NATIONAL
CENTER *for* ANALYSIS *of* LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

AIR

The Effect of Banning Affirmative Action on Human Capital Accumulation Prior to College Entry

Kate Antonovics and Ben Backes

WORKING PAPER 114 • JULY 2014

The Effect of Banning Affirmative Action on Human Capital Accumulation Prior to College Entry

Kate Antonovics
University of California, San Diego

Ben Backes
American Institutes for Research, CALDER

Contents

Acknowledgements

We thank Rick Sander, David Card and Julian Betts for their help obtaining the data we use. We also thank Eric Furstenberg for sharing the details of his analysis of the effects of Prop 209 on SAT scores. We thank Terrance August, Julie Cullen, Zoltan Hajnal, Rick Hanushek, Valerie Ramey, and seminar participants at UC San Diego for their helpful comments. Ben Backes was supported in part by grants from both the Bill & Melinda Gates Foundation and the Smith Richardson Foundation to CALDER at the American Institutes for Research.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

The Effect of Banning Affirmative Action on Human Capital Accumulation Prior to College Entry

Kate Antonovics and Ben Backes

CALDER Working Paper No. 114
June 2014

Abstract

This paper examines how banning affirmative action in university admissions affects both overall academic achievement and the racial gap in academic achievement prior to college entry. In particular, focusing on college-bound high school students, we use a difference-in- difference methodology to analyze the impact of the end of race-based affirmative action at the University of California in 1998 on both the overall level of SAT scores and high school GPA, and the racial gap in SAT scores and high school GPA. Our primary conclusion is that academic achievement changed very little after the ban.

1. Introduction

Universities in the United States are increasingly limited in their ability to practice race-based affirmative action. In the last two decades, public universities in a growing number of states have stopped practicing race-based affirmative action due to various court rulings, voter initiatives, and administrative decisions.[1] In addition, the United States Supreme Court's 2013 decision in *Fisher v. Texas* makes it more difficult for universities to justify using race as a factor in admissions, and its 2014 ruling in *Schuette v. Coalition* leaves the door open for additional states to implement statewide bans on affirmative action. The implications of eliminating race-based affirmative action in college admissions are far-reaching and have been the subject of considerable legal, political, and scholarly debate. In this paper, we focus on the potential effects of banning affirmative action on academic achievement prior to college entry.

Economic theory suggests that eliminating affirmative action could have important implications for human capital accumulation. There are a number of channels through which this effect could operate. First, the removal of racial preferences directly affects admissions probabilities, which in turn affects the return to studying prior to college application. Second, banning affirmative action could lead underrepresented minorities to feel institutionally discouraged from attending college, and they could respond by putting forth less effort in preparing for postsecondary education. Finally, if colleges react to bans on affirmative action by changing their admissions process more broadly, then this too could lead students to shift the focus of their efforts in high school.

While economic theory clearly predicts that ending affirmative action will affect human capital investment, it does not yield definitive predictions about whether doing so will lead the racial gap in human capital investment or the overall level of human capital investment (regardless of race) to rise or

[1] Establishing the list of states in which race-based affirmative action has been prohibited is complicated by ambiguities in case law, but arguably includes Alabama, Arizona, California, Florida, Georgia, Louisiana, Michigan, Mississippi, Nebraska, Oklahoma, Texas and Washington. See Blume and Long (2014) for a nice discussion of the policy environment surrounding affirmative action.

fall.[2] Thus, the effect of banning affirmative action on human capital investment is largely an empirical question.

In an attempt to answer this question, this paper explores the impact of Proposition 209, which prohibited public universities in California from practicing race-based affirmative action, on both the SAT scores and high school GPA of college-bound high school students. In particular, using data from the College Board, we examine how these measures of academic achievement changed in California relative to other states (a difference-in-difference estimate), and how the racial gap in these measures changed in California relative to other states (a triple-difference estimate). Our primary conclusion is that SAT scores and high school GPA changed very little after the ban on race-based affirmative action. This finding is consistent with previous research documenting a limited behavioral response to California's ban on affirmative action in terms of both college application behavior and college enrollment behavior (see, for example, Antonovics and Backes (2013) and Antonovics and Sander (2013)).

Our paper proceeds as follows. Section 2 discusses the related literature, and Section 3 presents an overview California's affirmative action ban. Section 4 discusses our empirical strategy and gives a short overview of our data. Section 5 presents our results, and Section 6 concludes.

2. Related Literature

Several theoretical papers have examined the effect of affirmative action on human capital investment.[3] The theoretical literature draws a distinction between ``color-sighted affirmative action", wherein there are explicit racial preferences in admissions, and ``color-blind affirmative action", wherein colleges implicitly favor minorities by using admissions rules that favor students who possess characteristics that are positively correlated with being a member of a targeted racial group (see, for

[2] We discuss the theoretical literature in greater detail below

[3] See Holzer and Neumark (2000) for a more comprehensive review of the theoretical and empirical literature on affirmative action.

example, Fryer et al. (2008) and Ray and Sethi (2010)).[4] Both forms of affirmative action stand in contrast to laissez-fair admission regimes in which race is not considered either explicitly or implicitly. In the case of California, it seems clear that Prop 209 shifted most University of California (UC) schools from color-sighted to color-blind affirmative action. That is, the UC administration openly acknowledged that diversity remained a high priority even after Prop 209, and Antonovics and Backes (2014) provide evidence that, after Prop 209, UC schools changed their admissions process to implicitly favor minorities.

Building a model of college admissions with endogenous human capital investment, Fryer et al. (2008) establishes that moving from color-sighted to color-blind affirmative action alters students' incentives to invest in human capital, but does not provide definitive predictions about whether investment levels will go up or down. Hickman (2012) and Hickman (2013) also model the link between different admissions regimes and human capital investment, but do not consider color-blind affirmative action.[5]

There is a substantial empirical literature examining the effect of affirmative action on academic achievement (for a recent survey, see Arcidiacono and Lovenheim (2014)). Of these, only a handful use data from before and after a ban on affirmative action, with most focusing on academic achievement after college entry (see, for example, Arcidiacono et al. (2014), Backes (2012), Cortes (2010) and Hinrichs (2012)). To our knowledge, only two previous studies have directly examined the effect of a ban on affirmative action on academic achievement prior to college entry, Furstenberg (2010) and Caldwell (2010). Both find evidence that banning affirmative action increases the racial test score gap. We discuss the weaknesses of these papers in greater detail when we present our results in Section 5.

3. Background on California's Ban on Affirmative Action

The first threat to affirmative action in California was in July 1995, when the Board of Regents of the

[4] In this paper, we use the terms ``color-sighted affirmative action'' and ``race-based affirmative action'' interchangeably.

[5] Hickman's color-blind admission policy (in which universities do not consider race at all) differs from color-blind affirmative action (in which colleges implicitly favor minorities).

