.41	9.66	9.39	9.54
Test: Bottom tercile = Top tercile						
p-value [F-stat]			0.0753 [3.1850]			
School-year FE	X	X	X	X	X	X
Ethnicity FE	X	X		X	X	X
Demographic controls	X	X		X	X	X
Student FE			X			
Grade-year FE			X			
Grade-ethnicity FE			X			

Notes: This table reports the difference-in-differences estimates of the effect of affirmative action on math grades in a large urban school district. An observation is a student, and the sample consists of repeated cross-sections of 11th graders. “Treated” is the coefficient on the interaction between being a URM and being observed post 2003. Ability terciles are assigned based on 8th grade average school grades. Standard errors are clustered at the school level.

Table A8: Effect of AA on English Grades for URMs Relative to Whites

	All students			Ability distribution		
				Bottom tercile	Middle tercile	Top tercile
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	1.1597***	1.5617***	1.6601***	1.4907**	0.7296	1.3799***
	(0.4225)	(0.4414)	(0.4032)	(0.6882)	(0.5912)	(0.5035)
Lagged dep. var. (grade 8)		0.3521***				
		(0.0098)				
Observations	58649	43522	87197	15058	15255	14503
R^2	0.200	0.234	0.713	0.188	0.195	0.169
Mean dependent variable	79.03	79.93	81.76	76.02	79.90	83.61
S.D. dependent variable	10.38	9.47	8.90	9.66	8.95	8.35
Test: Bottom tercile = Top tercile						
p-value [F-stat]			0.8893 [0.0194]			
School-year FE	X	X	X	X	X	X
Ethnicity FE	X	X		X	X	X
Demographic controls	X	X		X	X	X
Student FE			X			
Grade-year FE			X			
Grade-ethnicity FE			X			

Notes: This table reports the difference-in-differences estimates of the effect of affirmative action on English grades in a large urban school district. An observation is a student, and the sample consists of repeated cross-sections of 11th graders. "Treated" is the coefficient on the interaction between being a URM and being observed post 2003. Ability terciles are assigned based on 8th grade average school grades. Standard errors are clustered at the school level.

Table A9: Effect of AA on Stanford Test Scores for URMs Relative to Whites

		Ability distribution		
	All students	Bottom tercile	Middle tercile	Top tercile
	(1)	(2)	(3)	(4)
	Dependent variable: Stanford Test Scores (grade 11)			
Treated	4.7801***	4.2109***	4.6267***	7.3731***
	(1.1352)	(1.2879)	(1.5648)	(1.4314)
Observations	58096	15486	15347	14620
R^2	0.444	0.455	0.487	0.464
Mean dependent variable	49.40	42.24	50.49	59.99
S.D. dependent variable	25.74	23.38	24.00	23.76
Test: Bottom tercile = Top tercile				
p-value [F-stat]		0.0981 [2.7535]		
School-year FE	X	X	X	X
Ethnicity FE	X	X	X	X
Demographic controls	X	X	X	X

Notes: This table reports the difference-in-differences estimates of the effect of affirmative action on mean Stanford test scores in a large, urban school district. An observation is a student, and the sample consists of repeated cross-sections of 11th graders. "Treated" is the coefficient on the interaction between being a URM and being observed post 2003. Ability terciles are assigned based on 8th grade average school grades. Standard errors are clustered at the school level.

Table A10: Effect of AA on College Applications Behavior – Excluding Houston & Dallas

