# Affirmative Action's Effect on Educational and Wage Outcomes for Underrepresented Minorities

Vishnu G. Arul

ECON H195B – Senior Honors Thesis

Written under: Professor Sydney Caldwell

UNIVERSITY OF CALIFORNIA, BERKELEY

April 29, 2022

**Abstract**

Affirmative action has been a hot-button political issue for decades in the United States since it was first implemented into law via Presidential Executive Order. The policy intends to rectify past injustices done unto certain minority communities by taking their race into context when deciding university admissions. Educational opportunity is not as meritocratic as people would like to believe as the quality of k-12 schools, the potential for standardized testing tutoring, as well as home conditions are largely determined by familial wealth, born out of past racial injustices. This study finds that race based affirmative action bans decreased the quality of universities attended by underrepresented minority students, as well as decreased the proportion of underrepresented minority college-aged people that held college degrees. The study does not directly find statistically significant results on effects on wages, however, it estimates based on prior studies.

# 1 Introduction

Systematic racism is a persistent issue in the United States. In the 1960s, one way that the United States Federal government attempted to address this issue was through educational institutions via Executive Order 11246, later amended to Executive Order 11375 in 1965. These orders required federal agencies to take "affirmative action" in ensuring not only that members of protected groups (such as female, black, non-white Hispanic), would not be discriminated against in the hiring process, but also that those groups had increased employment for federal contracts and government agencies. Controversy immediately emerged surrounding the passage of these executive orders. Some politicians claimed these laws were overly intrusive measures taken by the federal government, and others even claimed that these laws were at odds with the 14$^{th}$ Amendment of the United States Constitution which provided equal protections to all regardless of race or sex ("History of Affirmative Action").

Challenges to this executive order manifested in many ways over the years, and this paper will focus on the state challenges to the Executive Order. Since the 1990s, eight states have changed their affirmative action policies pertaining to undergraduate university admissions. Through general ballot measures, voters have banned affirmative action in California (1996), Washington (1998), Florida (1999), Michigan (2006), Nebraska (2008), Arizona (2010), and Oklahoma (2012). In Texas (1996) and New Hampshire (2012), affirmative action was banned through the courts and state legislature. However, in 2003, Texas's statewide ban on affirmative action was reversed via the Supreme Court case Grutter v Bollinger (2003) ("History of Affirmative Action"). This paper recognizes these multiple different policy adoption dates and will draw its conclusions by collating all effects from each policy change.

It is challenging to estimate the impact of affirmative action on outcomes due to issues surrounding marginal student identification, changing application behaviors, differing effects on different university admission processes, and unique intrastate conditions.

In university admissions, not all students are equally affected by affirmative action in terms of final admission result, regardless of race or university in question. It is unlikely in a given year that affirmative action would end up affecting the result of a candidate whose qualifications rank them in the 90th percentile of applicants if the policy was banned or kept in place. What this means is that any measurement of changes in outcomes due to affirmative action will be a measurement of how these policies affected the marginal student rather than analyzing how it affected the entire student body. The marginally admitted student, could be a 10th percentile admit if the ban is put in place, or not be admitted at all if the policy is kept, and vice versa. Their effect on the student body, and subsequently outcomes is what this and other studies attempt to measure, but their affect on outcomes could be muted depending on how drastically the policy change affects the student body composition from the number of marginal candidates affected by the policy. Card and Kruger's paper in the literature review explores this issue a bit (Card et, al., 2004).

How applicants' behavior in applying to universities change in response to policy changes is another noisy factor when estimating the effect that affirmative action has on outcomes. Applicants are operating with not as much information as universities as to where they rank in the applicant pool for a specific university. Despite this, many schools require a fee and unique essays in order to be eligible for admission. Thus, if a student believes that they are a marginal candidate for a university, they may change their decision to apply or not based on whether an affirmative action policy is in place or not. In the case of a ban, this could lead to URM numbers being depressed not necessarily because of the actual effects that the policy had on admission decisions but because

URM applicants chose to apply less due to discouragement from the ban. Likewise a policy implementation could see higher numbers of URM at universities despite the policy having a limited effect on real admission decisions, solely due to URM applicants being encouraged to apply (Card, et al,. 2004).

Different schools would be affected by affirmative action on different magnitudes. We would expect that less selective schools to be affected less by affirmative action in general due to high admission rates. More selective universities at times experience a phenomena where the marginal accepted and marginal rejected candidate are nearly indistinguishable from each other barring a few superfluous qualities. This is where affirmative action's influence on outcomes could be seen at a higher magnitude, whereas at a non-selective university, the effect would not be seen as much, building off the logic from the marginal candidate point earlier (Smith, 2020).

The final issue in estimating the impact of affirmative action on outcomes is accounting for unique intrastate conditions. California's UC and CSU systems are able to demonstrate both scenarios for a state's public college system. Both UCs and CSUs are geographically diverse, with campuses across the state, in the valley and coast, and in Northern, Southern, and Central California. However, the enrollment and application behavior by applicants varies between systems. The UC system follows more of a tier-list type system, where many students aim to attend the highest ranked school they can get into, usually UC Berkeley or UCLA, regardless of their original geographic origin. These schools are the most selective of the UCs due to having better outcome metrics than the rest in terms of salary, retention/graduation rates, and graduate school prospects. The CSUs also vary in selectivity and outcomes, but many will opt to apply and attend a specific CSU based on geographic qualities, such as being near their hometown, regardless of differences in quality that the CSU has to other CSUs. Different states see different phenomena within their borders occurring regardless of affirmative action's presence, and thus it is difficult to determine a nationwide trend when different states' conditions are relatively isolated from each other based on their college admission metagame.

