The Obama Effect: Effect of Black Electoral Victory on Racial Prejudice and Inequality*

Jung Sakong†
Federal Reserve Bank of Chicago

February 13, 2020

Click here for the latest version

Abstract

Following the Obama presidency, pundits and researchers have asked how having a black leader affects white Americans' attitude toward black Americans. Given theoretical ambiguity, I test for causal impact of a black leader on racial attitudes using local elections of black politicians at the municipal level. Using Race Implicit Attitude Test (IAT) scores as a measure of racial prejudice and close election regression discontinuity design for causal inference, I find that electoral victory of a black leader leads to a rise in racial prejudice among white Americans against black Americans. Following a close electoral victory, the IAT score rises by about 0.03, or 7% of the average black-white difference. Simultaneously, using the same discontinuity design, black politicians' electoral victory causes lower employment and higher mortgage denial for black Americans relative to white Americans. Interpreting close electoral victory of black politicians as an instrument, I argue that the rise in racial prejudice caused black-white economic inequality to widen.

*These views are those of the authors and do not reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System. I am extremely grateful to my committee chair Amir Sufi and members Marianne Bertrand, Raghuram Rajan and Luigi Zingales for their continuous guidance and support. I also thank Lancelot Henry de Frahan, Lars Hansen, Paymon Khorrami, Canice Prendergast, as well as seminar participants at the University of Chicago Applied Microeconomics Lunch. All errors are my own.

†Email: jung.sakong@frbchi.org

1 Introduction

The Barack Obama presidency has motivated discussions as to how having a black leader affects white Americans' attitude toward black Americans in general. Empirical analysis of black leaders has been challenging due to two issues: 1) relative to population, black politicians are disproportionately few in numbers, and 2) black politicians are not elected randomly. This paper seeks to overcome those issues by looking at a much broader set of local elections of black politicians, and – aided by the larger sample size – by using close elections to establish causality from electing black leaders to racial attitudes and actual racial economic inequality.

Black electoral victory can affect whites' racial attitudes through multiple mechanisms, with differing direction of effect. First, a black politician could dispel negative stereotypes associated with black Americans. Second, having elected a black politician, white Americans can feel justified in holding and expressing racial bias via the self-licensing effect (Monin and Miller (2001)). In particular, Effron et al. (2009) found that endorsing Obama led participants to express views that favored whites over blacks. The idea that Obama election has led to a resurgence in racial prejudice has been circulated in popular media as well.[1]

In this paper, I estimate a causal impact of a black politician's electoral victory on white Americans' racial prejudice and black Americans' relative economic outcome, using local elections and a close election regression discontinuity design.

Data on local elections come from Our Campaigns, a Wikipedia-like website compiling electoral information. I use data on any US election with sub-federal constituency. This includes US federal House representatives, mayors, city council members, county executives and county council members. Given the disproportionately low prevalence of black politicians, extending the set of politicians beyond the commonly considered representatives and mayors drastically increases the sample size. The race of the candidates is classified using: 1) tags supplied on Our Campaigns, 2) candidates' surnames and corresponding racial distribution from the Census, and 3) facial recognition of the candidates' photos.

In order to measure racial prejudice in observation data, I use Race IAT score data from Project

[1] For example, see Blake, John (2016, November 19) "This is what 'whitelash' looks like." CNN. https://www.cnn.com/2016/11/11/us/obama-trump-white-backlash/index.html; "A reflection on Barack Obama's presidency." (2016, December 24) The Economist. https://www.economist.com/christmas-specials/2016/12/24/a-reflection-on-barack-obamas-presidency.

Implicit Database, compiled by Xu et al. (2014). Since its development in Greenwald et al. (1998), the IAT has been a widely used test of subconscious or implicit racial bias. Project Implicit Database collects voluntarily completed online IAT surveys, along with demographic and other respondent information. Since the online IAT surveys are voluntary, the pool of respondents is both self-selected and highly unrepresentative of the US population. I project out the demographic information to get at variation in local IAT scores not driven by composition. The residual scores are then aggregated to the county level. Given about quarter million completed surveys per year, the resulting panel data have informative local variation. Finally, the results are unchanged when using the raw scores or composition-adjusted scores, suggesting that selection is not affecting the identification strategy.

The empirical strategy starts with a standard difference-in-difference estimator, comparing the jurisdiction of black winners with surrounding areas. The difference-in-difference estimate shows that there was no differential change in racial prejudice in areas affected by the election relative to surrounding areas. However, the difference-in-difference is likely biased, for example because black politicians are more likely to be elected in areas where prejudice against blacks is falling.

To overcome this identification challenge, I use a standard close election regression discontinuity, looking at the 3-year period after an election. The 3-year period before an election is used for placebo tests. I only look at elections where the top two candidates include one black and one white candidate. Results in this draft are obtained by defining "close election" as those with less than 10% vote margin between the winner and the runner-up. Optimal bandwidth (for example using Imbens and Kalyanaraman (2012)) is wider, and results obtained using optimal bandwidth are not qualitatively different.

I find that a black politician's victory causes racial prejudice to rise as measured by the IAT score. The average effect size is around 0.03, which is about 7% of the raw average gap in IAT scores between white and black Americans, where higher IAT score indicates more racial bias. There is no difference in IAT scores in areas with a black winner versus a white winner, before the election.

Turning to economic inequality measures using the same regression discontinuity design, black politicians' electoral victory causes black workers to transition more into non-employment (flow), fewer blacks to be in employment (stock), originate less mortgage amount and be denied more mortgages, relative to white counterparts.

In the last part of the paper, I interpret close election of a black politician as a time-varying instrumental variable for local racial prejudice. I highlight how this approach of identifying discrimination is different from what has been done in the literature so far.

The benefits of "shocking" prejudice goes beyond just having another instrument for identifying the effect of race. It also gets at racial prejudice directly, as opposed to lumping in the effect of statistical discrimination (of the kind where forecasts are made with full information including race).

1.1 Literature Review

The election of Barack Obama in 2008 has spurred an explosion of literature looking at the two-way interaction between black representation in politics and attitudes toward black Americans (see Parker (2016) for a comprehensive review). Academic study of the causal direction from salient black political leaders to racial attitudes is smaller but growing fast, including in economics (for example, DellaVigna (2010)). There are multiple theoretical channels by which visible figures of a minority group can affect views toward that group in general, and theories differ on even the direction of the effect.

On one hand, seeing a black exemplar may reduce bias toward blacks by dispelling negative stereotypes and showing that black Americans are capable. In contemporary American, black males are stereotyped as criminals and black females as undeserving recipients of welfare (Hurwitz and Peffley (1997)). Seeing a black American rise to a position of prominence and competence may dispel such stereotype. In studying gender attitudes in India and female politicians in India, the authors find that randomized rise of female politicians do indeed reduce gender bias and lead to positive outcomes for girls (Beaman et al. (2009); Beaman et al. (2012); and Chattopadhyay and Duflo (2004)).

On the other hand, seeing a black leader may also increase bias toward blacks by convincing majority whites that they need no longer watch out for racism or by raising the threat they perceive as the dominant group. Monin and Miller (2001) argues that one act of racial tolerance endows the individual with a moral license to express biased views in other dimensions, exhibiting the socalled self-licensing effect. Effron et al. (2009) finds that endorsing Barack Obama during the 2008 presidential election made his supporters more likely to display racial bias against black Americans. Kouchaki (2011) further finds that even knowing of others' un-prejudiced behavior leads individuals to freely express biased views. Theories of intergroup bias based on social dominance argue that black politicians may raise the perceived threat to white Americans' racial dominance and motivate them to engage in backlash (Olzak (1990); Sidanius et al. (1992)).

Against this backdrop of ambiguous theoretical prediction, empirical studies examining racial attitudes following the election of Barack Obama find results in both directions, depending on sample and methodology (for example, Goldman and Mutz (2014)[2]; Kinder and Chudy (2016); Tesler and Sears (2010) ;Tesler (2016)). A key challenge is that the Obama election is one event, limiting the statistical power of empirical tests. The innovation of this paper is to extend the set of elections to local and municipal levels, where there are more black electoral victory to potentially establish statistical power. The key tradeoff is that local elections are less salient.

At the mayoral level, there is a large political science literature examining the impact of black mayors, although the focus is predominantly on actual policy impact. The key question is whether active political participation matters for voicing the preferences of minority constituents, as opposed to passive participation in the form of voting. These studies often seek to test whether black mayors enact more liberal policies (Keller (1978) ; Abney and Hutcheson Jr (1981); Saltzstein (1989); Marschall and Ruhil (2006); Hopkins and McCabe (2012); Nye et al. (2014); Eggers et al. (2015); Brollo and Troiano (2016)). Most recent studies find that the racial identity of mayors has no direct policy impact due to constraints that municipal executives face, except in a few policing practices. Studies looking at party or gender find similar results (Ferreira and Gyourko (2014); Ferreira and Gyourko (2009)). These studies suggest that if election of black politicians has economic impact, it is less likely to be through actual public policy changes. And that elections with black candidates receive more attention as measured by higher white and black voter turnout suggests a possible psychological channel if black electoral victory affects economic outcomes of black and

[2] Using a survey of 20,000 respondents, they find, "From the summer of 2008 through Obama's inauguration in 2009, there was a gradual but clear trend toward lower levels of white prejudice against blacks." "... this change in attitudes did not last."

white constituents (Washington (2006)).

2 Data

This study draws from multiple sources of data: 1) data on political elections and candidates from Our Campaigns, 2) data on local measures of racial prejudice from Project Implicit (Implicit Association Test scores), Google searches and interracial crimes, and 3) data on interracial economic disparity from Quarterly Workforce Indicators (for employment) and Home Mortgage Disclosure Act (for mortgage origination). The following sub-sections describe each in detail. To make sense of the various data components, it is helpful to think of political elections and candidates as data on the forcing variable or instruments (i.e. Z), racial prejudice measures as data on the independent variable of interest or the outcome of the first-stage (i.e. X), and interracial economic disparity measures as data on the ultimate or second-stage outcome variable (i.e. Y).

2.1 Election data

Political elections provide the ideal setting for studying the effect of a salient exemplar for both substantive and methodological reasons. Substantively, politicians are highly visible, partly because they have to campaign to attract votes. Methodologically, election outcomes are uncertain and provide an identification strategy to study the causal effect of having an exemplar.

Data on local electoral outcomes come from Our Campaigns, a Wikipedia-like website that aggregates electoral information. I use data on any US election with sub-state constituency. This includes US federal House representatives, mayors, city council members, county executives and county council members. Given the disproportionately low prevalence of black politicians, extending the set of politicians beyond the commonly considered representatives and mayors drastically increases the sample size. Table 1a shows the in the years since 2003, there were 202 black congressional victories in US House, 12 black gubernatorial victories and 215 black mayoral victories in the Our Campaigns data using the racial identification scheme to be described below. If I wanted to use close elections between black and white candidates, defining close election as those with a vote margin of 10% or less between the top two candidates, that leaves 68 elections and 39 black victories among the three commonly studied, most visible offices. Including state legislatures, county councils and city councils, along with other elected municipal offices (e.g. county president) increases these numbers six- to sevenfold, with 501 close black-white elections and 247 black victories in such elections. Of course, such sample size gain is traded off against the lower visibility and salience of these local offices in determining statistical power. The tradeoff will be explored and exploited in the empirical specifications below.

To compile the election data used, I scrape the Our Campaigns website for the following information: most recent map of the jurisdiction (map in Figure 1a and the associated longitude-latitude coordinates), history of electoral races with date, type, candidates and their vote counts and shares (Figure 1b), and candidate information, in particular the user-supplied tags for race and photos for facial identification of candidate race (Figure 1c). Since I am expanding the set of politicians studied beyond what is commonly done in the literature, identifying candidates' races is a challenge, which I address below.

2.1.1 Identifying race of candidate

Candidates' race information is provided for a small subset of those on the website, so I use two other methods to identify candidates' races. Whenever the website contains direct information on race, I use that information. This small subset is also useful for judging the accuracy of the other two methods. Table 1b shows how the candidates classified according to tags are classified using the other two methods of surnames and facial recognition. Note that the user-supplied tags can be wrong too.

The second source of racial identification is the candidates' last name. Surnames are a widely used source of demographic information, even outside of the US (for example, Neggers (2018) uses surnames in India to identify religious and caste identity). Using the Census 2000 and 2010 surnames files (Word et al. (2008); Comenetz (2016)), I classify a candidate as belonging to one racial group, if 80% or more of people with that candidates' last name belong to that race. For example, candidates with the surnames Little or Smalls are classified as being black, while candidates with surname Hansen are classified as being white. Surnames are highly informative for Hispanic and Asian Americans, but black and white Americans tend to share common surnames. This can be seen in the low success rate of classifying black and white candidates to their races in Table 1b.

Also note that for surnames that are common among black and white American, using a common threshold makes it more likely to classify a surname to be white. Since there are more white Americans in general, the fraction white is high overall. So even if relative to base population fraction, a surname might be "more black," I use the absolute probability that a person with a given surname belongs to each race.

Audit studies have exploited strong racial associations of first names (Bertrand and Mullainathan (2004)). But to be used in observational data, I need a comprehensive list of first names that are predominantly of one or another race. One recent list of black and white first names uses Census sample (Tzioumis (2018)). But since this database is based on a sample, there are only 17 first names that are 80% or more black. The most populous of these, Latoya, is based off of 93 observations. Given this data restriction, I do not exploit the information content of candidate first names, although information is being lost.

The last source of racial identification is the candidates' photo. Whenever a candidate has an associated mug shot, I use facial recognition to infer the candidate's race. To this end, I use free software provided by Face++. Their algorithm classifies race to Asian, black and white. I can check the accuracy of facial recognition by comparing the race from face recognition with the tag information. The algorithm classifies about 92% of white candidates (according to tags) as being white, while it classifies only about 68% of black candidates (according to tags) as being black (Table 1b). Given the noise, I use the other sources of information first if they are available, before resorting to facial recognition.

Facial recognition is a mature field of computer science, and while younger than other fields of facial recognition, extracting racial identity from faces too has a large literature behind it (see Fu et al. (2014) for summary of techniques). The value of facial recognition can be seen in the last column of Table 1b. Among the vast majority of candidates with no user-supplied tag, 4% of them can be classified as being black American using facial recognition. Even with the noise in classification, the sample gain is useful when creating a database of politicians by race.

When racial information from these three sources disagree, I give precedence to the tag information, surname, then facial recognition.