	All students	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
		Percentile of grade 6 test score distribution				
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Applications to selective colleges					
Partial treatment	0.0094***	0.0025	0.0008	0.0020	0.0146**	0.0295***
	(0.0028)	(0.0018)	(0.0024)	(0.0036)	(0.0070)	(0.0087)
Full treatment	0.0213***	0.0022*	0.0041*	0.0162***	0.0381***	0.0495***
	(0.0028)	(0.0013)	(0.0025)	(0.0038)	(0.0052)	(0.0093)
Observations (cells)	96281	18212	20513	20806	19792	16958
R^2	0.911	0.470	0.623	0.720	0.790	0.835
Mean dependent variable	0.1598	0.0094	0.0356	0.0910	0.2092	0.4422
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.0029 [8.9553]		
Full treatment: p-value [F-stat]				0.0000 [25.4628]		
	Panel B: Application to any college					
Partial treatment	0.0079***	0.0096***	0.0037	0.0013	0.0092*	0.0226***
	(0.0027)	(0.0027)	(0.0036)	(0.0046)	(0.0055)	(0.0073)
Full treatment	0.0295***	0.0106***	0.0118***	0.0280***	0.0449***	0.0581***
	(0.0034)	(0.0027)	(0.0035)	(0.0051)	(0.0054)	(0.0082)
Observations (cells)	96281	18212	20513	20806	19792	16958
R^2	0.913	0.792	0.817	0.807	0.799	0.779
Mean dependent variable	0.2808	0.0768	0.1564	0.2480	0.3700	0.5345
Test: Bottom quintile = Top quintile						
Partial treatment: p-value [F-stat]				0.1097 [2.5643]		
Full treatment: p-value [F-stat]				0.0000 [35.7148]		
Demographic controls	X	X	X	X	X	X
District-cohort-ability FE	X	X	X	X	X	X
District-ethnicity-ability FE	X	X	X	X	X	X

Notes: This table reports difference-in-differences estimates of the effect of affirmative action on URMs' college applications behavior. The regressions use the TEA data, and an observation is at the district-cohort-race-ability quintile level, where ability quintile is assigned based on 6th grade (pre-AA) test scores on the state standardized test. Cells are weighted by the number of student-years in a cell. The sample excludes the Houston Independent School District and the Dallas Independent School District. Partial treatment is the coefficient on the interaction between an indicator for being a URM and an indicator variable for entering high school after 2001 and before 2003. Full treatment is the coefficient on the interaction between entering high school after 2003 and being a URM. The outcome variable in Panel A is the average number of selective schools to which students applied. For Panel B, it is the probability students applied to any college. Standard errors are clustered at the district level.

Appendix A: Robustness of SAT Results to Accounting for Other Policy Changes

Accounting for AA Bans. Several states implemented affirmative action bans in university admissions during our study period. Washington, Michigan, and Nebraska passed affirmative action bans through ballot initiatives in November 1998, 2006 and 2008, respectively. Governor Jeb Bush issued an executive order banning affirmative action in Florida in November 1999.[33] We do not use these bans to estimate the effects of affirmative action because, as Hinrichs (2012) points out, the effect of affirmative action bans cannot be disentangled from the effect of percent plans; these two policies were almost always enacted concurrently. This would certainly be an issue for estimating the effect of bans in our SAT sample. In this sample, Florida would drive most of the variation in the AA ban indicator because of its large population size and the fact that Michigan and Nebraska are both ACT states and have fewer SAT test-takers. Florida implemented a very aggressive "percent plan" under which students in the top 20% of their high school graduating class were guaranteed admission to a state public university shortly after the ban.

While we cannot use the bans to identify the effects of affirmative action, in this subsection, we verify that our difference-in-differences estimates of the effect of *Grutter v. Bollinger* on students in Texas, Louisiana and Mississippi are robust to controlling for these affirmative action bans. The robustness tests for math SAT scores are reported in Appendix Table A11 and, for completeness, corresponding results for verbal scores are reported in Appendix Table A12. Column (1) reports our baseline estimates. In column (2), we explicitly control for the effect of affirmative action bans by including an indicator variable that takes a value of 1 for post-AA ban years in the associated states in the regression.[34] Our estimates of the effect of *Grutter v. Bollinger* on whites (4 points) and on URMs (8 points) are unaffected by the inclusion of an indicator for AA bans as a control variable. Controlling for AA bans yields a slightly higher triple-differences estimate of 4.2 points. The coefficient on the AA ban indicator is statistically insignificant in both difference-in-differences specifications. In the triple-differences model, the coefficient on the AA ban indicator interacted with a URM dummy is positive and significant (5.7 points), which suggests that URM students' SAT scores increased relative to white students' in the states that implemented bans on affirmative action. However, as discussed above, we strongly caution against interpreting this

[33]Following our study period, Arizona (2010), New Hampshire (2011) and Oklahoma (2012) also banned affirmative action in college admissions.

