A Natural Experiment on Taste-Based Racial and Ethnic Discrimination in Elections*

David E. Broockman Evan J. Soltas

October 18, 2017

Abstract

We exploit a natural experiment to plausibly identify taste-based racial and ethnic discrimination in elections. In Illinois Republican presidential primaries, voters vote for delegate candidates bound to particular presidential candidates. Delegate candidates' names signal their race, but voters' incentives for statistical discrimination against nonwhite delegates are negligible. Examining delegate vote totals from 2000–2016, we estimate about 10 percent of voters avoid voting for nonwhite delegates. Due to the primary's structure, voters' discrimination undermines their preferred presidential candidates' nomination prospects and we estimate has altered political outcomes. Our estimates are robust to several possible confounds and vary as taste-based theories predict.

Keywords: Taste-Based Discrimination, Racial Discrimination, Voter Behavior
JEL Codes: D72, J15

*Broockman: Political Economy Group, Stanford Graduate School of Business (dbroockman@stanford.edu). Soltas: Department of Economics, University of Oxford (esoltas@sjc.ox.ac.uk). We are grateful to Dave Wasserman for the article that served as inspiration for this study. For helpful feedback, we thank Hunt Allcott, David Autor, Tim Besley, Kate Casey, Stefano DellaVigna, Dana Foarta, Matthew Gentzkow, Don Green, Dan Ho, Saumitra Jha, Ilyana Kuziemko, Michael Luca, Jon Levin, Seth Stephens-Davidowitz, Ebonya Washington, Ismail White, and seminar participants at the National Bureau of Economic Research, Stanford University, the University of Zurich, and the University of Southern California. For their able research assistance, we thank Hadil Almowafak, Julia Espero, Corey Garcia, Chris McConnell, Leela Srinivasan, KS Tey, and Yao Yang. All remaining errors are our own.

1 Introduction

Racial and ethnic minorities are often underrepresented in elected office. For example, nonwhites[1] are severely underrepresented among elected officials in federal, state, and local government in the United States.[2] A large body of research indicates that the underrepresentation of racial and ethnic minorities in elected office contributes to racial and ethnic disparities in a variety of economic and political outcomes.[3] One potential explanation for the underrepresentation of racial and ethnic minorities in elected office is voter discrimination, wherein voters are less likely on average to vote for a racial or ethnic minority candidate than a non-minority candidate with similar attributes.

Understanding whether voters discriminate against racial or ethnic minorities and the mechanisms for any such discrimination are central questions in political economy and would inform significant policy debates, but credible evidence on these questions is limited. Existing evidence indicates that individuals engage in racial and ethnic discrimination in product and labor markets (for review, see Bertrand and Duflo, 2016), but little work has assessed whether they do so when voting in elections. Moreover, the mechanisms for any such racial or ethnic discrimination in elections are also unclear. Theories of discrimination outline two broad mechanisms that might generate such discrimination. First, discrimination may be taste-based—wherein voters behave as if they prefer non-minority candidates inferior to them on non-racial dimensions to minority alternatives superior to them on non-racial dimensions (Becker, 1957). It may also be statistical—wherein voters accurately use candidates' race or ethnicity as a signal of non-racial dimensions such as ideology (Phelps, 1972; Akerlof, 1976). Taste-based discrimination is of particular interest because it would reduce the appeal of election winners to voters on non-racial dimensions.

In this paper we exploit a natural experiment to identify voter discrimination against nonwhite political candidates in the U.S. in an institutional environment where discrimination can be plausibly attributed to taste. This natural experiment occurred in four recent

[1] For brevity, throughout the paper we refer to non-Hispanic whites only as whites, and all other groups as nonwhites. We often use "racial discrimination" as a shorthand for racial and ethnic discrimination.

[2] As of 2015, 38 percent of the U.S. citizen population was a race or ethnicity other than non-Hispanic white alone, compared to 17 percent of the U.S. Congress (Manning, 2016) and 18 percent of state legislators (Kurtz, 2015). Only two U.S. states had nonwhite governors. Of the 100 most populous U.S. cities, 62 of which are majority-nonwhite, only 25 had nonwhite mayors, according to data compiled by the authors.

[3] Consistent with other research on the importance of politicians' personal backgrounds for economic and political outcomes (Besley et al., 2011), electing individuals who are members of underrepresented groups to public office has been found to reduce gaps in public goods provision (e.g., Pande, 2003; Chattopadhyay and Duflo, 2004), resulting in relative gains in health, education, and criminal justice (Clots-Figueras, 2012; Iyer et al., 2012; Fujiwara, 2015). Politically, elected officials who are members of these groups receive more information about these groups' political preferences (Broockman, 2014) and are more likely to represent these preferences, even when poorly incentivized to do so (Clots-Figueras, 2011; Broockman, 2013).

Illinois Republican presidential primary elections: 2000, 2008, 2012, and 2016.[4] For Illinois Republican voters to fully express their preferences in these contests, they must vote for multiple individual candidates for delegate to the Republican National Convention who appear on the primary ballot. In each congressional district in Illinois, there is a fixed number N of delegate candidates who appear on the ballot for each presidential candidate, and voters can cast votes for up to N delegates.[5] The top N vote-getting delegates in each district win and are bound to vote for their presidential candidate at the convention. To maximize the value of their ballot, voters should use all their N delegate votes to vote for all N of their preferred presidential candidates' delegates. However, the delegates' names appear on ballots, and the delegates' names sometimes imply that they are not white.

To identify racial and ethnic discrimination, we exploit the fact that we observe the vote totals of multiple delegate candidates with the same platform—that is, bound to the same presidential candidate—on the same ballots, and voted on by the same voters, all of whom those voters should select to maximize the value of their ballot. Our identification strategy is to examine variation in delegate vote totals by delegate race and ethnicity within such groups of delegates. For example, suppose the delegate candidates for Donald Trump in Illinois' first congressional district were named Tom, Dick, and Ahmed. To maximize the value of their ballot, a Trump supporter should cast their three votes for Tom, Dick, and Ahmed. However, to the extent Trump supporters engage in racial discrimination, some may vote for Tom and Dick but not for Ahmed, leaving Ahmed's vote totals lower than Tom and Dick's vote totals. We observe 816 unique natural experiments of this form.

Unique institutional features of the primary mean we can plausibly attribute any discrimination we observe against nonwhite delegates to taste. In this environment, voters face a trade-off between casting a full vote for their preferred presidential candidate and the "psychic costs" (Becker, 1957) of casting a vote for a nonwhite delegate in the process. Taste-based racial discrimination occurs when individuals behave as if they prefer alternatives inferior to them on non-racial dimensions over alternatives superior to them on non-racial dimensions in order to avoid incurring these "psychic costs." Voters must behave in precisely this manner to engage in discrimination in this environment: Delegate candidate's sole relevant non-racial dimension is the presidential candidate to whom they are bound,[6] which is clearly printed on ballots. If voters opt not to vote for their preferred

[4] We use these years because relevant data was not available prior to 2000 and the 2004 Republican presidential primary was not contested, as President George W. Bush was running for renomination unopposed.

[5] N is always set ex-ante. $N \in \{2, 3, 4\}$. In most cases, $N = 3$.

[6] In Section 5.2, we consider other dimensions voters may value besides the presidential candidate to whom delegates are bound, such as if that delegate is an existing elected official, is a "local notable," or is listed higher on the ballot, and find that these dimensions appear uncorrelated with delegate race.

presidential candidates' nonwhite delegates, this undermines voters' preferred presidential candidate's nomination prospects and advantages delegates for less-preferred candidates.[7] Any statistical discrimination is likely to be trivial, as it is very unlikely that rational agents would perceive incentives for statistical discrimination against nonwhite delegates in this setting.[8] Under convention rules, delegates have essentially no discretion. To engage in statistical discrimination, a rational voter must believe that a nonwhite delegate bound to that voter's preferred presidential candidate is more likely to somehow circumvent the convention rules and vote for a dispreferred presidential candidate than is an alternative white delegate originally bound to a dispreferred presidential candidate to circumvent convention rules and vote for that voter's preferred presidential candidate.[9]

Analyzing variation in delegate vote totals for delegates bound to the same presidential candidate and who appear on the same ballots in front of the same voters, we find that delegates receive approximately 10 percent fewer votes when their names indicate they are not white. Examining the results by delegate race and ethnicity, we find clear evidence of discrimination against delegates whose names signal they are Hispanic, East Asian, Middle Eastern, or Indian. Our point estimates for discrimination against black delegates are similar, although fewer names clearly signal that the delegates are black, making our estimates of anti-black taste-based discrimination less precise. We present a simple expository model that shows that these results indicate that approximately 10 percent of Illinois Republican primary voters have racially-discriminatory tastes stronger than their presidential-candidate preference. Moreover, we estimate that in several cases voter's racial discrimination altered political outcomes: It appears very likely that discrimination against several presidential candidates' nonwhite delegates reduced nonwhite delegates' vote totals sufficiently that del-

[7] To continue the above example, if sufficiently large numbers of Trump voters do not vote for Ahmed, a delegate for one of Trump's opponents, such as Ted Cruz, may be elected instead.

[8] Becker's (1957) definition of taste-base discrimination is with respect to how one would interpret observed behavior as if it were undertaken by rational agents with correct beliefs: "An employer may refuse to hire a [black person] solely because he erroneously underestimates their economic efficiency. His behavior is discriminatory...[A] taste for discrimination incorporates both prejudice and ignorance" (p. 16–17). Ascribing actions that arise out of false or irrational beliefs to taste-based discrimination allows one to retain Becker's (1957) interpretation of taste-based discrimination as lowering the appeal of election winners to voters on non-racial dimensions but of statistical discrimination as not doing so.

[9] Under convention rules, even if delegates fail to appear or fail to cast their vote for the presidential candidate to whom they are bound, their vote is counted for the presidential candidate to whom they are bound regardless. We discuss these restrictions and the plausibility of statistical discrimination in more detail in Sections 2 and 5.2. It is also unlikely that behavioral voters with mistaken beliefs about the primary and delegates would perceive incentives for a broader definition of statistical discrimination that included actions based on incorrect beliefs. To continue our example from the text, a Trump supporter would need to believe that Ahmed was less likely to vote for Trump at the convention than an alternative white delegate listed as bound Ted Cruz would be to desert Cruz and vote for Trump.

egates for other presidential candidates won and served instead.[10]

We present a variety of robustness checks on these results. We show that the results are similar when we use each of three different strategies to measure the racial and ethnic signals delegates' names send voters: a measure based on the background of others with their last name in data provided by the U.S. Census, a measure based on anthropological data about the etymology of their full names, and a measure based on Americans' subjective perceptions based on their full names. We further show that ballot-order effects do not drive the results. In addition, we construct three measures of possible prior information voters could have about delegates and find that these measures are also uncorrelated with delegate race; our results are robust to excluding delegates about whom voters might have had other information and to controlling for the presence of this information. We also discuss several potential alternative interpretations of the general results, including the possibility of other unobserved confounds, residual incentives for statistical discrimination, voter indifference across presidential candidates, and voters making inferences about presidential candidates on the basis of nonwhite delegates.

Consistent with our interpretation of the discrimination we observed as taste-based, we present several secondary results that examine heterogeneity in the magnitude of discrimination across candidates, elections, and geography. Most importantly, we estimate that voters who select into voting for nonwhite presidential candidates (e.g., Ben Carson) harbor significantly weaker racially-discriminatory tastes than voters who select into voting for white presidential candidates. Consistent with Becker (1957), we also find suggestive evidence that discrimination appears to decrease somewhat when it is more costly to voters' preferred presidential candidates: Voter discrimination appears to decline when voters are more likely to be decisive, but we find that discrimination still appears to persist even in the most competitive elections. The geographic areas where discrimination appears strongest is also in line with expectations from prior research about where racial tastes are strongest.

Our findings contribute to the literature in two main ways. First, our study is among the first to plausibly identify racial discrimination in elections, shedding light on an important and largely unanswered question in political economy. Racial and ethnic prejudices clearly may influence voter behavior (e.g., Kuziemko and Washington, 2015), but little work has convincingly identified voter discrimination against racial or ethnic minority candidates for elected office. On average minorities run in different electoral districts, at different times, on different platforms, with different party affiliations, for different positions, with different pre-election experience and campaign resources, and so on (Washington, 2006), making credible identification of voter discrimination difficult in most electoral settings (Stephens-

[10] In Section 2.5 we discuss why presidential campaigns may nominate nonwhite delegates despite this cost.

Davidowitz, 2014). Previous research in settings outside of politics, such as in product markets, labor markets, and criminal justice has identified the existence of discrimination with a variety of credible research designs.[11] However, it is challenging to adapt many of these strategies to the political realm, as researchers are legally constrained in how they can intervene in elections. In the setting we study, white delegate candidates running alongside nonwhite delegate candidates on the same platform, for the same office, on the same ballots, in front of the same voters provide a naturally occurring control group that allow us to rule out the alternatives that would confound other research designs.

Second, ours is among few studies where discrimination can be causally identified in a real-world setting but where the naturally-occurring institutional environment forces individuals engaged in taste-based discrimination to make trade-offs on other dimensions they value and where incentives for statistical discrimination are naturally largely absent. Our paper is among the first we are aware of to study a real-world setting with both these features. As reviews of the discrimination literature note (Guryan and Charles, 2013; Bertrand and Duflo, 2016), the absence of these features in nearly all real-world settings means that to what extent any discrimination that existing studies identify is taste-based or statistical in nature is difficult to determine.[12]

We explain the context and natural experiment in greater detail in Section 2. Section 3 introduces the main data sources we use to examine whether voters engage in this behavior. To fix ideas, Section 4 presents a simple expository model of taste-based racial discrimination by voters and its consequences for election outcomes in this setting. Section 5 presents our empirical results as well as a variety of robustness checks. Section 6 discusses the implications for nonwhite political representation. Simulating the partial-equilibrium outcomes of Illinois presidential primaries, we estimate that if racial tastes were zero the number of nonwhite Republican presidential delegates in Illinois would increase by 20 percent. We conclude by discussing potential policy implications of our results.

[11] These include audit studies (e.g., Pager et al., 2009), correspondence studies (Bertrand and Mullainathan, 2004), natural experiments (Goldin and Rouse, 2000; Price and Wolfers, 2010; Abrams et al., 2012), and a host of lab and field experiments (for review, see Bertrand and Duflo, 2016).

[12] For example, many studies seek to identify taste-based discrimination by analyzing endogenous variation in incentives for engaging in statistical or taste-based discrimination. However, such research wrestles with the issue that, in nearly all real-world settings, agents have incentives to engage in both statistical and taste-based discrimination that may also vary with unobserved confounds (Bertrand and Duflo, 2016; Guryan and Charles, 2013). To avoid this problem, other research has created artificial settings where incentives for statistical discrimination can be largely eliminated (e.g., List, 2004).

2 The Illinois Republican Presidential Primary

2.1 Design of the Primary

The "delegate loophole primary" design of the Illinois Republican presidential primary is unique within the United States.[13] Voters vote separately for some number of delegates—usually three—who are bound to represent a given presidential candidate at the Republican National Convention. If elected, the delegate candidates have essentially no discretion in the votes they cast at the Republican National Convention. Rule 16(a)(2) of the 2012 and 2016 Republican National Conventions specifies that delegates who do not vote for the candidate to which they are bound have that vote canceled, and the Secretary of the Convention records the vote as one for the candidate to which the delegate was bound. In all the elections we study, about 80 percent of the Illinois delegation is allocated in this manner.[14]

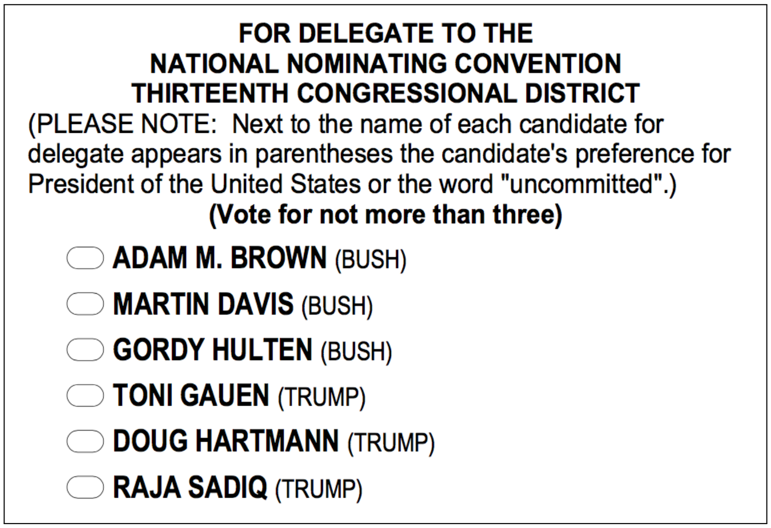
Direct votes for these delegates occur at the congressional-district level as follows. Illinois currently has 18 congressional districts and, in 2016, each district was allocated three delegates to the Republican nominating convention. Before the election, each presidential campaign nominates three candidates for delegate in each of the 18 congressional districts. The ballot instructs voters to vote for up to three delegate candidates, who need not be bound to support the same presidential candidate. For example, Figure 1 provides part of the relevant delegate-candidate section of the ballot from McLean County in the 13th congressional district of Illinois in the 2016 election. While the Jeb Bush campaign nominated three delegates with names voters likely perceived as white, the Donald Trump campaign nominated two likely-white delegates and one, Raja Sadiq, who voters likely perceived as nonwhite. A Trump voter with sufficiently strong racial tastes could vote for only Trump's two likely-white delegate candidates and not for Raja Sadiq. The performance of Raja Sadiq versus his two white counterparts on the same ballot thus represents one "natural experiment" of the 816 such "natural experiments" we observe, within which we analyze variation in delegate vote totals.[15]

[13] Delegate loophole primaries were once common in the U.S. but were largely replaced by candidate-based primaries in the 1970s contemporaneously with a variety of other reforms intended to empower voters in primary elections (Shafer, 1988). Difficulty in locating election returns, as well as biographical information on delegate candidates, in other states prior to the 1970s, and the low number of nonwhite delegate candidates who are likely to have run at that time limit our ability to extend our analysis to these earlier primaries. Some other states also list delegates on ballots, but in none of these other states does the institutional environment permit the same inferences as we can make here. For example, in some other states, the delegate vote totals only determine which particular delegates have the opportunity to represent a particular candidate and not the total number of delegates that a presidential candidate gets.