This paper will estimate the effects that affirmative action bans had on wages, quality of university attended, and college degree attainment for underrepresented minorities. Underrepresented minorities as defined by this paper refers to self-identified black and non-white Hispanic people. I will estimate these effects by using staggered adoption and synthetic controls across the time period in which several states had banned affirmative action within their public colleges. States will be weighed specifically in order to induce parallel trends in the outcome variable prior to the treatment period, so that in the post-treatment period, all effects can be attributed to the policy change.

When directly estimating effect that affirmative action policies had on wages, this paper can not determine a statistically significant point estimate due to the 95% confidence interval containing negative and positive estimates. This paper does find, however, a modest negative relationship between affirmative action bans and attainment of a bachelor's degree, as well as a negative relationship between affirmative action bans and quality of university attended by underrepresented minorities. Using relationships found in other papers regarding the wage premium for attending college, this paper estimates that the economic loss for underrepresented minorities is an average of $600.

The rest of the paper will be outlined here. Section II will focus on existing literature regarding this topic, how they got to their conclusions, and if they are at odds with the hypothesis. Section III focuses on the existing data that was used for this study, including race, educational levels, wage levels, and metrics for university quality. Section IV will describe the empirical strategy this study undertakes in order to measure the true effect of affirmative action policies on wages, quality of university attendance, and educational attainment. Section V will focus on the numerical results and robustness checks for the study. Section VI will conclude the paper and contextualize the results in terms of the research question, and real-world implications.

# 2 Literature Review

Arcidiacono, Aucejo, Coate, and Hotz used data from the University of California Office of the President (UCOP) to determine the effects of Proposition 209, the proposition that ended AA in California (Arcidiacono, et al., 2012). Their analysis was conducted to test the "mismatch hypothesis" which followed the logic that AA policies created a "mismatch" between students and universities creating unintended outcomes such as lower graduation rates, and lower academic preparation for students admitted under AA. They analyzed student-level data and aggregated it based on underrepresented minority (URM) status as well as family income and found that there was a decrease in enrollment for URM from higher-income families by 10%, and an increase in enrollment for lower-income URM by 2.5%. This paper only focuses on California and does not compare to national trends, something my paper intends to rectify by broadening the scope to include universities from several states.

Lincove and Cortes use data from the University of Texas admissions to determine the "undermatching" or "overmatching" of students given the unique Texas Top 10% Plan (Lincove et al., 2016). This plan automatically admits the top 10% of a public Texas high school's graduating class irrespective of other academic qualifications. They found that race-blind automatic admission policies favored lower-income students both in and outside of the top 10% of their graduating class. They reasoned that this was because of asymmetric information in holistic admissions that causes many high-achieving low-income students to match themselves at schools with admissions criteria lower than their academic achievements, while higher-income students tend to match themselves at schools with admissions criteria higher than their academic portfolio suggests. While this paper could indicate that race-neutral policies favor lower-income students, the conclusion is drawn from primarily exam-based admission criteria, like the Ellison and Pathak paper, rather than holistic admissions, something my paper will be focusing on (Ellison et al., 2012).

Bleemer uses a differences-in-differences design to see what effect California's ban on race-based affirmative action via Prop 209 had on general economic outcomes for graduates out of the university system from California after the ban. He finds that the "mismatch hypothesis" fails to explain the phenomena of cascading URM into lower quality universities as at these universities, degree attainment declined, and wages declined for their 20s and 30s especially among Hispanic students. One important note is Bleemer's belief that there was a deterrence effect as a result of the policy that affected application decision by applicants, independent of real effects of the policy itself (Bleemer, 2021).

Card and Kruger's 2004 study focuses on the application behaviors of highly qualified minority students, as denoted by their SAT/ACT scores. Their study was contained to Texas and California, whose bans were in 1996 and 1998 for affirmative action. They wanted to see if these applicants, whose ultimate admission result would likely not be affected by the ban on affirmative action, would change their application behabior in response to a perceived disadvantage, or an expectation of a less diverse campus. They, however, found no such changes in the application decisions made by highly qualified minority applicants in response to the policy bans (Card, et al., 2004).

In summary, my paper differs from all of the others because it will use synthetic controls with staggered policy adoption in an attempt to measure aggregate effects around the country from

2001 – present day. This will uniquely allow me to determine the effects of the policy on a nationwide scale, not being contained to a specific state. Texas's and California's public university systems specifically have a lot of unique mechanics surrounding admissions including a multi-tiered university system and automatic admission given a graduating class rank threshold. The Arcidiacono, and Lincove papers both obtain their data on a state or city level, and thus have issues when trying to extrapolate a nation-wide conclusion about the effects of affirmative action (Arcidiacono, et al., 2012; Ellison et al., 2012; Lincove et al., 2016). Bleemer as well as the Card and Kruger paper are the most similar to my study as they use differences in differences to determine the policy effects, however, they are contained to California and Texas (Bleemer, 2021; Card, et al., 2004). Thus, my research will allow for more informed discussions about race-based affirmative action, as voters and representatives will have a better idea as to who benefits and who is harmed in the policy

# 3 Data

I largely pulled from IPUMS datasets taken from the American Community Survey from 2001 – 2019. These data sets contained information regarding wages, education level, racial identification, all by state. While IPUMS does not have a specific denotation for underrepresented minority, I was able to construct that myself by creating a binary using black data points as well as non-white Hispanic and Native American data points. This was all aggregated into a URM status column. All data is on a nationwide basis with only a few states taken out. Another filter that was applied was to only analyze the age group of 22 – 28 as this age group would be the most likely candidates to have their earnings most affected by college attendance or not. Those who are older could have their wages influenced by other means, and those younger would largely not have had the opportunity to attend and graduate college. With the augSynth package in R, states that have been 'treated' with the policy must have several data points where they do not have the treatment in order for the regression to be run (Ben-Michael 2020). Thus, all states with limited pre-treatment data points as well as those without any pre-treatment data points were dropped from the data entirely. These states are California, Texas, Washington, and Florida. In this regression, states are only considered 'treated' 4 years after the policy change as to recognize the fact that the graduating class 4 years post-treatment would be the first ones to be affected by the policy change. With this data I was able to construct the data tables for estimating the effects that these policies had on wages as well as educational attainment.