University of California passed a resolution (SP-1), which stipulated that the UC would discontinue considering race in admissions by the beginning of 1997. The implementation of SP-1, however, was delayed. Then, in November 1996, California voters approved Prop 209, which banned the use of racial preferences at public universities.[6] Prop 209 underwent various legal challenges until the Supreme Court denied further appeals in November 1997. Thus, the incoming class of 1998 was the first to be admitted under the statewide ban on affirmative action. Table 1 presents a timeline of the events leading up to the ban.

Although the prohibition against affirmative action applied to all public universities in California, in practice it only affected UC schools because at the time most Cal State schools were not selective, admitting the vast majority of applicants, regardless of race. Given that one in six Californian high school graduates apply to at least one UC campus, however, it is reasonable to think that major policy changes at the UC could affect the human capital investment for the state as a whole.[7]

Evidence that students' human capital investment decisions respond to the incentives created by college admissions policies is also evident in Cortes and Friedson (2010) and Cullen et al. (2012), both of which find evidence of students moving to lower quality high schools after Texas introduced its top ten percent plan, which guaranteed admission to any public university in Texas for students who graduated in the top ten percent of their high school class.

In addition, it is important to recognize that Prop 209 had an enormous impact on the admission rate of underrepresented minorities (or URMs), especially at the more selective UC schools.[8] For example, the fraction of URM applicants admitted to Berkeley plummeted from 52 percent to 25 percent between the three-year period immediately before and the three-year period immediately after the ban. In contrast, the admission rates of non-minority applicants to Berkeley only fell from 32 percent to 28

[6] Searching the LexisNexis article database gives the first mention of Prop 209 in July 1996.
[7] Calculation made using UC application data (http://www.ucop.edu/news/archives/2001/applications_2001/table1.pdf) and California high school graduation data (http://www.cpec.ca.gov/completereports/2003reports/03-09/Display1.PDF).
[8] The UC uses the term underrepresented minority to refer to blacks, Hispanics and Native Americans.

percent over that same time period.

Nonetheless, isolating the effect of the end of race-based affirmative action at the UC is complicated by a number of concomitant policy changes. First, in an effort to minimize the effects of Prop 209 on minority enrollment, the UC system substantially increased minority outreach efforts.[9] It is unclear, however, how much of an effect the new outreach programs had on the applicant pool in the years immediately following Prop 209 since many of these programs took years to fully develop and were long-term in nature. To the extent that increased outreach had an effect on academic achievement, our estimates represent the net effect of ending race-based affirmative action and the accompanying change in outreach.

In addition to the increase in outreach, there were two major policy changes in California around the same time as Prop 209. First, the 1999 passage of the Public Schools Accountability Act brought about substantial changes to the public K-12 schools in California. In particular, beginning in 2000, schools were eligible for rewards if all ethnic subgroups within schools either scored above a certain threshold or met targets for test score growth. In addition, schools with low test scores could opt into an intervention program designed for schools not meeting growth targets. Second, in 2001, the UC implemented Eligibility in a Local Context (ELC), guaranteeing any student in the top four percent of his or her high school class admission to at least one UC school (conditional on completing specified coursework). The new policy was designed to attract students from high schools that did not typically send many students to the UC, giving the UC a way to potentially increase minority enrollment. Since both of these policy changes are likely to have affected human capital accumulation, we perform robustness checks by removing observations from 2000 and later from our analysis and find no substantial impact on our main results.

---

[9] For example, "In an attempt to improve minority access to UC without the help of affirmative action, the university's investment in kindergarten-through-12th-grade outreach has rocketed from about $60 million in 1995 to $180 million last year and a planned $250 million this year" (2000, January 21). UC Regents Urged to Step Up Minority Outreach at Schools. *The San Francisco Chronicle*.

4. Estimation and Data

Antonovics and Backes (2014) provide evidence that Prop 209 led to widespread shifts in the entire admission process at many UC schools. In particular, the more selective UC schools appear to have decreased the weight placed on SAT scores and increased the weight given to high school GPA and family background characteristics in order to boost minority admission rates after Prop 209 went into effect. This is important because the apparent reduced emphasis on SAT scores and increased emphasis on high school GPA might have led to a shift away from human capital investments that increase SAT scores towards those that increase high school GPA. For this reason, we separately examine the effect of Prop 209 on these two measures of academic achievement.

In addition, we measure the effects of the policy change in two ways. First, we explore the reaction of Californians relative to the rest of the country. Second, we document how the gap between whites and underrepresented minorities (URMs) changed in California relative to the rest of the country.[10] Each of the two measures is important. Since Californians of all races were affected by the changes in the admissions policies at UC schools, they may be thought of as one treated group. Comparing Californians to those in other states reveals the extent to which Prop 209 changed academic achievement for all Californians. On the other hand, affirmative action policies are generally thought of as a way to address the gap between white and minority students. Viewed in this way, it is natural to ask how the removal of explicit racial preferences affected the racial gap in academic achievement.

[10] At the UC, URMs include Hispanics, blacks and Native Americans.

A. Empirical Strategy

To examine the effect of Prop 209 on academic achievement, we estimate the following:

$$\begin{aligned}\text{Outcome}_{ist} = {} & \beta_0\text{CA}_s + \beta_1\text{Post}_{st} + \beta_2\text{URM}_i \\ & + \beta_{01}\text{CA}_s\text{Post}_{st} + \beta_{02}\text{CA}_s\text{URM}_i + \beta_{12}\text{Post}_{st}\text{URM}_i \\ & + \beta_{012}\text{CA}_s\text{Post}_{st}\text{URM}_i \\ & + \beta X_i + \epsilon_{ist},\end{aligned} \tag{1}$$

where Outcome_{ist} is the outcome (SAT score or high school GPA) for student *i* in state *s* in year *t*. Post_{st} is an indicator for whether the affirmative action ban was in place in state *s* in year *t*, CA_s is an indicator for whether the student resides in California, and URM_i is an indicator for an individual's race. Finally, X_i includes controls for parental income, education, gender, whether english is the student's first language and citizenship status. Following the discussion at the beginning of this section, there are two coefficients of interest. First, to the extent that Californians had a common response to Prop 209, it would be captured by β_1, which represents the change in the dependent variable for Californians relative to the rest of the country. Second, β_{012} represents the change in the minority-white test score gap in California relative to the rest of the country.

There are two important choices that must be made in estimating equation (1). The first is which cohorts to include. As discussed above, there was an extended period of uncertainty surrounding the end of race-based affirmative action at the UC. Table 1 shows the timing events leading up to the implementation of Prop 209 along with the timing of when various graduation cohorts are likely to have taken the SAT. As the table indicates, the first major threat to affirmative action at the UC came in July 1995 when the Regents of California passed SP-1, which committed the UC system to an eventual ban on the use of racial preference in admissions. It was not until over two years later, however, in November 1997, when Prop 209 officially became law that the UC stopped the use of explicit racial preferences.