2.1.2 Creating jurisdiction cross-walks

The website Our Campaigns contains interactive maps, with the most recent constituency highlighted using coordinates. I additionally scrape these set of coordinates to construct cross-walks from constituencies to zip codes and counties. The website supplies only the most recent jurisdiction, but political office jurisdictions change over time. While using the most recent jurisdiction for past terms will create measurement error in which zip codes and counties were affected by those offices, in this paper I use only the most recent jurisdictional boundaries, due to the difficulty of acquiring historical jurisdictions for local offices.

2.2 Racial prejudice data

Several proxies for racial prejudice have been used in the literature, almost always as a static measure. In this paper, I am interested in examining what changes the racial attitudes of a local area. The idea that the underlying racial attitude can change over time is not new. Becker (1971) wrote, "Another proximate determinant is geographical and chronological location: discrimination may varyfrom country to country, from region to region within a country, from rural to urban areas within a region, and from one time period to another. Finally, tastes may differ simply because of differences in personality." Gladwell (2007) wrote, "I had a student who used to take the IAT every day. It was the first thing he did, and his idea was just to let the data gather as he went. Then this one day, he got a positive association with blacks. And he said, 'That's odd. I've never gotten that before,' because we've all tried to change our IAT score and we couldn't. But he's a track-and-field guy, and what he realized is that he'd spent the morning watching the Olympics."

Below, I introduce data sets that can be used to produced local panels of racial attitudes.

2.2.1 Implicit Association Test (IAT)

The Race IAT is a widely used test of racial bias (Greenwald et al. (1998)). Respondents are shown either pictures of faces (black or white) and words (good or bad). They use the same set of buttons on the keyboard to classify the faces into black or white categories, and words into good or bad categories. IAT is based on the premise that if a respondent has a stronger association in his mind between being white and being good, the classification exercise will take longer when he has to use the same button to classify a face as being black and a word as being good. The main measures, the D score, is the difference in time it takes to classify, when black faces and good words are paired together (i.e. use the same button) versus when black faces and bad words are paired together.

Since its inception in Greenwald et al. (1998), there has been an explosion of studies using the IAT to test various implicit attitudes. Recent meta-studies find that the Race IAT is a good predictor of racial discrimination (Oswald et al. (2013)). Criterion measures (official term for the intergroup behaviors) include brain activity, response time, microbehavior, interpersonal behavior, person perception, and policy/political preferences. Explicit measure utilize includes one separate category for "feeling thermometer." Key object in meta-studies is "IAT criterion-related correlation (ICC)." Main results for IAT in Tables 1 & 2. Weaker explicit measure results (explicit-criterion correlation) in Table 5.

Most recent meta-study of the IAT's predictive power (Kurdi et al. (2018)): 1) implicit measures work well regardless of ICC moderator (p26) - "absence of theoretical predictors" (p22); 2) standard IAT is superior (p29). p50 - tighest range of ICC for Race IAT vs. other IATs. Three concepts for which use IAT: attitude stereotype, identity. Key conceptual moderators: social sensitivity, controllability (i.e. automatic activation). Found IAT correlated with behaviors thought to be controllable. "... univariate meta-regressions showed that implicit measures were equally associated with measures of intergroup behavior irrespective of social sensitivity, controllability, conscious awareness, or target concept. In fact, contrary to the widespread notion that implicit measures are not associated with highly controllable behaviors, the present meta-analysis found a sizable number of large ICCs for such behaviors, including self-reported enrollment intentions in mathematics classes (), self-reported career aspirations (), and voting behavior ()."

Also Greenwald et al. (2009): "for socially sensitive topics, the predictive power of self-report measures was remarkably low and the incremental validity of IAT measures was relatively high." "IAT measures had greater predictive validity than did self-report measures for criterion measures involving interracial behavior and other intergroup behavior."

I get data on online IAT scores from the Project Implicit Database (Xu et al. (2014)). The data span years 2003-2017, with roughly a quarter million completed tests per year. The online survey also collects democragraphic variables (age, sex, race, education), political ideology and religious information, explicit racial bias questions comparable to those asked in the General Social Survey (GSS), and self-reported zip code. Figure 2b shows the distribution of the raw IAT scores, where higher numbers indicate more racial bias.

The IAT has been used widely in economics as well as other social sciences as a prominent measure of racial bias (Reuben et al. (2014); Chetty et al. (2018)). Chetty et al. (2018) in particular also uses the Project Implicit Database. Another recent paper that uses the IAT is Carlana (2019).

Since the online surveys are voluntary, the sample is self-selected and highly unrepresentative of the US population. To adjust for selection, I project the IAT scores on age, race, education (9 buckets), gender, and experimental variables (month, hour, weekday as well as order of experiment), before aggregating them to the race-county-month level. Figure 2a plots the time series average of this composition-adjusted responses, for white respondents. The issues introduced due to the voluntary nature of data collection are addressed more fully below.

The rich demographic and other details self-reported in the Project Implicit Database show how racial prejudice as measured by the IAT score co-varies with demographics. By far the most salient is race: black respondents' average IAT score is -0.04 against 0.39 for non-Hispanic whites (Figure 3a). Asians & Hispanics are in-between but closer to whites. The other important individual characteristic is political ideology: 0.39 for conservative against 0.28 for liberals (Figure 3e). By geographic region, the average score is higher in Northeast and Midwest (0.33) than in South and West (0.30) if including everybody; among only white respondents, the average score is highest in South (0.40) and lowest in West (0.35) (Figure 3f).

Two widely cited co-variates for racial bias are gender and education (for cognitive ability, see Hodson and Busseri (2012)). By gender, the average IAT score is 0.34 for males against 0.30 for females (Figure 3d). By education, including everyone, the average IAT score is 0.32 for those with college degree or higher and also 0.32 for up to high school graduate (Figure 3c). Among only white respondents above age 25, the average score is 0.37 for those with college degree or higher, and higher at 0.426 for up to high school graduate. The pattern by age is highly nonlinear, first falling from age 20 to 40, then rising again. By religiosity, among white respondents, those that are not at all religious have the lowest average score (0.35); among all races, those that are strongly religious have the lowest average score (0.27), largely because black respondents are more religious.

A major concern with using the online IAT database is that individuals self-select into taking the test online. I take several approaches to address potential issues that arise as a result. Such self-selection is a feature of most measures of amorphous entities such as racial prejudice. For example, Google searches for racial slurs is another measure of racial bias that is affected by who chooses to use Google and for that purpose (Stephens-Davidowitz (2014)). Yet, the self-selection issue is potentially larger in the case of the IAT because it is a test explicitly designed for racial bias.

First, potential issues can be classified into four categories. Self-selection by survey-takers introduces two sets of issues, which can be broadly summarized as the level and change in selection. The level of selection makes the sample not representative of the broader population, such that external validity claims are impeded. The change of selection is potentially more problematic in that any results I find may be driven not by actual treatment effect of the shocks I consider but by changing selection. Both issues can be further broken out into observable and unobservable kinds. Below, I describe each issue in more detail and how I address each.

Lack of representativeness along the observable dimension can be fully addressed by assigning weights to observations. To check the degree of unrepresentativeness and to assign weights, I use distributions by gender, age and education buckets from the 2008-2012 5-year American Community Survey (ACS), accessed via Integrated Public Use Microdata Series (IPUMS). Table 2a shows the demographic distribution in the 2008-2012 ACS and the 2008-2012 IAT data in the first two columns. The IAT sample is highly over-represented in the 19-29 age range, and highly under-represented in the ages above 40, for example. To make the sample more representative of the US population along gender, age and education, I impute weights as $\frac{\text{ACS population share}}{\text{IAT sample share}}$ for gender $\times$ age $\times$ education bins, fully interacted. Single-year ages were used, and educational attainment has been grouped into categories shown in Table 2a. Figures 4a and 4b show the weighted distribution of IAT scores for the entire sample 2003-2017 for all races and for white respondents only (red lines) against unweighted distributions (blue lines). The distributions are largely the same, with a slight shift towards the right. The main regression results will be replicated using this re-weighted score.

Lack of represenativeness along the unobservable dimension is more difficult to fully deal with. This issue makes external validity difficult to establish. If estimates do not differ between potentially more and less self-selected subsamples, that reduces the concern that there is much selection along the unobservable dimension.

Selection along observables can be fully addressed by including controls on the right-hand side. Along with the raw IAT score, I will use a composition-adjusted average IAT score, when I project out dummies for age (single-years), nine education buckets and gender. Regression estimates with this composition-adjusted average IAT scores will also be reported below.

Selection along unobservables can be dealt with by running placebo tests by putting observable demographic variable on the left-hand side with the treatment on the right-hand side. The idea is that if the sample changes along unobservable dimensions in response to treatment, such sample changes should partly be reflected along some observable variables. Alternatively, the same regression specifications can be run with and without these observable controls on the right-hand side. If observables change in response to treatment, the treatment effect on outcomes of interest with and without controls should be the same. The distinction between the solution to selection along observable and unobservable dimensions is: for the first, even if estimates change when controls are included, the estimate from the regression with controls is correct; for the second, difference in estimates between the regressions with and without controls already signifies selection along unobservable dimension that cannot be corrected. For all estimation results, estimates with both raw IAT scores and composition-adjusted scores (i.e. comparable to running regressions with controls) will be reported. Just to preview, the composition adjustment with sex, age and education has no effect on the estimates.

Second, I compare responses of those who were assigned to take the test against those who completely voluntarily took it. How responses differ between the more and less self-selected subgroups within the sample will be informative in evaluating how the self-selection of the sample as a whole may bias the results I find. Using the question, "What brought you to this website?" I classify a response as being mandatory if the respondent chose "Assignment for work" or "Assignment for school." I classify a response as as being voluntary if the respondent chose "Recommendation of a friend or co-worker," "Mention or link at a non-news Internet site," "Mention in a news story (any medium)" or "My Internet search for this or a related topic." About 62% of the respondents gave a response to this question since the question was asked in 2006. Relative to respondents who took the test as an assignment, those who took it entirely voluntarily are more likely to be white

(71% vs. 64%), more likely to have a bachelor's degree or higher (29% vs. 16%), less likely to be female (51% vs. 61%), more likely to be liberal (58% vs. 39%), more likely to be not religious (39% vs. 27%) and older (average age of 33 vs. 26).

Based on raw level and across all races, those who took the test completely voluntarily have an average Race IAT D score of 0.30, slightly lower than 0.31 for those who took it as a part of an assignment (Table 2b). Among white respondents, the average IAT D score is 0.36 among those who took it completely voluntarily, and 0.39 among those who took it as a part of an assignment. The IAT score gap between the assignment and completely voluntary groups remains at 0.01 for all races, even after projecting out demographic and experimental controls. The gap for white respondents also remains at 0.03 even after projecting out controls.

I basically make three adjustments to the raw scores by race: 1) projecting out dummies for sex, age and education; 2) assigning weights to be representative along sex, age and education; and 3) checking robustness with only mandatory respondents, i.e. those who replied "Assignment for work" or "Assignment for school" to the question "What brought you to this website?"

2.2.2 Other prejudice measures

To complement the IAT scores in measuring time-varying local racial prejudice, I bring in other proxies using multiple data sources. In addition to widely used survey responses from the General Social Survey (GSS), I use Google search trends for keywords associated with racial prejudice, crimes committed by white offenders against black victims, and school corporal punishment on black students relative to white students. These proxies are associated with other things in addition to racial prejudice; for example, the racial slur that I use for Google search trends also often features in the popular rap culture. The goal is to combine proxies that are related to racial prejudice albeit with much measurement error, so that the common component can be attributed to the racial prejudice of the time and place. I will describe each of the measures in turn.

Apart from the General Social Survey, the other proxies are not pure elicitation of preferences. In designating them as proxies for racial prejudice, I follow the criteria from Guiso et al. (2011) in their indirectment measurement of civic capital: "For an outcome-based measure to qualify as a good indicator of civic capital, the relationship between the input (civic capital) and the measured output should be stable and unaffected by other factors, such as legal enforcement."

Guiso et al. (2011) also discuss common movement as a criterion for judging the value of proxies: "Consistent with the idea that these measures are capturing the same underlying norms, they tend to be highly correlated... Hence if one were to rely on measurse of this sort in applied work, one could gain some insights by obtaining several indirect indicators and looking at common components (see Tabellini (2010))."

General Social Survey (GSS). Charles and Guryan (2008) create an index of local racial prejudice using the GSS surveys. They take a set of questions that I replicate the prejudice index constructed by Charles and Guryan (2008), by standardizing and summing over the responses to the same set of questions they use. Unlike all the other measures including the IAT scores, the GSS has the benefit of being a representative survey.

Google search trends. The two measures are Google search trends for a prominent racial slur (the n-word) and for "KKK." Google shares an index of search volume at the DMA-level since 2004 at the monthly frequency. Google searches for the the n-word has been shown to predict voting against Obama better than the GSS (Stephens-Davidowitz (2014)). Google search trends have several advantages over survey responses. For example, Googling can be done in secret and hence the anonymity suggests that Google searches can reveal racial prejudice more directly than survey responses, where the respondents may feel social pressure to give politically correct answers.

Google search trends also have some issues. First, there are other contexts in which the n-word and "KKK" can be searched for without involving prejudice. For example, the n-word sometimes appears in rap music, although more often rap usage uses the variant of the word that ends in an "a." Fortunately, Google Trends offers most common related searches, which suggest that many uses of these search terms do involve racial prejudice. Second, for privacy reasons, only keywords with sufficient search volume can be tracked on Google Trends. Other ethnic slurs whose usage is less ambiguous are also less widely used and cannot be tracked consistently over time. Third, the time series for the search trends cannot be taken meaningfully. The type of Google users has changed dramatically over the past decade as well as the fads that dominate Googling. Since Google Trends gives me an index, how it changes over time is influenced as much by what else is searched by whom. Cross-sectional comparisons and diff-in-diff comparisons are meaningful and form the backbone of the analysis.

Using Google search trends for "KKK" is new. The Klan is still in operation, and this search proxies for multiple channels related to racial prejudice. Individuals interested in joining it would search for it, and for Google Chrome users, entering the first few characters of the Klan website and pressing enter pre-maturely would lead a user to search for that term on Google. Individuals who fear or are concerned about extremist groups may also search for the Klan as a notorious example, and such searches would also be related to local racial animus. The DMAs with the highest average searches over the 2004-2017 period are Presque Isle, ME, Greenwood-Greenville, MS and Parkersburg, WV. DMAs with the lowest average over the entire period are New York, NY, Washington DC and San Francisco-Oakland-San Jose, CA. The DMAs with the highest average searches over the 2004-2017 period are Parkersburg, WV, Twin Falls, ID and Greenwood-Greenville, MS. DMAs with the lowest average over the entire period are Macon, GA, Portland, OR and Salt Lake City, UT.