[14] About 20 percent of the delegation is independently allocated to the delegates determined in two other ways: delegates set aside for prominent party leaders and delegates bound to the winner of the statewide "beauty contest." These delegates do not appear on the ballot.

[15] For a related approach analyzing naturalization decisions in Switzerland but where voters do not face

Figure 1: Section of Sample Ballot


FOR DELEGATE TO THE
NATIONAL NOMINATING CONVENTION
THIRTEENTH CONGRESSIONAL DISTRICT
(PLEASE NOTE: Next to the name of each candidate for delegate appears in parentheses the candidate's preference for President of the United States or the word "uncommitted".)
(Vote for not more than three)

ADAM M. BROWN (BUSH)
MARTIN DAVIS (BUSH)
GORDY HULTEN (BUSH)
TONI GAUEN (TRUMP)
DOUG HARTMANN (TRUMP)
RAJA SADIQ (TRUMP)

Notes: This figure shows a relevant delegate-selection section of the 2016 Republican primary ballot from McLean County in Illinois' 13th Congressional District.

As Figure 1 shows, the names of each delegate candidate are printed in large, bold font, followed by the last name of the presidential candidate to whom they are bound. No other information is supplied on the ballot about delegate candidates. Delegate candidates are grouped by presidential candidate on the ballot, such that voters can easily identify the three delegate candidates bound to their preferred presidential candidate. By implication, the environment is structured such that voters discriminating by race are likely to be consciously aware they are doing so.[16]

This setting has three unique features we exploit to identify taste-based discrimination. First, in order to identify discrimination itself, the presence of white delegate candidates running alongside the nonwhite delegate candidates who appear on the same ballots, in front of the same voters, bound to the same presidential candidates provides a natural control group for estimating discrimination. Second, helping us attribute this discrimination to taste, there are essentially no incentives for rational voters to engage in statistical discrimination. Since delegates have no meaningful discretion if elected and are merely a mechanism for voting for a given presidential candidate, delegate names vary the "psychic cost" (Becker, 1957) of voting for a possibly-nonwhite delegate without varying relevant information about the

trade-offs and may statistically discriminate, see Hainmueller and Hangartner (2013).

[16] An identical process occurs for the selection of alternate delegates, and we pool the normal and alternate delegate candidates in our sample.

consequences of that vote, which are plausibly fixed within the individual "experiments" we analyze. In other elections, voters may rationally infer nonwhite candidates differ on other dimensions; here, the only relevant dimension is the presidential candidate to which the delegate is bound, which is printed on the ballot.[17] Third, any such taste-based discrimination entails trade-offs. Discrimination undermines voters' preferred presidential candidates' nomination prospects; to act on racial tastes, voters must act as if they prefer a less-preferred presidential candidate in order to avoid voting for a nonwhite delegate—just as, to be said to be engaged in taste-based discrimination in the labor market, an employer must act as if they prefer hiring a white who is less productive instead of a more-productive nonwhite (Becker, 1957).[18] In summary, to the extent that delegate candidates' names signal their race, the primary design therefore creates opportunities for credibly identifiable taste-based discrimination by voters in an environment where incentives for statistical discrimination are naturally absent.

2.2 Elections in the Dataset

A contested Illinois Republican presidential primary took place on March 21, 2000; February 5, 2008; March 21, 2012; and March 15, 2016. With the exception of the 2000 election, these primaries occurred relatively early in their respective primary seasons, before a "presumptive nominee" was clear but nevertheless with a clear front-runner. The median Congressional district contest was decided by only 2,541 voters.[19]

2.3 Candidates for Presidential Convention Delegates

2,386 unique delegate candidates stood for election in Illinois across the four presidential primaries included in this study.

Presidential candidates have discretion over who will serve as their delegate candidates. One possibility this introduces is that candidates may use delegate candidacies to help secure support from politically-influential individuals some voters may be familiar with or to advertise their support from such individuals. But if voters are more likely to vote for delegates whose names they recognize and white delegate candidates are especially likely to hold other political positions that would generate name recognition, we may uncover a spurious

[17] See Section 5.2.3 for more discussion of the plausibility of incentives for statistical discrimination.

[18] To the extent voters receive expressive utility for voting for their chosen presidential candidate (DellaVigna et al., 2017), engaging in racial discrimination also denies them this utility.

[19] For the statewide results, see Appendix Table A21 for the final ex-ante polling-average data from the polling aggregator website Real Clear Politics and the actual ex-post statewide "beauty-contest" results. For further context on the elections we study, see Appendix F.

relationship between delegate candidates' race and ethnicity and their vote totals.

To assess this possibility, we obtained information on delegates' backgrounds that some voters might plausibly know. In particular, we instructed research assistants to complete "background checks" on all the delegates in our sample, searching for whether they have ever held public office, a local Republican Party leadership position, or other prominent non-political posts and recorded the offices held. We then divided offices into four categories: major office (e.g. U.S. Representative), state legislature (e.g. State Senator), minor office (e.g. county board member, town mayor), or no office. Appendix D details our data-collection procedure.

Table 1 reports the number of unique delegates in each category. 40.3 percent of delegates in our sample have held public office or similar leadership positions, the vast majority of whom were minor officials, such as members of their local Republican party committee or members of town councils.[20] Crucially for our identification strategy, white delegate candidates are no more likely to be officials than nonwhite delegate candidates, with the exception that a few delegates were sitting members of the U.S. Congress or Illinois governors and all these were white.[21] We present robustness checks using this data in Section 5.2.

2.4 The Illinois Republican Primary Electorate

Mirroring national patterns, Illinois Republican primary voters appear to be almost entirely non-Hispanic whites. Two separate data sources suggest this same conclusion. First, Appendix Table A1 presents demographic summary statistics on Republican primary voters in Illinois from administrative voter file data and commercial data gathered by the voter file firm Catalist. Over 95 percent of the people that official records indicate voted in the 2008, 2012, or 2016 Republican primaries Catalist estimates to be non-Hispanic white based on their names and neighborhood racial composition.[22] Second, of Illinois Republicans who self-report voting in the 2008 primary on a large-sample survey that year, the Cooperative Congressional Election Study (Ansolabehere, 2010), 97 percent indicated they were white. Given the paucity of nonwhite voters, the data provide little opportunity to distinguish bias towards coethnics from bias against nonwhites by whites and nonwhites, but also largely eliminate the risk of attenuation of our estimates due to opposite-sign own-race biases.[23]

[20] See Appendix Table A4 for descriptive statistics on the number of unique delegates holding each office broken down by office type.

[21] See Appendix Table A5 for correlations between officeholding and our measures of delegate race. Female delegates, however, are 10.6 percentage points less likely to be officials.

[22] Data from the year 2000 is not available.

[23] This source of attenuation is negligible in magnitude. Denote the estimate of the voter net discrimination coefficient $\hat{\beta}$. If the nonwhite share of population is n, and whites and nonwhites have tastes for discrimination against the other of size β, then $\hat{\beta} = (1 - 2n)\beta$. Since $n \approx 0.04$, a plausible estimate of the magnitude of

Table 1: Delegate Officeholding by Race of Delegate Candidates

	Count and Percentage of Delegate Candidates		
Level of Office	White	Nonwhite	Total
Major Office	22	0	22
	1.13	0.00	0.92
State Legislature	153	27	180
	7.88	6.15	7.56
Minor Office	615	136	751
	31.68	30.98	31.55
No Office	1,148	274	1,422
	59.14	62.41	59.75
Missing	3	2	5
	0.15	0.46	0.21
Total	1,941	439	2,380

Notes: This table reports the distribution of the delegate population across four office categories. Major offices include the governorship and U.S. House of Representatives membership; minor offices include City Council or Board of Education seats. We dichotomize the PC1 race measure at 0.5. Percentages are calculated as a share of column totals. The $\chi^2(4)$ statistic is 8.64 ($p = 0.07$). This result differs from that in Appendix Table A5 due to the detail on the level of office held; nonwhite delegates are no less likely to hold office but hold offices of lesser stature than white delegates. See Appendix D for a discussion of the data-collection process, and see Appendix Table A4 for a detailed tabulation of delegates by type of official and race.

The Catalist data also indicate voters were split evenly by gender, their mean age was about 60, and they lived in Census block groups where one third of residents were college graduates and median annual per capita income was about $70,000. These averages are consistent with national data which finds that voters tend to be more white, aged, educated, and higher-income than nonvoters, and Republican voters especially so.

Propitiously for the external validity of our findings, available data suggest Illinois is not far from the median U.S. state in terms of racially-discriminatory tastes. While the strength of racial tastes do not lend themselves to easy quantitative measurement, Appendix Table A2 reports data on rates of racially-charged Google searches (Stephens-Davidowitz, 2014), 2.4 million results of self-administered Race Implicit Association Tests (Xu et al., 2014), the per-capita number of active hate organizations identified by the Southern Poverty Law Center (2015), and the per-capita rate of race-related hate crimes counted by the Uniform Crime Report of the Federal Bureau of Investigation (2000–2015). None of these measures

attenuation is 8 percent.

identify Illinois as a state with unusually high or low levels of racial animus.

Illinois Republican voters appear to place utility on voting for their preferred presidential candidates in these contests. Significantly more voters engage in the costly effort required to turn out to vote when a presidential nominating contest is on the primary ballot versus when only state and local primaries are on the ballot. Administrative records show that Republican primary election turnout was 34 percent higher on average in presidential than in non-presidential election years from 1998 to 2016, even though the non-presidential years during this period featured contested Illinois governor primaries and the presidential years did not.

2.5 Where and Why Do Campaigns Nominate Nonwhite Delegate Candidates?

Since our identification strategy only relies on variation within groups of delegates running to represent the same area and bound to the same presidential candidate, the set of cells that contribute to identification may differ from Illinois on average. In Appendix B we evaluate how the identifying set differs from Illinois in general. We do not find evidence of strong selection into the identifying set; the racial mix of the delegate-candidate population varies only modestly with district-level demographic variables.

Examining where nonwhites are nominated as delegates does suggest that search costs may be an explanation for why campaigns nominate nonwhite delegates in the first place if these delegates receive fewer votes: Campaigns are less likely to nominate nonwhites when the supply of plausibly eligible white Republican voters is less constrained, as it would be in congressional districts that are more white and more Republican.[24] These apparent supply constraints also suggest delegate candidate status is not particularly valuable for the delegate candidates, another reason why voters should have little incentive to engage in statistical discrimination.

3 Data

3.1 Vote Totals

We observe official vote counts by delegate candidate at the county–congressional district level for every delegate candidate and county–district in Illinois in 2000, 2008, 2012, and 2016.

[24] Conditional on the share of a district that is white, for every 10 percentage points more Republicans there are in a Congressional district, campaigns are 2.8 percentage points less likely to nominate a nonwhite delegate in that district ($t = 5.16, p < 0.001$).

A "county–congressional district" refers to the intersection of a county and congressional district: Some congressional districts span multiple counties and we observe vote totals separately within these county intersections. Our unit of observation is each delegate at the intersection of the county and congressional district. As described later, our standard errors are clustered at the delegate level, as this is the level at which the "treatment," a delegate's race, is assigned. A stylized diagram of the county–district concept can be found in Appendix Figure A1.[25] These county–district intersections are mutually exclusive and exhaustive of Illinois voters and geography. Importantly, we do not observe voting at the ballot level, and so we cannot directly observe individuals' joint voting decisions.

Our sample spans 2,380 unique delegate candidates and 19,711 vote-count observations, since we observe how a delegate candidate did in multiple county-congressional district intersections, representing a total of 22.3 million votes.[26] The mean (median) delegate candidate received 1,133 (306) votes in each county-congressional district area. All vote count data were drawn and are publicly available from the online database of the Illinois State Board of Elections. The data also include the name of the delegate candidate as printed on the ballot, which is fixed at the congressional district level.

Throughout this paper, we refer to delegate candidates who run in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions as in the same "cell." Recall that to maximize the value of their vote voters should vote for all their presidential candidates' delegate candidates in the same cell and have no votes left after that. Delegates in the same cell are the most suited to comparison, in that factors related to geography, year, and presidential candidate are all constant within cells and that all remaining variation is between delegate candidates. In all our regression specifications we introduce fixed effects for each individual cell.

Our quantity of interest is the share of voters who did not vote for their preferred presidential candidate's nonwhite delegates in order to express racially-discriminatory tastes. Our specification of the dependent variable is thus the logarithm of the number of votes a delegate candidate receives in their county–district unit and election year.[27] We derive our exact empirical specification in Section 4.

[25] If a congressional district spans multiple counties, we observe each candidates' vote total within each county. However, it is the total number of votes that delegates receive across all the counties that determines election outcomes. Likewise, when a county contains multiple congressional districts, we observe the vote totals for each district separately. A delegate candidate can only run in one congressional district.

[26] We drop from the sample six delegate candidates who ran as uncommitted to a presidential candidate.

[27] For example, suppose that, in a given county–district in 2016, a white delegate candidate received a total of 1,000 votes, and a nonwhite delegate candidate for the same presidential candidate received a total of 850. From these two data points, we would estimate that approximately 16.3 percent of voters for the specified presidential candidate in the specified county–district acted on a racial taste, as $\ln(850) - \ln(1,000) = -0.163$.

3.2 Racial and Ethnic Signals in Delegate Candidates' Names

We measure the perceived race and ethnicity of all delegate candidates in three distinct ways: using 2000 U.S. Census data, using a proprietary anthropological database of full-name frequencies, and using guesses of workers on Amazon Mechanical Turk (MTurk). These three sources yield two objective measures of likely race and ethnicity and one measure of subjective racial and ethnic perception. For all three measures, we distinguish between delegates who are white, black, Hispanic/Latino, and Asian; for two measures, we further make further distinctions among Asians.

3.2.1 Objective Measure 1: Census Data on Last Names

Public-use tabulations from the 2000 U.S. Census report the count and racial and ethnic composition of individuals with a specified last name for all last names occurring 100 or more times in the 2000 Census returns. The tabulations include data for 151,671 last names and define racial categories as non-Hispanic white only, black only, Hispanic only, Asian or Pacific Islander only, and several other smaller categories. We use only the specified four categories.[28]

Similar to Fryer and Levitt (2004), we match all delegates in our sample with the Census racial-composition data. About 87 percent (2,073 of 2,380) of our matches are exact. The remaining last names are necessarily less common in the U.S. but are often relatively similar to common last names. As a robustness check, we also present results where we identify the nearest match in the Census data for each delegate-candidate last name by the maximizing the Jaro–Winkler distance, a common measure of string similarity in record-matching applications, between the delegate-candidate last name and each Census last name.[29] To show our results are unaffected by a potential sample-selection bias generated by omitting delegates with rare names, we present estimates with these inexact matches in Appendix Table A6. The results remain similar.[30]

[28]There is significant missingness of racial-composition data for the smaller racial categories in the less-common last names in the tabulation and no delegate-candidate last names in our sample matched with last names that Census data identified as substantively (10 percent or more) linked to these smaller racial categories. These racial categories we omit are: American Indian or Alaska Native only, Native Hawaiian or Other Pacific Islander only, and Two or More Races.

[29]The Jaro–Winkler distance is a function of the share of matching characters between two strings and the number of transpositions required to order the input string's matching characters in the order of the output string's matching characters. Cohen et al. (2003) reports some evidence of the strong performance of the Jaro–Winkler distance in name-matching. A recent application in economics is Feigenbaum (2015).

[30]One advantage of the Census measure's reliance on last names only is that the racial signals last names send are not contaminated by class signals individuals' parents might have sent when choosing racially distinctive first names (Fryer and Levitt, 2004). This is also less of a problem for our sample in general, as Asian and Hispanic names send opposite class signals, yet we find discrimination against both groups.