Thus, students who took the SAT between July 1995 and November 1997 (those in the 1996-1998 graduation cohorts) did so during a time of considerable uncertainty about the future of race-based affirmative action. As a result, we experiment with two different pre and post period definitions: (i) 1994-1996 vs 1998-2001 and (ii) 1994-1995 vs 1999-2001. The first choice of sample years drops only the 1997 cohort, which took the SAT during the election cycle in which Prop 209 was on the ballot. The second choice of sample years takes a more conservative approach by dropping any cohort that could have taken the SAT during a period of uncertainty about Prop 209. However, it leaves only two years in the pre period and three in the post period.

The second important choice is whether to include an interaction term between the post period and demographic characteristics. By not including such an interaction, time-varying changes in the relationship between demographic characteristics and the outcome variable would be picked up by the coefficient on Post*URM, due to the correlation between race and demographic characteristics. However, if the effects of the policy operated in part through changes in the relationship between demographic characteristics and the outcome variable (due to, for example, changes in the weights placed on different student characteristics in determining admissions), the researcher may be interested in omitting the interaction between the post period and demographics to allow the changes to load onto the URM*Post coefficient. We present results from both specifications.

In practice, we also expand our estimating equation to include the full set of interactions for blacks, Hispanics, and Asians, with whites as the excluded group. We also use year fixed effects and state fixed effects, and we include state-specific linear time trends. We drop observations from Louisiana, Mississippi, Texas, and Washington, which were affected by their own affirmative action policy changes during our sample period. We also drop North Dakota and Wyoming due to small sample sizes of minority SAT test takers. Finally, we normalize test scores and GPA to be mean zero with standard deviation one.

Asians, who constitute a large portion of the college-going population in California, are not considered URMs for the purposes of admissions to the UC. We estimate effects on Asians separately but generally do not focus on their results because blacks and Hispanics were the intended beneficiaries of affirmative action policy at the UC and because the outcomes for Asians and whites are generally similar.[11]

B. Estimation of Standard Errors with Limited Treated Units

A growing number of papers have documented the inadequacy of typical methods of obtaining standard errors when the number of treated units is small (see Moulton (1990), Wooldridge (2006), Donald and Kang (2007), Abadie et al. (2010), and Buchmeuller et al. (2011)). To illustrate the problem, consider SAT scores as the outcome variable and suppose we are interested in $\beta_{01,}$ which measures the change in test scores for Californians relative to the rest of the country. Since SAT scores naturally fluctuate from year to year within a state even in the absence of a policy change, it is important to distinguish these fluctuations from the true policy effect. This is done in the typical DD setting by assuming that these state-specific, year-to-year fluctuations average out to a mean of zero over a large number of treatment and control states. In our case, there is only one treated unit, so there is little reason to believe this assumption holds.

We follow an established method of dealing with the problem of only one treated state by using the remaining untreated states to conduct a permutation test in which we construct the empirical distribution of β_{01} by estimating $\hat{\beta}_0$ for each of the control states, treating each control state as the treated state (see Abadie et al. (2010) and Buchmeuller et al. (2011)). Thus, the control states are used to estimate the variability of beta that is driven by year-to-year variation in test scores. The null hypothesis -- that Prop 209 had no effect on academic achievement in California -- is rejected when $\hat{\beta}_{01}{}^{California}$ is

[11] Adding Asians to the excluded group in the DDD regressions gives similar results.

large relative to the estimated empirical distribution of $\hat{\beta}_{01}{}^{j}$ for the control states (*j* indexes the control states). This procedure tests whether the change in test scores in California is large relative to the naturally occurring variation in test scores observed in other states. We also estimate $\hat{\beta}_{012}{}^{California}$ and $\hat{\beta}_{012}{}^{j}$ for each of the states in our data to measure whether the change in the minority-white gap in California was extreme relative to the change in states that did not ban affirmative action in the same time period. This correction reveals substantial within-state year-to-year variability in SAT scores.

Of course, one solution to the above problem would be to include more treated states. However, most states that banned affirmative action have done so too recently to be contained in our data.[12]

Two states that banned affirmative action during the period for which we have data are Texas and Florida; however, each introduced top x% plans in which the top x% of students within a high school were guaranteed admission to an in-state public university.[13] Louisiana and Mississippi were both affected by the court ruling that ended affirmative action in Texas, but both states were under de-segregation orders, so they may not have been under pressure to comply with the ruling.[14] In addition, neither state's public universities are as selective as the institutions typically affected by bans on racial preferences (see Blume and Long (2014)). Of the affirmative action banning states, the sole remainder is Washington, which voted to ban affirmative action in 1999. In principle, Washington could be included with our analysis; however, using two policy change states instead of one would still be insufficient for classical estimation of standard errors. Finally, there are several advantages to using California. First, it is a large state with a significant URM population. In addition, it has been well-established that the more selective UC schools practiced significant race-based affirmative action prior to Prop 209, and finally, the

[12] These states include Alabama (2002), Arizona (2010), Georgia (2002), Michigan (2006), Nebraska (2008), and Oklahoma (2012).

[13] In Texas, students could attend a university of their choice while in Florida, admission was guaranteed to at least one public university.

[14] For example: Healy, Patrick (1998, April 24). Affirmative Action Survives at Colleges in Some States Covered by Hopped Ruling. *The Chronicle of Higher Education*. Retrieved from http://www.chronicle.com.

measurement of how admissions rules changed at the UC after Prop 209 has been documented by Antonovics and Backes (2014).

C. Data

Our College Board data consist of SAT test takers who are expected to graduate from high school between 1994 and 2001.[15] The sample consists of all black and Hispanic test takers nationwide, all Californian test takers, and a 25 percent random sample of the rest of the country. The College Board includes a range of descriptive variables that are generated when students fill out the Student Descriptive Questionnaire before taking the exam. These include race, gender, parental characteristics, college aspirations, high school GPA, and many other variables. Unfortunately, our data do not reveal the date on which students took the SAT, but it is likely that students in a given graduation cohort took the SAT between the spring and fall of the preceding year.

One advantage of using the College Board sample is that nearly all SAT takers are interested in going to college, so they should be the ones most readily affected by a ban on affirmative action. On the other hand, a potential problem with using the College Board is that a student's decision about whether or not to take the SAT could itself be affected by the ban, leading to sample selection bias. Dickson (2006) finds that removal of affirmative action in Texas led to a decline in the percentage of minority high school graduates who take either the ACT or SAT, though the magnitude of this decline is small (roughly 3-4 percent). In addition, Furstenberg (2010) shows that the demographic characteristics of SAT takers are generally uncorrelated with the introduction of the bans on affirmative action in both California and Texas. In addition, Long (2004), Card and Krueger (2005) and Antonovics and Backes (2013) all find that Prop 209 did not have a large impact on score-sending patterns. Taken as a whole, this evidence suggests that Prop 209 did not have a large impact on the pool of SAT takers.

