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

Kate Antonovics†, Ben Backes‡

Preliminary Draft. Do Not Cite Without Permission.

September 25, 2013

Abstract

This paper examines how the end of race-based affirmative action at the University of California in 1998 affected the overall level of human capital investment and the racial gap in human capital investment prior to college entry. We proxy for human capital accumulation using standardized test scores from three different datasets. Using a difference-in-difference methodology, we examine how human capital accumulation changed in California relative to other states after California's ban on affirmative action, and we examine how the racial gap in human capital changed in California relative to other states. We point out the weaknesses of two previous empirical studies that find evidence that banning affirmative action increases the racial test score gap. Our primary conclusion is that once we adjust our standard errors to account for the natural within-state variation in test scores over time (an adjustment not made in previous studies), we can say very little about the effect of California's ban on affirmative action on either the racial test score gap or test scores more generally.

*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 Julie Cullen, Terrance August, Zoltan Hajnal, and seminar participants at UC San Diego for their helpful comments.

†UC San Diego, *kantonovics@ucsd.edu*

‡American Institutes for Research / CALDER, *bbackes@air.org*

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 in admissions as the result of various court rulings, voter initiatives, and administrative decisions.[1] In addition, the United States Supreme Court's recent decision in *Fisher v. Texas* will make it more difficult for universities to justify using race as a factor in admissions, and many believe the Supreme Court will place further limits on affirmative action when it issues its ruling in *Schuette v. Coalition* next year. 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 human capital accumulation 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 under-represented 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 the weight given to different student characteristics in order to to implicitly favor minorities, then students could alter their behavior in response to the new admissions incentives. Antonovics and Backes (2013a), for example, provide evidence that after California's ban on affirmative action, selective schools in the University of California (UC) system boosted minority admission rates by reducing the weight placed on SAT scores and increasing the weight given to high school grades, and such policy changes might lead students to shift the focus of their efforts in high school.

Unfortunately, while economic theory is clear that ending affirmative action will affect human capital accumulation, the theory does not yield definitive predictions about

[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 (2013) for a nice discussion of the policy environment surrounding affirmative action.

whether the racial gap in human capital investment will increase or decrease and whether the overall level of human capital investment (regardless of race) will go up or down.[2] Thus, the effect of affirmative action on human capital accumulation is largely an empirical question.

This paper attempts to answer this question by examining how the end of race-based affirmative action at the University of California in 1998 affected the overall level of human capital investment and the racial gap in human capital investment prior to college entry. To do so, we proxy for human capital accumulation using SAT scores and high school GPA from the College Board (CB), as well as standardized test scores from both the Children of the National Longitudinal Survey of Youth 79 (CNLSY) and the National Assessment of Educational Progress (NAEP). Using a difference-in-difference methodology, we examine how human capital accumulation changed in California relative to other states after California's ban on affirmative action, and we examine how the racial gap in human capital changed in California relative to other states. Along the way, we point out the weaknesses of two previous empirical studies (Furstenberg (2010) and Caldwell (2010)) that find evidence that banning affirmative action increases the racial test score gap.[3] Our primary conclusion is that once we adjust our standard errors to account for the natural within-state variation in test scores over time (an adjustment made in neither Furstenberg (2010) nor Caldwell (2010)), we can say very little about the effect of California's ban on affirmative action on either the racial test score gap or test scores more generally.

Our paper proceeds as follows. In Section 2 we discuss the related literature and present an overview California's affirmative action ban. Section 3 discusses our empirical strategy and gives a short overview of our data sources. Section 4 then uses data from the College board, NAEP, and CNLSY to examine the implications of Prop 209 on human capital investment prior to college entry. Section 5 concludes.

2 Background and Related Literature

The first threat to affirmative action in California was in July 1995, when the Board of Regents of the University of California passed a resolution (SP-1), which stipulated

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

[3] We discuss these papers when we present our empirical results.

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 Proposition 209 (Prop 209), which banned the use of racial preferences in university admissions.[4] 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.[5]

Given that one in six Californian high school graduates apply to at least one UC campus, it's reasonable to think that major policy changes at the UC could affect human capital accumulation.[6] Evidence that students' human capital investment decisions respond to the incentives created by college admissions policies is also evident in Cortes and Friedson (2011) 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.

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.[7] It's 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. Nonetheless, to the extent that increased outreach had an effect on human capital accumulation, 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 Cal-

[4]Searching the LexisNexis article database gives the first mention of Prop 209 in July 1996.

[5]For a complete summary of the events of the ban, see Table 1.

[6]UC application data: http://www.ucop.edu/news/archives/2001/applications_2001/table1.pdf. California high school graduation data: http://www.cpec.ca.gov/completereports/2003reports/03-09/Display1.PDF

[7]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*.

ifornia. 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 in our College Board data and find no substantial impact on our main results. Unfortunately, we cannot drop these later years from our other data sets because we would not have a post policy change period.

As mentioned above, several theoretical papers have examined the effect of affirmative action on human capital investment.[8] 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 example, Fryer et al. (2008) and Ray and Sethi (2010)).[9] 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 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 as mentioned above, Antonovics and Backes (2013a) provides evidence that after Prop 209 UC schools changed the weights given to SAT scores, high school GPA and family background characteristics in order to implicitly favor minorities in the admissions process.

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 ac-

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

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

tion 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.[10]

3 Empirical Strategy and Overview of Data

Antonovics and Backes (2013a) provide evidence that UC campuses reacted to Prop 209 in part by changing the weights given to test scores, high school GPA and family background characteristics. Since about one in six Californian high school graduates applies to at least one UC campus, it is plausible that Prop 209 could have led to changes in pre-college behavior. Given the apparent reduced emphasis on SAT math scores relative to SAT verbal scores and high school GPA at UC campuses, we might expect to see a shift in human capital investment away from investments that increase SAT math scores and towards investments that increase both SAT verbal scores and high school GPA.

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.[11] Each of the two measures is important. Since Californians of all races were affected by the change in admissions policy for UC campuses, they may be thought of as one treated group. Comparing Californians to those in other states reveals the extent to which Prop 209 changed investment incentives 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 student quality.

[10]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).

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

3.1 Empirical Strategy

We now turn to the basic model we use for our main results:

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

where Post is an indicator variable for whether the affirmative action ban is in place (1999 and after) and URM is an indicator for an individual's race, with separate indicators for blacks and Hispanics. 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_{01}, 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 after Prop 209. Since affirmative action policy is largely driven by achievement gaps between minority and white students, it is important to understand how Prop 209 affected these gaps.