White-on-black non-pecuniary crime. The next measure captures more extreme expressions of racial animus. I measure anti-black sentiments using crimes committed by white offenders against black victims, relative to those committed by white offenders against white victims. The data come from the National Archive of Criminal Justice Data (NACJD). Among crime categories, I take three codes only: simple assault, intimidation and destruction/damage/vandalism of property (the uniform common reporting codes are 132, 133 and 290). These crimes are non-pecuniary and therefore less likely to be linked to economic downturns for directly financial reasons. They are also the most common categories of hate crime. In fact, more than 10% of crimes in these categories are hate crimes. I do not use hate crimes directly because they are too few in number and the categorization into hate crime is subjective and potentially influenced by the fact that hate crimes are sometimes punished more severely. Finally, I scale the crime ratio by black-to-white population ratio in the DMA; otherwise, areas with higher black population would mechanically see more crimes committed against blacks. The exact variable definition is:

$$\left[\frac{\text{count of crime by white offender against black victim}_{it}}{\text{count of crime by white offender against white victim}_{it}} \Big/ \frac{\text{total blacks in area}_{it}}{\text{total whites in area}_{it}}\right]$$

for DMA i in year t. . The DMAs with the highest average over the 1991-2014 period are Bangor, ME, Alpena, MI and Sioux City, IA. DMAs with the lowest average over the entire period are

Wilkes Barre-Scranton, PA, Tucson, AZ and Jackson, MS.

School spanking. The last measure is the frequency with which school teachers corporally punish (i.e. spank) students of other races relative to white students. Corporal punishment is still legal in 19 U.S. states, and relevant statistics are released by the U.S. Department of Education as a part of the Civil Rights Data Collection. In some sense, this measure best captures changes in racial prejudice. Decision to spank a student may be impulsive and thus reflective of the underlying prejudice, and it is less likely to be affected by economic conditions directly. While a downturn can cause teachers to spank more overall, it is more difficult to think of why they would spank black students more than white students. When dealing with corporal punishment, I use ratio of rates of punishment, as there are large differences in the base rates. Black students are roughly twice as likely as white students to be spanked, whereas Asian students have much lower likelihood.

Crimes and spanking are less direct measures of racial prejudice, less like surveys and more like mortgage denial and non-employment. Conceptually, I imagine that there is a latent average racial prejudice associated with each local area and time. The prejudice measure I compile are a function of the latent racial prejudice and other factors that I claim are otherwise orthogonal to the local economic conditions and household finance outcomes of interest. As I accumulate more of these measures, my proxies will converge onto the latent racial prejudice.

2.2.3 Validation of IAT

The Race IAT D score will be the main measure of racial prejudice in this paper, mainly given its straightforward interpretation and the rich auxiliary demographic information the database provides. In this section, I validate the IAT as a proxy for racial prejudice at the local level, in three ways: comparison to black-white economic disparity (which some argue is partly an outcome of racial discrimination), comparison to the other racial prejudice proxies at both the county and DMA levels, and comparison to historical slavery, which Acharya et al. (2016) argue is an instrument for today's racial prejudice level.

The comparisons are cross-sectional. Panel comparisons (i.e. difference-in-difference) are coming. Note that none of the correlations below should be interpreted causally, with the possible exception of the historical slavery instrument a la Acharya et al. (2016).

First, Table 8b regresses economic disparity measures against the county's average IAT score.

On the employment side, areas with higher average IAT scores have blacks earning 54% even less than whites, and blacks transitioning to non-employment at a rate 11% higher than whites. On the mortgage side, areas with higher average IAT scores have black households' mortgage applications being denied at 36% higher rate than white households'.

Second, Table 3c regresses the other prejudice proxies against the area's average IAT score, at either DMA or county level. With the exception of the white-on-black crime measure, average IAT score is higher in areas where the other prejudice proxies are also higher: those areas search more for the n-word and "KKK" on Google, respond to the GSS questionnaire in a potentially more racially biased way, and spank black students more than white students. The opposite correlation for white-on-black crime will be explored further.

Lastly, I compare my measures to the cross-sectional prejudice instrument proposed in Acharya et al. (2016), who use as the instruments 1860 slave share and cotton production conditional on state fixed effects (see also Nunn (2009) for a discussion of the legacy of slavery). They argue that a legacy of slavery passes down over generations, and show that the prevalence of slavery among southern counties predicts survey responses today. The first column of Table 3a regresses the county's average IAT D score today against the area's 1860 slave share of the population from Census 1860, accessed via IPUMS. All regressions include state fixed effects following Acharya et al. (2016).

A strong correlate of this instrument is the contemporary black population share. This correlate is problematic for their interpretation because of an alternative hypothesis: the "racial threat" hypothesis argues that it is the current prevalence of black Americans that raises anti-black prejudice. This correlate is problematic for me, because more black Americans may proxy for prevalence of lower-skilled workers, hence raising employment cyclicality, for example. Their first solution is to simply control for current black population share. This makes instrument potentially not valid (i.e. over-control), but their coefficients are not affected much. In my validations, I also control for the contemporary black population share in the second colummn.

Finally, following Acharya et al. (2016), the third column regresses my average prejudice measures against slave share instrumented using per capita cotton production in 1860.

To give more background on the 1860 data on slavery intensity, 1,117 counties in slave states with non-missing data on slave share of the population. Weighted by 1860 population, average slave population share is 0.32, with a median of 0.30. The highest is 0.925, with the following percentiles: 1% is 0.03, 25% is 0.109, 75% is 0.509, and 90% is 0.65. Counties in the same state sometimes do have different slave population share. For example, in Georgia, Cherokee county (FIPS code 13057) had 0.106, while Clarke county (FIPS 13059) had 0.505. Both counties are in Atlanta CBSA. For another example of larger areas, in Alabama, Mobile county (FIPS 1097) had 0.277, while Montgomery county (FIPS 1101) had 0.66. This difference is comparable to the interquartile difference of around 0.4. States with the most intense slavery by population share are: SC (0.57), MS (0.55), LA (0.47), AL (0.45), FL (0.44), and GA (0.44).

Lastly, I validate my measures by showing that they capture something in common, by extracting a latent state.

3 Empirical methodology and prejudice results

This section lays out the empirical specifications designed to examine the reduced form impact of black electoral victory on both racial prejudice measures and racial economic disparity. To summarize, I use three sets of designs: 1) standard difference-in-difference around the election date, comparing counties affected to neighboring counties, 2) differential exposure to most salient black mayoral elections, defining exposure using ex ante level of racial prejudice, and 3) a regression discontinuity design (RDD) using close election of black winners over white contenders.

All the estimation techniques start from a difference-in-difference, using pre- & post- period of 3 years and compared against surrounding geographies in the same state. Exposure analysis compares these estimates against the pre-period level of racial prejudice, with the conjecture that areas that have a higher level of racial bias against blacks will respond more sensitively to the election of a black officeholder. This conjecture is later verified in heterogeneity analysis using regression discontinuity. Close election RDD compares the difference-in-difference estimates against the corresponding vote margin of the election, and then takes the discontinuity where vote margin is 0. These RDD estimates can also be plotted against the pre-period level for heterogeneity. All these regressions can be run with all elections or just the mayoral (or the set of most salient) elections, to examine the tradeoff between sample size and signal-to-noise ratio based on how salient a set of elections are.

While a difference-in-difference estimator is the most transparent, there is a clear identification issue that black electoral victory is not random. In particular, black politicians are more likely to win if white voters' racial prejudice is in decline. This creates a negative bias in the difference-in-difference estimate. Close election regression discontinuity is meant to overcome this identification challenge, relying on a vast literature arguing that winning a closely contested election is as good as random (Eggers et al. (2015)).

3.1 Difference-in-difference design

I start with the standard difference-in-difference estimator. I first compile a data set at the election-county-year level following Gormley and Matsa (2011) for difference-in-difference estimators with multiple events. Using only elections with a black winner, and for election i, county j and year t, I estimate

$$\begin{aligned} Y_{ijt} = & (\text{in jurisdiction } i)_{ij} + (\text{after election } i)_{it} \\ & + \Big((\text{in jurisdiction } i)_{ij} \cdot (\text{after election } i)_{it}\Big) + \eta_{ijt} \end{aligned}$$

With fixed effects, following Gormley and Matsa (2011):

$$Y_{ijt} = \alpha_{ij} + \alpha_{it} + \Big((\text{in jurisdiction } i)_{ij} \cdot (\text{after election } i)_{it}\Big) + \eta_{ijt}$$

For each election i where a black candidate won the election, I include the 3-year window before and after the election, and include all counties in the same state as the jurisdiction associated with the election.

Of course, black politicians do not get elected randomly. Previous literature has documented factors that predict black politicians' election (Marschall and Ruhil (2006)). The main concern is negative selection: black politicians are more likely to be elected in areas where dominant white voters' racial bias against blacks is decreasing. This will bias my estimate to find that areas where a black politician comes into office will experience a decrease in racial prejudice.

3.2 Regression discontinuity design

The empirical strategy is a standard close election regression discontinuity, looking at the 3-year period after an election. The 3-year period before an election is used for placebo tests. I only look at elections where the top two candidates include one black and one white candidate. Results in this draft are obtained by defining "close election" as those with less than 10% vote margin between the winner and the runner-up. Optimal bandwidth (for example using Imbens and Kalyanaraman (2012)) is wider, and results obtained using optimal bandwidth are not qualitatively different.

For observations at election i, geography (e.g. county) j, and event time (e.g. month or year) t, I run

$$Y_{ijt} = \alpha + \gamma_1 1\{\text{vote margin} > 0\}_{it} + \delta_0 [\text{vote margin}]^-_{it} + \delta_1 [\text{vote margin}]^+_{it} + \eta_{ijt}$$

When estimating heterogeneous treatment effect by election type or location characteristic, the specification is:

$$Y_{ijt} = \sum_k \left\{ \alpha^k + \gamma_1^k 1\{\text{vote margin} > 0\}_{it} + \delta_0^k [\text{vote margin}]^-_{it} + \delta_1^k [\text{vote margin}]^+_{it} \right\} 1\{\text{in sub-group}\}_{ij} + \eta_{ijt}$$

As with any reduced form identification scheme, the estimate γ_1 from this RDD is a local average treatment effect (LATE). Generalizing it to all elections is problematic, since close elections are likely to be different from other elections in many dimensions. The bigger external validity issue is if areas that have close elections between black and white candidates are systematically different from those that do not. I explore this issue further by looking at heterogeneity of treatment effects.

Since Lee et al. (2004) used close elections to test the median voter theory of Downs (1957) by getting at the causal impact of incumbency on future policy, there has been a large literature utilizing close election RDD as an identification scheme to estimate the causal impact of political victory, with well-established econometrics methodology (Imbens and Lemieux (2008); Calonico et al. (2014)).

Table 5 summarizes the main RDD results for white Americans' racial attitude. The first two columns of Table 5 show that in the 3-year window before the election, there is no difference in IAT scores between areas where the black candidate will narrowly win and those where she will lose (this is shown graphically in Figure 7a).

After the election, having a black winner causes IAT score among whites to increase by about 0.03 (last two columns of Table 5 and Figure 7b). The estimate is similar whether we use raw IAT score or the composition-adjusted one. This estimate of 0.03 corresponds to about 7% of the raw average gap in the IAT score between all black and white respondents in the Project Implicit Database.

Figures 8a and 8b plot the discontinuity estimates for each quarter relative to the election event at 0 (Figure 8a for the raw IAT and Figure 8b for the composition-adjusted IAT). There is stable zero difference leading up to the election, but discontinuity rises in the quarters following the election, to come back down eventually. Both the pre-period placebo and the time series of discontinuity plots suggest that the RD strategy is picking up a causal estimate.

3.3 Heterogeneity

This section explores the heterogeneity in treatment effects. The first dimension of heterogeneity is the level of racial prejudice. While the initial conjecture is that in areas with higher level of prejudice, seeing a black leader would lead to stronger backlash among whites (a point also made in assessing external validity in Beaman et al. (2009)), the major confound is that more prejudiced areas also tend to be more black and hence are more likely to be familiar with black politicians. Given the still low sample size of black-white races, I examine each area characteristic separately.

Another interesting dimension of heterogeneity is the economic condition of the area. In a downturn, white workers may feel less economically secure and perceive higher threat from black electoral victory (Bobo (1988)).

Table 7a presents heterogeneity results using

$$Y_{ijt} = \sum_k \left\{ \alpha^k + \gamma_1^k 1\{\text{vote margin} > 0\}_{it} + \delta_0^k [\text{vote margin}]_{it}^- + \delta_1^k [\text{vote margin}]_{it}^+ \right\} 1\{\text{in sub-group}\}_{ij} + \eta_{ijt}$$

where the group k is defined by the below- and above-median areas sorting by average IAT score level, black population share, or average income. Columns (3)-(6) show that treatment effect is not heterogeneous by either the black population share or income.

Columns (1)-(2) of Table 7a shows that the rise in prejudice is entirely concentrated in areas with above-median level of racial prejudice. In fact, the treatment effect in below-median prejudice level areas is insignificant and negative, consistent with Beaman et al. (2009).

Table 7b repeats the heterogeneous treatment effect analysis using the level of IAT, with black-white economic disparity variables. While statistical power drops as expected with heterogeneous treatment effects, point estimates consistently suggest that the negative effect of black electoral victory on black relative economic outcome is concentrated in high-prejudice areas.

4 Effect on racial gaps in employment and credit

4.1 Data

Data on employment gap come from Quarterly Workforce Indicators (QWI) and refer to all private employment at the county-level. The two main measures are likelihood of transitioning to non-employment among those employed (a flow measure), and employment ratio (a stock measure).

Data on mortgage denial come from the Home Mortgage Disclosure Act (HMDA) data. I calculate rejection rate as the number of mortgage applications for owner-occupied housing that were rejected, divided by all loans that were originated or denied. Since rejection rate is affected by who applies, I also look at mortgage origination per capita. The highest frequency of the HMDA data is year. Geographically, in one version I aggregate up to the county using the state and county variable; in a more disaggregated version, I use the tract variable to aggregate to zip codes. The tract variable is missing only for about 30 million out of about 500 million loan observations (i.e. the tract variable has value of "NA"). The HMDA TS file has variable "rzip" but that is likely the lender's zip code.

For all of these measures, I take the difference between black and white individuals in a given county as my economic gap measures.

For main measure, I will use the one that can be adjusted most promptly to reflect whatever change in underlying attitudes. For employment, this is the relative transition to non-employment.