3.2.2 Objective Measure 2: Onolytics Classifier

We also use a software package called Onolytics to estimate the races of the delegate-candidate population. Onolytics is a commercial software package developed in Mateos (2014) that classifies names by a proprietary international database of over 1 million last names and 500,000 first names. While Onolytics provides detailed ethnicity categories, we collapse these to seven: black, Asian, Hispanic/Latino, Indian, Middle Eastern, non-Hispanic white ethnic, and non-ethnic white.[31] We define the nonwhite categories to correspond as closely as possible to those in our other two measures of delegate race and ethnicity. The measure is dichotomous.

3.2.3 Subjective Measure: MTurk Perceptions of Full Names

Objective race measures may, in some cases, poorly reflect voters' perceptions of the racial content of last names. For example, the modal MTurk worker perceived the delegate "Julie Mercadante" as Hispanic, but the U.S. Census indicates that over 96 percent of Americans with this last name are non-Hispanic whites; the last name, per Onolytics, is Italian in origin.

To measure voter-perceived race of delegates, we paid MTurk workers to guess the race of delegates from their provided full names. We followed the procedure of Kuziemko et al. (2015) to ensure high-quality guesses, in particular limiting the sample of potential participants to "master" MTurk workers who live in the United States.[32] We paid for 30 guesses for each delegate name to yield reasonably precise estimates of perceived race. In Appendix G we estimate to what extent measurement error attenuates our estimates, both narrowly due to the sampling of a finite number of MTurk raters for each name and, more broadly, due to the fact that perceived race is a latent variable that manifest variables measure with error.

Another advantage of the MTurk-based perceived race measure is that we could ask MTurk workers to provide their perceptions in finer ethnic categories than the U.S. Census makes available. MTurk workers coded the race of each delegate as one of six categories: white, black, Hispanic, Asian, Indian, or Middle Eastern. By comparison, the Census definition of "Asian or Pacific Islander" spans individuals with East Asian, Indian and Middle Eastern racial heritage, in addition to Pacific Islanders, and it is plausible that there is

[31] See Appendix C for the details on the collapse. We define the white-ethnic category motivated by historical evidence for discrimination against "white ethnics" and for the attenuation of social distinctions among whites in the U.S. in the 20th century (e.g., Roediger, 2005). We define white ethnics in terms of the Onolytics classification for names of Southern European, Eastern European, and Jewish origin. The non-ethnic white category therefore includes names that are of Northern European, Central European, English, or Celtic origin. We generally collapse "white ethnic" and white names but also present them separately as a robustness check.

[32] See Appendix E for detail on the MTurk survey.

significant heterogeneity in voter taste-based discrimination within the Census definition of Asian.

Throughout the paper we present estimates using all our measures of the racial and ethnic signal voters' names send to demonstrate the robustness of our results.

3.2.4 Descriptive Statistics

To illustrate the racial and ethnic signal in names, Table 2 reports the five names identified as whitest and least white using the Census and MTurk whiteness measures; the dichotomous Onolytics measure does not enable such a ranking. Both measures identify names such as "Carol Hornickle" and "Mike Marron" as white and "Baba Padmanabhan" and "Noella Chung" as nonwhite. In Appendix Figure A2, we plot kernel density estimates for these measures. Most names in the sample are identified as very likely belonging to non-Hispanic whites, with a heavy left tail of names that likely belong to nonwhites.[33]

Table 2: Whitest and Least White Names of Delegate Candidates

	Census		MTurk	
	Name	Whiteness	Name	Whiteness
Whitest	Carol Hornickle	0.9946	Mike Marron	1
	Brian Milleville	0.9956	Jill Bess	1
	Sherry Hellmuth	0.9958	Helen Manson	1
	Ralph Baahlmann	0.9972	William S. Graham	1
	Gregory Musinski	0.9942	David L. Snyder	1
Least White	Baba Padmanabhan	0.0141	Noella Chung	0
	Ji Chung	0.0234	Gustavo Gonzalez	0
	Neil V. Patel	0.0155	Angel Garcia	0
	Noella Chung	0.0234	Rafael Rivadeneira	0
	Steve H. Kim	0.0260	Raja Sadiq	0

Notes: This table lists the five whitest and least white delegate names using the two continuous race measures in this paper. For the Census data, whiteness is defined as the proportion of U.S. citizens with the delegate's last name who are non-Hispanic white. For the MTurk data, whiteness is defined as the proportion of Turkers who perceive the full name as non-Hispanic white. The categorical definition of Onolytics race signal means there is no equivalent ranking of names by signal strength. For further detail by race category, see Appendix Table A7.

To increase power over any individual race measure, we use an index constructed by estimating the first principal component (PC1) of the three race measures, rescaled to the unit interval to permit interpretation, as our baseline measure of delegate-candidate race

[33] Appendix Table A7 gives examples of highly suggestive names for all racial and ethnic categories. Appendix Figure A3 provides analogous kernel density plots.

throughout the results presented in Section 5. We report the results from the principal component analysis, including for the detailed race categories, in Appendix Table A10.[34] In general, we find that same-race, different-measure correlations—that is, for instance, the MTurk black measure's correlation with the Census black measure—are robustly positive, and that different-race, different-measure correlations are almost entirely negative.[35] Our interpretation of these results is that all three measures capably differentiate between white and nonwhite names, and among detailed nonwhite categories, but with substantial noise that will bias our estimates toward zero, motivating the use of principal component analysis to extract the common signal.

Using the modal guesses of delegate-candidate race from the MTurk data, the delegate-candidate population is 94 percent non-Hispanic white, 1 percent black, 4 percent Hispanic, and 1 percent Asian. Appendix Table A8 reports further summary statistics on delegate candidates. Appendix Table A9 shows that plausibly-nonwhite delegate candidates are nominated by nearly all the presidential candidates in our sample and that our results are not driven by a particular presidential candidate's voters.

4 Taste-Based Racial Discrimination by Voters

To motivate our empirical analysis and fix ideas, in this section we present a simple expository model of how taste-based discrimination would manifest in Illinois Republican delegate elections.

4.1 Model

Suppose there are a finite number of presidential candidates $p = 1, 2, \ldots, Q$, each of whom nominates three delegate candidates in each congressional district, each of which divides into county–districts. A function $\phi(\cdot)$ maps from delegate candidate to the corresponding presidential candidate to which that delegate is bound. All delegate candidates are either white and thus members of the set W or nonwhite and thus members of the set N.

Voters, indexed by v and residing in a single county–district, vote for three unique delegate candidates and cannot vote for any given delegate more than once.[36] We denote their choices

[34] The first principal component of the three continuous measures explains 57 percent of the total variance, significantly more than 33 percent under the null hypothesis of independence.

[35] Different-race, same-measure correlations are negative in part by construction, since for all names, the probabilities across the race categories must sum to unity.

[36] In the actual election, voters can vote for up to a fixed number, usually but not necessarily three, of unique delegate candidates. Voters can choose not to cast all their votes. It suffices to conceive of non-votes as going to a placeholder candidate for whom voters cannot otherwise vote.

$c_v = \{c_{1v}, c_{2v}, c_{3v}\}$, noting that the ordering of delegates within the set is immaterial. The preferences of voter v are defined by (1) a $Q \times 1$ appeal vector α_v, which expresses how much voter v likes presidential candidate p independent of a delegate candidate's race or ethnicity, with $\alpha'_v = [\alpha_{1v}, \ldots \alpha_{pv}, \ldots \alpha_{Qv}]$, such that $\alpha_{iv} \neq \alpha_{jv}$ for any $i, j : i \neq j$, and by (2) a discriminatory-tastes parameter $\delta_v \geq 0$, which is the "psychic cost" to voter v of voting for a candidate $i : i \in N$.[37] In particular, we express the incremental instrumental and intrinsic utility a voter gains from voting for a delegate candidate i as $\alpha_{\phi(i)v} - \delta_v \cdot \mathbf{1}(i \in N)$.

The voter v's problem is thus

$$\max_{i,j,k:i\neq j\neq k} \left\{ \sum_{c\in\{i,j,k\}} [\alpha_{\phi(c)v} - \delta_v \cdot \mathbf{1}(c \in N)] \right\}, \tag{1}$$

and we denote the solution $c_v^* = \{c_{1v}^*, c_{2v}^*, c_{3v}^*\}$.

Proposition 1. *A solution to* (1) *always exists.*

For any finite number of delegate candidates, the set of feasible choices $\{c_v^*\}$ is finite. There exists a maximum of any finite set in $\mathbb{R}^n$, thus existence always.

Proposition 2. *For any* $i : c_{iv} \in c_v^*$ *and any* $i' : c_{i'v} \notin c_v^*$*, the solution to* (1) *satisfies*

$$\alpha_{\phi(i)v} - \delta_v \cdot \mathbf{1}(i \in N) \geq \alpha_{\phi(i')v} - \delta_v \cdot \mathbf{1}(i' \in N). \tag{2}$$

This follows immediately from the weak axiom of revealed preference.

Proposition 3. *For any* $i, j : c_{iv}, c_{jv} \in c_v^*$,

$$\phi(c_{iv}) \neq \phi(c_{jv}) \implies \delta_v > |\alpha_{\phi(i)v} - \alpha_{\phi(j)v}|. \tag{3}$$

A voter can always select $c_v : \phi(c_{1v}) = \phi(c_{2v}) = \phi(c_{3v})$. Thus if $\phi(c_{iv}^*) \neq \phi(c_{jv}^*)$, then $\exists i' : \phi(c_{i'v}) = \phi(c_{jv}^*)$ and $\phi(c_{i'v}) \neq \phi(c_{iv}^*)$ for $c_{i'v} \notin c_v^*$. Then, by (2),

$$\delta_v[\mathbf{1}(i \in N) - \mathbf{1}(i' \in N)] \geq |\alpha_{\phi(i)v} - \alpha_{\phi(i')v}|, \tag{4}$$

and thus (3), since $|\alpha_{\phi(i)v} - \alpha_{\phi(i')v}| > 0$ by $\alpha_{\phi(i)v} \neq \alpha_{\phi(i')v}$ for any $\phi(i) \neq \phi(i')$.

[37] The assumptions of uniqueness of α_{iv} and non-negativity of δ_v imply respectively that, absent race, voters would have strict preferences over candidates and that no voters prefer nonwhites to whites. As we discuss in Section 5.2.3, if some voters are indifferent between presidential candidates or wish to signal dissatisfaction with the presidential candidate and therefore choose to split their votes between multiple presidential candidates, this would not bias our estimates given our fixed effects specification.

Proposition 4. *When the nonwhite share of delegate candidates is small, or the number of presidential candidates is large, the mean difference in log vote counts between otherwise-comparable white and nonwhite delegate candidates equals the share of voters who expressed a racial taste.*

Suppose first that voters can only express a racially-discriminatory taste by abstaining from voting for a delegate candidate and cannot vote for the delegates of other candidates. Then, by Proposition 2, the difference in the number of votes between a white and nonwhite delegate candidate who are otherwise comparable equals the number of voters who expressed a racial taste. The log difference in vote counts yields a first-order approximation of the share of voters who expressed a racial taste.[38]

4.2 Estimating Equations

By Proposition 4, the share of voters for whom $\delta_v > |\alpha_{\phi(i)v} - \alpha_{\phi(j)v}|$ is true, where $\phi(i)$ and $\phi(j)$ are respectively their first and second preferences in presidential candidates, is well approximated by

$$\hat{\beta} = \frac{1}{|C||P||T|} \sum_{c\in C,\, p\in P,\, t\in T} \left[\left(\frac{1}{|N_{pct}|} \sum_{i\in N_{pct}} \ln(\text{Votes}_{ipct}) \right) - \left(\frac{1}{|W_{pct}|} \sum_{i\in W_{pct}} \ln(\text{Votes}_{ipct}) \right) \right] \tag{5}$$

for delegate i, presidential candidate $p = \phi(i) \in P$, county–district $c \in C$, and year $t \in T$. $N_{pct} \subset N$ and $W_{pct} \subset W$ are respectively the subsets of nonwhite and white delegate candidates for presidential candidate p, in county–district c, in year t running in cells with both white and nonwhite delegate candidates. $|S|$ denotes the cardinality of the set S.

This value can be estimated by the fixed-effects regression

$$\ln(\text{Votes}_{ipct}) = \beta \cdot \text{Nonwhite}_i + X'_{ipct}\gamma + \alpha_{pct} + \varepsilon_{ipct}, \tag{6}$$

where Nonwhite_i, is a binary indicator of whether voters believe delegate candidate i is

[38] If voters can instead vote for another presidential candidate's delegate candidate, then a given white–nonwhite vote-total difference is inflated by voters who shift to the given presidential candidate in order to avoid a nonwhite delegate candidate of another presidential candidate. When the number of nonwhite delegates is small, or the number of presidential candidates is large, the number of such inbound voters is small relative to the number of voters already voting for a given presidential candidate's delegates. The magnitude of the upward bias from inbound voters is likely small. If voters' second preferences in presidential candidate matches actual statewide vote shares, and the MTurk binary nonwhite indicator is without error, then the bias is $0.04/0.96 = 0.041 = 4.1$ percent, which would, for example, increase a true coefficient of 8.5 percentage points to 8.8 percentage points.

nonwhite, β estimates taste-based discrimination, α_{pct} is a vector of fixed effects, and ε_{ipct} is a disturbance term. We introduce a vector of covariates X_{ipct} in some specifications. In the continuous specification, we replace the binary nonwhite independent variable with the continuous measure of perceived race. Under the assumption that the continuous measure reflects the share of voters who believe a delegate is nonwhite, the coefficient β_t estimates the share of voters who, conditional on knowing the candidate was nonwhite, would express a racial taste. Throughout our analysis, we weight by the maximum number of votes won by a delegate candidate within the cell. This ensures that the estimate can be interpreted as the share of all voters who act upon racially-discriminatory tastes.

To estimate discrimination by delegate race and ethnicity—distinguishing within the nonwhite category—let δ_v^r be the "psychic cost" voter v feels for voting for a delegate of race or ethnicity r. We estimate the share of voters for whom $\delta_v^r > |\alpha_{\phi(i)v} - \alpha_{\phi(j)v}|$ by

$$\ln(\text{Votes}_{ipct}) = R_i'\beta_t^r + X_{ipct}'\gamma + \alpha_{pct} + \varepsilon_{ipct}, \tag{7}$$

where R_i is a vector containing either binary or continuous measures of perceived race for delegate i.

5 Results

5.1 What Share of Voters Expressed Racial Tastes?

Table 3 presents our main regression results. In all regressions the dependent variable is the logarithm of the vote count, and the unit of observation is the county-district-delegate. Standard errors are clustered at the delegate level, as this is the level at which the "treatment," a delegate's race, is assigned.[39] Fixed effects at the cell level are present in all regressions; the coefficients reported in the tables only capture variation in the performance of delegate candidates within the same cell—that is, among delegates who appear on the same ballots in front of the same voters to represent the same presidential candidate.[40] Since all our race measures are scaled to the unit interval, the coefficients represent the estimated difference between a certainly-white and a certainly-nonwhite candidate.

The first column of Table 3 reports estimates of equation (6) using a continuous measure of the treatment, the share of MTurkers who perceived each delegate as white. The coefficient

[39]The implied randomized experiment is that the same individual delegate candidates were randomly assigned to switch races/ethnicities with the other delegates running in their same cell. We present a permutation test later in the paper that implements this implied experiment under the sharp null hypothesis.

[40]Note that the fixed effects mean that voters' inferences about presidential candidates from their nomination of nonwhite delegates cannot affect our results. Section 5.2.3 considers this point at greater length.

Table 3: Effect of Delegate-Candidate Race on Log Votes

	MTurk Measure				Census Measure				Onolytics Measure		Standardized PC1	
	Continuous		Dichotomous		Continuous		Dichotomous		Dichotomous		Continuous	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nonwhite	-0.088***		-0.046**		-0.060**		-0.060***		-0.036***		-0.102***	
	(0.022)		(0.019)		(0.024)		(0.012)		(0.011)		(0.025)	
Black		-0.048		0.145**		0.024		-0.120***		-0.040**		-0.122***
		(0.041)		(0.060)		(0.038)		(0.043)		(0.018)		(0.056)
Hispanic/Latino		-0.048*		-0.043**		-0.076***		-0.055***		-0.041***		-0.093***
		(0.025)		(0.017)		(0.029)		(0.016)		(0.016)		(0.029)
Any Asian						-0.075***		-0.077***				-0.113***
						(0.020)		(0.023)				(0.031)
East Asian		-0.093***		-0.064***						-0.055***		
		(0.032)		(0.023)						(0.010)		
Indian		-0.174***		-0.136***						-0.070**		
		(0.048)		(0.033)						(0.029)		
Middle Eastern		-0.175***		-0.114***						-0.085*		
		(0.044)		(0.026)						(0.049)		
White Ethnic										-0.018		
										(0.013)		
F-test of equality		4.97***		5.64***		2.06		1.24		1.43		0.93
N	19,703	19,703	19,703	19,703	17,850	13,590	17,855	17,855	19,345	19,345	17,587	13,035
R^2	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.989

Notes: This table presents the results of estimating equations (6) and (7), yielding estimates of the share of Illinois Republican presidential primary voters whose racially-discriminatory tastes are stronger than their preference of presidential candidate. "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. Column 11 is our preferred estimate throughout this paper. In all regressions the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

implies that if two delegates were running in the same cell but all MTurkers perceived one as white and the other as nonwhite, the latter would receive approximately 8.8 percent fewer votes in these elections. The second column breaks down these estimates by delegate race and ethnicity by estimating equation (7). Delegates with names MTurkers percieve as Asian and especially as Indian and Middle Eastern appear to perform much worse than their counterparts running to represent the same candidates. The largest coefficient, which is for Middle Eastern candidates, suggests nearly 1 in 5 Illinois Republican primary voters would not vote for a delegate candidate for their preferred presidential candidate if they perceived the delegate candidate as Middle Eastern.