For estimating the effect that these policies had on wages, I constructed a data set that denoted the median wage by race by state for every year. I then broke this down by education level as well as having a table that aggregated across education levels. The regression used in the study only estimates using college graduates in the state. This is done in order to have another way of determining the effect that the policy change has on quality of university and thus wage outcomes. Theoretically, declining wages among URM aged 21 – 28 with college degrees in this time period in the post-period would indicate that the policy had a negative effect on the quality of university and education that the graduates in questioned received, and as a result, their wages declined. Education levels were set on a binary with people either having or not having at least a Bachelor's education. This was the threshold because that was the most important for the study rather than post-graduate education or some college. For educational attainment, the data set was created using the same IPUMS base data, with educational attainment being a binary of whether the person had obtained a Bachelor's education or not. Educational attainment as an outcome variable was measured on an aggregate level by race by state by year. The measure was the proportion of those in the group that had obtained the degree represented by a percent.

Neither of these datasets are free of noise as both are susceptible to a few issues. Wages specifically are a noisy outcome variable because of all of the inputs that can factor into changing the outcome outside of obtaining a degree from a specific college. Furthermore, the "by state" data simply records the people currently residing in the state, who we assume the majority of attended universities in the state or are from the state. This creates noise as we do not know how many people in the state with or without a degree were indeed affected by the policy change inside of that particular state. Without a state of origin dataset, this is the best estimations that we can make. Summary statistics for median wages and college attainment can be found in table 1.

To determine university quality, several data points for universities were taken including median SAT/ACT scores for admitted students and graduation rates. Card and Kruger both used SAT/ACT scores as a measure of high-quality applicants, so I find it a good measure of a school's quality independent of metrics that media uses (Card, et al., 2004). In the formula for quality score, the greater of relative SAT or ACT scores was used in calculation because not all schools will have data on both, as some will only accept one test, or some regions of the country only have one of the tests be popularly administered. Graduation rate is the best available outcome measure that can be used as a metric of university quality. Retention rates are also valuable due to indicating the university's ability to keep its students year over year. Wages and graduate school admissions would be good outcomes to factor in university quality, however, those data points are unavailable on the aggregate level. Using this data, I constructed a "quality score" for each eligible college in a state. The formula for calculating the score is as follows:

$$0.4 \times \text{retention rate} + 0.4 \times \text{graduation rate} + 0.3 \times (\max(\text{ACT Median}/36, \text{SAT Median} / 1600))$$

To indicate this ranking system's soundness, I have attached a sample of the table in Tables 2 and 3, in which some of the higher ranked schools are listed. These include some of the top public universities in the country including University of Virginia, William & Mary, UC Berkeley, and UCLA. Private schools were not included in this because they are not subject to any changes made at the state level for affirmative action policy. They act independently and are free to implement the policy or not regardless of what the state wants to do for public colleges.

The ranking was fixed for 2006, as that is the first policy change recorded in the study. Allowing the ranking to change over time as these factors change would add more noise to the study as it is entirely likely that the policy change could have had an effect on these rankings, thus it is fixed. The outcome measure here over time is how the racial demographics change in these universities, specifically, the proportion of URMs at these universities, and then breaking it down further into African American and non-white Hispanic students. Thus, the figure generated is a weighted average of university quality for URM in a given state for a given year. How this figure changes over time will be how the point estimate for that regression is generated.

# 4 Empirical Strategy

I estimate the effect of affirmative action policies on university quality, educational attainment and median wages for underrepresented minority students using a staggered events study with synthetic controls. The key identifying assumption is that, in the absence of treatment, the outcomes for the treatment group and the synthetic control group would have evolved similarly after the policy was implemented. We assess this assumption by examining whether the outcomes for the two groups exhibit parallel trends prior to the policy implementation. This is because this regression method can mimic an experiment by comparing the effect of a policy on a treatment versus control group.-In order to establish this, parallel trends do not need to hold for levels, but rather growth rates. Parallel trends does not require the levels in outcomes to be the same; it requires that changes over time in the pre-treatment period, both growth and decline, are the same between the control and treatment groups.

For this study, instead of pairing states with other states that have similar parallel trends on a one on one comparison, I will be using synthetic controls. Non-treated states should not simply be used as controls for two reasons. The first is that there will likely not be parallel trends, or if there are, they will be imperfect comparisons to the treated states. Without this key assumption holding, the regression cannot generate a reasonable point estimate. Secondly, this partially addresses the concern about unique intrastate qualities within each state. By pooling together the outcome variables of all states, each state's unique admission metagame is averaged out in the grand scheme of the regression. Synthetic controls are where using a nationwide dataset, all control states are weighed specifically to match the treatment states more closely in the pre-treatment period to establish parallel trends. This allows the differences found in the post-treatment period to be the estimated effect that the policy change had on the outcome. Each of the regressions has a different outcome variable but similar methodologies.

Beyond a simple synthetic controls, I will be running a synthetic controls regression with differences in differences with staggered policy adoption. This is because each of the states that banned affirmative action did so independently of each other largely, and had different implementation dates. This is done using the augSynth package in R developed by Eli Ben-Michael (Ben-Michael, 2020). This will give a regression coefficient and standard error for every year pre and post-treatment. The regression coefficients presented in the results section will utilize the average regression coefficient in the post-treatment periods. It is important to have one synthetic control group for this regression. If not, the alternative scenaior would be to have different control groups and estimate different regressions for each treated state. There is no apparent reason why different states would have different effects from affirmative action bans, thus I can increase the power of the regression by pooling all of the states together.