For mortgage, this would be relative denial rate.

4.2 Difference-in-difference

Table 6a shows the difference-in-difference estimates for employment and mortgage variables. As with the prejudice measures, economic gaps show no change with the election of a black politician, most likely given that black politicians are more likely to be elected where white voters' racial bias is declining.

Turning to identified regression discontinuity estimates, however, Table 6b shows patterns consistent with a negative causal effect on racial prejudice.

4.3 Regression discontinuity

4.3.1 Racial economic inequality

Then, looking at economic inequality measures, Table 6b shows that black politicians' electoral victory causes black workers to transition more into non-employment (flow), fewer blacks to be in employment (stock), originate less mortgage amount and be denied more mortgages, relative to white counterparts.

4.3.2 Employment effect by firm size

Equal Employment Opportunity Commission (EEOC) of Civil Rights Act of 1964 applies to any firm with 15 or more employees. The QWI data divide firm size into the following groups: 0-19, 20-49, 50-249, 250-499, and 500 or more.

5 Causal identification of racial discrimination

5.1 Instrumental variable (IV) estimator

I start from the standard regression discontinuity exclusion restriction that whether the black candidate wins or loses in a close election is uncorrelated with local conditions, except through its effect on local racial prejudice:

$$E\left[1\left\{\text{vote margin}>0\right\}_{it}\varepsilon_{ijt}|\text{IAT}_{ijt}\right]=0$$

where ε_{ijt} is the unobservable factors in

$$Y_{ijt}=\beta_0+\beta_1\text{IAT}_{ijt}+\varepsilon_{ijt}$$

for some black-white economic disparity measure Y_{ijt} in county j in year t, surrounding an election i.

This exclusion restriction motivates the following two-stage least squares (2SLS) estimator.

$$\begin{aligned}
Y_{ijt} &= \beta_0+\beta_1\hat{\text{IAT}}_{ijt}+\tilde{\delta}_0\left[\text{vote margin}\right]_{it}^{-}+\tilde{\delta}_1\left[\text{vote margin}\right]_{it}^{+}+\varepsilon_{ijt}\\
\text{IAT}_{ijt} &= \alpha+\gamma_1 1\left\{\text{vote margin}>0\right\}_{it}+\delta_0\left[\text{vote margin}\right]_{it}^{-}+\delta_1\left[\text{vote margin}\right]_{it}^{+}+\eta_{ijt}
\end{aligned}$$

The linear controls on either side of the vote margin enter both first stage and second stage regressions.

5.2 IV estimation result

The main IV estimation results are reported in Table 8. Given the exclusion restriction, the first row estimates show the causal impact of 1-point increase in the composition-adjusted IAT score among whites in the county. To interpret the magnitude of the estimate for mortgage rejection rate, a 0.1 increase in racial bias among whites as measured by the IAT would lead to the black-white rejection rate gap to widen by 6.6 percentage points. Most of the estimates are borderline insignificant at the 5% level, largely due to a lack of instrument strength as can be seen in Figure 7b.

For a naive and surely wrong back-of-the-envelope calculation, the mortgage rejection rate estimate of 0.658 along with the average IAT score of 0.39 implies that if whites' racial bias against blacks as measured by the IAT fell to 0, the rejection rate gap would fall by roughly 25 percentage points, equivalent to the actual rejection rate gap. Such extrapolation likely over-estimates the true counterfactual given endogenous responses and that my IV estimate is a LATE, but nevertheless demonstrates that the IV estimate is sizable.

5.3 Relation to previous literature

Example papers that show racial discrimination in various real life settings: Bertrand et al. (2005); Donohue III and Levitt (2001); Giuliano et al. (2011); Price and Wolfers (2010); Stoll et al. (2004)

In this section, I describe conditions under which I can interpret visible black politicians' close electoral victory as an instrumental variable for racial prejudice. This methodology allows for a different way to identify racial discrimination, under potentially more plausible exclusion restrictions.

The economics literature on racial discrimination using observational data mostly runs regressions of the following form (Charles and Guryan (2011)):

$$Y_i = \beta \text{black}_i + X_i \Gamma + \tilde{\varepsilon}_i$$

where Y_i is some outcome in the market in which we are studying discrimination (e.g. mortgage rejection), black_i is individual i's race, and X_i is a vector of controls such as income. In terms of difference between black and white individuals and taking within-group averages, an equivalent expression is:

$$\text{gap} \equiv E\left[Y_{i,\text{black}} - Y_{i,\text{white}}\right] = \beta + E\left[X_{i,\text{black}} - X_{i,\text{white}}\right]\Gamma$$

In either expression, the discrimination coefficient $\hat{\beta}$ is a residual from controlling for other characteristics. There is a well-known over-controlling problem here: if an area is highly prejudiced in both the mortgage market and the labor market, blacks may both earn less than whites and be rejected mortgages more frequently. In such a setting, it is possible for the specification to estimate $\hat{\beta} = 0$, even though mortgage market too practices racial discrimination. In other words, $\hat{\beta}$ from such specifications can only tell us how discriminatory the market for Y_i is relative to the markets for the controls in X_i.

Audit studies are a way to get around this issue by essentially "shocking" black_i by experimentally varying the perceived race of the applicant. By submitting identical applications but only varying the name of the applicant, these studies estimate

$$Y_i = \beta \hat{\text{black}}_i + \varepsilon_i$$

where $\hat{\text{black}}_i$ can be thought of as the perceived probability that an applicant is black given that the name is Lakisha and Jamal as opposed to Emily and Greg (Bertrand and Mullainathan (2004)).

A handful of papers take a different approach. Starting from the insight that taste for discrimination is a combination of the target of discrimination and the intensity of the racial prejudice, they estimate regressions of the form:

$$Y_{ij} = \delta \text{black}_i + \beta \text{black}_i \times \text{prejudice}_j + \varepsilon_{ij}$$

by using local measures of prejudice_j. Charles and Guryan (2008) uses General Social Survey responses to get geographical variations in prejudice_j. Acharya et al. (2016) instruments for prejudice_j using historical prevalence of slavery driven by cotton production intensity in 1860.

In this paper, I extend this second approach. I associate whole time-places with a level of racial prejudice and find proxies for it. Then, without including other controls,

$$Y_{ijt} = \delta \text{black}_i + \beta \text{black}_i \times \hat{\text{prejudice}}_{jt} + \varepsilon_{ijt}$$

or equivalently in black-white differences,

$$\text{gap}_{jt} = \delta + \beta \hat{\text{prejudice}}_{jt} + \tilde{\varepsilon}_{jt}$$

That is, if black households' mortgage rejections rise in an area as people in that area exhibit higher racial bias as measured by the Race IAT, I associate that co-movement with racial discrimination. Conceptually this method compares different geographical areas as opposed to different markets (for example, mortgage market against labor market as above). As a result, I cannot conclude which market in a geographical area exhibits prejudice; I can only say that city X has more discrimination than city Y. By instrumenting for prejudice_{jt} using close electoral victory of black politicians, I can make a causal claim for β. To my knowledge, this is the first time-varying instrument for racial prejudice in observational data.

6 Conclusion

The Obama presidency has motivated questions as to how having a visible black leader affects white Americans' attitude toward black Americans. Given the theoretical ambiguity, I test for causal impact of a black leader on racial attitudes using local elections of black politicians at the municipal level. Using Race Implicit Attitude Test (IAT) scores as a measure of racial prejudice and close election regression discontinuity (RD) design for causal inference, I find that electoral victory of a black leader leads to a rise in racial prejudice among white Americans against black Americans. Following a close electoral victory, the IAT D score rises by about 0.03, or 7% of the average black-white difference. Simultaneously, using the same RD design, black politicians' electoral victory causes lower employment and higher mortgage denial for black Americans relative to white Americans. By ruling out other channels by which electoral victory could adversely affect black Americans' relative economic outcome, I argue that the rise in prejudice caused black-white economic inequality to widen.

References

Abney, F. G. and Hutcheson Jr, J. D. (1981). Race, representation, and trust: Changes in attitudes after the election of a black mayor. *Public Opinion Quarterly*, 45(1):91–101.

Acharya, A., Blackwell, M., and Sen, M. (2016). The political legacy of american slavery. *The Journal of Politics*, 78(3):621–641.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: does exposure reduce bias? *The Quarterly journal of economics*, 124(4):1497–1540.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586.

Becker, G. S. ([1957] 1971). *The Economics of Discrimination*. University of Chicago press.

Bertrand, M., Chugh, D., and Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, 95(2):94–98.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.

Bobo, L. (1988). Group conflict, prejudice, and the paradox of contemporary racial attitudes. In *Eliminating racism*, pages 85–114. Springer.

Brollo, F. and Troiano, U. (2016). What happens when a woman wins an election? evidence from close races in brazil. *Journal of Development Economics*, 122:28–45.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

Charles, K. K. and Guryan, J. (2008). Prejudice and wages: an empirical assessment of becker's the economics of discrimination. *Journal of Political Economy*, 116(5):773–809.

Charles, K. K. and Guryan, J. (2011). Studying discrimination: Fundamental challenges and recent progress. *Annu. Rev. Econ.*, 3(1):479–511.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2018). Race and economic opportunity in the united states: An intergenerational perspective. Technical report, National Bureau of Economic Research.

Comenetz, J. (2016). Frequently occurring surnames in the 2010 census.

DellaVigna, S. (2010). The obama effect on economic outcomes: Evidence from event studies. Technical report, Citeseer.

Donohue III, J. J. and Levitt, S. D. (2001). The impact of race on policing and arrests. *The Journal of Law and Economics*, 44(2):367–394.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2):135–150.

Effron, D. A., Cameron, J. S., and Monin, B. (2009). Endorsing obama licenses favoring whites. *Journal of experimental social psychology*, 45(3):590–593.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and Snyder, J. M. (2015). On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science*, 59(1):259–274.

Ferreira, F. and Gyourko, J. (2009). Do political parties matter? evidence from us cities. *The Quarterly journal of economics*, 124(1):399–422.

Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? the case of us mayors. *Journal of Public Economics*, 112:24–39.

Fu, S., He, H., and Hou, Z.-G. (2014). Learning race from face: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 36(12):2483–2509.

Giuliano, L., Levine, D. I., and Leonard, J. (2011). Racial bias in the manager-employee relationship an analysis of quits, dismissals, and promotions at a large retail firm. *Journal of Human Resources*, 46(1):26–52.

Gladwell, M. (2007). *Blink: The power of thinking without thinking*. Back Bay Books.

Goldman, S. K. and Mutz, D. C. (2014). *The Obama effect: How the 2008 campaign changed white racial attitudes*. Russell Sage Foundation.

Gormley, T. A. and Matsa, D. A. (2011). Growing out of trouble? corporate responses to liability risk. *The Review of Financial Studies*, 24(8):2781–2821.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6):1464.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., and Banaji, M. R. (2009). Understanding and using the implicit association test: Iii. meta-analysis of predictive validity. *Journal of personality and social psychology*, 97(1):17.

Guiso, L., Sapienza, P., and Zingales, L. (2011). Civic capital as the missing link. In *Handbook of social economics*, volume 1, pages 417–480. Elsevier.

Hodson, G. and Busseri, M. A. (2012). Bright minds and dark attitudes: Lower cognitive ability predicts greater prejudice through right-wing ideology and low intergroup contact. *Psychological science*, 23(2):187–195.

Hopkins, D. J. and McCabe, K. T. (2012). After it's too late: estimating the policy impacts of black mayoralties in us cities. *American Politics Research*, 40(4):665–700.

Hurwitz, J. and Peffley, M. (1997). Public perceptions of race and crime: The role of racial stereotypes. *American journal of political science*, pages 375–401.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Keller, E. J. (1978). The impact of black mayors on urban policy. *The Annals of the American Academy of Political and Social Science*, 439(1):40–52.

Kinder, D. and Chudy, J. (2016). After obama. In *The Forum*, volume 14, pages 3–15. De Gruyter.

Kouchaki, M. (2011). Vicarious moral licensing: the influence of others' past moral actions on moral behavior. *Journal of personality and social psychology*, 101(4):702.

Kurdi, B., Seitchik, A. E., Axt, J., Carroll, T., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., and Banaji, M. R. (2018). Relationship between the implicit association test and intergroup behavior: A meta-analysis.

Lee, D. S., Moretti, E., and Butler, M. J. (2004). Do voters affect or elect policies? evidence from the us house. *The Quarterly Journal of Economics*, 119(3):807–859.

Marschall, M. J. and Ruhil, A. V. (2006). The pomp of power: Black mayoralties in urban america. *Social Science Quarterly*, 87(4):828–850.

Monin, B. and Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of personality and social psychology*, 81(1):33.

Neggers, Y. (2018). Enfranchising your own? experimental evidence on bureaucrat diversity and election bias in india. *American Economic Review*, 108(6):1288–1321.

Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, 1(1):65–92.

Nye, J. V., Rainer, I., and Stratmann, T. (2014). Do black mayors improve black relative to white employment outcomes? evidence from large us cities. *The Journal of Law, Economics, & Organization*, 31(2):383–430.

Olzak, S. (1990). The political context of competition: Lynching and urban racial violence, 1882–1914. *Social Forces*, 69(2):395–421.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., and Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of iat criterion studies. *Journal of personality and social psychology*, 105(2):171.

Parker, C. S. (2016). Race and politics in the age of obama. *Annual Review of Sociology*, 42:217–230.

Price, J. and Wolfers, J. (2010). Racial discrimination among nba referees. *The Quarterly journal of economics*, 125(4):1859–1887.

Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.

Saltzstein, G. H. (1989). Black mayors and police policies. *The journal of Politics*, 51(3):525–544.

Sidanius, J., Devereux, E., and Pratto, F. (1992). A comparison of symbolic racism theory and social dominance theory as explanations for racial policy attitudes. *The Journal of Social Psychology*, 132(3):377–395.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics*, 118:26–40.

Stoll, M. A., Raphael, S., and Holzer, H. J. (2004). Black job applicants and the hiring officer's race. *ILR Review*, 57(2):267–287.

Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic association*, 8(4):677–716.

Tesler, M. (2016). *Post-racial or most-racial?: Race and politics in the Obama Era*. University of Chicago Press.

Tesler, M. and Sears, D. O. (2010). *Obama's race: The 2008 election and the dream of a post-racial America*. University of Chicago Press.

Tzioumis, K. (2018). Demographic aspects of first names. *Scientific data*, 5:180025.

Washington, E. (2006). How black candidates affect voter turnout. *The Quarterly Journal of Economics*, 121(3):973–998.