In Columns 3 and 4 we dichotomize the MTurkers' perceptions by coding each delegate as only one race, whichever race the plurality of MTurkers perceived them to be. The results are largely the same, although slightly smaller as would be expected given the attenuation bias dichotomization introduces. The coefficient for black delegates changes sign, although this is driven by names MTurkers also think are plausibly white; the plausibly-black names in our sample are less distinctively black than are other races' names distinctively nonwhite.

In Columns 5–8 we present our estimates using Census data to code the racial signal delegates' last names send. The sample is limited to delegates whose last names match the Census data exactly. Similar to the MTurk data, we find that nonwhite delegate candidates receive fewer votes. About 6 percent of Illinois Republican primary voters would not vote for a delegate candidate for their preferred presidential candidate who was objectively likely to be nonwhite. Column 5 uses a continuous measure, like Column 1: the share of Americans with the same last name as the delegate who are non-Hispanic whites. Column 6 breaks this out by race. Columns 7 and 8 dichotimize this measure by coding each name as the race it is most likely to be on the basis of the Census data. Results for the dichotomous Census measure are relatively similar to the continuous results; there is again a discrepancy for the black names between the continuous and dichotomous measures due to the presence of moderately-black last names in the sample. We cannot reject the null hypothesis of equal discrimination across nonwhite Census race categories.

In Columns 9 and 10 we present estimates using the dichotomous Onolytics race categories. As we found using the MTurk and Census race signals, we estimate that delegate candidates identified by the Onolytics algorithm as nonwhite receive fewer votes. Broken down by race, we find significant shares of voters do not vote for black, Hispanic/Latino, Asian, and Indian delegate candidates, although the penalty against Middle Eastern delegate candidates is insignificant using the Onolytics coding. We also break out a "white ethnic" category but find no significant difference in votes between "white ethnics" and "non-ethnic" whites. We cannot reject the null hypothesis of equal discrimination across nonwhite Onolytics race categories.

Columns 11 and 12 present estimates using the rescaled PC1 measure. In Column 11, which provides the preferred regression estimate of this paper, the coefficient implies that delegate candidates who are generally identified as nonwhite across the three measures receive 10.2 percent fewer votes. In Column 12, we break this result down by the three racial categories common across our three measures. We find significant taste-based discrimination against blacks, Hispanics/Latinos, and Asians but cannot reject the null hypothesis of equal discrimination across racial and ethnic categories. We also treat these estimates as our preferred results for the detailed categories throughout the rest of the paper. The PC1 index is missing when any of its constituent parts are missing, as the Onolytics or Census exact measures sometimes are, but Appendix Table A11 shows that imputing these missing values does not change the results.

In Appendix G, we correct our estimates for the attenuation introduced by measurement error, defined either narrowly in the sampling of the MTurk race measure or more broadly in using the first principal component of the three race measures to proxy for true perceived race, a latent variable.[41]

5.2 Robustness Checks

Our research design absorbs into fixed effects all attributes that affect voting behavior but do not vary among delegate candidates representing the same presidential candidate in the same county–district in the same election, such as attributes of presidential candidates. A potential confound must therefore cause some delegate candidates to receive more or fewer votes than same-race delegate candidates in the same cell and be correlated with delegate-candidate race.

[41]While we estimate that attenuation due to sampling is trivial, attenuation due to noisy proxies for perceived race may be substantial. A correction using Cronbach's (1951) α implies that being perceived as nonwhite may reduce the number of votes a delegate candidate receives by up to 11 percent, and that being perceived as black, in particular, may reduce the number of votes by up to 41 percent. Estimates for taste-based discrimination against Hispanics/Latinos and Asians, whose names more clearly indicate race, rise by comparatively less. These estimates, however, require the strong assumption that all disagreement among the race measures is measurement error and has no direct effect on votes, thus attenuating the estimate. To the extent that the non-common components of these measures still affect voting behavior with the same sign as the common component, this reliability correction will overstate the true share of voters who engage in taste-based discrimination. The estimates we report in Appendix G are therefore most reasonably viewed as upper bounds.

5.2.1 Racial Differences in Name Recognition

One such possible confound is a racial differential in name recognition and serves to illustrate the main empirical challenge to the robustness of our results: If white delegate candidates are better known to voters than nonwhite delegate candidates, and voters are more likely to vote for delegate candidates whose names they recognize, then the coefficient on candidate race would capture the effect of the racial differential in name recognition as well as the direct effect of delegate-candidate race. We take three approaches to addressing this possibility.

First, since delegate candidates who campaigns repeatedly decide to list across multiple elections are plausibly more likely to be public officials or have other unobservable qualities that would increase vote totals, we introduce a fixed effect for candidates who run in more than one election year in our sample as a lightweight test of whether voters have information about delegate candidates other than the name-implied race that they use to determine their votes. Columns 1 and 2 of Table 4, which use the PC1 race measure, show that our results are essentially unchanged when we control for repeat candidates. Appendix Table A5 shows that status as a repeat candidate is also not meaningfully correlated with delegate race or ethnicity.

Second, Columns 3–8 of Table 4 control for voters' delegate-level prior information using the results of our delegate "background checks," as described in Section 2.3, which collected information on other public offices held by delegates. In Columns 3 and 4 we use the four-level specification of the officeholding categorical variable. We find significant returns to officeholding, in line with the literature on candidate name recognition (e.g., Panagopoulos and Green, 2008).[42] However, although our measure of delegate prior officeholding may contain some measurement error, this measure does not correlate with delegate race and controlling for it does not change the point estimates. Appendix Table A5 shows that delegate outside officeholding is also not meaningfully correlated with delegate race or ethnicity.

To allow for a more detailed specification of the returns to officeholding, Columns 5 and 6 of Table 4 replace the four-level officeholding control with fixed effects for the 17 types of official we coded. Even this detailed control for officeholding leaves our result unchanged. In Columns 7 and 8, we drop from the sample every delegate for whom we were able to find had held any office. Among the delegates about which voters plausibly know nothing other than the information on the ballot—their names and the presidential candidate to which they are bound—nonwhite delegates still receive about 8 percent fewer votes than white delegates.

Third, to further test confoundedness by delegate officeholding, we conduct a placebo test of whether our empirical approach detects significant taste-based discrimination against

[42] Appendix Table A12 reports the coefficients on the covariate terms.

Table 4: Robustness Checks

	With Controls For:									
	Repeat Candidates		Officeholders						Ballot Order	
			Basic Control		Detailed Control		Non-Officeholders Only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Nonwhite	-0.095***		-0.101***		-0.101***		-0.080***		-0.114***	
	(0.025)		(0.024)		(0.024)		(0.029)		(0.022)	
Black		-0.124***		-0.114**		-0.096*		-0.104*		-0.096*
		(0.057)		(0.054)		(0.053)		(0.061)		(0.052)
Hispanic/Latino		-0.087***		-0.091***		-0.085***		-0.057**		-0.098***
		(0.028)		(0.027)		(0.026)		(0.029)		(0.019)
Asian		-0.099***		-0.101***		-0.098***		-0.158		-0.118***
		(0.033)		(0.031)		(0.030)		(0.128)		(0.035)
N	17,587	13,035	17,587	13,035	17,587	13,035	10,064	7,493	17,587	13,035
R^2	0.986	0.990	0.987	0.990	0.988	0.990	0.997	0.997	0.987	0.990

Notes: In all regressions above, the dependent variable is ln(votes) for the delegate candidates and the race measure used is the rescaled PC1 measure. The unit of observation is the county–district-delegate. For results that include the coefficients on the controls, see Appendix Table A12. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

another group: women. Since women are much less likely than men to hold local political office, both in general in the U.S. (Center for American Women and Politics, 2016) and in our dataset, gender represents a noisy signal of officeholding. Intuitively, if officeholding biases the results in favor of taste-based discrimination, then we should find that men receive more votes, as men are substantially more likely than women to be officeholders. A failure to reject the null hypothesis of zero taste-based discrimination by voters in favor of male delegate candidates would suggest that differential name recognition is not likely to be an important source of bias in our results for nonwhites, either.

We measure a delegate candidate's likely gender by their first name. American first names robustly predict gender.[43] We use the baby-name file of the Social Security Administration (SSA) from 1930 to 2012. Our estimate of the probability that a delegate is male is the proportion of U.S. citizens born with the same first name who are male at birth.

As for nonwhites, we estimate a regression of the form

$$\ln(\text{Votes}_{ipct}) = \beta \cdot \text{Male}_i + \alpha_{pct} + \varepsilon_{ipct}. \tag{8}$$

The causal effect of a female first name is estimated quite precisely at zero, with a standard error of about 1 percentage point, both when the male variable is specified as a continuous probability male and as a binary indicator.[44] On its own terms, this null finding is perhaps unsurprising, as female officeholding is common in the U.S. (Beaman et al., 2009) and the relevant electorate is split approximately evenly by gender.

In summary, these three tests of observable differences among candidates suggest that our estimates of taste-based discrimination are unlikely to be confounded by differences in prior information due to outside officeholding. We find that our estimates remain unchanged with controls for possible sources of this prior information and that these possible sources appear uncorrelated with delegate race and ethnicity. This also suggests a related alternative—that white delegates are more personally well-known among primary voters active in the party—is unlikely, as nonwhite delegate candidates appear just as likely to serve on party committees.[45] Below we discuss a related possibility about white delegates being more likely to have large unobservable social networks or be "local notables."

[43]Delegate candidates' first names strongly signal gender: 95 percent of delegate candidates have first names that are either more than 95-percent male or more than 95-percent female in the SSA data. See Appendix Figure A4 for a histogram of delegate candidates' first names by SSA percent female.

[44]Appendix Table A20 reports these results. Upon the inclusion of delegate officeholding and ballot-order controls, we find statistically significant but politically modest (1–1.5 p.p.) taste-based voter discrimination *against* men. This strongly suggests that unobserved officeholding is not driving our result for nonwhites.

[45]This implies nothing about whether nonwhites are more or less likely to serve in government or party offices in Illinois—only that, conditional on being nominated as a delegate, nonwhites and whites have similar officeholding profiles.

5.2.2 Ballot-Order Effects

Another potential confound stems from ballot-order effects (Miller and Krosnick, 1998), in which delegate candidates may receive more or fewer votes as a causal result of their ordinal position on the ballot. In particular, if presidential campaigns place nonwhite delegate candidates into their lowest positions on ballots, then the coefficient estimate on having a nonwhite name would be inflated by the indirect effect of ballot order. However, we find in Appendix Table A5 that, conditional on the number of ballot slots available (2, 3, or 4), delegate race and ballot order are nearly uncorrelated. In addition, we control for delegate-candidate ballot order with dummy variables for the rank (1–4) of a delegate candidate among those in the same and interact these with the maximum number of delegate candidates (2–4) for whom a voter may vote in a given congressional district and year. Columns 9 and 10 of Table 4 show report these results and indicate that our findings are if anything strengthened by controlling for ballot-order effects, perhaps because this control further increases the precision of our estimates. Our results are not driven by campaigns listing nonwhite candidates ordinally lower on the ballot.

5.2.3 Other Alternative Explanations and Threats to Interpretation

Inferences about presidential candidates. It is possible voters make inferences about the presidential candidates from the race and ethnicity of the delegates they nominate. However, our fixed effects mean that such inferences must affect the delegates of a given presidential candidate heterogenously to be a source of concern. If some voters select out of voting for a presidential candidate entirely after seeing that a candidate nominated a nonwhite delegate, these voters would entirely disappear from the fixed-effect cells. To the extent this behavior exists, it implies that those with the strongest tastes select out of the identifying set, leading us to underestimate discrimination.[46]

[46] Another variety of this concern is more subtle and involves voters casting votes for white delegates that appear first on a ballot before they realize a presidential candidate nominated a nonwhite. Suppose a presidential candidate's nonwhite delegate is listed third on the ballot, and a voter casts a vote for the first two white delegates before realizing that a nonwhite delegate is present. Upon reaching the nonwhite delegate, the voter may select out of the set because of the resultant inference about the presidential candidate, not because of any taste against voting for nonwhite delegates. Were this behavior to be responsible for the results, we would expect the effect to be driven by nonwhite delegates who appear lower on the ballot and be zero for nonwhite delegates who appear in the first position on the ballot, as all such voters would have selected out before casting any votes within a cell. However, we find that the coefficient estimates are essentially identical for candidates who appear in the first, second, and third positions on the ballot ($F = 0.01, p = 0.99$). This is the regression specification in equation (6) with an interaction of the nonwhite variable with ballot order. The relevant voters who do not vote for nonwhites, after seeing a nonwhite name, thus continue to vote for the presidential candidate's white delegates, suggesting little room for explanations in which voters make inferences from delegate race to presidential candidates.

Voter prior information in local social networks. Another possible concern is that white delegates may be more likely to be "local notables" or highly-connected individuals in particular areas in ways our "background checks" could not capture but that would increase their vote totals. To assess this possibility, we exploit the facts that some of the congressional districts in our sample span fairly large areas, often hundreds of miles across, and that we can observe outcomes by county within each congressional district. Based on the idea that "local notables" or highly-connected individuals should do especially well in particular counties where they are "notable," we examine the robustness of the estimates to four different trimming strategies: dropping counties where the overall estimates are largest, where white delegates do particularly well, where nonwhite delegates do particularly poorly, and where any particular delegate outperforms their overall performance. Appendix Table A13 shows that the estimates barely change with these trimming strategies. These results suggest that the class of confounds such as local social networks that would produce apparent taste-based racial discrimination by advantaging whites within particular counties do not drive our findings.

Unobservable confounds in general. It remains possible that there are other unobserved variables that vary by delegate race and increase delegate vote totals. Following Oster (forthcoming), we therefore evaluate the general plausibility of the claim that unobserved variables could explain our finding of taste-based racial discrimination by comparing the magnitude of the heterogeneity in vote totals we can explain with ballot order, etc., to the magnitude of remaining unobserved heterogeneity that would need to exist to drive the findings. We find that an unobserved confound would need to be very large in magnitude to explain our results. For the true level of taste-based discrimination to be zero, demeaning our data at the cell level and thus assuming that unobservables explain all remaining within-cell variance, the ratio of selection on unobservables to selection on observables would need to be larger than 4.5, well above the threshold of one Oster (forthcoming) recommends and more robust than nearly all studies in Oster's sample.

Residual incentives for statistical discrimination. An interpretation of the estimates as taste-based discrimination relies upon the assumption that rational voters have minimal incentives to engage in statistical discrimination. Consistent with the interpretation of Becker's (1957) definition of taste and statistical discrimination with reference to rational behavior, it is difficult to see why a rational agent in this setting would perceive incentives for statistical discrimination. To have incentives for statistical discrimination, a substantial fraction of rational agents would need to believe that nonwhite delegates selected by their candidate of choice would be less likely to vote for their candidate of choice at the convention than white delegates selected by the opposing candidates. For example, to drive our results in 2012, at least 10 percent of rational voters would need to believe both that delegates were able to change conventional rules and exert discretion and that, once doing so, the nonwhite delegates Mitt Romney had selected were less likely to vote for Romney at the convention than the white delegates Rick Santorum had selected were to desert Santorum in order to vote for Romney.[47] That convention rules do not allow delegates to engage in this behavior in the first place makes it all the more implausible.[48] That the point estimates do not meaningfully vary with the ideological positions of the candidates—as measured by ideology scores compiled by Bonica (2013), as we show in Section 5.3—is also inconsistent with voters inferring nonwhites are more likely to abandon conservative presidential candidates to vote for liberal presidential candidates; voters voting for more liberal Republican presidential candidates still discriminate.

Implications of indifference for the cost of discrimination. Another set of alternative interpretations arises from the possibility that voters split their votes across delegates for multiple presidential candidates out of indifference between the presidential candidates. If indifferent voters randomly split their votes across delegates for multiple presidential candidates without regard to delegate race, this would not bias our estimates away from zero, nor would it if they did so by always selecting candidates higher in ballot order, as we showed ballot order is uncorrelated with race.