The datasets prepared for the regressions were constructed similarly for the wages and educational attainment regressions. Each of these regressions had nationwide data and had outcomes for wages and educational attainment by state by race by year. Wages are represented in a nominal figure and are not adjusted for inflation. Educational attainment is represented in a % for all those who obtained at least a bachelor's degree. Weighted average university quality attended by URM students is represented by a figure from 0-1.

The regression equation used for the effect on wages in this study is as follows:

$$Y_{it} = \beta_0 + \sum_{t=2001} \gamma_t + \sum_{\mathrm{i}} \gamma_{\mathrm{i}} + \beta_{DiD}(AA_{it}) + \mathrm{X}_{it} + \varepsilon_{it}$$

The two subscripts are 'i' and 't'. 'i' is a denotation of each state in the US. 't' refers to each of the times from 2001 – 2019 which the data could be derived from. This study runs 3 regressions, and each have specific control and outcome variables. The other inputs for the regression are the same. $\beta_{DiD}$ measures the effect that the affirmative action policy change had overall by considering binary variables $\gamma_t$ and $\gamma_{\mathrm{i}}$. $\gamma_t$ is a binary variable for a given state depending on whether it is the post or pre-period. In the post period, it is represented with a 1, and in the pre-period it is represented with a 0. This figure is only relevant for states that received treatment at all, being states that had a policy change in the time period and are included in the study. $\gamma_{\mathrm{i}}$ is a binary variable that is either a 1 or a 0 depending on whether the state received treatment at all during the time period of 2001 – 2019.

$\mathrm{X}_{it}$ represents the different control variables run as part of the study. For the regression estimating median wages, this represents the general state average wages in the state which could affect these outcome variables independently of changing affirmative action policies. The outcome variable refers to the median income for URM in the state with a college degree, whereas this control intends to fix the effects of a state's general economy at the time. State average wages shifting independently of the policy would indicate independent state conditions that are driving the individual results of the state rather than any change being the result of the affirmative action policy change. $\varepsilon_{it}$ is the error term for the regression to capture anything not in the regression, as in any noise that the regression is unable to take into account. For the regression estimating changes in college degree attainment among URM, $X_{it}$ represents the control for states' high school graduation rates year over year. This figure changing could affect the college degree attainment figure independently of affirmative action policies changing, thus it should be controlled for.

Here $Y_{\mathrm{it}}$ represents the different outcome variables for each regression. There are three regressions that I will be running with identical independent variables, but varying outcome variables.

For the regression estimating the effects of affirmative action on median wages, $Y$ represents median wages for group i (URM or non-URM) in year t. T starts at 2001 because that is the first data point that is available for the study, hence why certain states, whose affirmative action policy changes occurred before 2001, or have limited data points are not included in this study.

For the regression estimating the effects of affirmative action on college degree completion represents the proportion of URM or non-URM status people that have at least a bachelor's degree. Once again this is taken on the state level by year. The other variables representing dummy variables for time and treatment still hold with the same mechanics. Similarly, to the regression about median wages, this regression's time for treatment does not start on the year in which the policy changed occurred, rather this metric refers to graduates, thus, the years affected by treatment would only start 4 years after the policy was implemented to ensure that the estimation is capturing the effects on a student body whose composition was affected by the policy change. Examples of this will be given later in this section.

For the regression estimating the effects of affirmative action on quality of university attended, $Y$ represents the weighted average university quality score attended by URM. This data is also done on a by state by year basis. This data does not have any control variables like state average wages or high school graduation rates, as these factors should not affect this regression to the same degree as the other regression estimations. Unlike the other two regressions, this will not take into account the 4-year delay that occurs in the full change of a student body. The outcome variable here directly targets the student body of the present, and effects should be seen immediately, not necessarily after a delay, although they will be muted initially.

For the purposes of the study, several caveats go beyond looking at the date of the policy change seen in figure (figure with the timeline table). First off, as previously stated, treatment is only considered given 4 years after the policy change to take into account the first graduating class on the job market for universities from that state. This means for a state like Michigan, whose affirmative action ban occurred for 2006 admissions year, their treatment date would really be considered for 2010 because that would be the first graduating class whose majority admission occurred after the policy change. Secondly, because of limitations of the augSynth package, several states had to be dropped because of limited pre-treatment data (Ben-Michael, 2020). While states like Washington and Florida have data in the pre-period for the study due to the 4 year delay in the policy taking effect for the regressions estimating effects on wages and college degree attainment, there is not enough pre-treatment data for the states' data to be taken account properly in the regression. For that reason, Florida and Washington's data are dropped from consideration in the regression results. California's data is dropped for not having any pre-treatment data in the dataset. Texas's data is dropped because while they did ban affirmative action from their admission process in at a valid point to have sufficient pre-treatment data, the state reversed its ban over the course of the study as well, but it was implemented on a voluntary basis. This means that while states like Texas A&M and UT-Dallas continued to not consider race in admissions, their state flagship, UT-Austin, reimplemented affirmative action in admissions. The difficulty in isolating universities from one another within the state of Texas is why their data will be dropped from consideration for the regression.

A final note to make is that in these cases, the treatment policy is not an implementation of affirmative action from a pre-treatment case of not implementing the policy. Rather, all of these states began the pre-treatment period by having affirmative action due to Executive Order 11375. As a result, any regression results should be read in the context of resulting from the ban, which was the real treatment. Thus, a potential negative regression coefficient effect for wages should be attributed to the ban and would demonstrate that affirmative action had a positive effect on wages, which is why they would decline in the post-treatment period. As the post-treatment period is the period that no longer has affirmative action.

# 5 Results and Discussion

The results show affirmative action bans were associated with a $2761 decrease in median wages for URM with a college degree aged 21-28 in the years 2001 – 2019 (Table 4; Median Wages Column). This figure was noisily estimated as the regression had high standard errors and, in several years of interest in the study, exceeded the regression coefficient indicating that there is a lot of 'noise' driving wages differences between treatment and non-treatment states. While the exact magnitude is difficult to determine because of noise, this point estimate is consistent with Bleemer's conclusion regarding the decline of URM earners over $100,000 in California after their affirmative action ban.