Word, D. L., Coleman, C. D., Nunziata, R., and Kominski, R. (2008). Demographic aspects of surnames from census 2000. *Unpublished manuscript, Retrieved from http://citeseerx. ist. psu. edu/viewdoc/download*.

Xu, K., Nosek, B., and Greenwald, A. (2014). Psychology data from the race implicit association test on the project implicit demo website. *Journal of Open Psychology Data*, 2(1).

Figure 1: Screenshots from Our Campaigns

The figures below are screenshots from the Our Campaigns website (www.ourcampaigns.com/). Panels (a) and (b) are for Chicago's fifth city ward, with panel (a) showing the political position's main page with the current officeholder and the map of the jurisdiction along with other details, and panel (b) showing the electoral race history, with the date, type, candidates and their count and shares of votes received for each electoral race. Panel (c) is the candidate page for Harold Washington, who was the first black mayor of Chicago. Among other personal details, "Tags" contains his race.

(a) Political office example

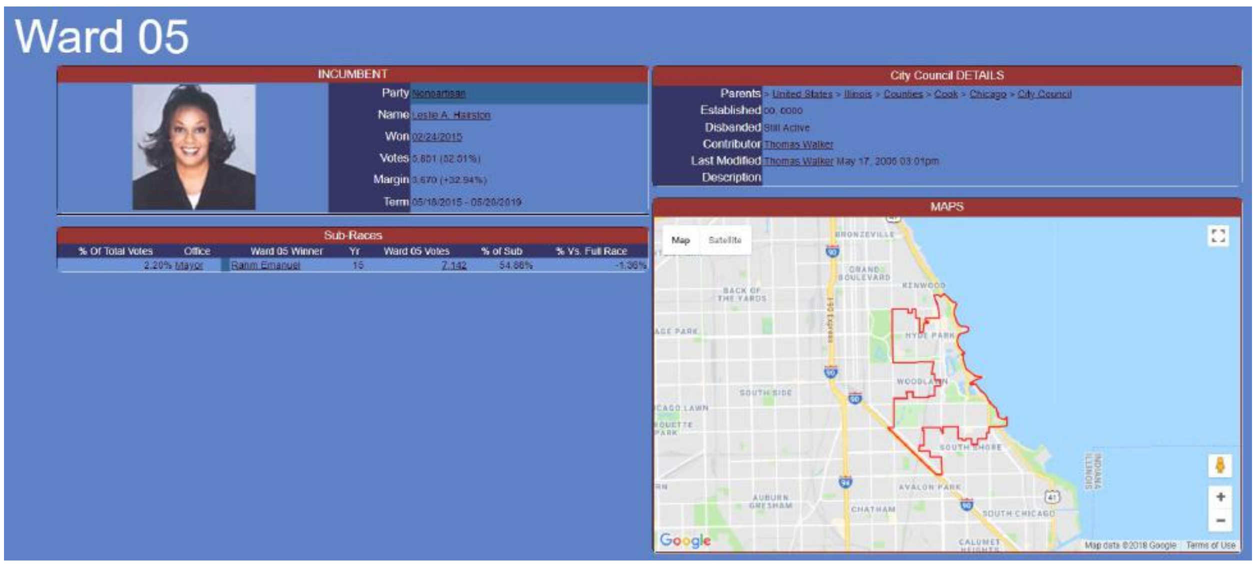


Figure 1: Screenshots from Our Campaigns (continued)

(b) Electoral race history example

RACES [Show Primaries]

Date	Type	Results						
Feb 24, 2015	General Election	Leslie A. Hairston(I) 5,851 52.51%	Anne Marie Miles 2,181 19.57%	Tiffany N. Brooks 891 8.00%	Jocelyn Hare 821 7.37%	Jedidiah L. Brown 792 7.11%	Robin Boyd Clark 599 5.38%	Write-In (W) 8 0.07%

Date	Type	Results				
Feb 22, 2011	General Election	Leslie A. Hairston(I) 7,217 61.77%	Anne Marie Miles 2,489 21.30%	Glenn Ross 826 7.07%	Carolyn Hightower Chalmers 701 6.00%	Michele A. Tankersley 451 3.86%
Feb 27, 2007	General Election	Leslie A. Hairston(I) 6,748 74.67%	Oscar Worrill 1,769 19.58%	Sylvester "Junebug" Hendricks 520 5.75%		
Feb 25, 2003	General Election	Leslie A. Hairston(I) 6,355 71.93%	Oscar Worrill 1,073 12.14%	Carolyn Hightower Chalmers 713 8.07%	Anthony T. Blair 694 7.86%	

(c) Candidate example

Washington, Harold

CANDIDATE DETAILS

Affiliation	Democratic
Name	Harold Washington
Address	Chicago, Illinois , United States
Email	None
Website	None
Born	April 15, 1922
Died	November 25, 1987 (65 years)
Contributor	Wishful Thinking
Last Modifed	RBH Jan 31, 2016 04:33am
Tags	Black - Divorced - Methodist -
	Harold Washington (1922-1987) was the first African-American mayor of Chicago.

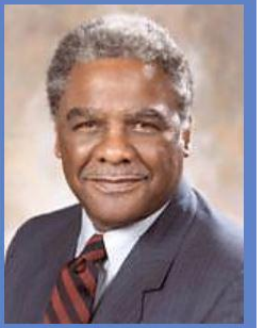
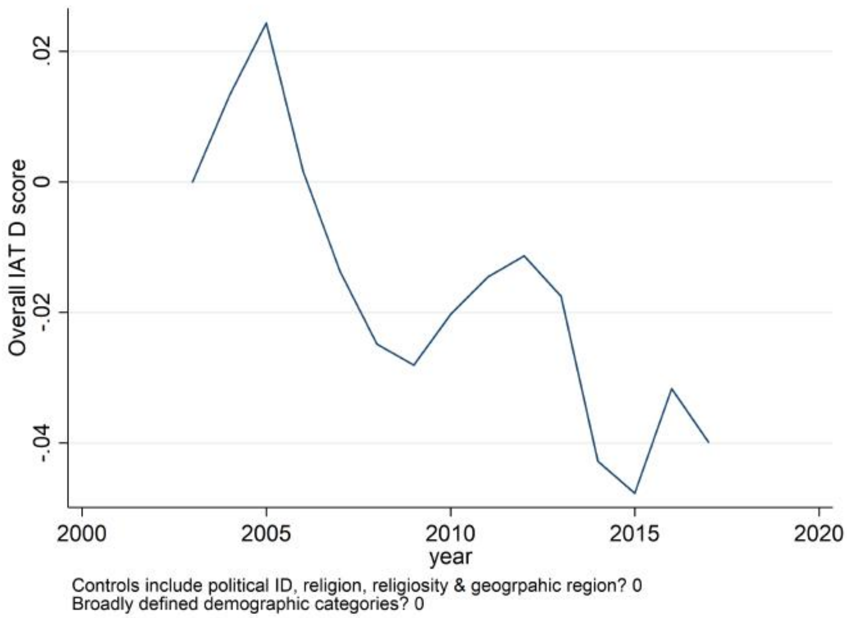


Figure 2: IAT data from Project Implicit Database

Panel (a) plots the average Race Implicit Association Test (IAT) D score for white respondents from Project Implicit Database. The average has been taken after projecting out dummies for age (each whole number), nine education buckets and gender. Panel (b) plots the distribution of raw Race IAT D score from Project Implicit Database for 2002-2017. The D score has a possible range of -2 to +2, where higher number indicates bias against black Americans. The vertical red lines indicate break points for a common description of pro-white bias. [CHECK which ranges indicate which colloquial descriptor of bias]

(a) Composition-adjusted, for white respondents



(b) Distribution (raw)

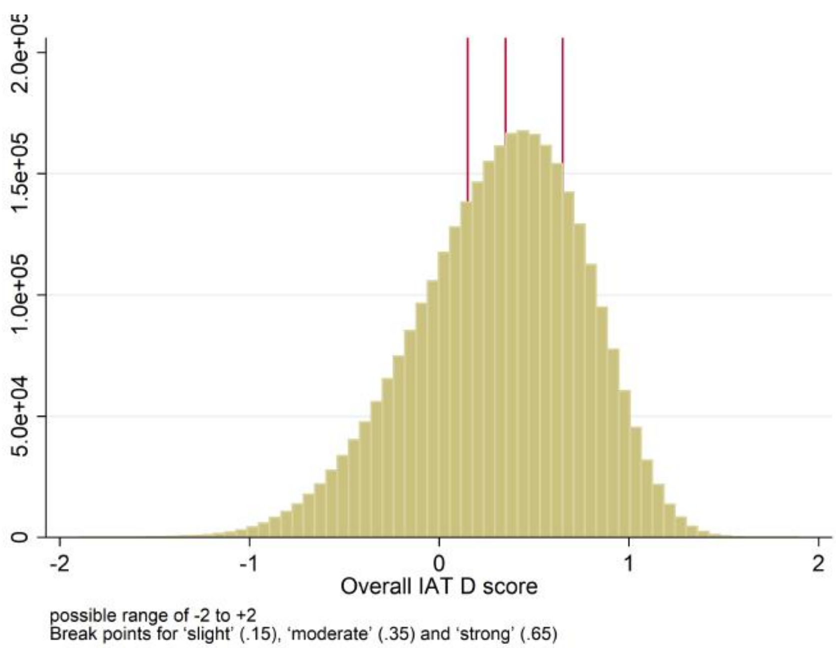
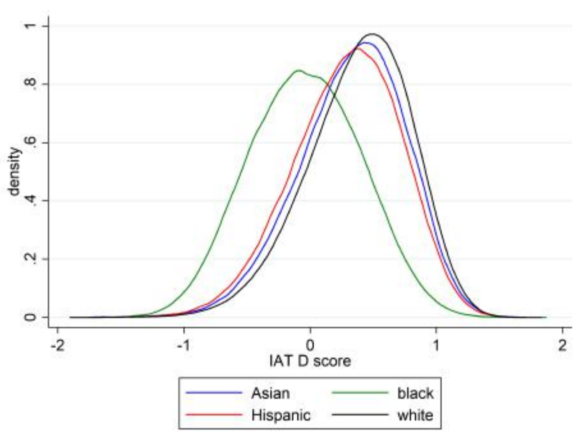


Figure 3: IAT distribution

Panel (a) plots the distribution of raw Race IAT D score from Project Implicit Database for 2002-2017. The D score has a possible range of -2 to +2, where higher number indicates bias against black Americans. The vertical red lines indicate break points for a common description of pro-white bias. [CHECK which ranges indicate which colloquial descriptor of bias] The next five panels plot the raw Race IAT D score by demographic sub-groups for 2002-2017. Panel (b) splits the Race IAT scores by the response to the question, “What brought you to this website.” “Voluntary” responses in green line include “Recommendation of a friend or co-worker,” “Mention or link at a non-news Internet site,” “Mention in a news story (any medium)” or “My Internet search for this or a related topic.” “Mandatory” responses in blue line include “Assignment for work” or “Assignment for school.” Responses of “other” and those without a response to this question were classified as “unknown/other” in the red line. Panel (c) splits the Race IAT scores by education and plots two groups: the green histogram is for those with up to a high school diploma, while the white histogram is for those with bachelor’s degrees or higher. Panel (d) splits the Race IAT scores by gender: the green histogram is for females, while the white histogram is for males. Panel (e) splits the Race IAT scores by self-stated political ideology along a 7-point scale: the red histogram is for those who responded with any degree of “conservative”, the blue histogram is for those who responded with any degree of “liberal”, while the white histogram is for those who responded “neutral.” Panel (f) splits the Race IAT scores by self-stated race: the blue line is for Asians, the green line is for black Americans, the red line is for those of Hispanic origin, and the black line is for white Americans.

(a) By race



(b) By where respondents came from

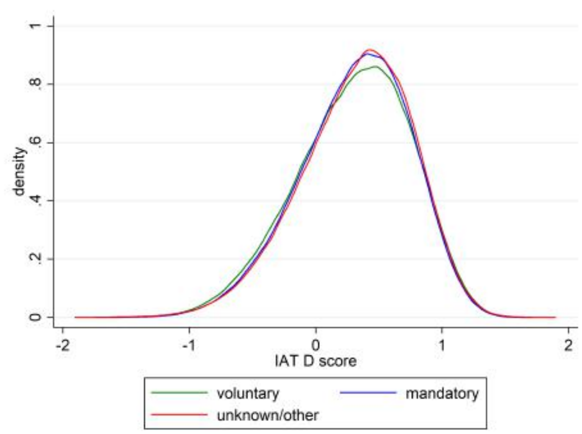
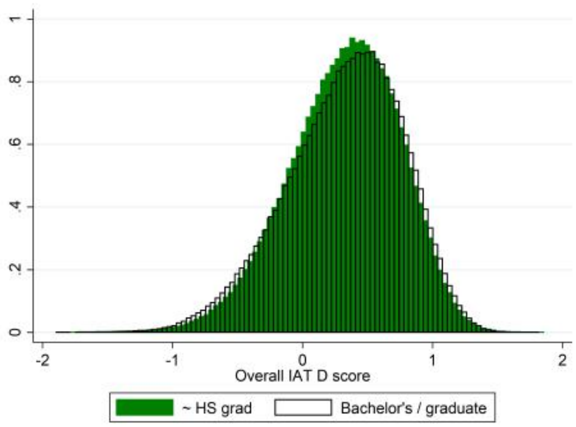
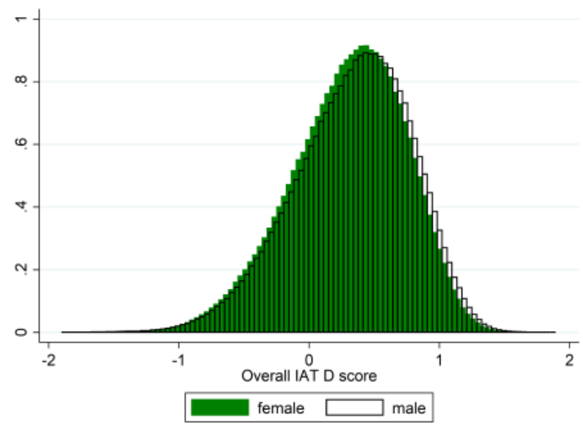


Figure 3: IAT distribution (continued)