If some voters were exactly indifferent between their favorite two presidential candidates, however, it remains possible that these voters could lexicographically choose white delegates over nonwhite delegates due to arbitrarily-weak racial tastes against nonwhites. While this behavior would allow us to retain our interpretation of the estimate as the share of voters who have racial tastes stronger than their presidential candidate preference, these racial tastes could then be weak in absolute terms. Few voters in these primaries, however, are

[47] A corollary of this observation is that even if one were to adopt a nontraditional, broader definition of statistical discrimination that includes choices stemming from mistaken beliefs, it is unlikely that behavioral voters with such beliefs about the primary and delegates would perceive incentives to discriminate. With this said, retaining Becker's (1957) definitions allows one to retain his interpretation of taste-based discrimination as lowering the appeal of equilibrium election winners to voters on non-racial dimensions but of statistical discrimination as not doing so.

[48] The sole case in which delegates do have discretion is if the convention is contested and goes to a second round of voting. One may wonder whether voters made their choices with these subsequent rounds in mind. For example, in the 2016 presidential nominating contest, some observers anticipated a possibility that no presidential candidate would receive a majority of votes on the first ballot at the convention, "un-binding" delegates for subsequent rounds of voting. This potentially creates incentives for statistical discrimination if voters in 2016 believed subsequent rounds of convention ballots were likely to occur. However, as reviewed in Section 2.2 and Appendix F, the primary race had progressed sufficiently by the time Illinois voted in 2000, 2008, and 2012 that subsequent rounds of balloting were not possible, and our estimates remain largely unchanged when examining these years only. Moreover, the last contested Republican convention was in 1952, and the last one close to contestation was in 1976. Given the elections in our dataset, it seems a priori unlikely that voting behavior in Illinois was informed by the possibility of second-round convention balloting.

likely to be indifferent between the presidential candidates: As reviewed in Section 2.4, a substantial share of voters appear to turn out to vote in these primary elections specifically to cast a vote for their favored presidential candidate, as voter turnout in Illinois Republican primaries is much higher when presidential candidates appear on the ballot. This suggests, consistent with intuition, that voters are very likely to have strong preferences in presidential primary elections and would not select into turning out for these elections if they did not.[49] Moreover, voters are more likely to have strong preferences in presidential elections than in other elections: A general feature of presidential elections is that many voters vote for presidential candidates and leave the rest of the ballot blank, and only between 0.23 percent and 0.75 percent of presidential election voters indicate they are indifferent, versus an order of magnitude larger share for many other elections (Tomz and Houweling, 2003). This suggests, if anything, that our estimates may understate the extent to which voters' racial tastes change their votes in other elections where candidate preferences are weaker on average.

Is discrimination driven by delegate race and ethnicity or class? When voters vote against nonwhite candidates, are they expressing tastes against nonwhites per se, or tastes against those lower in socioeconomic status, which they infer nonwhites are more likely to be? Fryer and Levitt (2004) show that black individuals whose parents give them racially-distinctive first names tend to be lower in socioeconomic status. Our estimates are mostly driven by racial minorities other than blacks and remain robust when we use the Census measure, which is based on last names only, indicating that inferences about class from distinctively-black first names are unlikely to drive the results. Furthermore, our estimates for discrimination against East Asians and Indians—who have higher median income than whites on average both in the U.S. in general and in Illinois specifically—are still significant and negative.

Non-parametric standard errors. To assess the robustness of our standard error estimates we also conduct a permutation test in which we repeatedly re-randomize the PC1 race measure at the delegate level within district-presidential candidate-years. This allows us to non-parametrically evaluate the significance of our estimated effect of having a nonwhite name, with the implied experiment of randomly assigning race to delegates given a fixed distribution of racial signals among the candidates running in the same district to represent the same candidate in the same year. Appendix Figure A5 plots 10,000 draws from a Monte Carlo simulation of the main regression specification with treatment status permuted in this way. The estimate remains clearly significant.

[49] Because by the time Illinois voted in 2000, 2008, and 2012 the election had narrowed to only two plausible candidates, it is also unlikely that voters were indifferent between two candidates who they both preferred to a third.

5.3 Heterogeneity in Racial Discrimination

We next estimate how voters' racial discrimination varies along several dimensions: by presidential candidate, by competitiveness, by geography, and over time.[50] Although some of these tests have low statistical power and we caution that they are fundamentally observational in nature, they all have results that are consistent with what taste-based theories would predict. In particular, we estimate that voters who self-select into voting for nonwhite presidential candidates also have significantly weaker racial tastes against nonwhite delegates; that discrimination is largest in places where white income and education is lower; and that discrimination may decrease but does not disappear in more competitive elections where voters are more likely to be individually decisive. Propitiously for the generalizability of our results, we also show that discrimination does not appear to be decreasing over time.

5.3.1 By Presidential Candidate

The design of the primary allows us to separately estimate the magnitude of taste-based discrimination among voters for each presidential candidate. Presidential candidates may attract voter populations with different levels of racially-discriminatory tastes on average. In particular, if the discrimination we identified is taste-based in nature, we would expect voters who self-select into voting for nonwhite presidential candidates to have weaker racial tastes on average than voters who self-select into voting for white presidential candidates. In Appendix Table A15, this is precisely what we find. Estimating equation (6) with two interaction terms—(1) interacting the race and ethnicity of the presidential candidate with the race and ethnicity of the delegate candidate and (2) interacting the race of the delegate candidate with year dummies, so that the comparisons between presidential candidates are within-year—we estimate that voters for nonwhite presidential candidates display zero net taste-based discrimination against nonwhite delegates. Moreover, the coefficient on the interaction term, which captures the within-year difference in taste-based discrimination between white and nonwhite presidential candidates, remains significant even after controls for delegate officeholding and ballot order are introduced. This suggests that voters for nonwhite presidential candidates likely have, on average, weaker racial tastes than voters for white presidential candidates.[51] Meta-analytic procedures suggest that net expression

[50]See Anwar and Fang (2006), Abrams et al. (2012), and Alesina and La Ferrara (2014) for other studies of heterogeneity in racial discrimination.

[51]We also estimate the share of voters who express racially-discriminatory tastes by individual presidential campaign, although these estimates are relatively imprecise. Appendix Table A16 presents the results. The presidential campaigns for which we have the strongest statistical evidence of taste-discrimination among their voters are, in descending order, the 2016 Donald Trump campaign, the 2016 Rand Paul campaign, and the 2012 Ron Paul campaign. These patterns are generally consistent with expectations. For evidence that

of racially-discriminatory tastes against nonwhite delegates is likely to exist for most white presidential candidates.[52] We also examined the relationship between the strength of racially-discriminatory tastes of each candidates' supporters and estimates of presidential-candidate liberal-conservative ideology inferred from identities of campaign donors from Bonica (2013). These results do not display a significant relationship.

5.3.2 By Competitiveness and the Instrumental "Cost of Discrimination"

Taste-based theories of discrimination predict that individuals will discriminate less when it is more costly for them to do so. In this setting, following Becker (1957), voters face a trade-off between the psychic costs of voting for nonwhites and the costs of engaging in discrimination—one instrumental component of which is the possibility that withholding their vote for their preferred presidential candidates' nonwhite delegates would prove decisive, causing this delegate to lose and impairing their preferred presidential candidates' nomination prospects. Other components of the cost of discrimination in this setting—such as the intrinsic utility voters gain from the act of voting for their chosen candidate (DellaVigna et al., 2017)—do not vary with the probability voters will be decisive.

Across the elections in our dataset, we observe significant variation in the magnitude of these instrumental costs, as voters in some elections are much more likely to cast decisive votes than others. Outcomes would change in the median Congressional district delegate contest if only 2,541 voters changed their votes—an unusually small number in the context of U.S. elections. Given that some elections were decided by many fewer votes and some by many more, there is variation in the likelihood a voter with rational expectations would believe they might cast a deciding vote. When rational voters are voting in an election they expect to be close, those deciding whether to engage in taste-based racial discrimination may plausibly be more likely to "hold their nose" and prioritize the victory of their preferred presidential candidate over avoiding the "psychic cost" of voting for nonwhites.

To examine whether voters appear to discriminate less when they are more likely to be decisive—and whether any voters still engage in taste-based discrimination even when they are—we examine how the extent of discrimination varies by the expected competitiveness

Trump voters, had a high average level of racial animus relative to the median Republican voter, see Schaffner et al. (2017). In his campaigns, Ron Paul was criticized for racially inflammatory content in newsletters he published; the controversy may have selected Paul voters towards racial animus.

[52] In particular, we estimated a bootstrapped DerSimonian–Laird (BDL) model (DerSimonian and Laird, 1986; Kontopantelis et al., 2013) to estimate between-candidate variance in racially-discriminatory tastes; the analogy is that the candidate-level estimates of the fraction of voters for each candidate who express racially-discriminatory tastes are equivalent to separate studies of taste-based racial discrimination. The upper boundary of the centered 95-percent interval of the density is -0.067. In Appendix Table A14 we present evidence for candidate-level heterogeneity.

of the election. In particular, we estimate, for each delegate candidate, the logarithm of the minimum number of votes he or she would have needed to win (lose), relative to the actual vote outcome, in order to win (lose) a delegate position.[53] We then estimate a regression specification that modifies equation (6) by interacting the nonwhite-delegate variable with several transformations of this measure of election competitiveness. We transform this measure in multiple ways since it is highly skewed and because we do not know the function voters use to map it to the likelihood that their vote is decisive. We consider three transformations: (1) the percentile of the distribution of minimum distances to winning or losing a delegate position; (2) a binary indicator of whether the candidate is in the top 10 percent the competitiveness distribution; and (3) right-censoring the distance measure at unity, so that all elections in which a delegate's vote count would need to more than double or fall by more than half are treated the same.[54]

Appendix Table A17 reports coefficient estimates for all three specifications. It also reports estimates of equilibrium shares of voters who engage in taste-based racial discrimination under two election scenarios: a tied race and a race of median competitiveness. The coefficient estimates on the interaction of delegate race and all three measures of election competitiveness suggest that the equilibrium share of voters who engage in taste-based racial discrimination is lower in more competitive elections. These point estimates also suggest that some voters continue to discriminate even in situations when they are most likely to be decisive—that is, when voters would rationally expect a delegate candidate of their preferred presidential candidate to be tied with a delegate candidate of another presidential candidate for the marginal delegate position. We interpret these patterns as further suggestive evidence in favor of our interpretations of the discrimination we observe as taste-based in nature and as politically consequential.

5.3.3 By Geography

The average voter's racially-discriminatory tastes may vary geographically due to, among other factors, local racial histories and demography. For example, Stephens-Davidowitz

[53] This competitiveness measure is unfortunately ex-post of election outcomes and post-treatment to delegate race and ethnicity, but the signals voters would have used to estimate how likely their vote was to be decisive are likely not measurable. For example, we are not aware of any publicly available Congressional-district-specific polling.

[54] Each of these transformations has advantages and disadvantages. The first is clearly independent of the distribution but is cannot be interpreted directly in terms of minimum percentage change of the vote; the second provides a straightforward of election competitiveness at the cost of censoring potentially important variance in competitiveness; and the third, while allowing a continuous analysis of relatively competitive elections, accepts a crude right-censoring to address the skewness created by delegate candidates who are particularly far from winning or losing a delegate position.

(2014) finds that the strongest predictor of his Google-search-based measure of geographic anti-black racial animus is the college-educated white share of population.

We examine whether our area-level estimates of taste-based racial discrimination vary with several population attributes previous literature has identified: the share of the population that is white, the share of the white population that has a college degree, and the real median per capita income of whites. These are all measured by the U.S. Census at the county–district level. To estimate how these measures correlate with taste-based discrimination, we modify equation (6) by interacting the nonwhite-delegate variable with our county–district-level demographic measures. In an alternate specification, we also add delegate-level fixed effects; since we observe vote counts across geographies within delegates, adding delegate fixed effects allow us to vary electorate demography while holding delegate identities constant.

Appendix Table A18 reports these results. Column 1 finds no significant association between our county-level estimates of taste-based racial discrimination and the white share of population. Column 2 finds a significant positive association with the college-educated white share of population, consistent with expectations from the literature that individuals in high-education areas have weaker racial tastes. Columns 3 and 4 find that the per income of white is positively associated with discrimination, although significance is sensitive to specification. This finding is consistent with historical patterns wherein American political parties that explicitly appeal to prejudice against nonwhites tend to perform better in districts where white median income is lower (e.g., Mulkern, 1990). Columns 5 and 6 show there is little evidence that the effect varies with with the Stephens-Davidowitz (2014) or Xu et al. (2014) measures of anti-black bias, perhaps because most of the nonwhite delegates in our sample are not black.[55] As with candidate-level heterogeneity, we estimate a meta-analytic model which suggests that taste-based discrimination exists in the vast majority of counties and report results in Appendix Table A14.

5.3.4 Over Time

One potential concern is that our results may be driven by earlier years of data and that taste-based discrimination in elections is disappearing. To examine the variation in racial discrimination in this setting over time, we re-estimate equation (6) but interacting the

[55] The lack of correlation across these measures may be attributed variously to multidimensionality in the general concept of racial bias, to distinctions between anti-black bias and bias against the other nonwhite groups examined in this paper, to measurement error in all three variables, or to a number of other factors. While it would have been promising to find agreement across these three measures, these considerations suggest it is not particularly surprising we do not. The correlation between the Stephens-Davidowitz (2014) and Xu et al. (2014) measures, for instance, is only 0.1.

nonwhite variable with year dummies. Appendix Table A19 presents the estimates by year. A Wald test of the equality restriction across election years shows that we cannot reject the null hypothesis that the share of voters who express racially-discriminatory tastes is constant over time. We also impose a linear time trend in taste-based racial discrimination, starting from equation (6), by interacting the nonwhite-delegate variable with directly with the continuous year variable. The point estimate on this interaction term is essentially zero.

6 Discussion

6.1 Effects on Nonwhite Political Officeholding

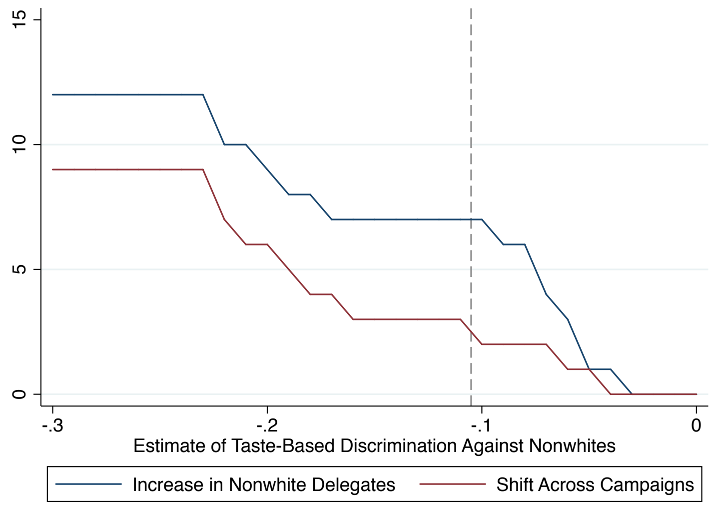
To interpret the economic and political significance of voter racial and ethnic discrimination in these elections, we simulate counterfactual election outcomes absent discrimination. Simplifying the simulation, we assume that delegate candidates who were more likely than not to be nonwhite according to the PC1 race measure all lost the same fraction of votes due to discrimination. We vary the estimated penalty from 0 to -0.3. Figure 2 shows that, at our preferred estimate, -0.102 as reported in Column 11 of Table 3, about 7 additional nonwhite delegates would have won their elections, relative to an actual baseline outcome of 30 nonwhite delegate winners over the entire period, implying an increase in nonwhite political representation of about 20 percent.

This analysis, however, is necessarily in partial equilibrium. An implication of voter taste-based discrimination for presidential-candidate behavior is that presidential candidates would nominate an inferior white delegate candidate over nonwhite delegate candidates, even if it were possible to fully inform voters about the nonwhite delegate candidates. Presidential candidates thus have an incentive to nominate fewer nonwhites for delegate races, relative to a counterfactual without voter taste-based racial discrimination. General-equilibrium effects may thus mean our estimate of the increase in nonwhite representation from the elimination of taste-based racial discrimination is materially understated.

In Figure 2, we also estimate the number of delegates who are shifted across campaigns as a result of taste-based discrimination against nonwhites.[56] We estimate that delegate allocation outcomes changed in several instances, shifting delegates across presidential candidates in approximately 4 percent of congressional district contests as a result of voter taste-based discrimination, even though only about 6 percent of delegate candidates are nonwhite.

[56]This is not identical to the increase in nonwhite representation, as, in varying the estimate, a nonwhite delegate candidate wins either at the expense of having a white delegate from the same presidential campaign or a white delegate from a different campaign, and only the latter results in a shift of a delegate position across campaigns.

Figure 2: Counterfactual Number of Nonwhite Delegates Without Voter Racial Taste-Based Discrimination



Notes: This figure plots the counterfactual increase in the number of nonwhite candidates would have won election to alternate and regular delegate positions if voter racial taste-based discrimination were eliminated, varying the assumed magnitude of racially-discriminatory tastes. 458 delegates and alternates won across the elections we study, only 30 of whom were likely not white.