[34]The indicator turns on after 1999 in Washington, after 2000 in Florida, after 2007 in Michigan, and after 2009 in Nebraska. It is zero in all years for all others states.

coefficient as the causal effect of AA bans. In column (3) of Appendix Table A11, we drop the four states that banned affirmative action between 1998 and 2010 from the estimating sample. Again, our estimates of the effect of the re-instatement of affirmative action are virtually unchanged.

Accounting for Potential Non-Compliance. There is some evidence that Louisiana and Mississippi may have continued to use race in university admissions to some extent in 1998-2003 despite the *Hopwood v. Texas* ruling due to pre-existing rulings that required them to de-segregate their institutions of higher education (Hinrichs, 2012). Thus, we also drop these two states from the sample and estimate the effects of the policy change on Texas alone relative to the control states in column (4) of Appendix Table. Dropping these two states has little effect on the estimates.

Table A11: Robustness of Effect of AA on Math SAT Scores to Controlling for AA Bans

	Baseline	Control for AA bans	Drop AA ban states	Drop Mississippi and Louisiana
	(1)	(2)	(3)	(4)
	Panel A: URMs			
DD coefficient	7.998***	8.009***	8.122***	8.082***
	(1.497)	(1.544)	(1.693)	(1.491)
AA Ban Indicator		0.312		
		(2.695)		
Observations (cells)	1904	1904	1748	1830
R^2	0.844	0.844	0.839	0.844
State, year and ethnicity FE	X	X	X	X
	Panel B: Whites			
DD coefficient	4.145***	4.048***	3.835***	4.552***
	(0.994)	(0.983)	(1.021)	(0.861)
AA Ban Indicator		-3.282		
		(2.666)		
Observations (cells)	663	663	611	637
R^2	0.968	0.969	0.967	0.968
State, year and ethnicity FE	X	X	X	X
	Panel C: Triple-Difference			
DDD coefficient	3.975***	4.155***	4.253***	3.881***
	(0.872)	(0.827)	(0.816)	(0.859)
AA Ban Indicator × URM dummy		5.714***		
		(1.152)		
Observations (cells)	2555	2555	2347	2455
R^2	0.998	0.998	0.998	0.998
State-year FE	X	X	X	X
State-ethnicity FE	X	X	X	X
Ethnicity-year FE	X	X	X	X

Notes: This table reports difference-in-differences and triple-differences effects of affirmative action on SAT scores. Each observation is a state-race-year group. In all specifications, cells are weighted by the number of test-takers in a group. In Panels A and B, the DD coefficient reports the interaction of an indicator variable for belonging to a treated state (Texas, Louisiana, Mississippi) and being tested after *Grutter v. Bollinger* (post 2003). In Panel C, the coefficient is on the interaction between being a URM, being tested post 2003, and belonging to a treated state. Standard errors are clustered at the state level.

Table A12: Effect of AA on Verbal SAT Scores, Controlling for AA Bans

	Baseline	Control for AA bans	Drop AA ban states	Drop Mississippi and Louisiana
	(1)	(2)	(3)	(4)
		Panel A: URMs		
DD coefficient	-0.779	-0.634	-0.170	-0.793
	(1.755)	(1.783)	(1.929)	(1.767)
AA Ban Indicator		4.131*		
		(2.263)		
Observations (cells)	1901	1901	1745	1828
R^2	0.795	0.796	0.788	0.793
State, year and ethnicity FE	X	X	X	X
		Panel B: Whites		
DD coefficient	0.0247	0.0342	0.0008	-0.0352
	(0.878)	(0.888)	(0.955)	(0.877)
AA Ban Indicator		0.321		
		(3.180)		
Observations (cells)	663	663	611	637
R^2	0.971	0.971	0.970	0.970
State, year and ethnicity FE	X	X	X	X
		Panel C: Triple-Difference		
DDD coefficient	1.083	1.260	1.440**	1.156
	(0.825)	(0.753)	(0.633)	(0.833)
AA Ban Indicator $\times$ URM dummy		5.604***		
		(1.478)		
Observations (cells)	2552	2552	2344	2453
R^2	0.998	0.998	0.998	0.998
State-year FE	X	X	X	X
State-ethnicity FE	X	X	X	X
Ethnicity-year FE	X	X	X	X

This table reports difference-in-difference and triple-differences effects of affirmative action on SAT scores. Each observation is a state-race-year group. In all specifications, cells are weighted by the number of test-takers in a group. In Panels A and B, the DD coefficient reports the interaction of an indicator variable for belonging to a treated state (Texas, Louisiana, Mississippi) and being tested after *Grutter v. Bollinger* (post 2003). In Panel C, the coefficient is on the interaction between being a URM, being tested post 2003, and belonging to a treated state. Standard errors are clustered at the state level.