In practice, we estimate the above model using OLS and expand our estimating equation to include the full set of interactions for blacks, Hispanics, and Asians, with whites as the excluded group. We also add year fixed effects, state fixed effects, state-specific linear time trends, and use the sampling weights provided by each dataset. The X term includes available demographic controls, which differ by dataset, and a constant term. We drop observations from Louisiana, Mississippi, Texas, and Washington, which were affected by their own affirmative action policy changes during our sample period.

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.[12]

[12] And adding Asians to the excluded group in the DDD regressions gives similar results.

3.2 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 Lang (2007), Abadie et al. (2010), and Buchmueller et al. (2011)). To illustrate the problem, consider SAT scores as the outcome variable and suppose we are intested in β_{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 $\hat{\beta}_{01}$ for each of the control states, treating each control state as the treated state (see Abadie et al. (2010) and Buchmueller et al. (2011)). Thus, the control states are used to generate an estimate of the natural year-to-year variability in test scores. The null hypothesis – that Prop 209 had no effect on human capital investment in California – is rejected when $\hat{\beta}_{01}^{\text{California}}$ is large relative to the estimated empirical distribution of $\hat{\beta}_{01}^{j}$ of the control states (j indexes the control states). This procedure tests whether the change in test scores California's post period is large relative to the naturally occurring variation in test scores observed in other states. We also estimate $\hat{\beta}_{012}^{\text{California}}$ 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 states which did not ban affirmative action in the same time period. We find that this correction is extremely important when considering student effort because of substantial within-state year-to-year variability in three different standardized tests – the SAT, the NAEP, and the PIAT.

Of course, one solution to the above problem would be to include more treated states. However, most states which banned affirmative action have done so too recently to be contained in our data.[13] Two states which banned affirmative action during the period

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

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.[14] 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.[15] In addition, neither state's public universities are as selective as the institutions typically affected by bans on racial preferences (see Blume and Long (2013)). 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 measurement of how admissions rules at the UC after Prop 209 has been documented by Antonovics and Backes (2013a).

3.3 Overview of Datasets

To test for the effects of Prop 209 on human capital investment, we make use of data from the College Board, which contain information on both SAT scores and self-reported high school GPA. In addition, we utilize data from the National Assessment of Educational Progress (NAEP) on a representative sample of eighth graders. Finally, we use the Children of the National Longitudinal Study of Youth 79 (CNLSY), which administers the PIAT to children in the sample up to 14 years old. While none of the three data sets are ideal, they allow us to conduct three different tests of whether there were shifts in human capital investment in response to Prop 209. To ease the comparison of our results across datasets and outcome measures, we normalize test scores and GPA to be mean zero with standard deviation one.

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

[15]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.

4 Results

We now explore the results from each of the three data sets we employ, and we relate our findings to previous studies examining the impact of banning affirmative action on human capital investment.

4.1 College Board (SAT Scores and High School GPA)

Our College Board data consist of SAT test takers who are expected to graduate from high school between 1994 and 2001.[16] Our 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.

Defining the period before and after the ban on racial preferences in the College Board data is complicated by the timing of when students took the SAT. 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. With this in mind, Table 1 shows the timing of important events related to the passage 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 did so during a time of considerable uncertainty about the future of race-based affirmative action. To ensure a clean before and after comparison, we drop observations in the 1996-1998 cohorts.[17]

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 the

[16] 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.

[17] Our main conclusion – that there is too much variation in the data to detect a policy effect – is not sensitive to the inclusion of these years.

affirmative action bans. On the other hand, a potential problem with using the College Board is that students decide whether or not to take the SAT, which could lead to sample selection bias. Indeed, Dickson (2006) finds that removal of affirmative action in Texas led to a decline in the share of minority students taking either the ACT or SAT. However, Furstenberg (2010) shows that the demographic characteristics of SAT takers are generally uncorrelated with the introduction of the bans in California and Texas. In addition, Antonovics and Backes (2013b) shows that changes in SAT score-sending by URMs to the UC system after Prop 209 were minimal relative to changes in admissions rates and concentrated at the most selective UCs.

Basic summary statistics for our College Board sample are displayed in Table 2. Blacks and Hispanics tend to score lower on the SAT, have lower GPAs, and have parents with lower levels of education. Not surprisingly, a high fraction of SAT test takers report hoping to obtain at least a BA: about 80 percent. Of the remainder, nearly all report their degree goal as "Other" or "Unknown". The small remaining share are those who hope to obtain certificates or associates degrees.

As a first pass at gauging the effects of Prop 209 on SAT scores and high school GPA, we plot normalized average SAT scores by race and year in Figure 1. Panel (a) shows normalized SAT math scores for Californians and the rest of the US. Although whites tend to score higher than URMs, the gap appears to be stable over time. Panel (b) shows normalized 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.

We next estimate our basic difference-in-difference-in-difference equation with results shown in Table 3. Each coefficient shows two sets of standard errors: one generated using standard errors clustered at the state level, and another generated by using the permutation tests mentioned above. To generate the second set of standard errors, we iterate over each state to estimate coefficients and then take the second most extreme values of the 47 states (all 50 states plus the District of Columbia, excluding Louisiana, Mississippi, Texas, and Washington due to their own affirmative action policy changes) to form a confidence interval. A statistically significant estimate for California would entail falling outside this interval.

Our point estimates are small, showing no change in Californian SAT math scores and about a 4 percentage point increase in Californian SAT verbal scores relative to the rest of the country after Prop 209. In terms of the racial gap in test scores, we estimate a small reduction in the black-white SAT math gap and a small increase in the Hispanic-white SAT math and verbal gap. In addition, the point estimates suggest a reduction in the minority-white high school GPA gap. The largest estimated change in the racial test score gap – an increase in the Hispanic-white test score gap of .05 standard deviations – represents less than 7 percent of the Hispanic-white SAT test score gap in California. For high school GPA, the coefficient estimates are somewhat larger in magnitude. Although the traditional standard errors obtained by clustering at the state level indicate that some of these point estimates are significant, the confidence intervals obtained from the permutation tests are very large.

The information in Table 3 is presented in Figure 2, which plots histograms of the placebo coefficient estimates for each of the 47 states in our data. The red line indicates the coefficient estimate for California. Our point estimates for California are generally extremely small and fall well within the set of estimates obtained from other states. However, the confidence intervals are so large that we would be unable to detect even a large change in California.

To investigate the imprecision of our estimates, we plot average SAT math scores by state and year for the first eight states alphabetically.[18] The plot is displayed in Figure 3, with California represented by the dashed line.[19] A naive look at the graphic would suggest a small increase in the Californian SAT math scores in the post period. However, two factors prevent being able to make the claim of an increase in Californian performance due to Prop 209. First, the small increase in SAT math scores was part of a nationwide upward trend (see Panel (a) of Figure 1), which is why the estimated coefficient in Table 3 is zero. 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 of the control

[18]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.