6.2 Policy Implications

In the United States, credible evidence about discrimination against nonwhite candidates in elections has significant policy implications: Both Congress and the Supreme Court routinely rely on academic assessments of voter racial discrimination when crafting and reviewing American election laws. Our results are particularly relevant to the debate over the Voting Rights Act of 1965, which was passed in part to facilitate the election of nonwhites to political offices. In the subsequent decades, Congress and the Court have evaluated the continuing necessity of the Act in part by attempting to examine whether nonwhite candidates are still discriminated against in elections. In decisions in 2009 and 2013 (*Northwest Austin Municipal Utility District No. 1 v. Holder* and *Shelby County v. Holder*), the Supreme Court struck down major portions of the Act, finding that Congress had only marshaled "decades-old data" in support of its claims that nonwhites continued to face discrimination in elections in many areas. Our results—from a setting where the Act's main provisions never applied, and without a particularly distinctive history of discrimination—are consistent with the concern that the racial tastes of voters in majority-white electorates could represent a continuing barrier to the election of nonwhites in general.

Taking our point estimates at face value, this barrier appears considerable in magnitude.

Illuminating comparisons of the effect of a nonwhite name may be to effects of campaign spending and incumbency, with the caveat that all such comparisons rely upon estimates from other contexts. How much more would a nonwhite candidate's campaign need to spend, for instance, in order to offset votes lost due to taste-based discrimination? Gerber (1998) estimates that increasing campaign expenditure per capita by 10 percent changes a candidate's two-party vote share by 0.6 percentage points; a nonwhite candidate would thus need to spend $2.60 to every $1 spent by a white candidate to overcome voter taste-based discrimination of the magnitude we estimate. Likewise, the effect is about 40 percent larger than the personal incumbency advantage in U.S. congressional elections, an effect considered large (Erikson and Titiunik, 2015). These comparisons suggest that taste-based racial discrimination is plausibly a major barrier to minority political officeholding. Although eliminating the names of candidates from the ballot is an unreasonable proposal in the context of nearly any other election than this primary, our findings suggest that efforts to attenuate taste-based racial discrimination or that minimize its political consequence—such as the drawing of majority-minority districts—will continue to be necessary. Policy responses that operate only on informational margins to reduce statistical discrimination, however effective (Casey, 2015), are likely to leave intact a substantial amount of discrimination against racial and ethnic minorities seeking elected office.

References

Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan, "Do Judges Vary in Their Treatment of Race?," *Journal of Legal Studies*, 2012, *41* (2), 347–383.

Akerlof, George, "The economics of caste and of the rat race and other woeful tales," *The Quarterly Journal of Economics*, 1976, pp. 599–617.

Alesina, Alberto and Eliana La Ferrara, "A Test of Racial Bias in Capital Sentencing," *The American Economic Review*, 2014, *104* (11), 3397–3433.

Ansolabehere, Stephen, "CCES, Common Content, 2008," 2010.

Anwar, Shamena and Hanming Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *The American Economic Review*, 2006, *96* (1), 127–151.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, Petia Topalova et al., "Powerful Women: Does Exposure Reduce Bias?," *The Quarterly Journal of Economics*, 2009, *124* (4), 1497–1540.

Becker, Gary S, *The Economics of Discrimination*, Chicago: University of Chicago Press, 1957.

Bertrand, Marianne and Esther Duflo, "Field Experiments on Discrimination," 2016. NBER Working Paper 22014.

_ **and Sendhil Mullainathan**, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *The American Economic Review*, 2004, *94* (4), 991–1013.

Besley, Timothy, Jose G. Montalvo, and Marta Reynal-Querol, "Do Educated Leaders Matter?," *The Economic Journal*, 2011, *121* (554), F205–227.

Bonica, Adam, "Database on Ideology, Money in Politics, and Elections," Stanford, CA: Stanford University Libraries 2013.

Broockman, David E, "Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives," *American Journal of Political Science*, 2013, *57* (3), 521–536.

_ , "Distorted Communication, Unequal Representation: Constituents Communicate Less to Representatives Not of Their Race," *American Journal of Political Science*, 2014, *58* (2), 307–321.

Casey, Katherine, "Crossing Party Lines: The Effects of Information on Redistributive Politics," *The American Economic Review*, 2015, *105* (8), 2410–2448.

Center for American Women and Politics, "Women in Elective Office 2016," Eagleton Institute of Politics, Rutgers University 2016.

Chattopadhyay, Raghabendra and Esther Duflo, "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, 2004, *72* (5), 1409–1443.

Clots-Figueras, Irma, "Women in politics: Evidence from the Indian States," *Journal of Public Economics*, 2011, *95* (7), 664–690.

_ , "Are Female Leaders Good for Education? Evidence from India," *American Economic Journal: Applied Economics*, 2012, *4* (1), 212–244.

Cohen, William, Pradeep Ravikumar, and Stephen Fienberg, "A comparison of string metrics for matching names and records," in Subbarao Kambhampati and Craig A

Knoblock, eds., *Proceedings of IJCAI-03 Workshop on Information Integration on the Web*, Vol. 3 AAAI 2003.

Cronbach, Lee J, "Coefficient alpha and the internal structure of tests," *psychometrika*, 1951, *16* (3), 297–334.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao, "Voting to tell others," *The Review of Economic Studies*, 2017, *84* (1), 143–181.

DerSimonian, Rebecca and Nan Laird, "Meta-analysis in clinical trials," *Controlled clinical trials*, 1986, *7* (3), 177–188.

Erikson, Robert S. and Rocío Titiunik, "Using Regression Discontinuity to Uncover the Personal Incumbency Advantage," *Quarterly Journal of Political Science*, 2015, *10*, 101–119.

Federal Bureau of Investigation, "Agency Hate Crime Reporting by State," 2000–2015.

Feigenbaum, James J, "Intergenerational Mobility during the Great Depression," 2015.

Fryer, Roland G and Steven D Levitt, "The Causes and Consequences of Distinctively Black Names," *The Quarterly Journal of Economics*, 2004, pp. 767–805.

Fujiwara, Thomas, "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil," *Econometrica*, 2015, *83* (2), 423–464.

Gerber, Alan, "Estimating the Effect of Campaign Spending on Senate Election Outcomes using Instrumental Variables," *American Political Science Review*, 1998, *92* (02), 401–411.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians," *The American Economic Review*, 2000, *90* (4), 715–741.

Guryan, Jonathan and Kerwin Kofi Charles, "Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots," *The Economic Journal*, 2013, *123* (572), F417–F432.

Hainmueller, Jens and Dominik Hangartner, "Who gets a Swiss passport? A natural experiment in immigrant discrimination," *American Political Science Review*, 2013, *107* (1), 159–189.

Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova, "The Power of Political Voice: Women's Political Representation and Crime in India," *American Economic Journal: Applied Economics*, 2012, *4* (4), 165–193.

Kontopantelis, Evangelos, David A Springate, and David Reeves, "A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses," *PloS One*, 2013, *8* (7), e69930.

Kurtz, Karl, "Who We Elect: The Demographics of State Legislatures," *State Legislatures Magazine*, 2015.

Kuziemko, Ilyana and Ebonya Washington, "Why Did the Democrats Lose the South? Bringing New Data to an Old Debate," Technical Report, National Bureau of Economic Research 2015.

_ , Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva, "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments," *The American Economic Review*, 2015, *105* (4), 1478–1508.

List, John A, "The nature and extent of discrimination in the marketplace: Evidence from the field," *The Quarterly Journal of Economics*, 2004, pp. 49–89.

Manning, Jennifer E, "Membership of the 114th Congress: A Profile," CRS Report R43869, Congressional Research Service 2016.

Mateos, Pablo, *Names, Ethnicity and Populations*, Springer, 2014.

Miller, Joanne M and Jon A Krosnick, "The impact of candidate name order on election outcomes," *Public Opinion Quarterly*, 1998, pp. 291–330.

Mulkern, John, *The Know-Nothing Party In Massachusetts: The Rise and Fall of a People's Party*, Northeastern University Press, 1990.

Nathan, Max, "Same difference? Minority ethnic inventors, diversity and innovation in the UK," *Journal of Economic Geography*, 2015, *15* (1), 129–168.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Validation," *Journal of Business Economics and Statistics*, forthcoming.

Pager, Devah, Bruce Western, and Bart Bonikowski, "Discrimination in a Low-Wage Labor Market: A Field Experiment," *American Sociological Review*, 2009, *74* (5), 777–799.

Panagopoulos, Costas and Donald P. Green, "Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition," *American Journal of Political Science*, 2008, *52* (1), 156–168.

Pande, Rohini, "Can Mandated Political Representation Increase Policy Influence for Disadvantaged minorities? Theory and Evidence from India," *The American Economic Review*, 2003, *93* (4), 1132–1151.

Phelps, Edmund S, "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 1972, *62* (4), 659–661.

Price, Joseph and Justin Wolfers, "Racial Discrimination Among NBA Referees," *The Quarterly Journal of Economics*, 2010, *125* (4), 1859–1887.

Roediger, David R, *Working Toward Whiteness: How America's Immigrants Became White*, Basic Books, 2005.

Schaffner, Brian F, Matthew MacWilliams, and Tatishe Nteta, "Explaining White Polarization in the 2016 Vote for President: The Sobering Role of Racism and Sexism," 2017.

Shafer, Byron E, *Bifurcated Politics: Evolution and Reform in the National Party Convention*, Harvard University Press, 1988.

Southern Poverty Law Center, "Active Hate Groups in the United States in 2015," 2015.

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

Tomz, Michael and Robert P. Van Houweling, "How Does Voting Equipment Affect the Racial Gap in Voided Ballots?," *American Journal of Political Science*, 2003, *47* (1), 46–60.

Washington, Ebonya, "How Black Candidates Affect Voter Turnout," *The Quarterly Journal of Economics*, 2006, pp. 973–998.

Xu, Kaiyuan, Brian Nosek, and Anthony Greenwald, "Psychology Data from the Race Implicit Association Test on the Project Implicit Demo Website," *Journal of Open Psychology Data*, 2014, *2* (1).

Appendices for Online Publication

A Tables and Figures

Table A1: Demographic Summary of Illinois Republican Primary Voters

	Directly Observed	Inferred from Individual & Neighborhood Characteristics		Inferred from Neighborhood Characteristics	
Year	Mean Age	% White	% Male	Median Income	% College Ed.
2008	62	96.3	51.5	$70,135	34%
2012	61	96.4	51.6	$70,342	34%
2016	56	95.2	51.0	$69,441	33%

Notes: This table reports demographic summary statistics on the population of voters in the Illinois Republican primary in the 2008, 2012, and 2016. The data source is the Catalist voter file, from which reliable data are not available before 2008. Neighborhoods are defined at level of Census block group. Income is per capita in nominal dollars.

Table A2: Racial Animus in Illinois Versus Other U.S. States

	Illinois	Median	Percentile 25th	Percentile 75th
Racially-charged Google search rate	65.34	62.81	50.86	70.31
Average Race IAT score	0.415	0.402	0.385	0.417
Active hate groups per million	1.79	2.89	1.79	4.49
Race-related hate crimes per million	16.00	17.95	9.55	23.87

Notes: This table reports data on measures of racial animus in Illinois in comparison to other U.S. states. The racially-charged Google search rate, from Stephens-Davidowitz (2014), is measured as the share of searches of the state-level total, scaled so that the maximum state = 100. Data on the Implicit Association Test (IAT) scores come self-administered tests offered by Project Implicit (Xu et al., 2014); a high score indicates greater implicit bias. Active hate groups were counted by Southern Poverty Law Center (2015). Hate crime data come from Federal Bureau of Investigation (2000–2015), reflect 2000–2015, and cover race-, ethnicity-, and ancestry-related hate crimes.

Table A3: Selection into the Identifying Set

	(1)	(2)	(3)	(4)	(5)
White Share	-0.030***				-0.038***
	(0.008)				(0.009)
College Share		0.005			0.043
		(0.029)			(0.040)
Per-Capita Income			0.040**		-0.001
			(0.016)		(0.021)
Rep. 2-Party Vote Share				-0.138***	-0.143***
				(0.031)	(0.033)
N	5,714	5,714	5,714	5,714	5,714
R^2	0.010	0.000	0.009	0.020	0.034

Notes: This table presents two ways of analyzing how the Illinois population that contributes to the identification of the causal effect differs from the full state population as a result of potentially endogenous behavior by presidential campaigns. The dependent variable is the within-cell standard deviation of the PC1 race measure, which is a proxy for the relative contribution of the cell to identification of the overall causal effect. The independent variables are the county-level white share of population, the college-educated share of whites, white per-capita income, and, to proxy for the number of Republicans in a district, the Republican presidential candidates' share of the two-party vote in each district in the presidential election held that year. The unit of observation is the cell. Observations are weighted by the maximum number of votes for a delegate candidate in the cell. Standard errors are clustered at the county level.. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A4: Detailed Office by Race of Delegate Candidate

	Count of Delegate Candidates		
	White	Nonwhite	Total
Board of Education	10	4	14
City Council	61	8	69
County Office	102	13	115
Court Clerk	3	1	4
Governor	3	0	3
Judge	7	2	9
Local Official	70	15	85
Mayor	39	8	47
Other Notable	10	0	10
Party Office	219	54	273
Past Candidate	60	24	84
Sheriff	14	1	15
State House	99	17	116
State Senate	54	10	64
State's Attorney	12	2	14
Statewide Office	18	4	22
US House	19	0	19
Missing	3	2	5
No Office	1,138	274	1,412
Total	1,941	439	2,380

Notes: This table reports the distribution of delegate candidates, split by race, among detailed public-office categories. As in Table 1, which provides a higher-level summary of the officeholding data, we dichotomize the rescaled PC1 measure of race at 0.5. Office was marked as missing only when research assistants were found some information that plausibly, but not conclusively, indicated the delegate candidate held office. Delegate candidates are marked as past candidates only when they have been nominated for another office than convention delegate; the category is not collinear with repeat delegate status.

Table A5: Delegate Race Is Nearly Uncorrelated with Potential Confounds

	Correlation Coefficient
Repeat Candidates	0.0150**
Any Office	-0.0202***
Ballot Order	-0.0087

Notes: This table reports correlation coefficients of the PC1 race measure with three potential confounds. The ballot order correlation we report is a partial correlation conditional on the number of total ballot slots available in the district; recall districts vary from 2-4 in the number of maximum votes voters can cast and the number of delegates that are elected, with these two always being the same and determined ex ante by party rules. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A6: Inclusion or Exclusion of Inexact Census Matches Does Not Affect Results

	Exact Matches Only				All Matches			
	Continuous		Binary		Continuous		Binary	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nonwhite	-0.060***		-0.060***		-0.048***		-0.048***	
	(0.024)		(0.012)		(0.021)		(0.014)	
Black		0.002		-0.120***		-0.022		-0.078**
		(0.041)		(0.043)		(0.042)		(0.031)
Hispanic		-0.076**		-0.055***		-0.060**		-0.052**
		(0.029)		(0.016)		(0.026)		(0.022)
Asian		-0.075***		-0.077***		-0.084***		-0.050**
		(0.020)		(0.023)		(0.020)		(0.022)
N	17,850	13,120	17,855	17,855	19,677	13,590	19,703	19,703
R^2	0.986	0.989	0.986	0.986	0.986	0.989	0.986	0.986

Notes: This table re-estimates equations (6) and (7) using the Census race measure, first restricting the sample to delegates whose last names can be matched exactly in the 2000 Census tabulation, and second allowing for matches to similar last names according to the Jaro–Winkler distance. In all regressions the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A7: Most Strongly Signalling Names, By Nonwhite Racial Category

	Census		MTurk	
	Name	Score	Name	Score
Black	Pleas Honeywood	0.9172	Norris A. Washington	0.7813
	Norris A. Washington	0.8987	Devin Dante Johnson	0.7667
	Tonia S. Members	0.8776	Tondalaya Marie Nelson	0.7667
Hispanic/Latino	Melanie Alejandre	0.9611	Eduardo Fernandez	1.0000
	Carlos Saucedo	0.9515	Gustavo Gonzalez	1.0000
	Jesus E. Solorio	0.9509	Angel Garcia	1.0000
Asian	Baba Padmanabhan	0.9507		
	Steve H. Kim	0.9452		
	Sandra Yeh	0.9382		
East Asian			Noella Chung	1.0000
			Ji Chung	0.9677
			Sandra Yeh	0.9024
Indian			Pawel Hardej	0.8064
			Vasavi Krishnasri Chakka	0.7188
			Neil V. Patel	0.6875
Middle Eastern			Muneer Ahmad Satter	0.8667
			Habeeb Habeeb	0.7667
			Nabi Fakroddin	0.7000

Notes: This table lists, for each nonwhite racial category, the three delegate names that most strongly signal that the delegate belongs to this category, using the Census and MTurk race measures. Recall that the Census category does not differentiate within the Asian category. Similar results are unavailable for the Onolytics measure because it is dichotomous.