Basic summary statistics for our College Board sample are displayed in Table 2. Californians are quite

[15] In April 1995, the College Board recentered the SAT score scales to reestablish a mean score of about 500. To ensure consistency over time, we use College Board-provided recentered scores for all years.

similar to the rest of the country in terms of SAT scores and high school GPA, but, as might be expected, a smaller fraction are U.S. citizens and a larger fraction speak English as a second language. In both California and the rest of the US, blacks and Hispanics tend to score lower on the SAT, have lower GPAs, and have parents with lower levels of education.

5. Results

A. Overview of Trends in Test Scores

As a first pass at gauging the effects of Prop 209 on SAT scores and high school GPA, we plot normalized (mean zero and standard deviation one) average SAT scores by race and year in Figure 1. Panels (a) and (b) show normalized SAT math and verbal scores for Californians and the rest of the US. Although whites tend to score higher than URMs, the gap appears to be roughly stable over time. Panel (c) shows normalized high school GPA. The patterns in the figure underscore the importance of controlling for state-specific time trends: there was a gradual rise in GPA over time that begins before the implementation of the preference ban. By including state-specific time trends, we measure the deviation from these trends that accompanied the preference bans, rather than attributing pre-existing trends to the policy change.

B. Regression Results

Panel 1 of Table 3 presents results when including all but the 1997 cohort. The first three columns display our results when we include controls for demographic characteristics but not for the interaction between demographic characteristics and the post-Prop 209 indicator. In terms of the effect of Prop 209 on the overall level of academic achievement, for California relative to the rest of the country, we find that after Prop 209 there was a 0.002 standard deviation increase in SAT math scores and a 0.019 standard deviation increase in SAT verbal scores. We also find a 0.02 standard deviation relative increase in high school GPA. While the direction of these point estimates indicate an overall increase in academic achievement in California relative to the rest of the country, we note that the magnitudes are extremely small and, on balance, suggest that Prop 209 had no meaningful impact on the overall level of academic achievement.

In terms of the racial gap in academic achievement, for California relative to the rest of the country, we find a small reduction in the black-white SAT gap and a small increase in the Hispanic-white SAT gap. For high school GPA, the point estimates suggest a slight relative decrease in both the black-white GPA gap and the Hispanic-white GPA gap. Again, however, we note that the magnitude of our point estimates is extremely small. For example, the estimated 0.027 standard deviation increase in the Hispanic-white SAT math gap represents about 3.5 SAT points, or about 4 percent of the Hispanic-white SAT math gap in California. In addition, despite the estimated increase in the Hispanic-white SAT score gap, we find evidence of a decrease in the Hispanic-white high school GPA gap. Thus, our findings provide little support for the notion that banning affirmative action widens the racial gap in academic achievement.

Columns (4)-(6) present results when allowing the coefficients on the demographic controls to vary in the post period. Results are generally similar but smaller in magnitude since part of the effect which was previously captured by the URM*Post coefficient is now contained in the Demographics*Post coefficients.

Panel 2 further restricts the sample by dropping each of the 1996 through 1998 cohorts from the analysis. Many coefficients become larger in magnitude, especially when comparing Californians to the rest of the country. For example, the estimated performance gain of Californians on the SAT verbal section increases from .019 standard deviations to .046 standard deviations when dropping the additional years. Again, the estimates become somewhat smaller in magnitude when including Post*Demographic interactions as shown in columns (4) - (6).

The standard errors in Table 3 are generated by clustering at the state level. However, as discussed above, the presence of one policy change state may result in standard errors that are misleadingly small.

As a result, we conduct permutation tests in which we generate a coefficient `estimate' for each of the 45 states in our data using the specification which resulted in the largest point estimates -- columns (1) - (3) of panel 2. Figure 2 plots histograms of these placebo coefficient estimates, with the red line indicating the coefficient estimate for California. Our point estimates for California are generally extremely small relative to the other states. However, given wide the range of coefficient estimates from other states, we would be unable to detect even a large change in California.

In seeking to understand the imprecision of our estimates, we plot average SAT math scores by state and year for the first eight states alphabetically.[16] The plot is displayed in Figure 3, with California represented by the dashed line.[17]A naive look at the graphic would suggest a very small increase in the Californian SAT math scores in the post period. However, two factors prevent being able to make a definitive causal statement about the change in Californian performance due to Prop 209. First, the small increase in SAT math scores was part of a general upward trend in both California and the rest of the country (see Panel (a) of Figure 1). Second, the other states shown in Figure 3 generally have substantial year-to-year variability, which does not show up in the ``US'' panel of Figure 1a since all the states are averaged together. Thus, considerable year-to-year variation within each state makes it difficult to make definitive statements about the causal effect of the policy change despite the very large sample sizes in the College Board.

Despite the inherent problem with performing a difference-in-difference estimation with one policy change state, we believe that the small point estimates are noteworthy in themselves. As discussed above, Prop 209 led to substantial changes in URM admissions rates and in the UC admissions process. However, despite these changes, we find that the performance of Californians as a whole and of black

[16] The other control states are similar to those shown here but are not displayed in the figure due to lack of clarity when too many states are plotted at once.

[17] The relative ranking of the states shown in Figure 3 is largely driven by SAT participation rates, with low participation rates generally corresponding to high average scores. For example, Alabama has relatively high average SAT scores despite its poor performance on most standardized tests, such as the NAEP, because only about 10% of high school graduates in Alabama take the SAT.

students relative to white students was relatively stable. Even for Hispanics, who appear to have performed worse on the SAT in the post period, saw an increase in their high school GPAs. Thus, for neither Californians as a whole nor for any minority group within California do we find consistent evidence of lower academic achievement in response to Prop 209.

It is possible that our small estimated effect sizes are the result of including many students who would be unaffected due to small likelihoods of attending the top-tier UC campuses where the effects of Prop 209 were strongest. In our sample, only one in six Californians applies to a UC campus, and even fewer to a selective campus where the effects of Prop 209 were strongest. In an effort to isolate a sample of students likely to be most affected by the policy change, we predict which students would have a high likelihood of sending an SAT score to either Berkeley or UCLA. Specifically, for cohorts in the pre-Prop 209 period, we regress the likelihood of score-sending to at least one of Berkeley or UCLA on demographic characteristics that plausibly were not affected by Prop 209, which consist of parental education and income, race, gender, whether English was the first language spoken, and citizenship status. We then use the coefficients to predict score-sending for the entire sample, both in the pre and post periods. The density of predicted Berkeley or UCLA score-sending is shown in Figure 4, with the vertical lines denoting cutoff points between quartiles of score-sending likelihood. We keep the top quartile of predicted score-senders to generate a sample of students likely to be most affected by Prop 209.