[19]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.

states precludes the researcher from making definitive statements about the causal effect of the policy change despite the very large sample sizes in the College Board because it is impossible to tell whether changes within California were due to Prop 209 or simply reflect yearly variations inherent to the SAT. The results for our other measures of human capital accumulation contain similar variability.[20]

Unlike us, Furstenberg finds a statistically significant widening of the black-white SAT gap in California following Prop 209. However, his standard errors do not take into account the limited number of treated states and the large variability in SAT scores in control states. However, even if one were to take the standard errors used in Furstenberg's paper at face value, there are two other major issues with his findings. 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. In contrast to Furstenburg, our main results compare the 1994 and 1995 cohorts to the 1999, 2000, and 2001 cohorts to ensure a clean before and after comparison. Second, 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.[21] 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. To the extent that the College Board is prone to errors when generating random samples, it would affect Furstenberg's results more readily than ours, since our data include the entire population of minority and Californian test takers.

In summary, if one were to ignore the problems associated with estimating a policy effect from one treated state, the College Board results would provide some evidence

[20]We also broke the sample down by parental education and found similarly large variability in the estimates from the control states. In addition, we conducted an analysis of stated degree goal preferences and failed to find an effect.

[21]Furstenberg no longer has access to the version of the data used in his paper, so we were unable to explore further.

that Californian students reacted to the changes in UC admissions policy by investing slightly more in SAT verbal (relative to math) scores and no evidence of an increase in Californian high school GPA. However, these results are small in magnitude and using permutation tests to account for the presence of only one treated state generates much larger confidence intervals.

4.2 NAEP (8th Grade Math Scores and Homework)

We supplement our College Board analysis by examining the math test scores of a younger group of students using the eighth grade math sample of the National Assessment of Educational Progress (NAEP), administered by the National Center for Education Statistics.[22] Beginning in 2002, states wishing to receive a Title I grant must participate in the NAEP; however, state assessments have been given since 1990 on a voluntary basis, when 37 states, including California, participated. According to the NAEP administrators, the twin goals of the survey are to compare student achievement across states and to track changes in achievement over time, both of which are useful for this paper.[23]

There are two advantages to supplementing our College Board data with the NAEP. First, the NAEP is designed to be representative of the statewide population of eighth-grade students enrolled in public schools, so it does not suffer from the same selection issues of the College Board, which is nearly entirely comprised of students interested in attending college. Second, to the extent that Prop 209 altered incentives to invest in human capital, the effects should be apparent at earlier ages, allowing us to test whether there are patterns apparent in multiple datasets. However, the NAEP suffers from some 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 less precise than those from the College Board data.

Our sample consists of Grade 8 state mathematics results from 1990, 1992, 2000, and 2003. Since only public schools were sampled in 1990, we restrict our sample to students enrolled in public schools. In addition, to ensure a consistent group of control states, we include only states that were sampled in each of the four years.[24]

[22] The NAEP Reading and Writing portions for Grade 8 each began in 1998.

[23] For further details, see http://nces.ed.gov/nationsreportcard/faq.asp.

[24] The authors were also able to obtain Grade 4 NAEP Reading exams from 1992 and 1998 and Grade 4 NAEP Mathematics exams from 1992 and 2000. We did not find any significant effects on test scores

Summary statistics by race are shown in Table 4. The only individual level control variables available for our regression analysis are gender and parental education, measured in years of educated for the highest educated parent.[25] As with the other datasets, white students tend to have higher test scores than do URMs and have more educated parents.

As an initial exploration of the NAEP data, we plot mean test scores by race and year in Figure 4. Panel (a) shows the results from the Grade 8 mathematics assessment. It appears that the average Californian performed substantially worse in 2000 and 2003 relative to the rest of the country. As with the College Board test scores, it is difficult to detect a pattern in the URM-white test score gap in California.

Our main NAEP results are presented in Table 5. The first column shows our estimates for normalized NAEP math scores, comparing 1990 and 1992 (pre) to 2000 and 2003 (post). The coefficient on CA*Post indicates that relative to the rest of the country NAEP math test scores fell in California after Prop 209 by about 1/4 of a standard deviation. However, while this point estimate is very large, it does not fall outside the confidence intervals generated by the permutation tests, which have a distribution shown in Figure 5. In addition, as with the SAT, the point estimate suggests a reduction in the black-white math score gap, but the confidence intervals are very large.[26]

4.3 National Longitudinal Survey of Youth (PIAT Reading)

Finally, we examine test score data using the Children of the NLSY 79 (CNLSY). Caldwell (2010) also uses these data to argue that that Prop 209 increased the racial gap in academic achievement prior to college. We obtained the same restricted version of the CNLSY used by Caldwell and were able to replicate his qualitative findings. However, there are several problems with his analysis.

One issue is that the CNLSY is not designed to comprise a representative sample of the state population. From the NLS FAQ,[27]

The National Longitudinal Surveys are designed to represent specific birth

(for Californians relative to the rest of the country or for minority-white gaps within California), and the estimates are imprecise due to having few years of data.

[25] While mother's and father's education are separately available, using each results in a large number of missing values, so we use the composite measure provided by the NAEP.

[26] As with the College Board data, we attempted splitting the sample into students whose parents attended college and those that did not, with similarly imprecise estimates.

[27] http://www.bls.gov/nls/nlsfaqs.htm#anch14

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, there is no reason to believe that estimates of the effects of Prop 209 generated from the CNLSY are representative of the general population of Californian students. As a result, any differences between estimates obtained from the CNLSY and estimates obtained from the NAEP for students in the same age group could be caused by the non-representativeness of the CNLSY.

Compounding the non-representativeness of the sample are the relatively small sample sizes in the CNLSY 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 Californian 62 blacks, with 17 of these are 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.

One additional problem with the analysis in Caldwell is that some variables in the CNLSY were not collected until the post period, so there is no pre-post comparison. For example, the paper estimates a simple interaction between state and race in only the post period to argue that that black and Hispanic students in California and Texas spent less time doing homework as the result of the affirmative action bans in those states. However, it could be the case that black and Hispanic students in those states were already different from those in other states before the policy change.

With the above caveats in mind, we turn to our analysis of the CNLSY. The data are from children of the original NLSY sample. Since the CNLSY is not representative of the nation's population, we use the survey-provided sample weights. Again, we note that the CNLSY weights are designed to create a representative sample of the nation as a whole, rather than any individual state. Following Caldwell (2010), we restrict our sample to children living with their mothers at the time of the survey. To keep our estimation consistent across data sets, we treat the CNLSY as a repeated cross section and do not exploit the longitudinal nature of the data.[28]

[28]In one section of his paper, Caldwell uses individual fixed effects to estimate test score growth rates.