Table A8: Summary Statistics on Delegate-Candidate Demographics

Source	Type	Race/Gender Category	Mean	Std. Dev.	Min	Max	Obs.
MTurk	Continuous	White	0.846	0.168	0	1	19,711
		Black	0.088	0.109	0	0.781	19,711
		Hispanic	0.029	0.106	0	1	19,711
		Asian	0.008	0.053	0	1	19,711
		Indian	0.010	0.042	0	0.806	19,711
		Middle Eastern	0.019	0.052	0	0.867	19,711
	Dichotomous	White	0.960	0.196	0	1	19,711
		Black	0.016	0.124	0	1	19,711
		Hispanic	0.016	0.125	0	1	19,711
		Asian	0.005	0.067	0	1	19,711
		Indian	0.003	0.058	0	1	19,711
		Middle Eastern	0.004	0.059	0	1	19,711
Census	Continuous	White	0.841	0.176	0.014	0.997	19,685
		Black	0.112	0.130	0	0.917	16,291
		Hispanic	0.036	0.119	0	0.961	17,597
		Asian	0.015	0.074	0	0.951	16,069
	Dichotomous	White	0.965	0.184	0	1	19,711
		Black	0.013	0.114	0	1	19,711
		Hispanic	0.016	0.127	0	1	19,711
		Asian	0.007	0.085	0	1	19,711
Onolytics	Dichotomous	White (all)	0.945	0.228	0	1	19,769
		White (ethnic)	0.032	0.177	0	1	19,353
		Black	0.000	0.023	0	1	19,353
		Hispanic	0.012	0.107	0	1	19,353
		Asian	0.003	0.056	0	1	19,353
		Indian	0.001	0.056	0	1	19,353
		Middle Eastern	0.006	0.076	0	1	19,353
SSA	Continuous	Male	0.732	0.428	0	1	19,228
	Dichotomous	Male	0.739	0.439	0	1	19,228

Notes: This table reports the summary statistics on the population of Illinois Republican Party presidential convention delegate candidates in 2000, 2008, 2012, and 2016 using data from the Illinois State Board of Elections, Amazon Mechanical Turk (MTurk), the 2000 U.S. Census, Onolytics, and the Social Security Administration (SSA).

Table A9: Average of Delegate PC1 Race and Ethnicity Measure By Delegate Presidential Campaign

2000		2012	
Gary Bauer	0.123	Newt Gingrich	0.182
George W. Bush	0.161	Ron Paul	0.178
Steve Forbes	0.125	Rick Perry	0.026
Alan Keyes	0.145	Buddy Roemer	0.063
John McCain	0.129	Mitt Romney	0.186
		Rick Santorum	0.120
2008		**2016**	
Rudy Giuliani	0.165	Jeb Bush	0.175
Mike Huckabee	0.141	Ben Carson	0.155
John McCain	0.123	Chris Christie	0.129
Ron Paul	0.156	Ted Cruz	0.144
Mitt Romney	0.153	Carly Fiorina	0.158
Fred Thompson	0.148	Jim Gilmore	0.104
		Mike Huckabee	0.110
		John Kasich	0.136
		Rand Paul	0.238
		Marco Rubio	0.164
		Rick Santorum	0.111
		Donald Trump	0.181

Notes: This table reports the average PC1 race measure for delegates nominated by every presidential campaign.

Table A10: Reliability of Race and Ethnicity Variables

Race Category	R^2 of PC1	Reliability	
		MTurk: Bootstrap	All 3: Cronbach's α
White	0.5649***	0.9237	0.5698
Black	0.4455***	0.8563	0.3801
Hispanic	0.8029***	0.9542	0.8833
Asian	0.6626***	0.9688	0.7398
Indian	0.6957***	0.9085	0.5598
Middle Eastern	0.6474***	0.8591	0.4337

Notes: This table reports proportions of total variance across the three race measures explained by the first principal component. For all race categories, χ^2 statistics confirm that we can robustly reject the null hypothesis of independence among the three race measures. For the first four race categories, we include data from all three race measures; for the latter two, Census data are unavailable due to its broader definition of Asian. We further report two measures of reliability: (1) we boostrap the MTurk survey responses to estimate measurement error due to sampling, and (2) we estimate Cronbach's (1951)'s α using the three individual race measures.

Table A11: Results Robust To Imputation of Missing PC1 Race Measure Values

	(1)	(2)
Nonwhite	-0.104***	
	(0.022)	
Black		-0.083
		(0.067)
Latino		-0.055**
		(0.026)
Asian		-0.166***
		(0.028)
N	19,703	19,703
R^2	0.986	0.986

Notes: This table reports the results of estimating equations (6) and (7) while imputing missing values of the PC1 race measure using the MTurk race measure. Approximately 10.7 percent of values were missing and were predicted from the bivariate relationship between the PC1 and MTurk race measures. In all regressions the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A12: Results of Officeholding and Ballot Order Controls

	Office Level (1)	Detailed Office (2)	Ballot Order (3)	Detailed Ballot Order (4)	Office Level & Detailed Ballot Order (5)	All Controls (6)
Major Office	0.428***				0.383***	
	(0.080)				(0.082)	
State Legislature	0.105***				0.084***	
	(0.020)				(0.019)	
Minor Office	0.015*				0.013*	
	(0.008)				(0.008)	
Missing	0.198				0.200	
	(0.132)				(0.129)	
Board of Education		0.039*				0.045*
		(0.024)				(0.023)
City Council		-0.006				0.000
		(0.016)				(0.016)
County Office		0.029*				0.023
		(0.015)				(0.014)
Court Clerk		0.117***				0.155***
		(0.044)				(0.039)
Governor		0.358**				0.305*
		(0.179)				(0.170)
Judge		0.217				0.181
		(0.134)				(0.136)
Local Official		0.018				-0.001
		(0.016)				(0.015)
Mayor		0.015				0.018
		(0.024)				(0.022)
Missing		0.193				0.192
		(0.128)				(0.125)
Other Notable		0.040				0.023
		(0.027)				(0.025)
Party Office		-0.001				-0.001
		(0.010)				(0.010)
Past Candidate for Other Office		0.015				0.011
		(0.015)				(0.014)
Sheriff		0.150***				0.133***
		(0.040)				(0.043)
State House		0.093***				0.079***
		(0.021)				(0.020)
State Senate		0.137***				0.107***
		(0.031)				(0.030)
State's Attorney		0.116*				0.132**
		(0.066)				(0.059)
Statewide Office		0.066				0.069
		(0.077)				(0.046)
US House		0.456***				0.421***
		(0.083)				(0.085)
Ballot Order = 2			-0.055***	-0.108***	-0.092***	-0.095***
			(0.010)	(0.023)	(0.020)	(0.020)
Ballot Order = 3			-0.092***	-0.096***	-0.132***	-0.028*
			(0.010)	(0.016)	(0.022)	(0.017)
Ballot Order = 4			-0.129***	-0.132***	-0.062***	-0.056***
			(0.013)	(0.016)	(0.017)	(0.017)
Ballot Order = 2 × Number of Votes = 3				0.059**	0.054***	0.059***
				(0.023)	(0.023)	(0.023)
Ballot Order = 2 × Number of Votes = 4				0.050**	0.091***	0.101***
				(0.025)	(0.025)	(0.026)
Ballot Order = 3 × Number of Votes = 3				0.008	-0.043	-0.048**
				(0.021)	(0.023)	(0.024)
N	17,587	17,587	17,587	17,587	17,587	17,587
R^2	0.988	0.988	0.987	0.987	0.988	0.988

Notes: This table reports the results of estimating equation (6) with controls for delegate-candidate officeholding and ballot order, displaying coefficients associated with these controls. In all columns the dependent variable is ln(votes) for the delegate candidates. We omit the nonwhite coefficient in the above table but include it in the regression specification. The base level of both office level and detailed office is set to no office; for ballot order, it is = 1. Columns 1–2 show returns to officeholding, Columns 3–4 show ballot-order effects,and Columns 5–6 combine officeholding and ballot-order controls to show that some of the returns to officeholding come from appearing high on the ballot. The ballot-order interactions address the possibility heterogeneous effects for different maximum numbers of votes that voters may cast. The unit of observation is the county–district-delegate. All regressions include FEs at the cell level and two-way-clustered standard errors, at the delegate and cell level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A13: Local Prior Information Unlikely to Confound Results

	Dropping county with:			
	(1)	(2)	(3)	(4)
	Most Neg. Coef.	Largest White Overpeform.	Largest Nonwhite Underperform.	Likely Home
Nonwhite	-0.083***	-0.091***	-0.082***	-0.124***
	(0.027)	(0.027)	(0.028)	(0.031)
N	15,350	15,346	15,336	15,349
R^2	0.989	0.989	0.988	0.993

Notes: This table presents the results of several tests to assess whether local confounds, like social networks, could plausibly have generated our results. In Columns 1–3, we drop a county from each district-by-presidential-candidate set with a given characteristic that might be linked to local information. Column 1 drop the county where taste-based discrimination seems largest, as measured by the county-level coefficient estimate. Columns 2 and 3 drop the county in each set where, respectively, the whitest delegate outperforms the other candidates by the most and where the most nonwhite delegate underpeforms the other delegate candidates by the most. Column 4 drops the likely "home county" of each delegate—the one in which they most significant outperform their predicted log vote count. None of these sample restrictions meaningfully change our results. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A14: Meta-Analytic Estimates of Marginal Distributions of Racially-Discriminatory Tastes

Unit	N	Cochran's Q	I^2	τ	95% of Dist. (Normal)
County–District	102	138.59***	0.264	0.050	[-0.140, -0.095]
Presidential Candidate	24	37.67**	0.380	0.061	[-0.159, -0.067]

Notes: This table reports estimates of the variance in the share of Illinois Republican presidential primary voters who expressed racially-discriminatory tastes in voting. The subpopulation estimates come from estimating equation (6), interacting the PC1 race measure with group-level dummies. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A15: Effect of Delegate-Candidate Race on Log Votes,
by Presidential-Candidate Race

	Without Covariates		With Covariates	
	(1)	(2)	(3)	(4)
Nonwhite		-0.130*		-0.171**
		(0.068)		(0.067)
× White Pres. Cand.	-0.170***		-0.169***	
	(0.049)		(0.039)	
× Nonwhite Pres. Cand.	0.089	0.259**	0.047	0.215**
	(0.106)	(0.115)	(0.095)	(0.100)
N	17,587	17,587	17,587	17,587
R^2	0.986	0.986	0.988	0.988

Notes: This table reports estimates of the variance in the share of Illinois Republican presidential primary voters who expressed racially-discriminatory tastes in voting, split by the race of the presidential candidate. The subpopulation estimates come from estimating equation (6), interacting the PC1 race measure with a binary measure of presidential-candidate race. So that the interactions of delegate and presidential candidate race reflect within-year comparisons, all regressions include an interaction of the nonwhite variable with the year. The covariates introduced in Columns 3 and 4 are ballot order and detailed officeholding. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A16: Effect of Delegate-Candidate Race on Log Votes, by Presidential Campaign

Year	Campaign	Estimate	Std. Err.	t-stat	p-value	95% CI	
2016	Bush (Jeb)	-1.879	1.212	-1.55	0.121	-4.257	0.500
	Carson	-0.016	0.414	-0.04	0.969	-0.828	0.796
	Christie	-1.328	0.532	-2.50	0.013	-2.370	-0.286
	Cruz	-0.004	0.069	-0.06	0.951	-0.139	0.131
	Fiorina	-0.839	0.456	-1.84	0.066	-1.733	0.056
	Kasich	-0.093	0.032	-2.86	0.004	-0.156	-0.029
	Paul (Rand)	-0.415	0.113	-3.68	0.000	-0.636	-0.194
	Rubio	0.285	0.295	0.97	0.334	-0.294	0.864
	Santorum	2.577	2.998	0.86	0.390	-3.303	8.457
	Trump	-0.102	0.037	-2.77	0.006	-0.175	-0.030
2012	Gingrich	-0.135	0.112	-1.21	0.226	-0.354	0.084
	Paul (Ron)	-0.159	0.041	-3.85	0.000	-0.240	-0.078
	Roemer	1.596	3.068	0.52	0.603	-4.421	7.612
	Romney	-0.064	0.034	-1.87	0.061	-0.132	0.003
	Santorum	-0.123	0.039	-3.18	0.002	-0.199	-0.047
2008	Huckabee	-0.204	0.079	-2.59	0.010	-0.359	-0.049
	Giuliani	-0.179	0.479	-0.37	0.709	-1.118	0.760
	McCain	0.025	0.158	0.16	0.876	-0.286	0.335
	Paul (Ron)	0.012	0.069	0.17	0.863	-0.124	0.148
	Romney	-0.095	0.064	-1.49	0.137	-0.220	0.030
	Thompson	-0.432	0.480	-0.90	0.368	-1.373	0.509
2000	Bush (George)	-0.075	0.082	-0.92	0.358	-0.235	0.085
	Forbes	-0.779	0.368	-2.12	0.034	-1.501	-0.058
	Keyes	0.013	0.222	0.06	0.955	-0.422	0.448
	McCain	-0.243	0.121	-2.01	0.044	-0.480	-0.006

Notes: This table reports estimates of the share of Illinois Republican presidential primary voters by presidential campaign who expressed racially-discriminatory tastes in voting. Since many of these estimates are highly noisy, particularly for minor candidates, we provide full detail on the t-statistic, p-value, and 95-percent confidence interval. The subpopulation estimates come from estimating equation (6), interacting the PC1 race measure with campaign-level dummies, and including controls for delegate office level and ballot order. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A17: The Supply of Voter Taste-Based Racial Discrimination with Respect To Election Competitiveness

	Distance CDF (1)	Binary Indicator (2)	Censored Distance (3)
Nonwhite	-0.210*	-0.112*	-0.126*
× Competitiveness	(0.119)	(0.064)	(0.071)
Scenarios			
Tied	-0.021	-0.041	-0.058
	(0.046)	(0.051)	(0.040)
Median	-0.110***	-0.126***	-0.145***
	(0.045)	(0.033)	(0.053)
N	17,504	17,587	17,587
R^2	0.991	0.988	0.989

Notes: This table presents the results of several tests to assess whether voters engage in less discrimination when the costs of discrimination are higher. In all regressions the dependent variable is ln(votes) for the delegate candidates. Each column uses a different specification of the competitiveness variable. The unit of observation is the county–district-delegate. The PC1 race measure is included in all regressions, and we report only its interaction with the competitiveness variable. All regressions include cell-level FEs and controls for delegate office level and ballot order and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A18: Heterogeneous Treatment Effects by Geographic Attributes

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No Delegate FEs						
Nonwhite	-0.110***	-0.124***	-0.147***	-0.146***	-0.111***	-0.114***
	(0.024)	(0.025)	(0.034)	(0.035)	(0.021)	(0.024)
× % White	0.016			0.023		
	(0.065)			(0.066)		
× % College		0.185*		0.067		
		(0.111)		(0.103)		
× Income			0.122**	0.116*		
			(0.061)	(0.062)		
× Animus					0.000	
					(0.001)	
× IAT Score						0.305
						(0.403)
Panel B: Delegate FEs						
Nonwhite × % White	-0.006			-0.055		
	(0.166)			(0.083)		
Nonwhite × % College		0.299**		0.277*		
		(0.126)		(0.146)		
Nonwhite × Income			0.139**	0.049		
			(0.069)	(0.085)		
Nonwhite × Animus					-0.001	
					(0.001)	
Nonwhite × IAT Score						0.424
						(0.379)

Notes: In all regressions above, the dependent variable is ln(votes) for the delegate candidates and the race measure used is the rescaled PC1 measure. Income is defined as the mean per-capita income of non-Hispanic whites. The unit of observation is the county–district-delegate. In Panel A, all regressions include controls for ballot order, delegate officeholding, and cell-level FEs. In Panel B, all regressions include delegate- and cell-level FEs. All are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the county level due to the county-level interaction terms. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A19: Voter Taste-Based Racial Discrimination By Year

	PC1 (1)
2000	-0.129***
	(0.065)
2008	-0.106**
	(0.054)
2012	-0.098***
	(0.026)
2016	0.123***
	(0.039)
N	17,587
R^2	0.988
Wald test of equality restriction	
F-statistic (p-value)	0.13 ($p = 0.94$)