We then run the same regressions as previously on this selected sample. Results for these likely Berkeley or UCLA score-senders are shown in Table 4. Coefficient estimates continue to be small, with the possible exception of an estimated reduction in the Hispanic-white GPA gap.[18]

C. Why Our Findings Differ from Previous Work

Standing in contrast to the results presented here, Furstenberg (2010) finds a statistically significant widening of the black-white SAT gap in California following Prop 209. However, our analysis improves

[18] We also investigated children of parents who attended at least some college. Unsurprisingly, these results are similar to the full sample since about 80% of the sample has at least one parent who attended at least some college.

upon his in several important ways. First, his College Board sample only includes the 1996-2000 cohorts; his paper compares the 1996-1997 cohorts to the 1998-2000 cohorts. However, interpreting results from 1996-1998 is difficult since, as discussed above, the UC Regents first announced their intention to end their use of racial preferences in July 1995. Thus, it is possible that students began responding to the policy change long before 1998, his first post-policy change year. Second, Furstenberg's data consist of a 30% sample of SAT takers, while we have obtained a more comprehensive dataset containing all Californian test takers, all black and Hispanic test takers nationwide, and a 25% sample of the remaining non-Californian whites. As a result, our estimates are obtained from a much larger sample.

However, even when using our dataset with Furstenberg's sample years, definition of the policy change year, estimating equation, and set of controls (obtained by direct correspondence with Furstenberg), we are still unable to replicate his finding of an increase in the black-white gap in California.[19]

An additional previous empirical study of the impact of affirmative action bans on human capital accumulation, Caldwell (2010), examines PIAT math test scores using the Children of the NLSY 79 (CNLSY). Caldwell finds that Prop 209 increased the racial test score gap. We obtained the same restricted version of the CNLSY used by Caldwell and successfully replicated his findings. However, there are several issues with his analysis.

First, the CNLSY was not designed to yield representative sample at the state level. From the NLS FAQ,[20]

> *The National Longitudinal Surveys are designed to represent specific birth cohorts at the national level. The surveys cannot provide representative estimates for States [...] NLS data files with geographic variables are available on a restricted basis for authorized researchers to use, but the permitted uses do not include producing estimates for States.*

Thus, it is difficult to generalize the findings obtained from the CNLSY because it is not a randomly

[19] Furstenberg no longer has access to the version of the data used in his paper, so we were unable to explore further.
[20] http://www.bls.gov/nls/nlsfaqs.htm\#anch14

selected sample of Californians. While the College Board data used in this paper is also not representative of all Californian high school students, it does include all Californian SAT test takers and these are the most likely to respond to Prop 209.

Compounding the non-representativeness of the CNLSY sample are the relatively small sample sizes in the panel. For example, one of Caldwell's strongest results -- suggesting that the PIAT math scores of 13 and 14 year olds fell considerably for blacks in California relative to blacks in the rest of the country -- is estimated from only 62 Californian blacks, with 17 of these from the period after Prop 209. This is important because, as noted in Wooldridge (2006), small sample sizes exacerbate the problems inherent to estimation with only one policy change group.[21]

Finally, we urge caution in interpreting the results of both Furstenberg and Caldwell since neither takes into account the limited number of treated states in evaluating the statistical significance of their results.

6. Conclusion

Much of the popular debate surrounding affirmative action in higher education focuses on how it affects the allocation of students to universities, taking the achievement of high school graduates as fixed. However, the disparities in educational preparation which drive racial differences in enrollment at selective colleges arise early in the education process and are formed well before college admissions come into play. As affirmative action was originally conceived to mitigate these gaps in racial achievement, it is natural to ask whether and how the removal of racial preferences affects these gaps. In addition, a number of scholars have pointed out that since policies such as Prop 209 give colleges and universities an incentive to place a greater weight on non-academic factors in determining admissions, they could lower

[21] The authors also obtained restricted-use data from the eighth grade math sample of the National Assessment of Educational Progress (NAEP), which has the advantage of being designed to be representative at the state level. However, the NAEP suffers from two serious drawbacks. First, background information is relatively sparse: for example, parental education but not income is available. Second, smaller sample sizes result in estimates that are considerably less precise than those from the College Board data.

student quality by weakening all students' incentives to invest in their academic qualifications prior to college entry.

Our results suggest that the end of raced-based affirmative action at the UC did not have any sizable impact on academic achievement prior to college entry. There is no evidence that either the black-white SAT score gap or the black-white high school GPA gap widened in California after Prop 209, and although there is some evidence that the Hispanic-white SAT gap grew, we find a simultaneous narrowing of the Hispanic-white high school GPA gap. Moreover, the magnitude of these changes is generally very small. A finding of no effect is consistent with several previous studies that have shown a lack of a response to Prop 209 in terms of both application behavior and enrollment behavior.

References

Abadie, A., Diamond, A., Hainmueller, J.: Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association **105**(490), 493-505 (2010)

Antonovics, K., Backes, B.: Were Minority Students Discouraged from Applying to University of California Campuses After the Affirmative Action Ban? Education Finance and Policy **8**(2), 208-250 (2013)

Antonovics, K., Backes, B.: The Effect of Banning Affirmative Action on College Admissions Rules and Student Quality. Journal of Human Resources **49**(2) (2014)

Antonovics, K., Sander, R.H.: Affirmative Action Bans and the 'Chilling Effect'. American Law and Economics Review **15**(1), 252-299 (2013)

Arcidiacono, P., Aucejo, E., Coate, P., Hotz, J.: Affirmative Action and University Fit: Evidence from Proposition 209. Duke Workign Paper and NBER Workign Paper w18523 (2014)

Arcidiacono, P., Lovenheim, M.: Affirmative Action and the Quality-Fit Tradeoff. Duke Working Paper. http://public.econ.duke.edu/psarcidi/aa.pdf . Accessed 14 May 2014 (2014)

Backes, B.: Do Affirmative Action Bans Lower Minority College Enrollment and Attainment? Evidence from Statewide Bans and Judicial Rulings. Educational Evaluation and Policy Analysis **36**(2), 228-252 (2014)

Buchmueller, T.C., DiNardo, J.E., Valletta, R.G.: The Effect of an Employer Health Insurance Mandate on Health Insurance Coverage and the Demand for Labor: Evidence from Hawaii. American Economic Journal: Economic Policy **3**(4), 25-51 (2011)

Caldwell, R.: The Effects of University Affirmative Action Policies on the Human Capital Development of Minority Children: Do Expectations Matter? RePEc Working Paper. http://ideas.repec.org/p/kan/wpaper/200812.html. Accessed 14 May 2014 (2010)

Card, D., Krueger, A.B.: Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants? Evidence from California and Texas. Industrial and Labor Relations **58**(3), 416-434 (2005)

Cortes, K., Friedson, A.: Ranking Up by Moving Out: The Effect of the Texas Top 10% Plan on Property Values. IZA Working Paper No. 5026 (2010)