Summary statistics for the CNLSY sample are shown in Table 6.[29] We display both PIAT reading subtests, reading recognition (measuring word recognition and pronunciation ability) and reading comprehension (measuring ability to derive meaning from sentences that are read silently). The patterns are consistent with the previous datasets, with black and Hispanic students being relatively more likely to come from disadvantaged backgrounds and scoring lower on the provided tests. When conducting our regressions, we follow Caldwell and control for times taken the test, gender, a quadratic in age and grade, mother's education in years, an indicator for urban, and birth order.

We next conduct the same permutation exercise as performed on the previous two datasets by estimating coefficients for each state and comparing the California coefficient to the distribution of coefficient estimates from each state. To ensure adequate sample sizes, we only include states with at least three each of black and Hispanic students both before and after Prop 209.[30] Unfortunately, only 18 states meet this criteria. The limited number of comparison states make it relatively more likely for California to have one of the more extreme coefficient estimates, so it is not possible to form a 95% confidence interval. We proceed with the permutation tests in order to assess the extent to which changes in California differ from the estimated placebo changes in the other 17 states. The distribution of coefficient estimates for each state are shown in Figure 6, with the California estimate represented as a vertical red line. Point estimates are displayed in Table 7. The coefficient estimates for California are consistently in the middle of the distribution, with the exception of the estimated black-white gap within California. Of the 18 states, California has the third largest increase in the math achievement gap, behind only New Jersey and Nevada. Taken together, these results are not strongly indicative of a systematic difference between California and the rest of the country. The lack of significance when comparing California to 17 other states indicates that if data from additional states existed, the coefficients for California would not fall in the tails of the distribution of coefficient estimates. As with the College Board and NAEP surveys, the test scores contained in the CNLSY are not stable enough within states across years

His findings were consistent with his non-fixed effects analysis. We were able to replicate this results but they were not significant when accounting for within-state variation in test scores over time.

[29]While broadly similar to Caldwell's summary statistics table, our slight changes to sample restriction – such as dropping 1998 and not including Texas – result in some minor differences.

[30]When further restricting the sample by requiring larger numbers of black and Hispanic students, results are similar.

to detect a policy effect in California.

4.4 Comparing Results from All Three Datasets

A summary of results from our three datasets is displayed in Table 8. There is no consistent pattern across the datasets and in some cases results are starkly different, although none of them fall outside the confidence intervals generated by the permutation tests. For example, for mathematics exams, evidence from the NAEP indicates a decrease of .25 standard deviations for Californians relative to the rest of the country while the NLSY gives a .22 standard deviation increase and the College Board suggests no change in Californian SAT math scores. The disparate estimates could be due to different samples – the College Board data consist of SAT test takers who are near high school graduation, the NAEP is designed to be representative of 8th graders within each sampled state, and the NLSY's sample of Californians is small and non-representative. On the other hand, the difference in point estimates could simply reflect the variability inherent to each of the data sets, since none of the estimates are significantly different than zero when standard errors are adjusted to reflect the presence of a sole policy change state.

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

In an effort to answer these questions, we assembled data from three different sources on standardized tests, and examine how test scores changed in California after Prop 209 led to the end of race-based affirmative action at the University of California in

1998. We demonstrate that once we adjust our standard errors to account for the natural within-state variation in test scores over time, we can say very little about the effect of California's ban on affirmative action on either the racial test score gap or test scores more generally. Given the limitations of our data, we do not have the power to detect even large changes, although point estimates in our favored dataset, consisting of SAT takers, are generally extremely small.

Although we cannot measure how human capital investment changed in response to Prop 209, the effect of banning race-based affirmative action in college admissions on students' behavior prior to college entry is of central importance. As data becomes available from states that have recently banned affirmative action, more precise measurement of policy effects may become possible. On the other hand, given the amount of within-state variation in test scores over time, it may be difficult to ever utilize state-level policy changes to satisfactorily determine the effect of banning race-based affirmative action on human capital accumulation.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

Antonovics, K. and B. Backes, "The Effect of Banning Affrmative Action on College Admissions Rules and Student Quality," *Working Paper*, 2013.

_ **and** _ , "Were Minority Students Discouraged From Applying to University of California Campuses After the Affirmative Action Ban?," *Education Finance and Policy*, 2013, *8* (2), 208–250.

Blume, Grant and Mark Long, "Changes in Levels of Affirmative Action in College Admissions in Response to Statewide Bans and Judicial Rulings," *Working Paper*, 2013.

Buchmueller, Thomas C, John Enrico DiNardo, and Robert G Valletta, "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*, 2011, *3* (4), 25–51.

Caldwell, R., "The Effects of University Affirmative Action Policies on the Human Capital Development of Minority Children: Do Expectations Matter?," *Working Paper*, 2010.

Cortes, Kalena and Andrew Friedson, "Ranking Up by Moving Out: The Effect of the Texas Top 10% Plan on Property Values," *Working Paper*, 2011.

Cullen, Julie Berry, Mark C Long, and Randall Reback, "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*, 2006, *25* (1), 109–119.

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

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

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.

_ , "Pre-College Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions,," *Working Paper*, 2013.

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

Moulton, Brent R, "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit," *The Review of Economics and Statistics*, 1990, *72* (2), 334–38.

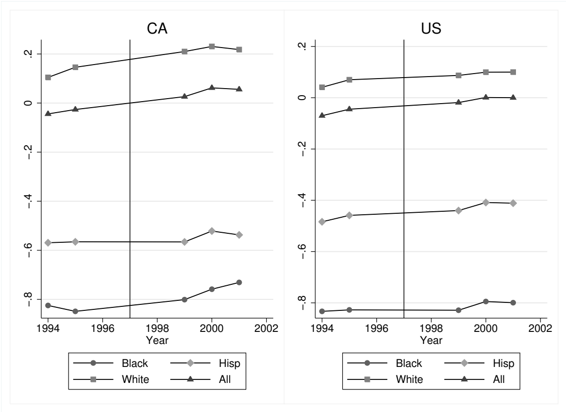
Ray, D. and R. Sethi, "A Remark on Color-Blind Affirmative Action," *Journal of Public Economic Theory*, 2010, *12* (3), 399–406.