The results show bans on affirmative action were associated with a -2.3% decrease in college degree attainment for URM and non-URM people aged 21 – 28 in the United States during the period 2001 – 2019. These results were significant at the 5% level (Table 4; Educational Attainment Column).

The results show bans on affirmative action were associated with a -0.113 unit decrease in university quality on average for URM attending universities from 2001 – 2020 (Table 4; University Quality Column). This is significant at the 1% level. For context, using this scale, that would equate to a downgrade of quality of university from UC Berkeley to California State University, San Luis Obispo (Tables 2 & 3).

There are some potential issues in interpreting these results. The largest issue iresults the large standard errors in the first two regressions. The large standard errors should not be interpreted that no relationship could exist between the outcome and explanatory variables, rather, it means that in the scope of the data used in the study, a potential relationship was noisily estimated, and while the magnitude is difficult to ascertain, the general direction can be more easily interpreted. The lack of statistically significant results are more important to interpret on its own in for the regression regarding educational attainment rather than wages. A smaller issue is the omission of various states due to a lack of data, or confusing policy changes. Florida, Washington, California, and Texas each suffer at least one of these issues. These diverse, populated states, would have provided an ample amount of data, and their exclusion does call into question the applicability of these results, as the data represents the United States far less without the 2 most populated states in the country having their data represented at all in the regressions.

As stated before, the wages regression would be extremely noisy given the nature of the data. There are already several large factors and events that would have potentially affected states in different ways over the course of this time period that are unrelated to affirmative action policies. The Great Recession, for example, had uneven effects on states causing many to have differing economic futures after the financial crisis (Tosci, et al., 2010). The IPUMS data set that was used in the regression only had data for the current residents at the time. While the age range used was intentionally chosen in order to attempt to target those in the state who would be affected by the admissions policy change, there is still no guarantee that the current residents in the state would be affected the policy. While young college graduates in a state could indeed be from the state, it is also entirely likely that any given person could have been educated from a different state and simply currently resides and works in their current state, rather than have been affected by that state's educational policies. Furthermore, it is entirely possible that wage gains or losses due to the policy change could only be realized at higher ages for URM, and this would not be captured by the regression result as it only focuses on recent college graduates. Despite the difficulty in interpreting this result, it is still possible to approximate an effect on wages using the educational attainment regression.

The educational attainment regression has the same issue with state of origin issues. Another major issue that could be driving noise here is the fact that the data is unable to distinguish between private and public schools. IPUMS reports educational attainment, but does not show whether the education, if occurred, was at a public or private university. This is important to note because affirmative action bans, if implemented, would have only affected public universities in each state. Private universities are free to maintain affirmative action in their admissions even if the state law changes. This issue does not occur in the university quality regression, which could be a reason the standard errors were smaller, as state of origin does not matter for university quality.

While the regression for affirmative action bans on median wages yielded noisy results, using the 2.3% decline in college degree attainment, it is still possible to determine the real wage loss overall for URM due to this. The Federal Reserve Bank of Cleveland finds that college degree holders out earn their high school diploma peers by $26,104. With the 2.3% decline in degree attainment, the total effect on wages that this had can be estimated to $600 per person year over year in this time period. The first assumption in this estimation is that the expected wages without college and with college are not affected by the number of degree holders, meaning that there is not a supply/demand relationship with college degree holders in the market and their expected wages earned. If this assumption is violated, the real wage effect would be smaller than a $600/year decline. The second assumption is that all those attending college are equally benefiting from college at the same wage premium, as in the marginal college student is not earning any higher or lower of a wage benefit than the top-tier admits at a school. This is a critical assumption as previously noted, the marginal admit is likely to be most affected by affirmative action policies. Without this assumption, this figure is not attributable to the policy change.

This should not replace the $2,761 income point estimate, as these economic losses refer to two different phenomena. The $600 loss is from not having a college degree versus having a college degree. The $2,671 loss is the loss attributed to attending a lower quality university as a result of affirmative action policy bans. This study is able to find a significant effect that affirmative action policies have on the quality of university attended by URM, however, among college graduates was unable to precisely estimate the real wage effects that would cascade into. The $2,761 figure is a noisy estimate, but still attributable to the decline in average university quality attended by URM, as this data only contains URM college graduates. Furthermore, while this is a noisy point estimate, the regression point estimate for the drop in university quality attendance is not. This is important to note because the drop in average quality of university attended by URM is supposed to be the factor causing wages to decrease. This relationship should not be ruled out, especially when viewing figures 1 and 3 as both follow the same trend despite being predicted from different data sets. These figures represent the regression plots for estimated changes in median income and quality of university attended by URM.

One issue that is exclusive to the university quality score regression, however, is the different standards in different states. Different states have different quality universities, and the drop from one state's flagship university to the second best could be a different quality score drop than a different state. Table 2 shows that the quality drop from UC Berkeley to UCLA is a drop of 0.011 points, however for Michigan, the quality drop from the University of Michigan, Ann Arbor to Michigan State University is a drop of 0.14 points. Therefore, in some states, the scale of the quality drop could be more intense than others, simply because of the inherent quality differences in their universities being of a small or large magnitude.

The secondary issue exclusive to the university quality score regression point estimate is the issue in interpretation. As previously stated, attending the #2 university in a state versus the #1 has differing impacts depending on the state, and this carries over to economic outcomes. Dropping the average of 0.113 points in university quality could have a small or large effect on median wages depending on several intrastate factors, as well as field of study. The most interpretable outcome of this is the proportion of URM that attend universities that are more likely to retain and graduate students who generally scored higher on standardized testing in high school.