Cortes, K.E.: Do Bans on Affirmative Action Hurt Minority Students? Evidence from the Texas Top 10% Plan. Economics of Education Review **26**(6), 1110-1124 (2010)

Cullen, J.B., Long, M.C., Reback, R.: Jockeying for Position: Strategic High School Choice Under Texas' Top Ten Percent Plan. Journal of Public Economics (2012)

Dickson, L.M.: Does Ending Affirmative Action in College Admissions Lower the Percent of Minority Students Applying to College? Economics of Education Review **25**(1), 109-119 (2006)

Donald, S.G., Lang, K.: Inference with Difference-in-Differences and Other Panel Data. The Review of Economics and Statistics **89**(2), 221-233 (2007)

Fryger, R.G., Loury, G.C., Yuret, T.: An Economic Analysis of Color-Blind Affirmative Action. Journal of Law, Economics, and Organization **24**(2), 319-355 (2008)

Furstenberg, E.: College Admissions Reform and Racial SAT Score Gaps: An Empirical Analysis. Working Paper (2010)

Hickman, B.R.: Using Auction Theory to Study Human Capital Investment in Matching Markets: a Look at Affirmative Action in College Admissions. Working Paper (2012)

Hickman, B.R.: Pre-College Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions. Working Paper (2013)

Hinrichs, P.: The Effects of Affirmative Action Bans on College Enrollment, Educational Attainment, and the Demographic Composition of Universities. Review of Economics and Statistics **94**(3) 712-722 (2012)

Holzer, H., Neumark, D.: Assessing Affirmative Action. Journal of Economic Literature **38**(3), 483-568 (2000)

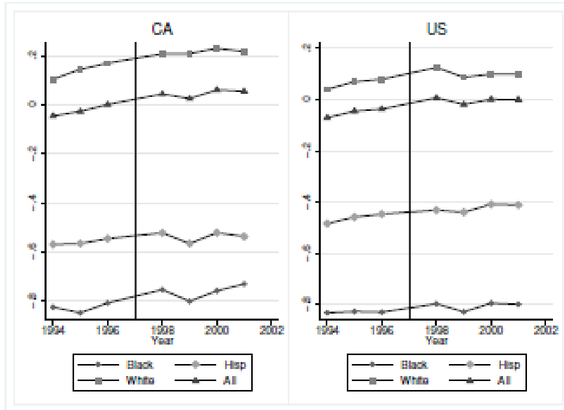
Long, M.C.: College Applications and the Effect of Affirmative Action. Journal of Econometrics **121**(1-2), 319-342 (2004)

Moulton, B.R.: An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit. The Review of Economics and Statistics **72**(2) 334-338 (1990)

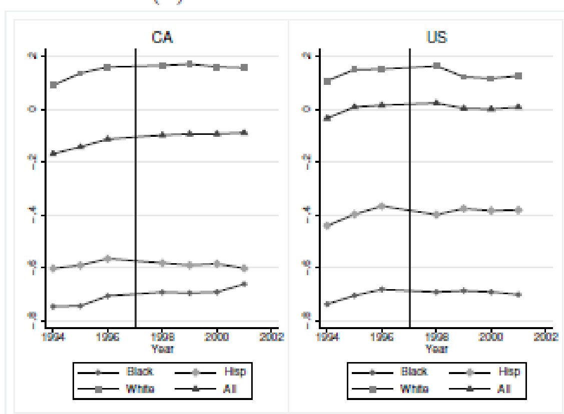
Ray, D., Sethi, R.: A Remark on Color-Blind Affirmative Action. Journal of Publics Economic Theory **12**(3), 399-406 (2010)

Wooldridge, J.W.: Cluster Sample Methods in Applied Econometrics: An Extended Analysis. Michigan State University. Unpublished Manuscript. (2006)

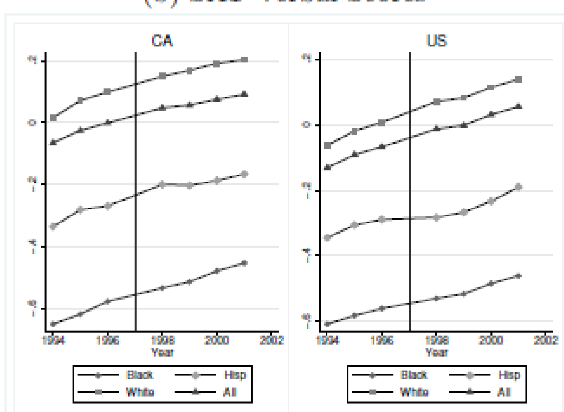
Figures



(a) SAT Math Scores

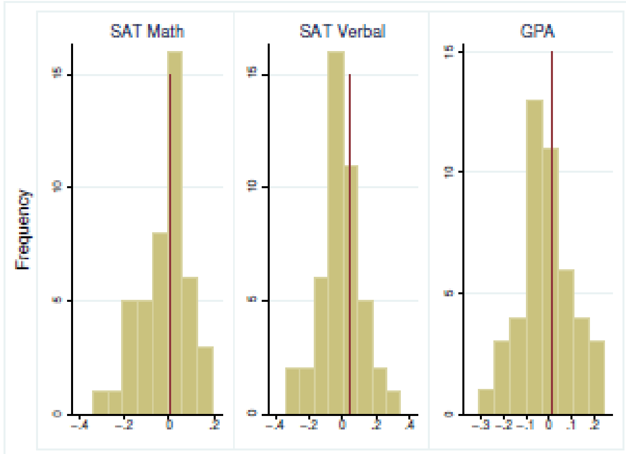


(b) SAT Verbal Scores

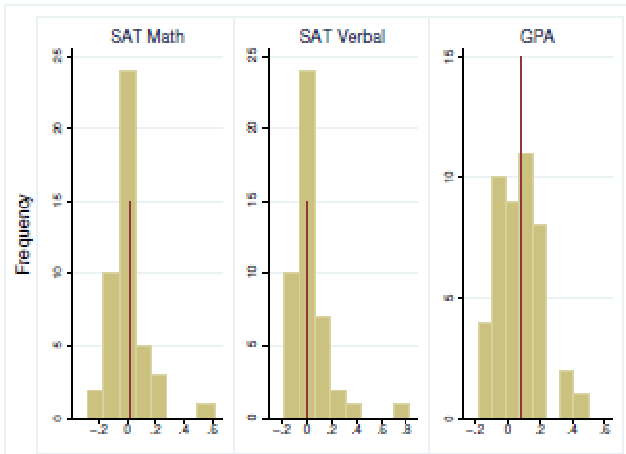


(c) High School GPA

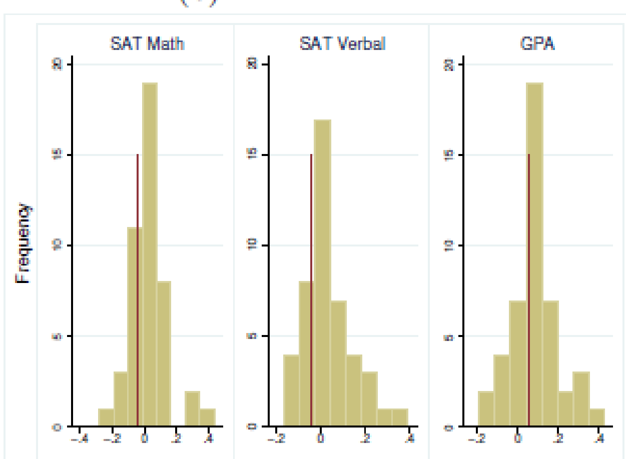
Figure 1: College Board outcome variables. Each panel shows the mean of the normalized outcome (mean zero and standard deviation one) by race and year. The US panel excludes California, Louisiana, Mississippi, Texas, Washington, North Dakota, and Wyoming.