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

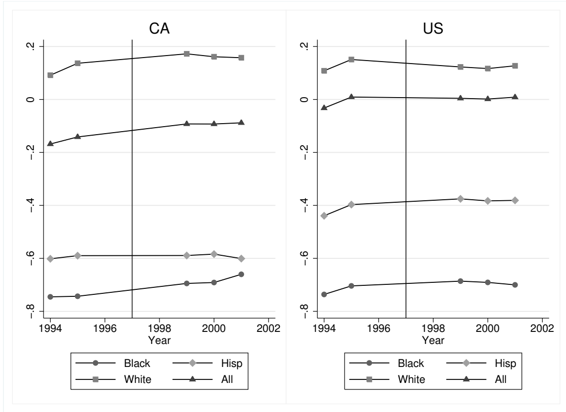
A Additional Data

Perhaps the best measures of human capital investment come from time-use diaries, which record the actual time invested in human capital accumulation. The American Time Use Survey (ATUS), however, did not begin until 2003, which is well after Prop 209 went into effect. In principal, one could use the ATUS to explore time use patterns in states which banned affirmative action after 2003 (for example, Michigan ended racial preferences in university admissions in 2004), but none of the states that banned affirmative action after 2003 have large enough minority populations to make such a study reliable.

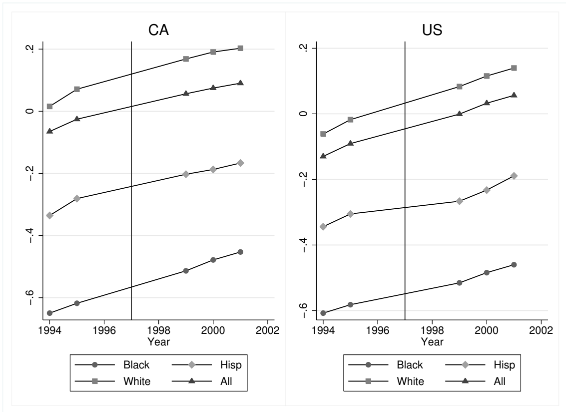
The Child Development Supplement of the Panel Study of Income Dynamics (PSID) also collected time use diaries of youths in 1997 and 2002/2003. However, it has the same problems as the CNLSY: it has relatively small sample sizes, and it is not designed to be representative at the state level. In addition, time diaries are only available in 1997 and 2002/2003, both of which are after the passage of Prop 209.



(a) SAT Math Scores

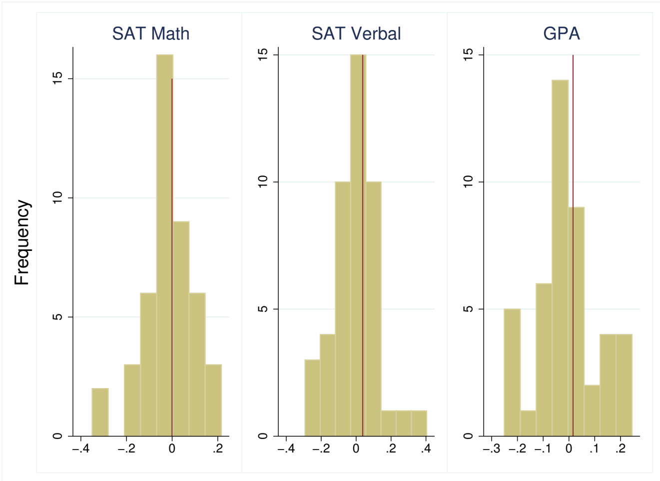


(b) SAT Verbal Scores

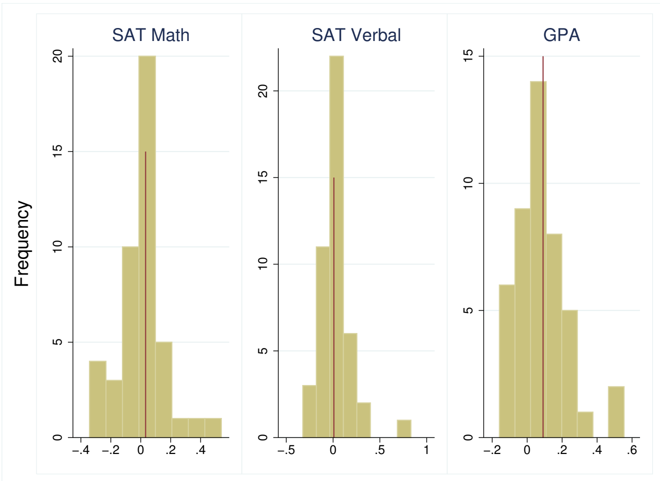


(c) High School GPA

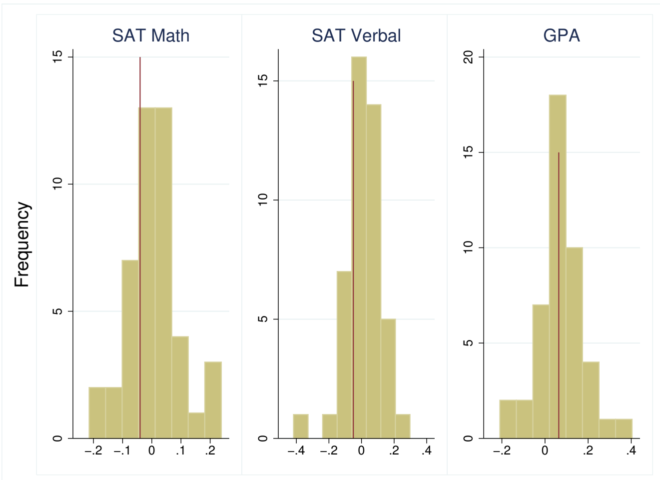
Figure 1: College Board outcome variables. Each panel shows the mean of the normalized outcome (SAT scores or GPA) by race and year. The US panel excludes California, Louisiana, Mississippi, Texas, and Washington.



(a) CA*Post



(b) CA*Post*Black



(c) CA*Post*Hispanic

Figure 2: Distribution of coefficient estimates for each state for each College Board outcome, with California represented by red line. See text.