This paper's overall conclusions fit the conclusions drawn by Bleemer's paper rather than Arcidiacono's. Like in Bleemer's paper this study finds the decline in quality of university attendance by URM to be a real effect of banning affirmative action. While Bleemer focused on California and the UC system, this study extended that to the country and all iterations of affirmative action bans.

In fact, it is entirely likely that the point estimate generated by this paper is a better representation of the United States' expectations for decline in university attendance quality for URM with affirmative action bans because California's universities are generally higher quality, while other states do not share that trait. Furthermore, even within California, Bleemer focused on the UC system rather than the CSU system, which would be more representative of the state college systems found in other states due to the wider differing quality between schools.

Arcidiacono's main conclusion in his paper is the mismatch theory in which URM actually benefit from attending universities in which they are "better suited" for in the absence of affirmative action policies. This, for him, drives the lack of change in degrees obtained within the UC and CSU system overall with Prop 209 banning affirmative action in the state. While the paper indeed finds and attributes the higher graduation rates and academic performance by URM to mismatch theory, it underestimates the total economic cost on average to URM (Arcidiacono, et al., 2012). While in California it is entirely possible that there are enough seats for all those interested in attending college to the point where even "lower quality" schools are able to graduate URM, this is not the case around the country, in which college degree attainment actually does decline. This is attributable to a $600 loss in income per year on average for URM overall.

There are a few reasons why California could have different outcomes than the rest of the country. The first is the large higher education infrastructure that exists in the state. Even without the UCs, the state has over 20 CSU campuses and many more Community Colleges which feed into the state's university system and have high URM enrollment. The second is a much different racial composition than the rest of the country. California is minority-majority, a trait shared by few other states in their younger population, notably, Arizona, Texas, and Florida which also banned affirmative action (Maciag, 2015). However, of these states, it is important to recall that Texas and

Florida were unable to be included in the study due data limitations. If these states were able to be included, perhaps the results would closer match Arcidiacono's paper, but that is unclear.

# 6 Conclusion

This study estimated the effects of affirmative action bans in the United States from 2001 – 2020 on several educational and economic outcomes for the population of interest, underrepresented minorities or URM. This was accomplished by running 3 regressions using a staggered events study with synthetic controls. Each eligible control state in the US was weighed specifically to match the outcome variable's trend to the treatment group of states prior to the affirmative action policy bans so the average effect could be estimated. Overall, the study found that affirmative action bans were associated with a -0.113 decline in average university quality attended via the index I created based on SAT/ACT scores as well as retention/graduation rates at universities, a -2.3% decline in college degree attainment, and a noisy estimate of a wage loss of $2,762. All of these figures pertain to URM in the United States, and the final figure is only for those with a college degree.

This study had several severe limitations in drawing conclusions. The largest one in estimating college degree attainment and wage effects is IPUMS data. This data does not have any information on migration among samples taken, nor does it have information on the origin of people whose data is in the datasets. This means that the data on who has a college degree, or who is earning what is merely attributable to who is in the state at the time, which may not be someone who was affected by the policy. It is entirely possible for someone to have gotten educated in a different state, be affected by that state's affirmative action policy, whatever it may be, and then work in another state after graduating college.

Another large issue is the lack of data from the 90s in general both from IPEDS and IPUMS. This lack of data made it impossible for this study to take into account several key states that banned affirmative action earlier: Washington, California, and Florida. The latter two are especially important to note because of their large minority populations.

While the university quality did not have the same issue, it had different issues. Different states have overall different university quality and could have wider or shallower differences between schools within their state. This means that the drop in quality could be overstate in its real effects state by state, depending on the public college system that exists there. Secondly, IPEDS data was lacking for several schools for several years. The estimation methodology for university demographics went through 3 changes over the course of the time period, which could lead to shifts in demographics purely from data issues, not policy change or anything else.

Future affirmative action studies should attempt to obtain higher quality data that does not have the same issues that the data in this study had, especially in relation to the issue for origin of college degree. If that data is able to be tracked, then point estimates for real wage losses due to lower quality universities and college degree attainment in general can be better estimated. A more accurate quality ranking that is consistent between states could be created in order for the results in the university quality attendance regression to be more accurate as well. Another interesting direction for future studies to take is to directly compare how affirmative action policies compare with alternative changes in admission policy. Texas's top 10% among other policies also have real effects on these outcomes, and seeing how they match up with the outcomes here can be informative for policy makers trying to make the best, most equitable policies for all high school graduates intending on applying to college.

# 7 Tables and Figures

Table 1: Summary Statistics

| Statistic | N | Mean | SD | Min | Median | Max |
|---|---|---|---|---|---|---|
| **Outcome Variable** | | | | | | |
| Median Wages | 968 | 18371 | 6472 | 5782 | 16950 | 11400 |
| College Attainment | 920 | 0.4136 | 0.0974 | 0.1478 | 0.4052 | 0.6826 |

*Table 2: Top 15 Schools by Quality Score*

| Institution Names | Quality Score |
|---|---|
| University of Virginia-Main Campus | 0.89825 |
| William & Mary | 0.89300 |
| University of California-Berkeley | 0.89175 |
| University of California-Los Angeles | 0.88050 |
| University of Michigan-Ann Arbor | 0.87775 |
| University of North Carolina at Chapel Hill | 0.86500 |
| United States Naval Academy | 0.86225 |
| University of California-San Diego | 0.85800 |
| The College of New Jersey | 0.85800 |
| University of Illinois Urbana-Champaign | 0.84625 |
| United States Military Academy | 0.84300 |
| Pennsylvania State University-Main Campus | 0.83700 |
| University of Maryland-College Park | 0.83600 |
| Georgia Institute of Technology-Main Campus | 0.83575 |
| University of Florida | 0.83450 |