(a) CA*Post



(b) CA*Post*Black



(c) CA*Post*Hispanic

Figure 2: Placebo Test Coefficients. Distribution of DDD coefficient estimates for each of the 45 states in the sample, with California represented by red line. See text.

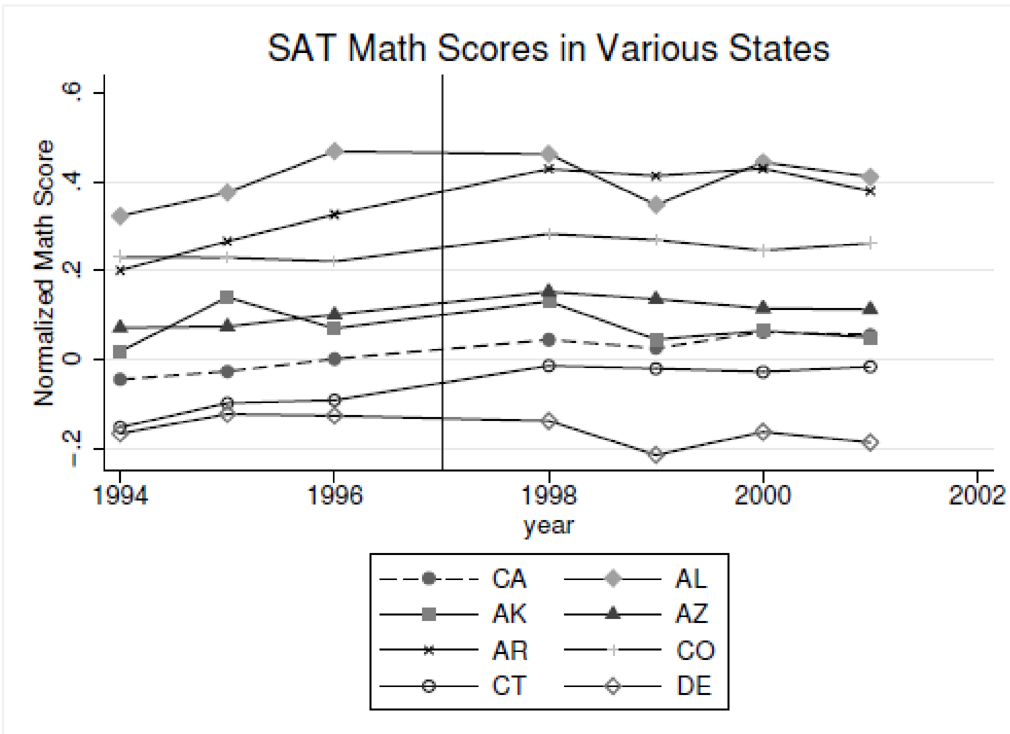


Figure 3: Mean of normalized SAT math scores. Mean zero and standard deviation one.

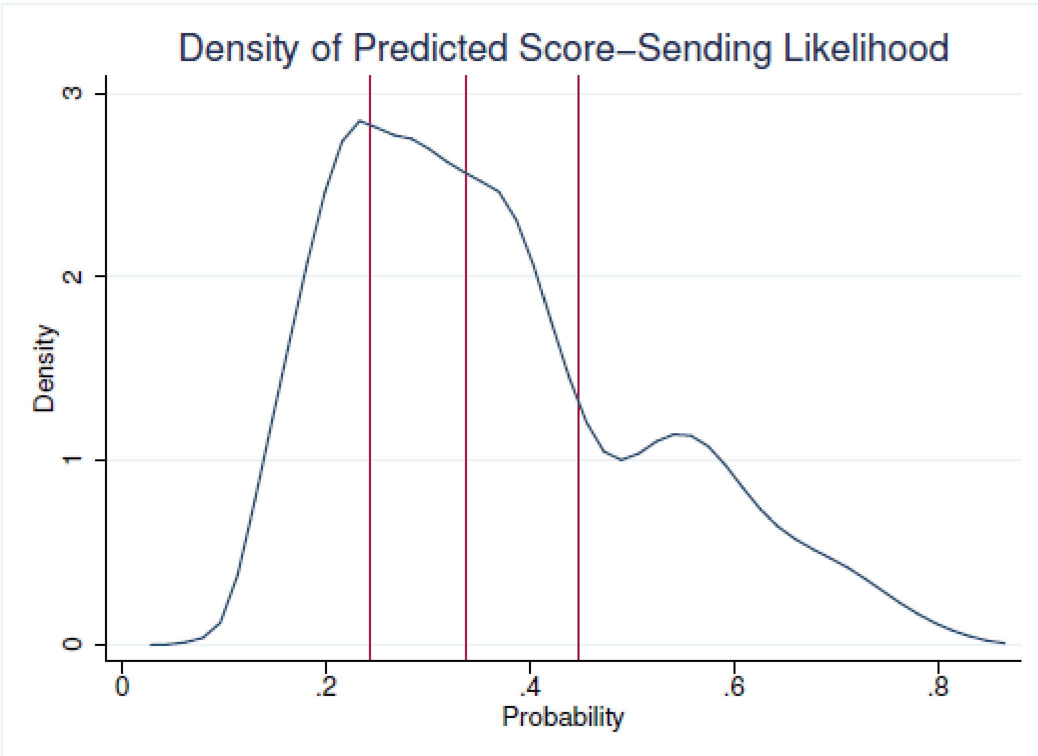


Figure 4: Density of predicted score-sending. Predicted score-sending obtained by regressing the likelihood of score-sending for Californian pre-period cohorts to at least one of Berkeley or UCLA on demographic characteristics consisting of parental education and income, race, gender, first language spoken, and citizenship status. Details in text. Each vertical line represents the boundary of a quartile.