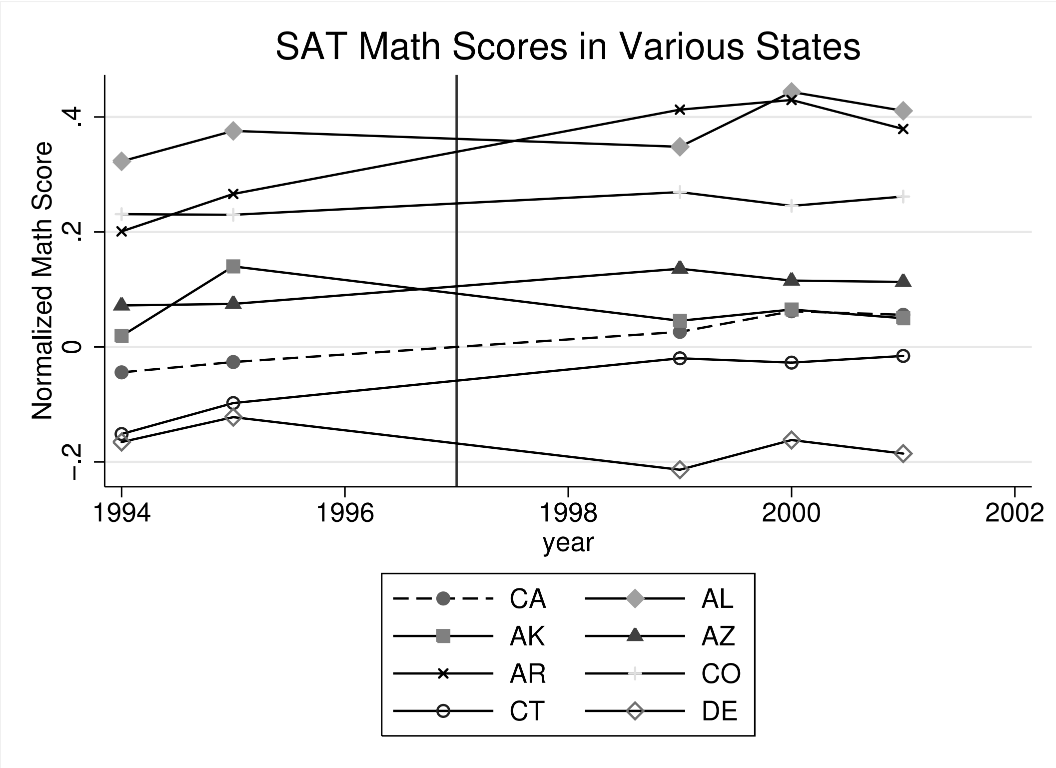


Figure 3: Variability of within-state year-by-year SAT math scores

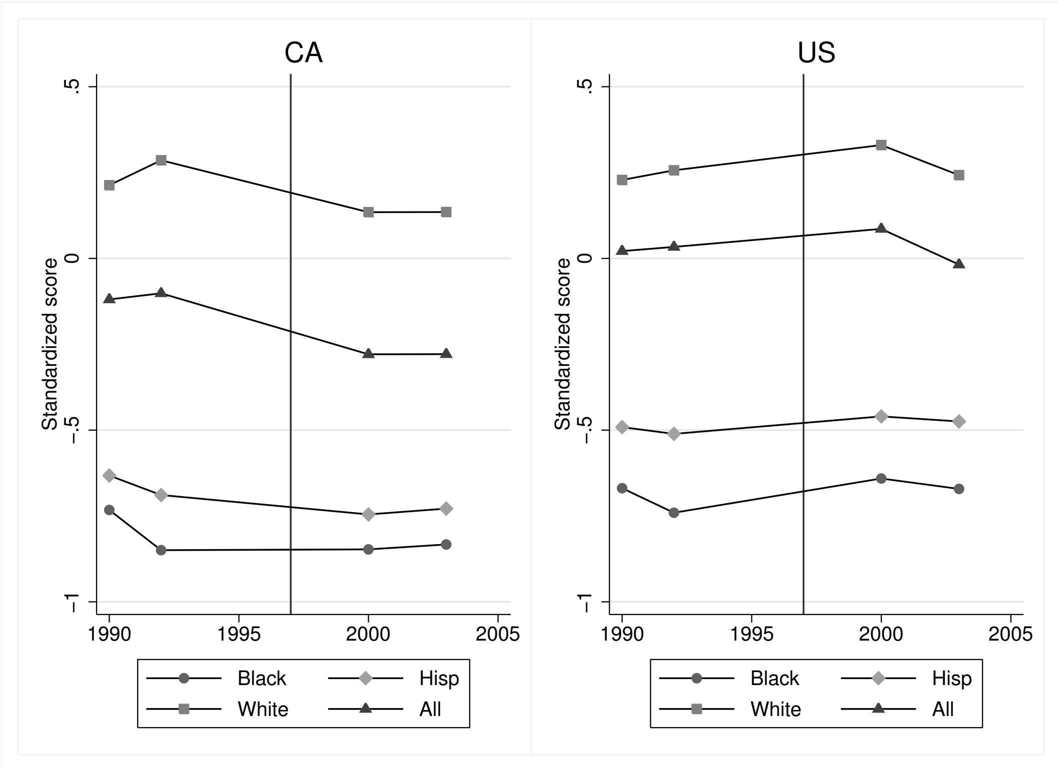
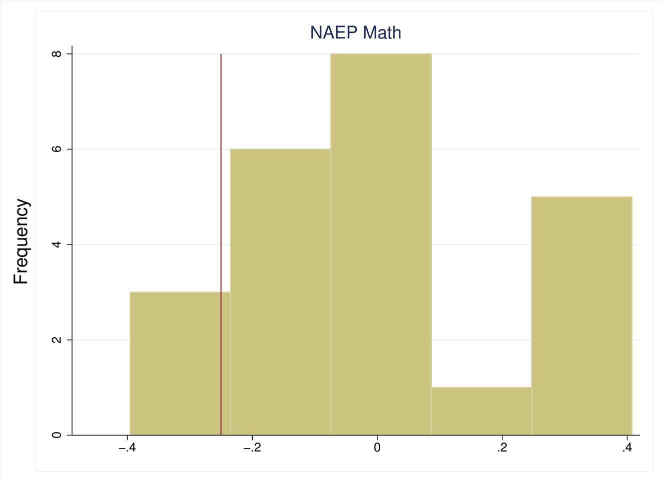
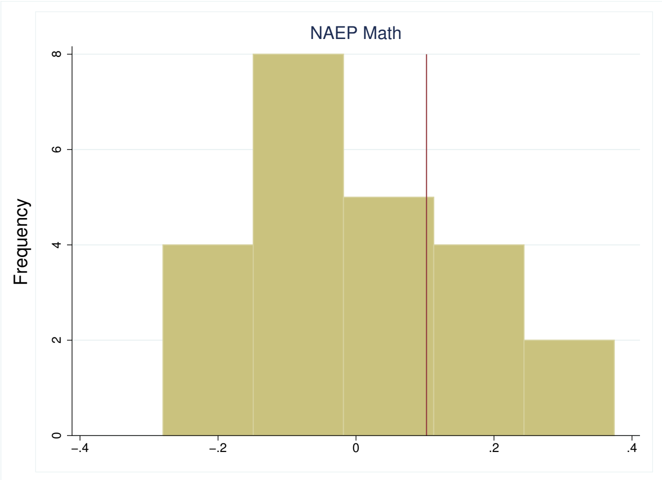