*Table 3: Bottom 15 Schools by Quality Score*

| Institution Names | Quality Score |
|---|---|
| University of Arkansas at Pine Bluff | 0.4756250 |
| Texas A & M International University | 0.4751111 |
| Indiana University-Kokomo | 0.4672500 |
| Indiana University-Northwest | 0.4602500 |
| Oklahoma Panhandle State University | 0.4596667 |
| University of South Carolina Beaufort | 0.4588750 |
| Northeastern Illinois University | 0.4575556 |
| Alabama State University | 0.4545000 |
| Kentucky State University | 0.4535000 |
| University of Alaska Southeast | 0.4482500 |
| Indiana University-East | 0.4387500 |
| Glenville State College | 0.4387500 |
| Chicago State University | 0.4301111 |
| West Virginia State University | 0.4215556 |
| Haskell Indian Nations University | 0.3963333 |

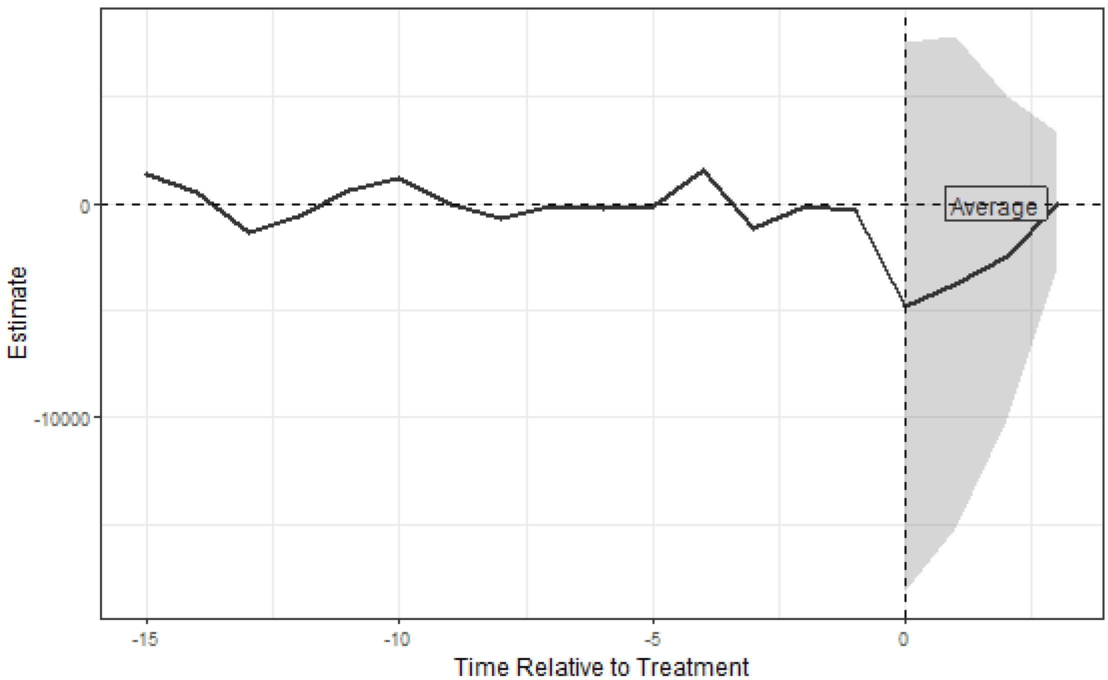
Table 4: Regression Results

| | *Dependent variables:* | | |
|---|---|---|---|
| | (Median Wages) | (Educational Attainment) | (University Quality) |
| $Affirmative\ Action\ Ban$ | -2761.86 | -0.023$^{*}$ | -0.113$^{**}$ |
| | (3742.71) | (0.015) | (0.052) |
| Observations | 968 | 920 | 969 |

$^{*}$p <0.1, $^{**}$p<0.05$^{***}$p<0.01

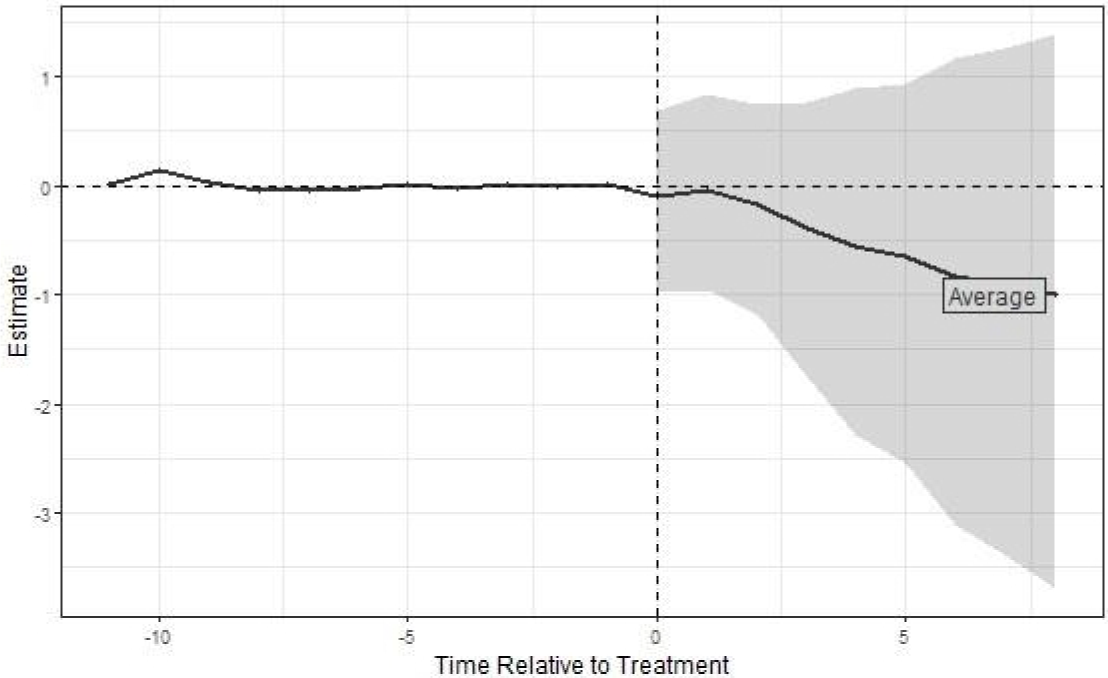
**Table Notes:** Median wages are in terms of USD, fixed to 2001 CPI; Educational Attainment interpreted as % of URM that have a college degree and estimation here in absolute scale, not relative % based on previous values; university quality scale reference seen in above in Tables 2 & 3; All regression estimations are outcomes from affirmative action bans