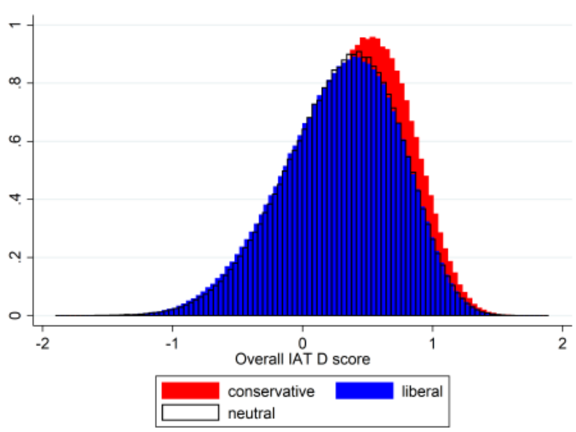
(c) By education



(d) By gender



(e) By political ideology



(f) By region (white only)

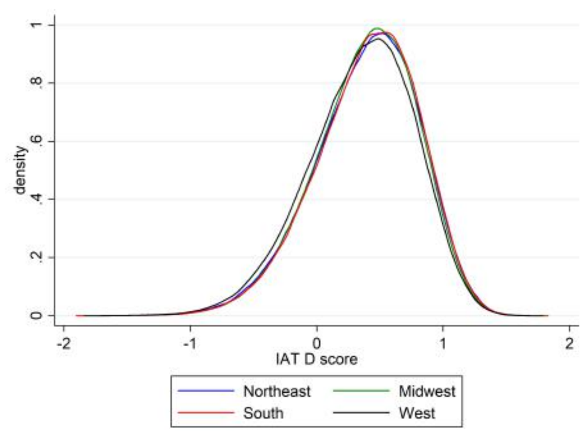
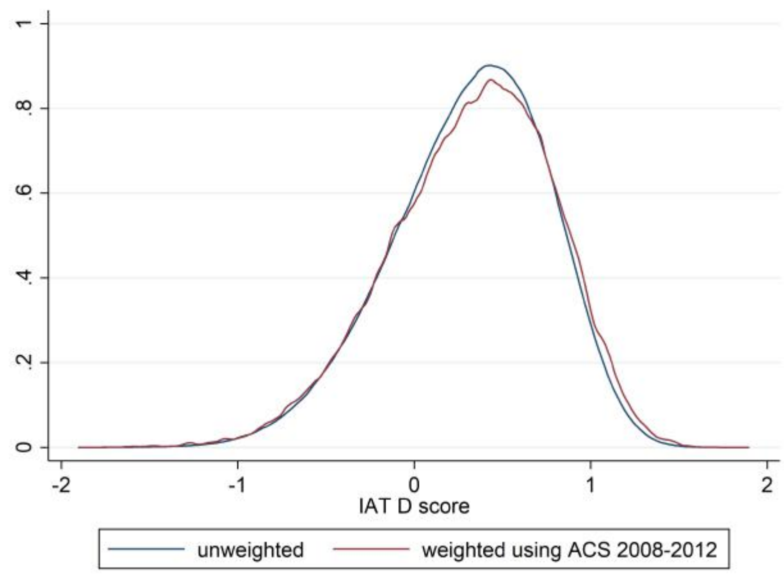


Figure 4: IAT re-weighted using ACS

Both figures plot the distribution of Race IAT D scores, using no weight (blue line) and using weights imputed using the 2008-2012 American Community Survey (ACS) to make the sample representative along the demographic variables of gender $\times$ age $\times$ education (red line). Panel (a) plots the distribution for all observation; panel (b) plots the distribution only for white respondents.

(a) All races



(b) White only

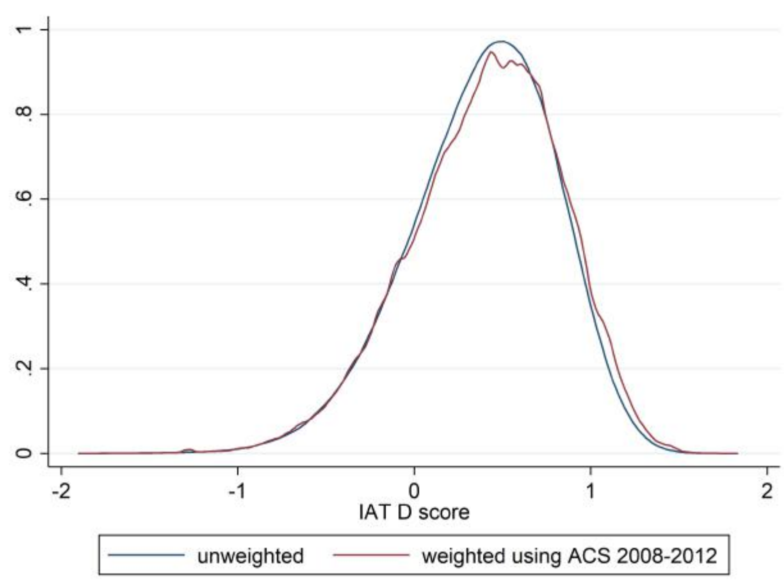
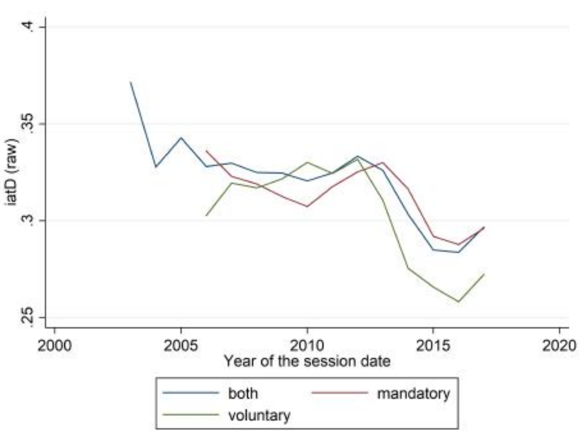


Figure 5: IAT time series by source

REPLACE SOME WITH REPRESENTATIVE RE-WEIGHTED (from time series version of iat-LocalPanel)

Panels (a), (b), (c), (d) plot average Race IAT D scores for sub-groups from the Project Implicit Database; panels (e) and (f) on the bottom row plot counts of responses for sub-groups. Panels (a), (c) and (e) on the left column are for all responses; panels (b), (d) and (f) on the right column are for responses with self-reported race of non-Hispanic white. Within each plot, separate lines are sub-groups based on the responses to the question, “What brought you to this website.” “Voluntary” responses in green line include “Recommendation of a friend or co-worker,” “Mention or link at a non-news Internet site,” “Mention in a news story (any medium)” or “My Internet search for this or a related topic.” “Mandatory” responses in red line include “Assignment for work” or “Assignment for school.” The blue lines are for both of these sub-groups. Plots (a) and (b) took sub-group averages of the raw Race IAT D score; plots (c) and (d) took the averages after projecting out dummies for age (each whole number), nine education buckets and gender.

(a) All: Raw



(b) White only: Raw

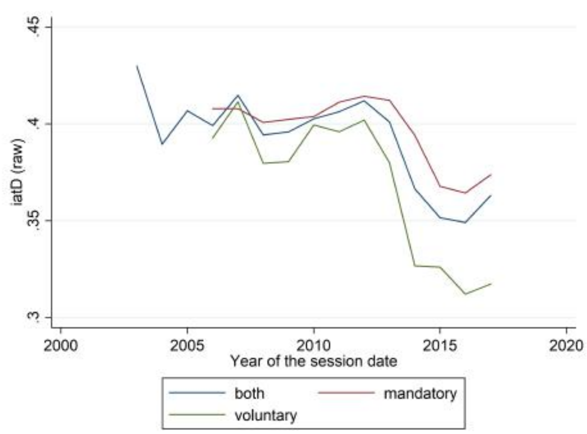
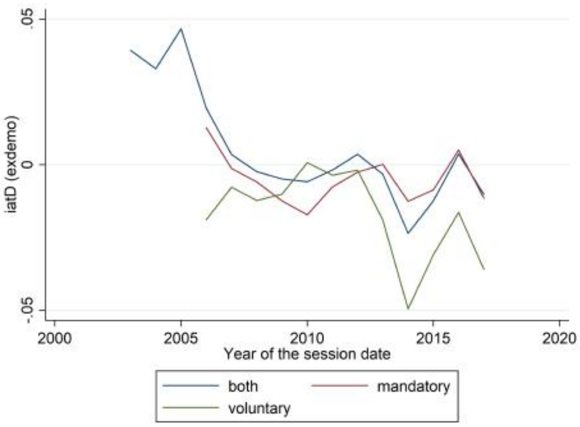
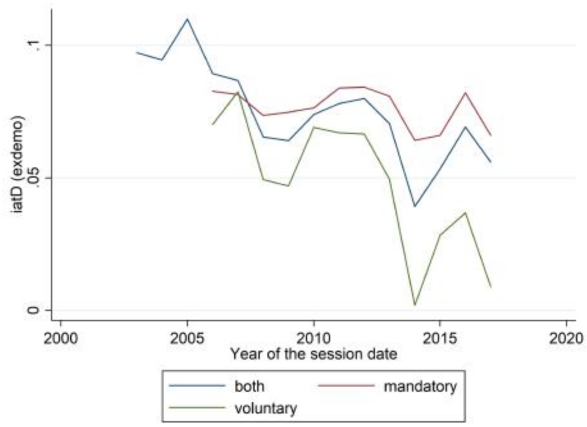


Figure 5: IAT time series by source (continued)

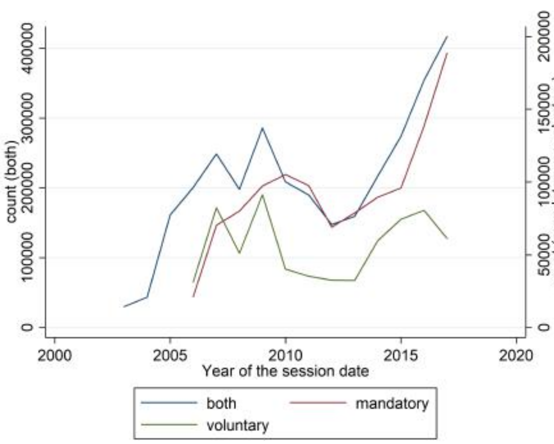
(c) All: Ex demo



(d) White only: Ex demo



(e) All: Count



(f) White only: Count

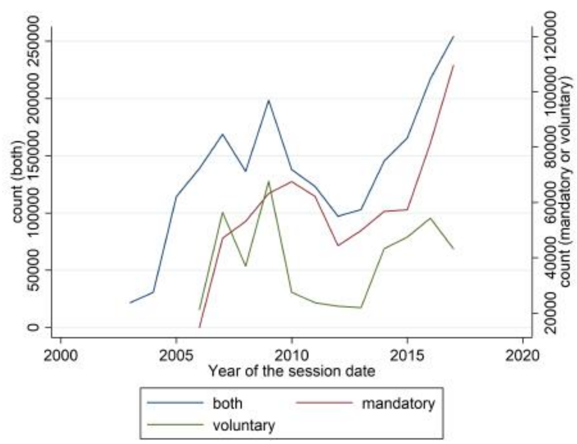
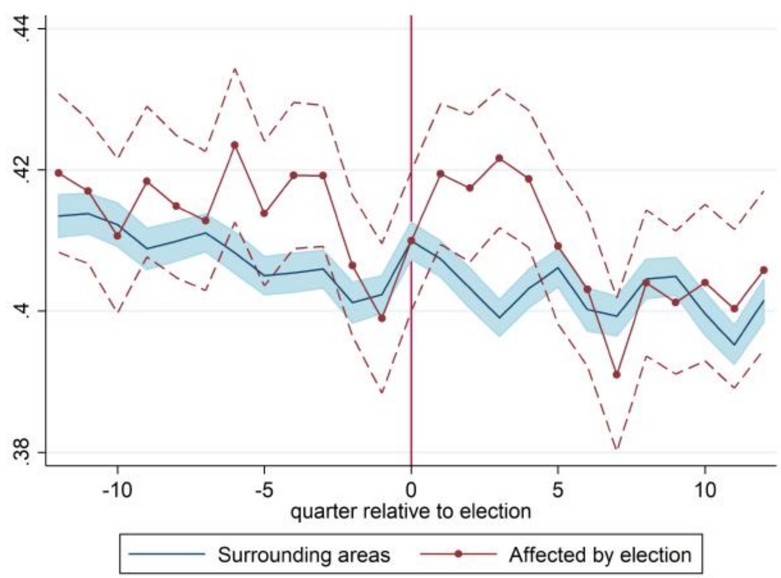


Figure 6: Difference-in-difference

All panels plot the average Race IAT D score for white respondents living in the areas affected by a set of elections in connected red dots, against surrounding areas in the same state in solid blue line. Panels (a) and (c) on the left are for elections where a black candidate won against a white runner-up; panels (b) an (d) on the right are for elections where a black runner-up lost to a white candidate. In the top panels (a) and (b), raw Race IAT D scores are used; in the bottom panels (c) and (d), the Race IAT D scores have been first residualized by dummies for age (each whole number), nine education buckets and gender, and then re-weighted by the same three demographic variables to be representative of the US population using the American Community Survey for 2008-2012. The dotted red lines and and light blue shaded area show standard errors for the areas affected and the surrounding areas, respectively.