Appendix B: Robustness of SAT Results to Alternative Synthetic Control Specifications

Matching on Fewer Pre-treatment Cohorts. In our baseline synthetic control results, we choose the control group by minimizing the mean squared prediction errors in 1998-2003 using the following set of variables as predictors: number of white SAT test takers, number of URM SAT test takers, white math SAT scores, URM math SAT scores, white verbal SAT scores, and URM verbal SAT scores. Each of these variables is averaged over 1998-2000 and over 2001-2003 in the matching

Here, we verify that our results are robust to using fewer pre-treatment years in the construction of the synthetic control group. Appendix Figure A12 shows time series of math SAT scores for synthetic control groups based on 4, 5 and 6 years of pre-treatment data.[35]

For whites, the results are insensitive to the model's specification, with the 3 synthetic groups tracking each other very closely. Reassuringly, when fewer years of pre-treatment data are used to construct the synthetic control group, the gap between treated and untreated states during pre-treatment years remains very small, even in years that were not used in the construction of the synthetic control group. The states included in the 5-year match synthetic control group are California (46.4%), Florida (39.5%), Indiana (6.1%) and Pennsylvania (7.9%). The states included in the 4-year match synthetic control group are California (44.9%), Florida (41.9%), Indiana (10.7%), North Carolina (1.8%) and Pennsylvania (0.6%).

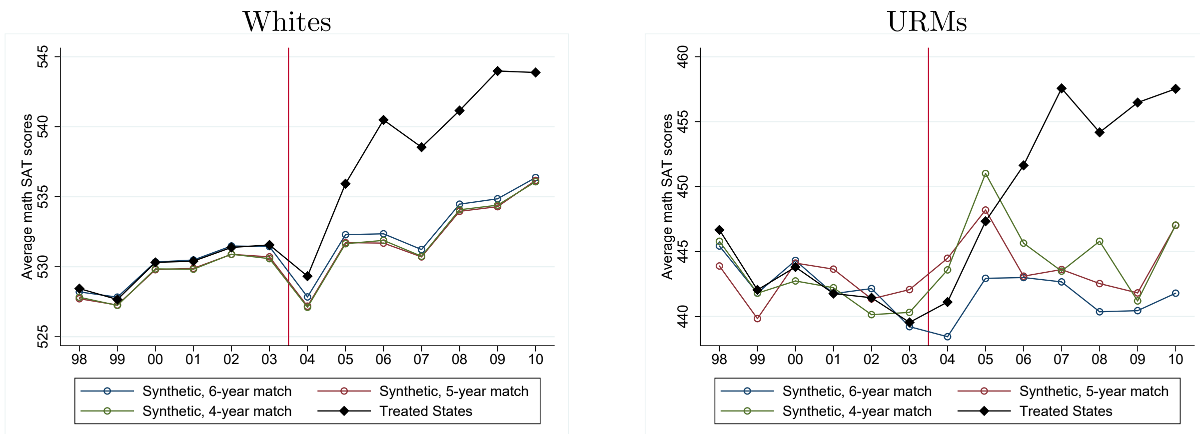
There are fewer URMs than whites taking the SAT in a state in a given year. As a result, yearly mean SAT scores for URMs are more volatile, and the composition of the synthetic control group is more sensitive to the number of pre-treatment years over which the RMSPE is minimized. While the SAT scores of the 6-year match synthetic control group track our treated group very closely for all pre-treatment years, the 4-year and 5-year match groups do not replicate the treated states' negative pre-trend as closely. In both cases, an upward movement in SAT scores of the synthetic group appears around 2002, whereas the treated states are still on a downward trend at that time. By 2007, however, the outcomes of the three synthetic control groups are similar, and in all cases, the control groups' SAT scores are significantly below those of the treated states. The states included in the 5-year match synthetic control group are California (84.3%), Pennsylvania (11.6%) and New Hampshire (4%). The states included in the 4-year match synthetic control group are California (82.4%),

[35]When minimizing the RMSPE over 4 years, we average the predictors over 1998-1999 and 2000-2001. When minimizing the RMSPE over 5 years, we average the predictors over 1998-2000 and 2001-2002.