Tables

Table 1: California's Proposition 209 Timeline

Date	Event
Spring 1995	1996 graduation cohort begins taking SAT
July 1995	Regents of UC pass SP-1
Fall 1995	1996 graduation cohort finishes taking SAT
Spring 1996	1997 cohort begins taking SAT
July 1996	First mention of Prop 209 in media
November 1996	Prop 209 passed by voters
Fall 1996	1997 cohort finishes taking SAT
Spring-Fall 1997	1998 cohort takes SAT
November 1997	Supreme Court declines to review case, Prop 209 becomes law
Fall 1998	First affected cohort (1998) enrolls

Notes: See text for description.

Table 2: College Board Summary Statistics

	California					Rest of Country				
	All	Black	Hispanic	White	Asian	All	Black	Hispanic	White	Asian
SAT Math	514	427	453	536	548	512	424	467	524	559
	(114)	(101)	(101)	(102)	(120)	(110)	(96.2)	(105)	(104)	(120)
SAT Verbal	495	434	445	529	485	510	435	469	524	501
	(114)	(103)	(103)	(102)	(123)	(107)	(98.4)	(106)	(101)	(128)
GPA	3.26	2.89	3.1	3.33	3.38	3.22	2.91	3.08	3.27	3.38
	(.627)	(.624)	(.624)	(.602)	(.612)	(.659)	(.66)	(.659)	(.645)	(.641)
Male	.451	.427	.419	.457	.473	.456	.406	.435	.463	.483
US citizen	.861	.956	.827	.952	.665	.947	.936	.815	.977	.674
ESL	.2	.0313	.36	.0505	.43	.0563	.0243	.316	.0201	.379
Income ($10,000s)	5.31	3.96	3.65	6.54	4.54	5.62	3.68	4.06	6.05	5.06
	(3.5)	(2.93)	(2.75)	(3.46)	(3.37)	(3.27)	(2.69)	(3.05)	(3.21)	(3.45)
Mom educ (yrs)	13.9	14.5	10.6	15.4	13.6	14.8	14.3	13.3	15	14
	(4.32)	(3.1)	(5.28)	(2.88)	(4.54)	(3.18)	(3.14)	(4.16)	(2.97)	(4.33)
Dad educ (yrs)	14.4	14.3	10.8	15.9	14.4	15.1	13.9	13.4	15.4	15.1
	(4.49)	(3.28)	(5.51)	(3.1)	(4.43)	(3.43)	(3.36)	(4.55)	(3.21)	(4.28)
Observations	829336	56636	172424	410050	190226	1765615	534226	240502	922409	68478

Notes: Standard deviations in parentheses. Rest of country excludes California, Louisiana, Mississippi, Texas, Washington, North Dakota, and Wyoming. Sample includes 1994-1996 and 1998-2001.

Table 3: College Board SAT and GPA DDD Estimates

	SAT M (1)	SAT V (2)	GPA (3)	SAT M (4)	SAT V (5)	GPA (6)
Panel 1: 1994-6 (pre) vs 1998-2001 (post)						
Post*CA	0.002	0.019***	0.020***	-0.000	0.016***	0.019***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)	(0.005)
Post*Black*CA	0.019*	0.007	0.065***	0.020**	0.010	0.063***
	(0.010)	(0.007)	(0.012)	(0.010)	(0.007)	(0.012)
Post*Hispanic*CA	-0.027***	-0.030***	0.058**	-0.016***	-0.018***	0.069***
	(0.005)	(0.006)	(0.023)	(0.004)	(0.005)	(0.022)
Observations	2648191	2648191	2743387	2648191	2648191	2743387
R-squared	0.252	0.235	0.133	0.252	0.235	0.133
Panel 2: 1994-5 (pre) vs 1999-2001 (post)						
Post*CA	0.001	0.046***	0.018*	-0.002	0.042***	0.017*
	(0.007)	(0.007)	(0.010)	(0.007)	(0.007)	(0.010)
Post*Black*CA	0.019	0.000	0.079***	0.020	0.004	0.077***
	(0.013)	(0.009)	(0.015)	(0.013)	(0.009)	(0.014)
Post*Hispanic*CA	-0.041***	-0.047***	0.057*	-0.028***	-0.034***	0.068**
	(0.004)	(0.006)	(0.030)	(0.004)	(0.005)	(0.030)
Observations	1965483	1965483	2033955	1965483	1965483	2033955
R-squared	0.254	0.237	0.135	0.254	0.238	0.135
Demographics	x	x	x	x	x	x
Post*Demographics				x	x	x

Notes: Each column shows the coefficient estimates from a regression of the outcome variable listed at the top of each column on an indictor for whether the student was from California, an indictor for whether they took the SAT in the post period, and an indicator for the student's race, along with the full set of interactions between these variables. Additional controls for parental education and income, gender, first language spoken, and citizenship status are also included in the regressions. All outcome variables normalized to have mean zero and standard deviation one. Included time trends are linear state-specific time trends. The excluded racial group includes white, other, and unknown. Excludes Louisiana, Mississippi, Texas, Washington, North Dakota, and Wyoming. Standard errors clustered at the state level.

Table 4: College Board SAT and GPA DDD Estimates for Students in Top Quartile by Likelihood of Score-sending to Berkeley or UCLA

	SAT M	SAT V	GPA	SAT M	SAT V	GPA
	(1)	(2)	(3)	(4)	(5)	(6)
1994-6 (pre) vs 1998-2001 (post)						
Post*CA	0.011	0.021	0.030**	0.020*	0.024*	0.037***
	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)	(0.014)
Post*Black*CA	0.007	0.022	-0.013	-0.011	0.023	-0.015
	(0.011)	(0.014)	(0.014)	(0.011)	(0.015)	(0.012)
Post*Hispanic*CA	0.011	0.010	0.074***	0.007	0.012	0.068***
	(0.012)	(0.008)	(0.018)	(0.011)	(0.009)	(0.017)
Observations	432578	432578	441475	432578	432578	441475
R-squared	0.184	0.260	0.089	0.188	0.264	0.091
1994-5 (pre) vs 1999-2001 (post)						
Post*CA	-0.015	0.033	0.043	-0.005	0.034	0.051*
	(0.023)	(0.024)	(0.026)	(0.024)	(0.024)	(0.027)
Post*Black*CA	-0.005	0.021	-0.010	-0.024	0.019	-0.010
	(0.014)	(0.018)	(0.018)	(0.017)	(0.019)	(0.017)
Post*Hispanic*CA	0.013	0.001	0.061***	0.004	0.005	0.053**
	(0.016)	(0.009)	(0.022)	(0.014)	(0.012)	(0.021)
Observations	314804	314804	320673	314804	314804	320673
R-squared	0.186	0.261	0.089	0.190	0.265	0.092
Demographics	x	x	x	x	x	x
Post*Demographics				x	x	x

Notes: Compares Californians in the top predicted quartile of score-sending to Berkeley or UCLA with students in other states who would have been predicted to be in that quartile had they been in California. See notes from Table 3.