(a) Black winner: Raw



(b) Black runner-up: Raw

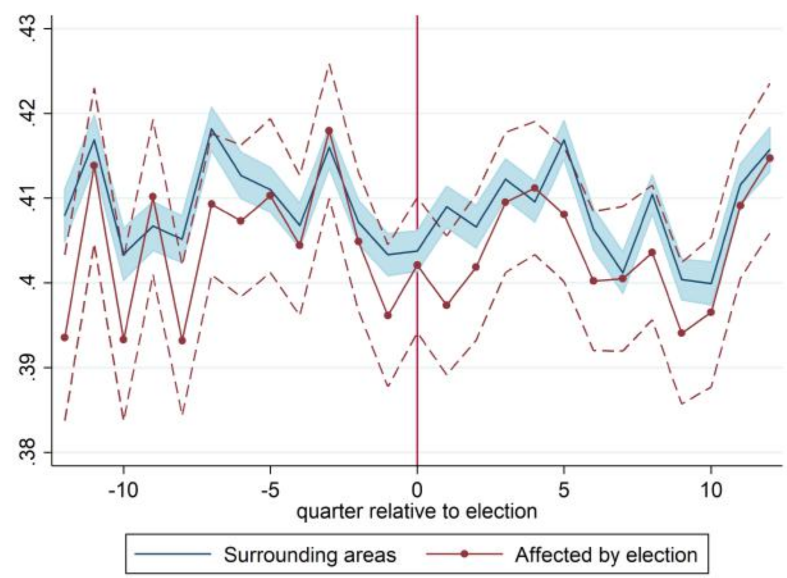
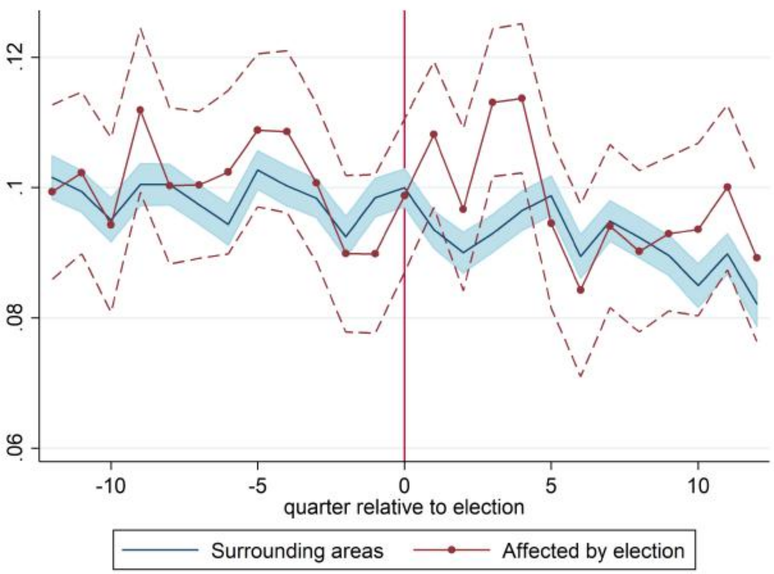


Figure 6: Difference-in-difference (continued)

(c) Black winner: Composition-adjusted



(d) Black runner-up: Composition-adjusted

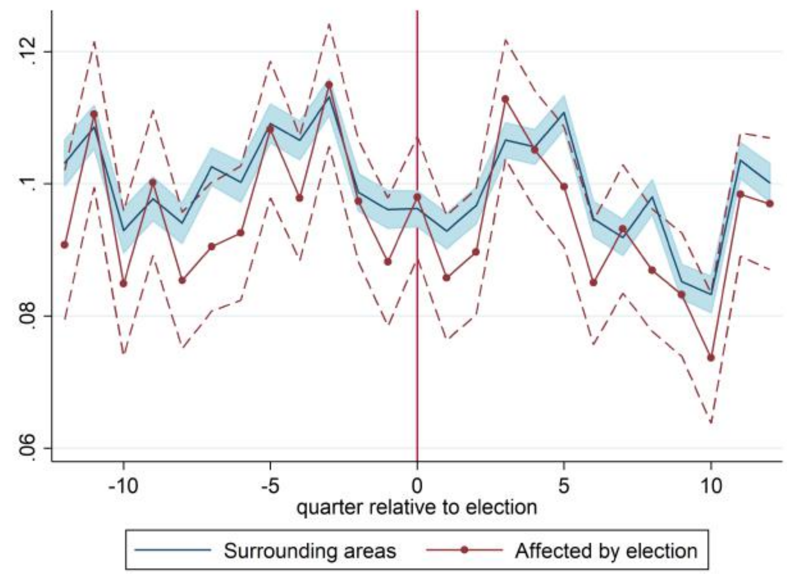
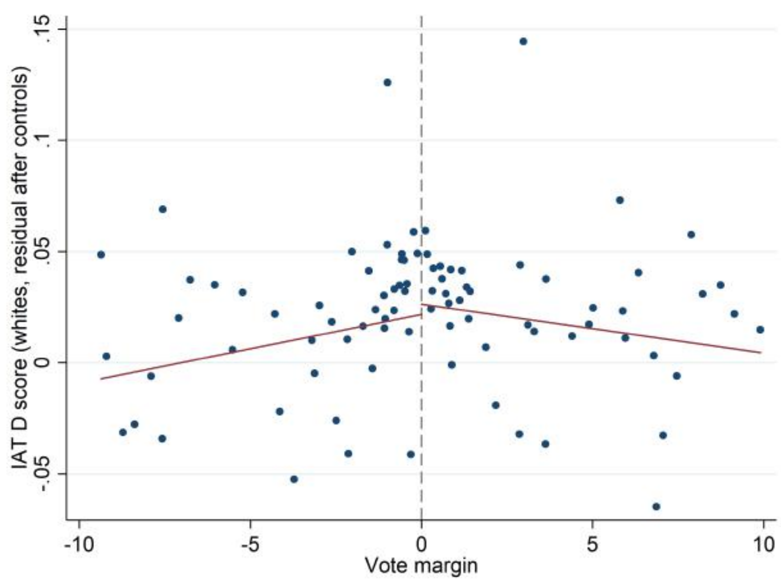


Figure 7: Regression Discontinuity

Both panels plot bin scatters of average Race IAT D score among white respondents from Project Implicit Database: panel (a) is for the 3-year window before an election; panel (b) is for the 3-year window after the election (i.e. placebo period). Each point represents within-bin averages of Race IAT D scores across election-county-month level observations), where the bins are defined by centiles sorted on the vote margin in the corresponding election. The sample has been restricted to elections where the winner and the runner-up include one black and one white candidate; vote margin is defined as the percentage point difference between the black candidate and the white candidate (i.e. positive vote margins represent elections where the black candidate won, while negative vote margins represent elections where the black candidate lost to the white candidate). The domain has been restricted to elections where the vote margin was within 10%. The averages were formed using Race IAT D scores after projecting out dummies for age (each whole number), nine education buckets and gender. The vertical dotted line divides black candidates' losses on the left to their victories on the right. The red lines are best linear fit on each side. The vertical distance along the vertical dotted line between the two red lines is a rough estimate of the regression discontinuity estimate.

(a) Before the election



(b) After the election

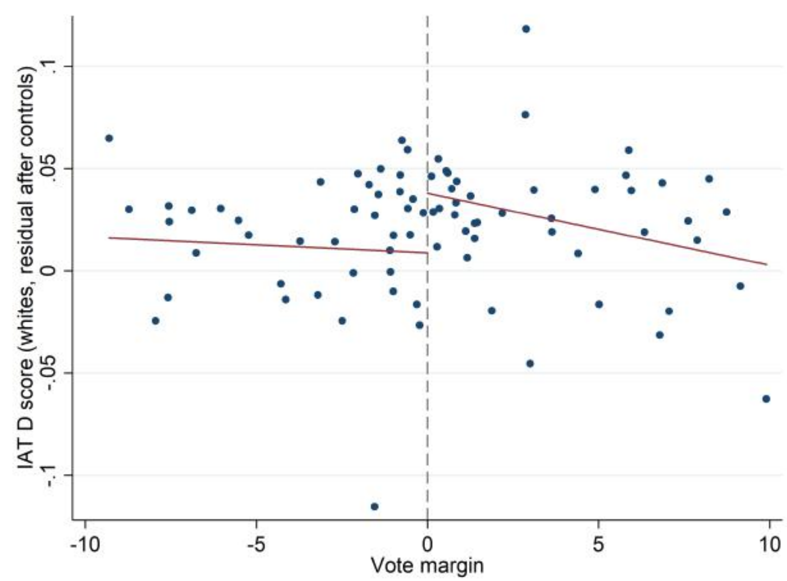
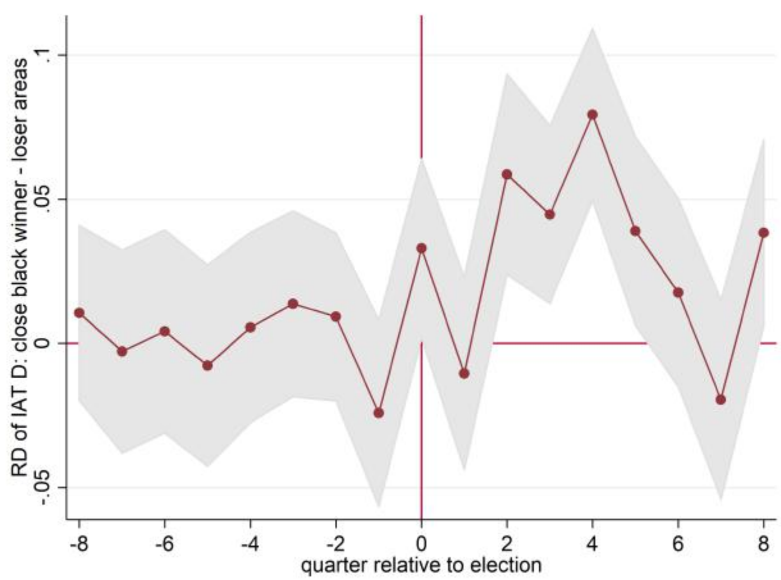


Figure 8: Time series of discontinuity

Both panels plot time series of regression discontinuity estimates of the Race IAT D score between areas with a close black winner and a close black loser, for non-overlapping 3-month windows relative to the election. Each red point is the estimate, with 95% confidence intervals shaded in gray. The vertical red line at 0 indicates the 3-month period starting from the month of election; points to the right of the vertical line represents non-overlapping 3-month periods after the election. The regression discontinuity estimates have been estimated from by using elections where the vote margin between the winner and the runner-up was within 10%, with linear controls on either side of the discontinuity. Regressions have been run on observations at the election-county-month level, with each observation weighted by the fraction of the county affected by the eleciton.

(a) Raw



(b) Controlling for demographic composition

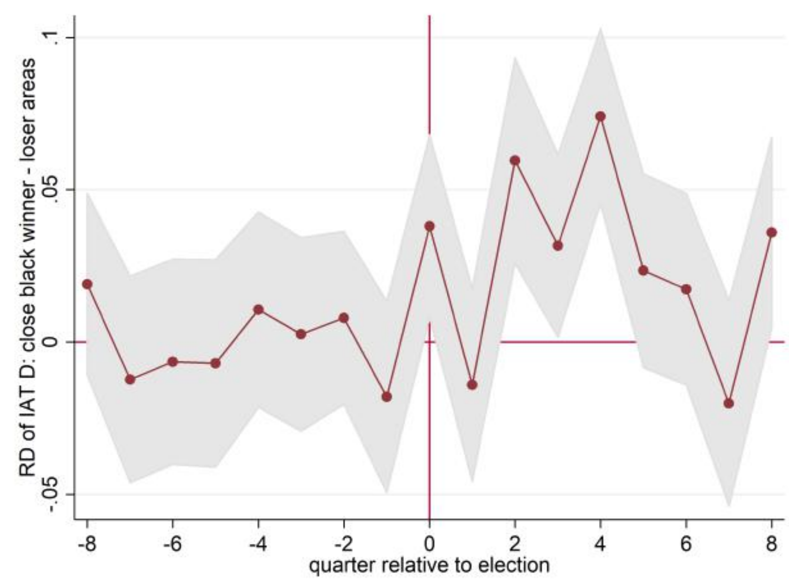


Table 1: Our Campaigns

Panel (a) shows the number of elections falling into each category, for years 2003 and onward. Row titles indicate which type of office the elections in that row are for. The first two columns are for all electoral races; the third and fourth columns are for electoral races in which the winner and the runner-up include one white and one black candidate; the last two columns are such electoral races between black and white candidates that are close. Close election is defined here as those with a vote margin between the winner and the runner-up of 10% or less of votes. Odd columns show counts of all such elections; even columns show the count of elections in which the winner is a black candidate. The last row with the total counts show that for close elections between a black and a white candidate, the probability that the black candidate wins is roughly half.
Panel (b) shows. For all candidates who make it among the top two candidates in the electoral races in panel (a), after 2002.

(a) Offices

		all races		black-white races		close black-white races	
Gov level	Office type	all	black winner	all	black winner	all	black winner
federal	house	2,933	202	111	78	11	4
state	governor	308	12	26	11	15	9
	senate	9,362	655	471	156	51	24
	house	33,046	2,426	1,518	486	145	78
	state other	1,428	92	147	49	88	37
local	county other	7,350	475	448	146	77	34
	city other	4,309	601	277	148	72	35
	mayor	2,755	215	166	74	42	26
Total		61,491	4,678	3,164	1,148	501	247

(b) Candidate race identification

		Our Campaigns tag				
Method	race	Asian	Black	Hispanic	White	no tag
	total count by tag	39	276	187	921	49,938
Facial recognition	Asian	0.67	0.06	0.10	0.03	0.03
	Black	0.15	0.68	0.06	0.01	0.04
	White	0.13	0.19	0.78	0.92	0.48
	unidentifiable	0.00	0.02	0.01	0.03	0.01
	no photo	0.05	0.05	0.04	0.01	0.45
Surname	Asian	0.49	0.00	0.00	0.00	0.01
	Black	0.00	0.02	0.00	0.00	0.00
	Hispanic	0.05	0.01	0.66	0.01	0.03
	White	0.00	0.12	0.11	0.58	0.53
	mixed	0.41	0.76	0.14	0.33	0.34
	uncommon/unmatched	0.05	0.08	0.09	0.08	0.09

Table 2: IAT distribution

Panel (a) shows the demographic distribution in the 2008-2012 5-year American Community Survey and 2008-2012 Race IAT data from Project Implicit Database, along gender, age bins and educational attainment. The ACS statistics have been accessed via th e Integrated Public Use Microdata Series (IPUMS), in order to get counts interacting the three demographic variables. Weights to make the IAT sample more representative of the American population were imputed as $\frac{\text{ACS population share}}{\text{IAT sample share}}$ for gender $\times$ age $\times$ education bins, fully interacted. Single-year ages were used, and educational attainment has been grouped into the categories displayed in panel (a). The last column of panel (a) displays the average weight for the broader demographic subgroup weighted using the IAT sample share; these weights are equivalent to the ratio of the first two columns.

(a) Distribution vs ACS

demographic variable		share 2008-2012		average
		ACS	IAT	IAT weight
gender	male	0.49	0.41	1.2
	female	0.51	0.59	.87
age	unknown	0.00	0.01	0
	under 19	0.25	0.23	1.1
	19-22	0.05	0.26	.2
	23-29	0.09	0.20	.47
	30-39	0.13	0.15	.89
	40-49	0.15	0.09	1.7
	50-59	0.13	0.05	2.7
	60+	0.18	0.02	9.5
education	no HS grad	0.27	0.14	1.9
	HS grad	0.29	0.09	3.2
	some college / associate's	0.17	0.46	.38
	bachelor's	0.13	0.13	.99
	graduate	0.07	0.06	1.3
	N/A	0.08	0.12	.65

Table 2: IAT distribution (continued)

Panel (b) show the counts and characteristics of survey responses, grouped by the response to the question, “What brought you to this website,” from the Project Implicit Database. For each subgroup, I also report the average age and the average Race IAT D score, both raw and residualized by demographic fixed effects.