West Virginia (14.8%), and Pennsylvania (2.7%).

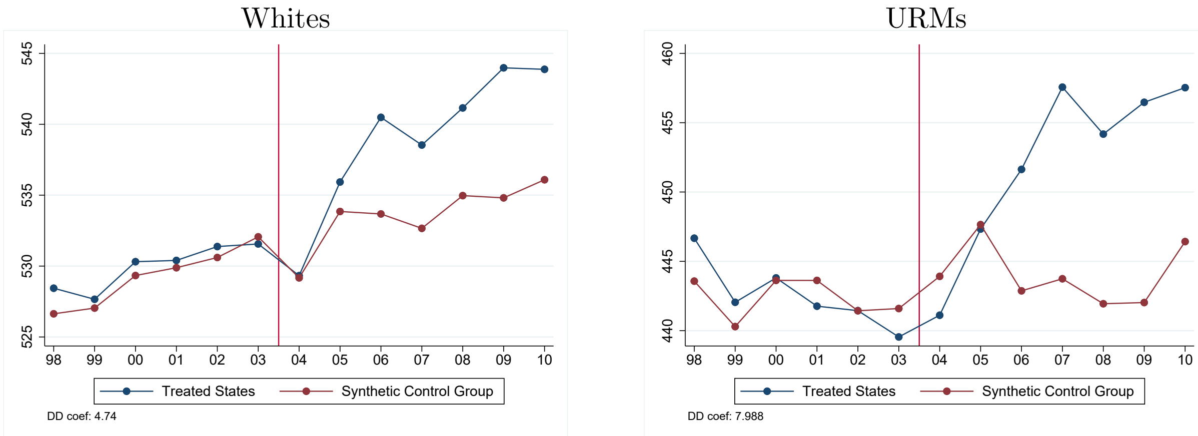
Excluding States That Banned AA During the Study Period. As a final robustness check, we drop the four states that banned affirmative action between 1998 and 2010 (Florida, Nebraska, Michigan and Washington) from the donor pool. The results are shown in Appendix Figure A13. In the pre-treatment years, the synthetic control group tracks the treated states fairly closely, albeit with more volatility than under our baseline approach. Overall, our results on the effect of *Grutter v. Bollinger* are not significantly affected by this sample restriction. The states included in the synthetic control group for whites are California (35.1%), Pennsylvania (38.7%), New York (15.2%), Utah (4.5%), New Hampshire (2.8%), Minnesota (2.5%), and Montana (1.1%). The states included in the synthetic control group for URMs are California (70.8%), Pennsylvania (20.1%), and New Hampshire (9.1%).

Figure A12: Alternative Synthetic Control Estimates of the Effect of AA on Math SAT Scores



Notes: This figure reports synthetic cohort analyses separately for whites and URMs. It shows SAT math scores for the treated states (Texas, Mississippi and Louisiana) and for the synthetic control group under alternative matching specifications. The control group "Synthetic, 6-year match" is obtained by minimizing the root mean squared prediction error (RMSPE) over the 1998-2003 period. For "Synthetic, 5-year match," the RMSPE is minimized over the 1998-2002 period, and for "Synthetic, 4-year match," it is minimized over the 1998-2001 period.

Figure A13: Synthetic Control Estimates of the Effect of AA on Math SAT Scores, Dropping AA Ban States



Notes: This figure reports synthetic cohort analyses separately for whites and URMs. It shows SAT math scores for the treated states (Texas, Mississippi and Louisiana) and for the synthetic control group. In constructing the control group, Florida, Nebraska, Michigan and Washington were omitted from the donor pool.