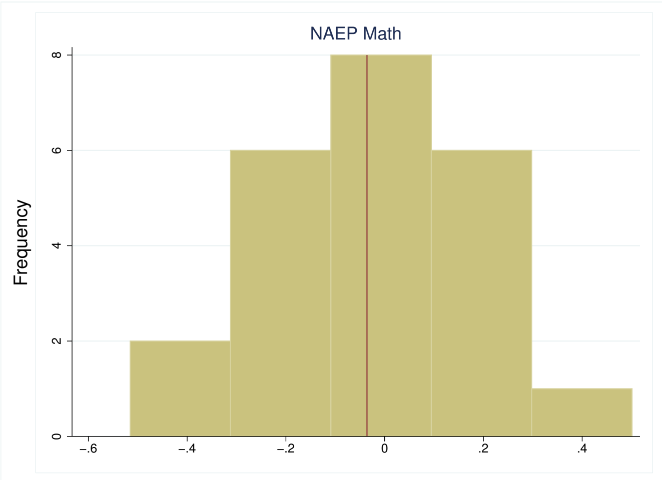
Figure 4: NAEP Math Grade 8 scores. Shows the mean of normalized NAEP math test scores by race and year. The US panel excludes California, Louisiana, Mississippi, Texas, and Washington.



(a) CA*Post

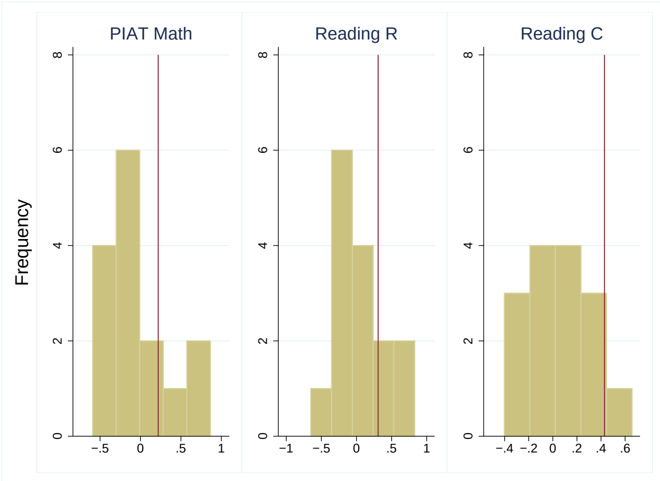


(b) CA*Post*Black

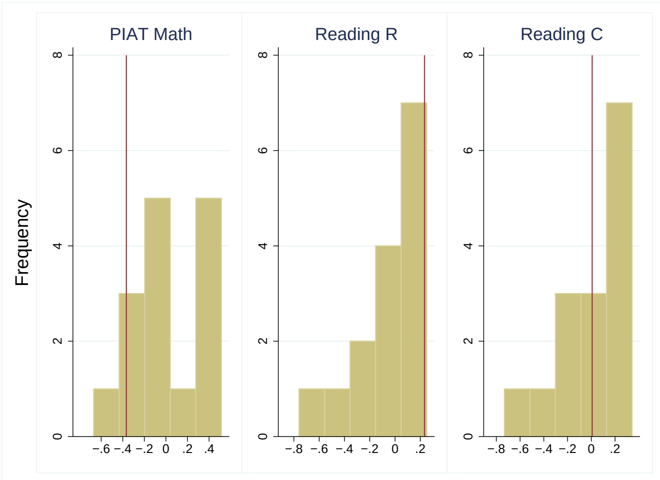


(c) CA*Post*Hispanic

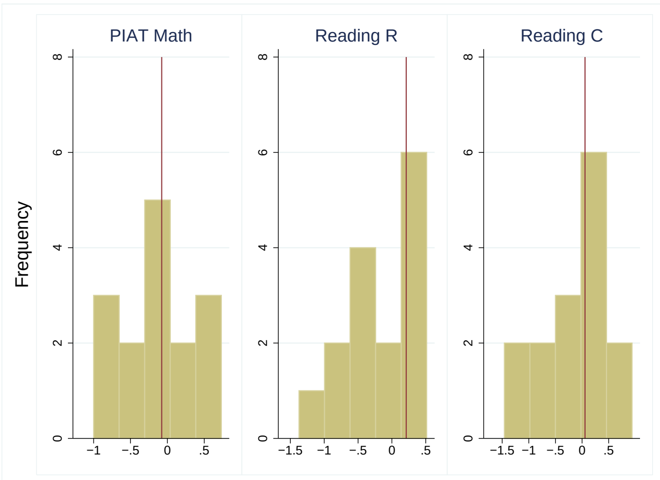
Figure 5: Distribution of coefficient estimates for each state for NAEP grade 8 mathematics exam, with California represented by red line. See text.



(a) CA*Post



(b) CA*Post*Black



(c) CA*Post*Hispanic

Figure 6: Distribution of coefficient estimates for each state for PIAT (from NLSY), with California represented by red line. See text.

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
Jan - March 1996	1996 NAEP sampled
July 1996	First mention of Prop 209 in media
November 1996	Prop 209 passed by voters
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			
	Black	Hispanic	White	Asian	Black	Hispanic	White	Asian
SAT	860	897	1064	1033	859	935	1047	1059
	(188)	(188)	(188)	(222)	(179)	(196)	(190)	(226)
SAT Math	426	452	535	548	424	466	523	558
	(101)	(101)	(103)	(120)	(96)	(105)	(104)	(120)
SAT Verbal	433	444	529	485	435	469	524	501
	(103)	(103)	(103)	(123)	(99)	(107)	(101)	(128)
Male	.43	.42	.46	.47	.41	.44	.46	.48
Income ($10k)	4	3.6	6.5	4.5	3.7	4	6	5.1
	(2.9)	(2.7)	(3.5)	(3.4)	(2.7)	(3.1)	(3.2)	(3.5)
Mom educ (years)	14	11	15	14	14	13	15	14
	(3.1)	(5.3)	(2.9)	(4.5)	(3.1)	(4.2)	(3)	(4.3)
Dad educ (years)	14	11	16	14	14	13	15	15
	(3.3)	(5.5)	(3.1)	(4.4)	(3.4)	(4.6)	(3.2)	(4.3)
Goal Bachelors	.85	.79	.79	.81	.82	.79	.76	.8
Degree goal unknown	.12	.18	.19	.17	.13	.17	.21	.17
Observations	60906	184937	420059	194819	568254	257419	943277	70022

Notes: Standard deviations in parentheses. Rest of country excludes California, Louisiana, Mississippi, Texas, and Washington. Sample includes 1994-1995 (pre- period) and 1999-2000 (post- period).