Figure 1: Regression Plot for Estimated changes in Median Wages for College Educated URM



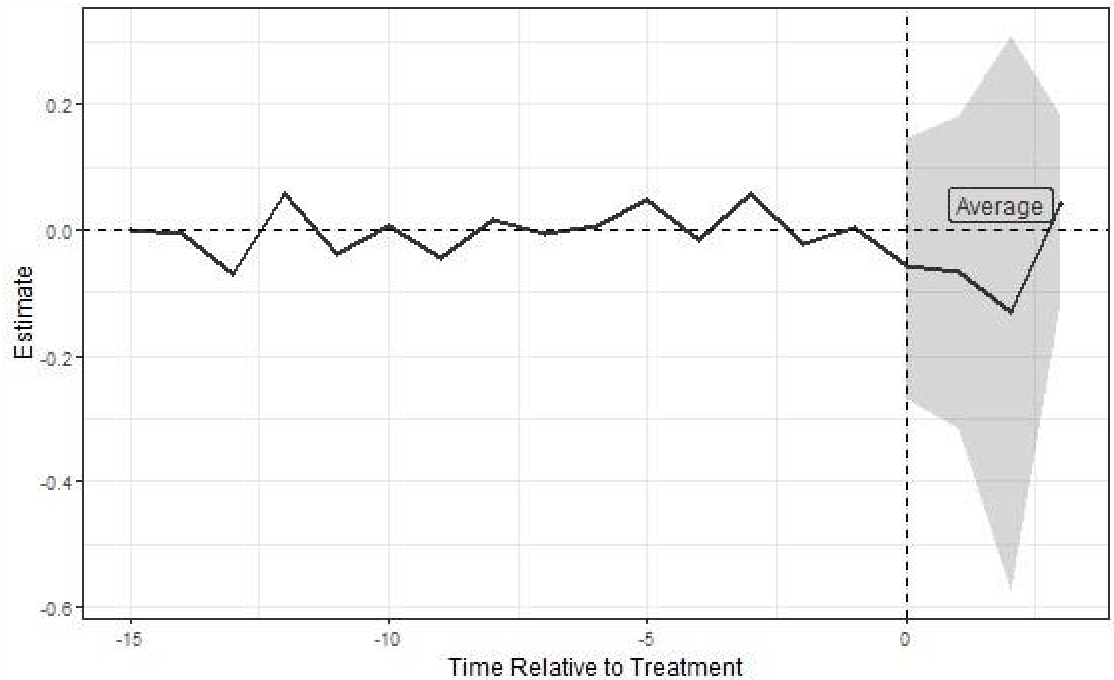

*Notes: Estimation is in terms of USD; Shaded region represents full range of outcomes and line represents average estimation for change*

Figure 2: Regression Plot for Estimated changes in Bachelor's Degree Attainment for URM




*Notes: Estimate is in absolute terms %*

Figure 3: Regression Plot for Estimated changes in Average University Quality Attended




*Estimation is in terms of the units of the quality score, where scale can be found in Tables 2 & 3*

# 8 References


Arcidiacono, Peter, et al., Nov. 2012. "Affirmative Action and University Fit: Evidence from Proposition 209." National Bureau of Economic Research, Working Paper, no. 18523.

Ben-Michael, Eli. "AugSynth: Augmented Synthetic Control Method." *Github*, Apr. 2020, https://github.com/ebenmichael/augsynth/blob/master/README.md.

Bleemer, Zachary. "Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209." *The Quarterly Journal of Economics*, vol. 137, no. 1, 2021, pp. 115–160., https://doi.org/10.1093/qje/qjab027.

Card, David, and Alan B. Kruger. "Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants?" National Bureau of Economic Research, vol. 78, no. 28, ser. 10366, Mar. 2004. 10366.

Ellison, Glenn, et al., Sept. 2012. "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools." National Bureau of Economic Research, Working Paper, no. 22589.

Leonard, Jonathan S., Apr. 1984. "Splitting Blacks?: Affirmative Action and Earnings Inequality within and Across Races." National Bureau of Economic Research, Working Paper, no. 1327.

Lincove, Jane Arnold, et al., August. 2016. "Match or Mismatch? Automatic Admissions and College Preferences of Low- and High-Income Students." National Bureau of Economic Research, Working Paper, no. 22559.

Maciag, Mike. "A State-by-State Look at Growing Minority Populations." *Governing*, Governing, 14 June 2021, https://www.governing.com/archive/gov-majority-minority-populations-in-states.html.

"History of Affirmative Action: American Association for Access Equity and Diversity." AAAED, www.aaaed.org/aaaed/history_of_affirmative_action.asp.

"Federal Pell Grant Program." Home, US Department of Education (ED), 4 June 2015, www2.ed.gov/programs/fpg/index.html.

Smith, Tyler. "Elite Education and Affirmative Action." *American Economic Association*, 4 Dec. 2020, https://www.aeaweb.org/research/affirmative-action-unintended-consequences-chicago.

Tasci, Murat, and John Linder. "The Great Recession and Its Impact on Different Industries." *Cleveland Fed*, Federal Reserve Bank of Cleveland, 10 Sept. 2010, https://www.clevelandfed.org/en/newsroom-and-events/publications/economic-trends/economic-trends-archives/2010-economic-trends/et-20100910-the-great-recession-and-its-impact-on-different-industries.aspx.