(b) By source of respondent

source		count	average age	IAT score	
grouping	actual response			raw	adjusted
mandatory	Assignment for work	142,367	37.6	0.27	-0.02
	Assignment for school	989,129	24.1	0.31	-0.00
	Sub-total	1,131,496	25.8	0.31	-0.01
voluntary	Recommendation of a friend or co-worker	304,153	30.2	0.30	-0.02
	Mention or link at a non-news Internet site	187,723	33.1	0.30	-0.01
	Mention in a news story (any medium)	132,086	37.6	0.30	-0.02
	My Internet search for this topic or a related topic	48,546	31.8	0.30	-0.01
	Sub-total	672,508	32.6	0.30	-0.02
unknown	.	1,238,011	26.1	0.33	0.02
	Other	93,791	32.3	0.30	-0.02
	Sub-total	1,331,802	26.5	0.33	0.01

Table 3: IAT validation

These tables report the cross-sectional and panel relationship between the local average of the Race IAT D score and the variables in column headers. The cross-section is either a county for panels (a) and (b), or a Nielsen Designated Market Area for panel (c). See text for definition and source for each variable used here. In panel (a), the sample is restricted to slave states in 1860, and state fixed effects are included, following Acharya et al. (2016).

(a) 1860 slavery (county)

	IAT D			Thermology white-black			Prefer white/black		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS	IV
slave share 1860	0.020**	0.029***	0.041***	0.224***	0.247***	0.222*	0.140***	0.139***	0.123**
	(2.08)	(2.58)	(2.92)	(3.23)	(2.76)	(1.94)	(3.89)	(3.04)	(2.09)
pop black		-0.020	-0.026*		-0.022	-0.009		0.007	0.015
		(-1.51)	(-1.86)		(-0.19)	(-0.07)		(0.11)	(0.22)
Adjusted R^2	0.039	0.039	0.039	0.101	0.078	0.078	0.154	0.141	0.141
Observations	1114	1111	1111	1114	1111	1111	1114	1111	1111

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: IAT validation (continued)

(b) Racial prejudice (explicit measures from Project Implicit Database)

	Thermology white-black		Prefer white/black	
	(1) cross-section	(2) panel	(3) cross-section	(4) panel
IAT D	1.299***	0.983***	0.705***	0.532***
	(26.52)	(22.79)	(26.67)	(20.91)
Adjusted R^2	0.135	0.208	0.184	0.252
Observations	3115	37806	3114	37357

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(c) Racial prejudice (other measures)

	DMA-level			county-level	
	(1) Google n-word	(2) Google KKK	(3) Spanking (b-w)	(4) crime w.o.b.	(5) GSS
IAT D	0.551***	0.299	0.013**	-3.175	0.319
	(2.62)	(1.57)	(2.18)	(-1.19)	(1.20)
Adjusted R^2	0.178	0.224	0.009	-0.001	0.061
Observations	204	207	105	1615	350

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Difference-in-difference

Using only elections with a black winner, and for election i, county j and year t, I estimate

$$Y_{ijt} = \alpha_{ij} + \alpha_{it} + \beta \, \{\text{in jurisdiction of winner}\}_{ij} \times \{\text{after election}\}_{it} + \eta_{ijt}$$

For each election i, I include the 3-year window before and after the election, and include all counties in the same state as the jurisdiction associated with the election.

(a) IAT scores

	Raw		Demo-adjusted		Both FE	
	(1) all	(2) mand	(3) all	(4) mand	(5) all	(6) mand
Treatment	0.006***	0.003	0.002	0.002		
	(2.81)	(0.91)	(0.83)	(0.46)		
Post	-0.005***	-0.001**	-0.006***	-0.001		
	(-10.78)	(-2.36)	(-10.90)	(-1.39)		
Treatment x Post	-0.001	-0.003	0.004	0.000	0.003	-0.003
	(-0.20)	(-0.68)	(1.00)	(0.11)	(1.06)	(-0.77)
Constant	0.408***	0.410***	0.098***	0.101***		
	(1210.98)	(904.88)	(261.59)	(201.58)		
Observations	2045102	1241535	1838082	1113151	1835660	1107872
Adjusted R^2	0.000	0.000	0.000	-0.000	0.020	0.021
FE demographic			O	O	O	O
ACS weights			O	O	O	O
FE year					O	O
FE county					O	O
Subsample?		mandatory		mandatory		mandatory

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Regression discontinuity: IAT D scores

Both panels display the main results from the regression discontinuity design. The main specification imposes a 10% bandwidth (i.e. regression sample only includes elections for which the vote margin is at most 10% between the winner and the runner-up). Optimal bandwidths are typically wider, and results with optimal bandwidths are reported in the appendix. For observations at election i, geography (e.g. county) j, and event time (e.g. month or year) t, the following regression specification is run:

$$Y_{ijt} = \alpha + \beta_1 1\{\text{vote margin} > 0\}_{it} + \gamma_0 [\text{vote margin}]^{-}_{it} + \gamma_1 [\text{vote margin}]^{+}_{it} + \varepsilon_{ijt}$$

where "vote margin" is the percentage point vote margin between the black candidate and the white candidate. The first indicator term denotes a dummy for elections in which the black candidate won. The next two terms are linear controls separately for elections where the white candidate won (i.e. negative vote margin) and where the black candidate won (i.e. positive vote margin). Observations are weighted by the fraction of the geography (e.g. county) affected by the election. Panel (a) displays the regression discontinuity results for Race IAT D scores from Project Implicit Database.

	Post							
	Raw		Demo-adjusted		Time FE		County FE	
	all	mand	all	mand	all	mand	all	mand
Black winner	0.025**	0.025**	0.027**	0.037***	0.023*	0.035***	0.012	0.013
	(2.50)	(2.16)	(2.22)	(2.98)	(1.91)	(2.92)	(1.34)	(1.25)
Margin (winner)	-0.004***	-0.003*	-0.002	-0.002	-0.002	-0.002	0.002	0.003
	(-2.61)	(-1.95)	(-0.97)	(-1.35)	(-1.19)	(-1.34)	(0.87)	(1.12)
Margin (loser)	0.001	-0.001	-0.001	-0.002	0.000	-0.002	-0.000	0.001
	(0.49)	(-0.56)	(-0.23)	(-0.99)	(0.05)	(-1.05)	(-0.26)	(0.50)
Observations	23590	18177	21462	16481	21462	16481	21444	16452
Adjusted R^2	0.002	0.001	0.001	0.002	0.026	0.033	0.075	0.070
	Pre							
Black winner	0.003	0.012	0.011	0.022	0.015	0.016	-0.003	-0.001
	(0.33)	(0.97)	(1.08)	(1.58)	(1.60)	(1.23)	(-0.29)	(-0.08)
Observations	23437	15319	22271	14421	22271	14421	22250	14390
Adjusted R^2	0.001	0.000	0.001	0.002	0.032	0.044	0.065	0.075
	Model specifications:							
FE demographic			O	O	O	O	O	O
ACS weights			O	O	O	O	O	O
FE year-month					O	O	O	O
FE county							O	O
only mandatory		O		O		O		O

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Economic gaps

Using only elections with a black winner, and for election i, county j and year t, I estimate

$$Y_{ijt} = \alpha_{ij} + \alpha_{it} + \beta \left\{\text{in jurisdiction of winner}\right\}_{ij} \times \left\{\text{after election}\right\}_{it} + \eta_{ijt}$$

For each election i, I include the 3-year window before and after the election, and include all counties in the same state as the jurisdiction associated with the election.

(a) Difference-in-difference

	(1) unemployment transition	(2) employment to pop	(3) rejection rate	(4) log origination to pop
Treatment x Post	-0.002	-0.000	-0.003	0.015*
	(-0.69)	(-0.13)	(-1.63)	(1.76)
Observations	546705	503597	427763	459933
Adjusted R^2	0.064	0.930	0.347	0.638

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Economic gaps (continued)

Panel (b) displays results from the regression discontinuity design. The main specification imposes a 10% bandwidth (i.e. regression sample only includes elections for which the vote margin is at most 10% between the winner and the runner-up). Optimal bandwidths are typically wider, and results with optimal bandwidths are reported in the appendix. For observations at election i, geography (e.g. county) j, and event time (e.g. month or year) t, the following regression specification is run:

$$Y_{ijt} = \alpha + \beta_1 1\{\text{vote margin} > 0\}_{it} + \gamma_0 [\text{vote margin}]^{-}_{it} + \gamma_1 [\text{vote margin}]^{+}_{it} + \varepsilon_{ijt}$$

where "vote margin" is the percentage point vote margin between the black candidate and the white candidate. The first indicator term denotes a dummy for elections in which the black candidate won. The next two terms are linear controls separately for elections where the white candidate won (i.e. negative vote margin) and where the black candidate won (i.e. positive vote margin). Observations are weighted by the fraction of the geography (e.g. county) affected by the election. Panel (a) displays the regression discontinuity results for Race IAT D scores from Project Implicit Database.

(b) Regression discontinuity

	labor		mortgage	
	(1) unemployment transition	(2) employment to pop	(3) log origination to pop	(4) rejection rate
Black winner	0.025**	-0.050	-0.256**	0.035**
	(2.24)	(-1.55)	(-2.32)	(2.38)
Margin (winner)	0.000	-0.008	-0.017	0.001
	(0.12)	(-1.58)	(-1.19)	(0.54)
Margin (loser)	-0.003**	0.012**	0.024*	-0.001
	(-2.24)	(2.27)	(1.83)	(-0.37)
Constant	-0.045***	0.025	-1.028***	0.178***
	(-4.96)	(1.00)	(-13.16)	(18.45)
Observations	4510	4376	4311	4166
Adjusted R^2	0.004	0.023	0.024	0.027

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Heterogeneity

Both panels display results from the regression discontinuity design, with heterogeneous effects. The specification imposes a 10% bandwidth (i.e. regression sample only includes elections for which the vote margin is at most 10% between the winner and the runner-up). For observations at election i, geography (e.g. county) j, and event time (e.g. month or year) t, the following regression specification is run and I report γ_1^k:

$$Y_{ijt} = \sum_k \left\{ \alpha^k + \gamma_1^k 1\left\{\text{vote margin} > 0\right\}_{it} + \delta_0^k \left[\text{vote margin}\right]_{it}^- + \delta_1^k \left[\text{vote margin}\right]_{it}^+ \right\} 1\left\{\text{in sub-group}\right\}_{ij} + \eta_{ijt}$$

where "vote margin" is the percentage point vote margin between the black candidate and the white candidate. The first indicator term denotes a dummy for elections in which the black candidate won. The next two terms are linear controls separately for elections where the white candidate won (i.e. negative vote margin) and where the black candidate won (i.e. positive vote margin). Observations are weighted by the fraction of the geography (e.g. county) affected by the election.

In panel (a), the outcome variable Y_{ijt} is the raw and composition-adjusted Race IAT D scores, and there are two groups k where the sample is split along the median of the variable in the column header. In panel (b), the groups k are for above-median and below-median Race IAT D score level averaged for 2003-2017, while column headers indicate the outcome variable Y_{ijt}.

(a) Heterogeneity by

	IAT		black pop		income	
	(1)	(2)	(3)	(4)	(5)	(6)
	ex-demo FE	raw	ex-demo FE	raw	ex-demo FE	raw
Black winner	-0.013	-0.007	-0.000	0.003	0.039**	0.041**
	(-1.05)	(-0.55)	(-0.03)	(0.24)	(2.07)	(2.13)
HIwinner	0.047***	0.043***	0.025*	0.024*	-0.026	-0.024
	(3.75)	(3.26)	(1.81)	(1.77)	(-1.46)	(-1.35)
Observations	23590	23590	23590	23590	23590	23590
Adjusted R^2	0.015	0.014	0.003	0.004	0.002	0.003

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Heterogeneity (continued)

(b) Heterogeneity by average IAT level

	labor		mortgage	
	(1)	(2)	(3)	(4)
	unemployment transition	employment to pop	log origination to pop	rejection rate
Black winner	0.013	-0.055	-0.131	0.011
	(0.93)	(-1.33)	(-0.96)	(0.52)
Black winner x high _____	0.020	0.012	-0.173	0.030
	(1.53)	(0.28)	(-1.13)	(1.30)
Observations	4510	4376	4311	4166
Adjusted R^2	0.005	0.025	0.032	0.060

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: IV estimates

This table displays the second-stage estimates of an instrumental variables regression, using the regression discontinuity (dummy for the black candidate's victory) as the instrument. The main specification imposes a 10% bandwidth (i.e. regression sample only includes elections for which the vote margin is at most 10% between the winner and the runner-up). For observations at election i, geography (e.g. county) j, and event time (e.g. month or year) t, the following regression specification is run, and I report β_1 for different outcome variable Y_{ijt}:

$$
\begin{aligned}
Y_{ijt} &= \beta_0 + \beta_1 \mathrm{IAT}_{ijt} + \tilde{\delta}_0 \left[\text{vote margin}\right]^-_{it} + \tilde{\delta}_1 \left[\text{vote margin}\right]^+_{it} + \varepsilon_{ijt} \\
\mathrm{IAT}_{ijt} &= \alpha + \gamma_1 1 \left\{\text{vote margin} > 0\right\}_{it} + \delta_0 \left[\text{vote margin}\right]^-_{it} + \delta_1 \left[\text{vote margin}\right]^+_{it} + \eta_{ijt}
\end{aligned}
$$

where "vote margin" is the percentage point vote margin between the black candidate and the white candidate. The first indicator term denotes a dummy for elections in which the black candidate won. The next two terms are linear controls separately for elections where the white candidate won (i.e. negative vote margin) and where the black candidate won (i.e. positive vote margin). Observations are weighted by the fraction of the geography (e.g. county) affected by the election.

(a) County-level IAT scores

	labor		mortgage	
	(1) unemployment transition	(2) employment to pop	(3) log origination to pop	(4) rejection rate
IAT (adj)	0.439*	-1.224	-4.367*	0.639*
	(1.88)	(-1.55)	(-1.70)	(1.80)
Margin (winner)	0.003	-0.012**	-0.032**	0.006***
	(1.39)	(-2.28)	(-2.25)	(2.62)
Margin (loser)	-0.002	0.010	0.024	-0.001
	(-1.33)	(1.52)	(1.33)	(-0.30)
Constant	-0.041***	0.015	-1.091***	0.174***
	(-3.57)	(0.47)	(-10.23)	(11.78)
Observations	3265	3188	3099	3015
Adjusted R^2	.	.	.	.

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(b) County-level IAT scores