Table 3: College Board SAT and GPA DDD Estimates

	SAT M	SAT V	GPA
CA*Post	-0.00	0.04	0.02
Uncorrected s.e.	(0.01)	(0.01)	(0.01)
Permutation c.i.	(-.29, .18)	(-.26, .25)	(-.24, .21)
CA*Post*Black	0.03	0.01	0.09
Uncorrected s.e.	(0.01)	(0.01)	(0.02)
Permutation c.i.	(-.27, .35)	(-.2, .27)	(-.11, .54)
CA*Post*Hispanic	-0.04	-0.05	0.06
Uncorrected s.e.	(0.00)	(0.01)	(0.03)
Permutation c.i.	(-.2, .19)	(-.18, .16)	(-.16, .28)
Observations	1614959	1614959	1681150
R-squared	0.24	0.21	0.12

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 gender, family income, and parental education are also included in the regressions. Confidence intervals generated by running permutation tests with each of the 47 states (see text). All outcome variables normalized to have mean zero and standard deviation one. Years include 1994-1995 (pre- period) and 1999-2001 (post- period). Included time trends are linear state-specific time trends. The excluded racial group includes white, other, and unknown.
Uncorrected standard errors consist of standard errors clustered at the state level. Permutation confidence intervals generated by estimating iterating over each of the states in the dataset (see text).

Table 4: NAEP Grade 8 Mathematics Summary Statistics

	CA				US			
	Black	Hispanic	White	Asian	Black	Hispanic	White	Asian
NAEP math score	-0.88	-0.70	0.22	0.38	-0.75	-0.49	0.28	0.32
	(0.89)	(0.94)	(0.90)	(1.03)	(0.88)	(0.92)	(0.87)	(1.07)
Male	0.51	0.52	0.50	0.47	0.53	0.51	0.50	0.50
Parental Educ (yrs)	14	13	15	15	14	13	14	15
	(1.8)	(2.3)	(1.7)	(1.9)	(2.0)	(2.3)	(2.0)	(2.0)
Observations	966	3303	4179	1095	33012	11648	141137	8659

Notes: Rest of country excludes California, Louisiana, Mississippi, Texas, and Washington. Sample includes 1990, 1992 (pre- period), 2000, and 2003 (post-period).

Table 5: NAEP Grade 8 Mathematics DDD Estimates

	Math
CA*Post	-0.25
Uncorrected s.e.	(0.05)
Permutation c.i.	(-.34, .34)
CA*Post*Black	0.10
Uncorrected s.e.	(0.05)
Permutation c.i.	(-.27, .27)
CA*Post*Hispanic	-0.04
Uncorrected s.e.	(0.04)
Permutation c.i.	(-.5, .29)
Observations	203999
R-squared	0.29

Each column shows the coefficient estimates from a regression of the normalized 8th grade NAEP math score on an indictor for whether the student was from California, an indictor for whether they took the NAEP 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 gender and parental education are also included in the regressions. Years include 1990, 1992 (pre-period), 2000, and 2003 (post- period). Included time trends are linear state-specific time trends. White students comprise the excluded racial group.
Uncorrected standard errors consist of standard errors clustered at the state level. Permutation confidence intervals generated by estimating iterating over each of the states in the dataset (see text).

Table 6: CNLSY Summary Statistics

	California			Rest of Country		
	Black	Hispanic	White	Black	Hispanic	White
Math	97	97	106	97	98	105
	(13)	(13)	(15)	(13)	(14)	(13)
Reading Recognition	99	101	107	100	103	108
	(17)	(15)	(16)	(14)	(15)	(14)
Reading Comprehension	95	98	103	97	99	105
	(14)	(14)	(14)	(13)	(14)	(13)
Birth Order	2.0	2.1	1.8	1.9	1.9	1.7
	(1.2)	(1.1)	(1.1)	(1.1)	(1.0)	(0.9)
Family Size	4.6	5.0	4.5	4.5	4.4	4.4
	(1.55)	(1.50)	(1.67)	(1.64)	(1.35)	(1.18)
Mom High Grade	13	11	13	12	11	13
	(1.6)	(3.1)	(2.4)	(1.9)	(2.7)	(2.3)
Most recent mom AFQT	22	24	49	21	23	51
	(20)	(23)	(26)	(18)	(22)	(26)
Mom age at birth	22.5	23.8	24.1	22.9	23.2	24.9
	(4.3)	(4.9)	(4.5)	(4.7)	(4.5)	(4.8)
Female	0.49	0.47	0.49	0.50	0.50	0.50
Two parent HH	0.36	0.66	0.70	0.33	0.48	0.79
Urban	0.95	0.92	0.80	0.87	0.94	0.65
Grade	4.6	4.7	4.7	4.6	4.7	4.4
	(2.2)	(2.3)	(2.3)	(2.2)	(2.2)	(2.3)
Times Taken	2.9	2.9	3.0	3.0	2.8	3.0
	(1.0)	(1.0)	(1.1)	(1.1)	(1.1)	(1.1)
Observations	249	821	430	1860	591	2809

Notes: Rest of country excludes California, Louisiana, Mississippi, Texas, and Washington. Sample includes 1988, 1990, 1992, 1994 (pre- period) and 2000 and 2002 (post- period).

Table 7: CNLSY DDD Estimates

	Math	Reading R	Reading C
CA*Post	0.22	0.31	0.43
Uncorrected s.e.	(0.08)	(0.09)	(0.06)
Permutation c.i.	–	–	–
CA*Post*Black	-0.3	0.23	0.01
Uncorrected s.e.	(0.07)	(0.05)	(0.07)
Permutation c.i.	–	–	–
CA*Post*Hispanic	-0.08	0.21	0.06
Uncorrected s.e.	(0.11)	(0.10)	(0.13)
Permutation c.i.	–	–	–
Observations	9301	9240	8175
R-squared	0.23	0.23	0.25

Notes: Standard deviations in parentheses. Rest of country excludes California, Louisiana, Mississippi, Texas, and Washington, and states with less than three black and Hispanic observations in both the pre and post periods. Sample includes 1988, 1990, 1992, 1994, (pre- period) and 2000 and 2002 (post-period).

Uncorrected standard errors consist of standard errors clustered at the state level. Permutation confidence intervals not displayed due to lack of sufficient number of comparison states.

Table 8: Summary of Results Across Datasets

	Math	Verbal	Reading R	Reading C	GPA
CA*Post					
College Board	0	.04			.02
NAEP	-.25				
NLSY	.22		.31	.43	
CA*Post*Black					
College Board	.03	.01			.09
NAEP	.10				
NLSY	-.30		.23	.01	
CA*Post*Hispanic					
College Board	-.04	-.05			.06
NAEP	-.04				
NLSY	-.08		.21	.06	

No coefficient estimates are significant under permutation tests. See text.