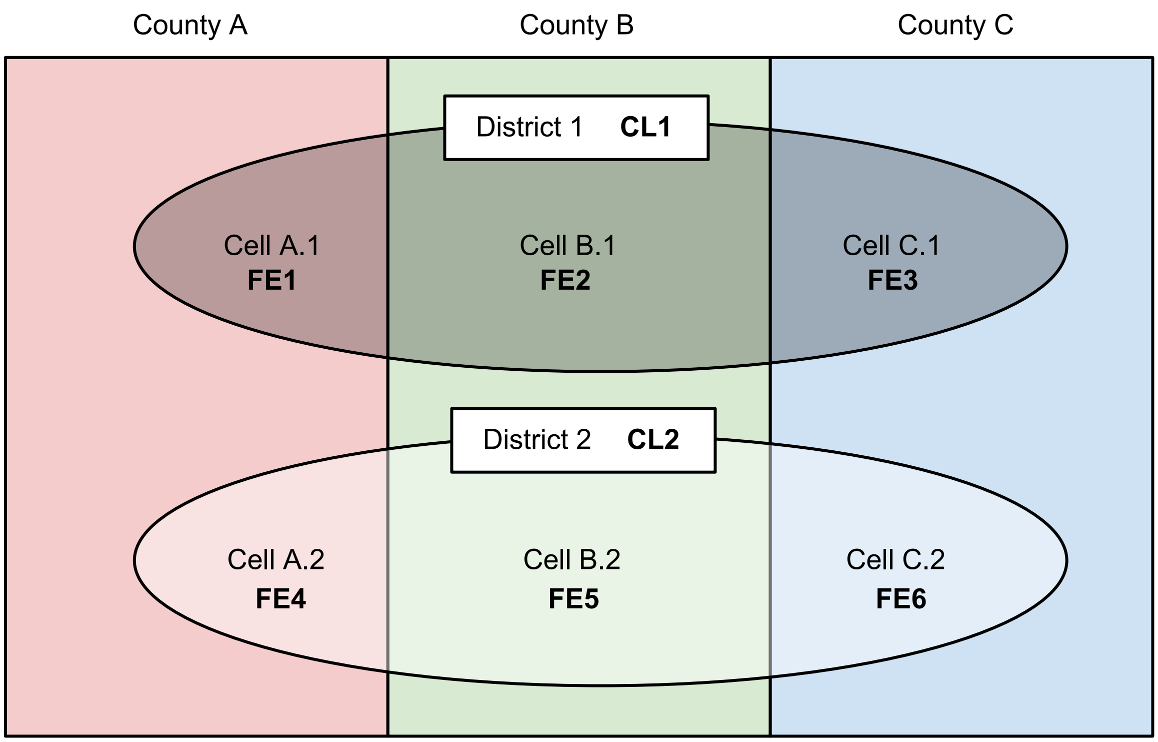
Notes: This tables presents the results of estimating equation (6), modified by interacting the PC1 race measure with year dummies. In all regressions the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level FEs, controls for delegate officeholding and ballot order, and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A20: Placebo Test: Voter Taste-Based Sex Discrimination

	No Controls (1)	No Controls (2)	Basic Officeholding (3)	Detailed Officeholding (4)	Ballot Order (5)	Detailed Officeholding & Ballot Order (6)
Male (continuous)	-0.001		-0.014*	-0.018**	-0.007	-0.015*
	(0.010)		(0.008)	(0.008)	(0.009)	(0.008)
Male (binary)		-0.002				
		(0.010)				
N	19,220	19,220	19,220	19,220	19,220	19,220
R^2	0.987	0.987	0.988	0.988	0.987	0.988

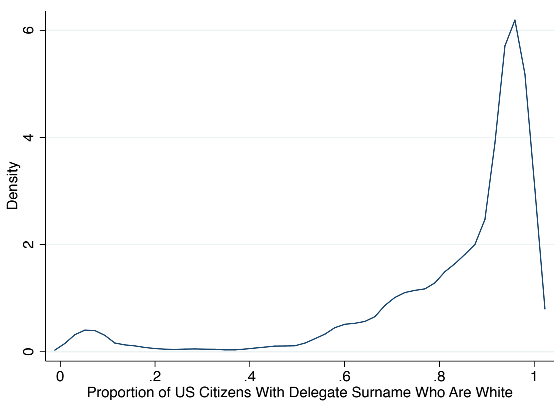
Notes: This table reports the results of estimating equation (8) using continuous and binary measures of delegate candidate sex and with varying sets of controls. Negative coefficient estimates indicate that women perform *better* than men, whereas if officeholding strongly affected election results we would expect men to perform better than women because men are much more likely to hold office. In all regressions the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

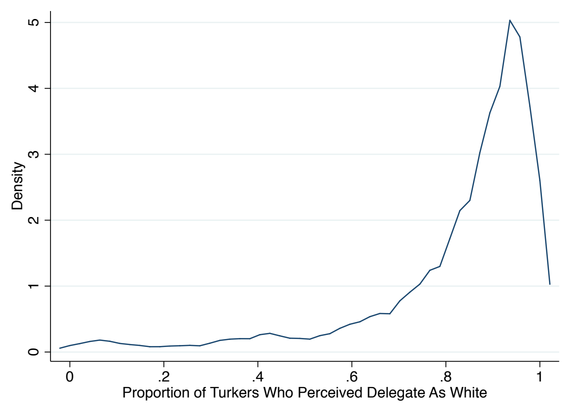
Figure A1: Definition of County–District



Notes: This figure depicts the definition of a county–district as the intersection of a county and a congressional district. "FE" denotes a fixed effect; "CL" denotes the level of clustering of standard errors, since delegates vary at the district level.

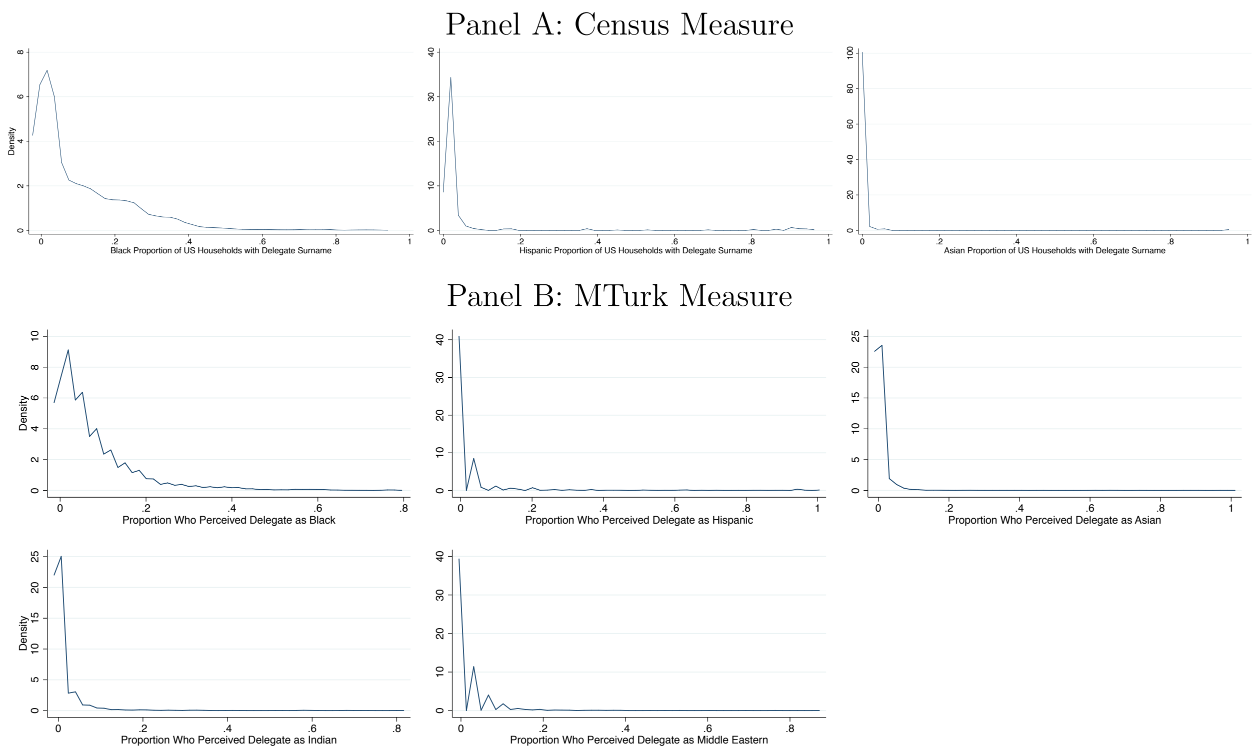
Figure A2: Kernel Density Plots of Delegate Candidate Names by Whiteness





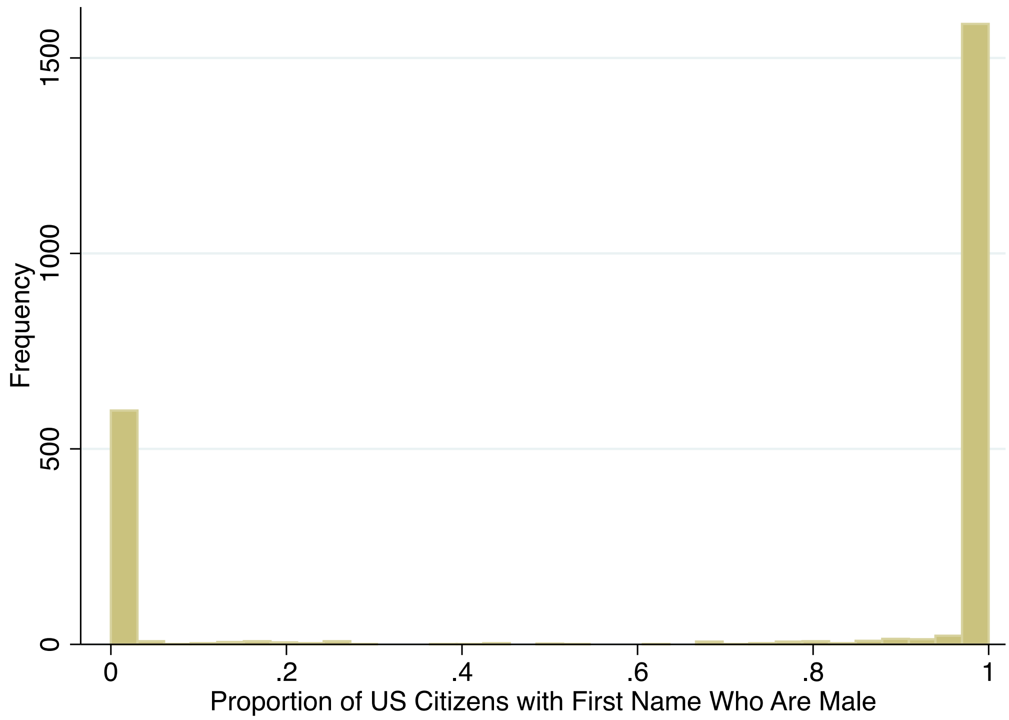
Notes: This figure plots the kernel density estimates of delegate-candidate name whiteness according to two distinct measures. The left panel matches delegate candidate last names to 2000 U.S. Census data on the percent of U.S. citizens with that last name who are non-Hispanic white. The right panel data are the proportion of Turkers who perceived delegate candidates, given their full names, as non-Hispanic white.

Figure A3: Kernel Density Plots of Delegate Candidate Names by Racial Category



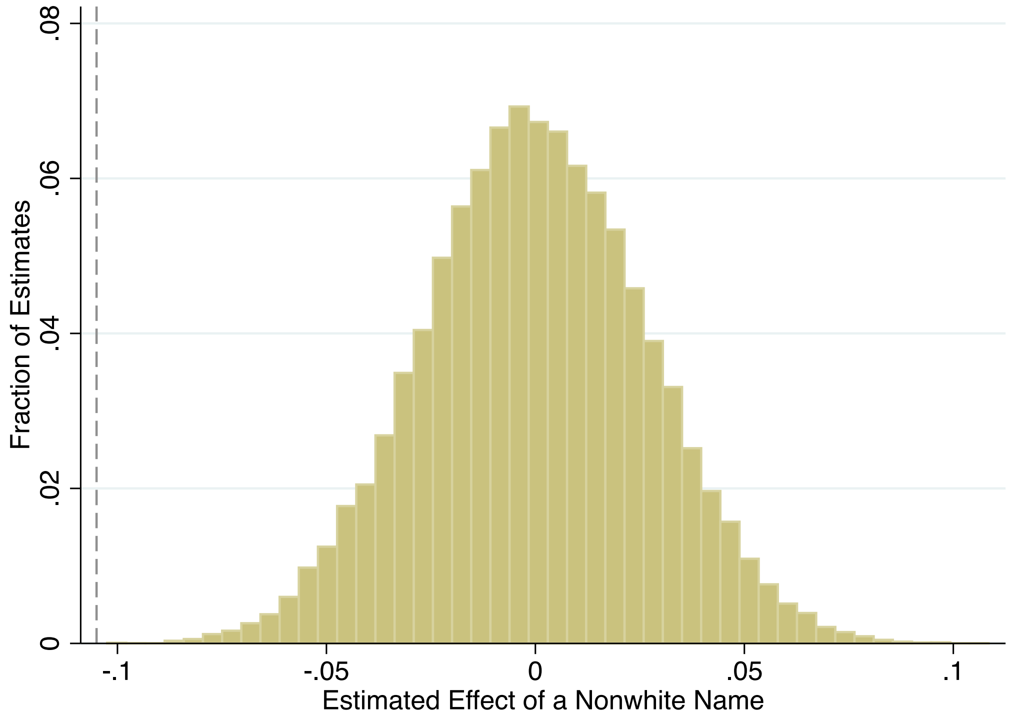
Notes: This figure plots the kernel density estimates of delegate-candidate name racial measures according to two distinct measures. Panel A matches delegate candidate last names to 2000 U.S. Census data on the percent of U.S. citizens with that last name who are black only, Hispanic only, and Asian or Pacific Islander only. Panel B is the proportion of Turkers who perceived delegate candidates, given their full names, as black, Hispanic, Asian, Indian, or Middle Eastern, respectively. Equivalent plots cannot be made for the dichotomous Onolytics measure.

Figure A4: Histogram of Delegate Candidate First Names by Gender



Notes: This figure plots a histogram of our measure of the perceived gender of delegate, which is the proportion of U.S. citizens with the same first name who are male at birth, according to the baby-name file of the U.S. Social Security Administration. Almost all delegate candidates' first names clearly signal their gender.

Figure A5: Permutation Test for Estimated Effect of a Nonwhite Name



Notes: This figure plots 50,000 estimates of the main regression specification with treatment status (nonwhite name) randomized at the delegate level. The coefficient estimated on the true treatment data using the specification in equation (6) and the PC1 race measure is far from the distribution of the test statistic under the sharp null hypothesis.

B Where Nonwhite Candidates Run

Since our identification strategy only relies on variation within groups of delegates running to represent the same area and bound to the same presidential candidate, the set of cells that contribute to identification may differ from Illinois on average. If cells with both white and nonwhite delegate candidates differ on average from cells with delegate candidates of only one race, then estimates from the identifying set may not generalize well to the full population. We evaluate this additional external-validity concern by testing for sample selection into the identifying set. In particular, we examine variation in the standard deviation of the measure within a given cell, as high within-cell variance means the cell contributes relatively more to identification.

Appendix Table A3 reports the results. The areas of Illinois that contribute relatively more to identification are modestly less white, higher income, and less Republican—as measured by the share of two-party vote in the general election the Republican presidential candidate received—than the state as a whole but are not significantly different in terms of the college-educated share of whites. More generally, the results are inconsistent with the possibility of strong selection on observables in the identifying set; the racial mix of the delegate-candidate population varies little with demographic variables.

The data also suggest some explanations as to why campaigns nominate nonwhite delegates in the first place if these delegates receive fewer votes. First, it is likely that individuals eager to serve as delegates can provide compensating differentials. Campaigns may use delegate candidacies to reward large donors or other party insiders who can benefit the campaign in other ways. As shown in Table 1 and Appendix Table A4, a considerable share of delegates hold minor leadership positions in local Republican parties. Another likely reason campaigns nominate nonwhites is search costs. There may be a relatively small number of people in many districts who are eligible, willing, and able to serve as delegate candidates. Only the people who live in a particular congressional district are eligible to serve as delegates in that district. An even smaller number of individuals from this eligible group are likely to support a particular presidential candidate, be willing to appear on the ballot, and be both willing and able to pay to attend the convention. If a campaign locates a nonwhite willing to serve as a delegate candidate and does not yet have a full slate of whites willing to run, they may choose to nominate the nonwhite to avoid the costs of continuing to search for a white willing to run. Consistent with search costs as an explanation for why nonwhites are nominated, we find that campaigns are less likely to nominate nonwhites when the supply of plausibly eligible white Republican voters is less constrained, as it would be in congressional districts that are more white and more Republican. Conditional on the share of a district that is white, for every 10 percentage points more Republicans there are in a Congressional district, campaigns are 2.8 percentage points less likely to nominate a nonwhite delegate in that district ($t = 5.16, p < 0.001$). These apparent supply constraints also suggest delegate candidate status is not particularly valuable for the delegate candidates, another reason why voters should have little incentive to engage in statistical discrimination.

C Onolytics

Onolytics is commercial software available at http://www.onolytics.com that implements anthropological research on the etymology of first and last names to estimate individuals' race and ethnicity from their full names (Mateos, 2014). Previous research on elections and in economics has used the software in a similar manner (e.g., Nathan, 2015). We use the Onolytics data to complement our other two sources of data on the racial signals each name sends. However, Onolytics provides much more detailed data on the estimated etymology of each name—for example, distinguishing between English and Scottish names, between Swedish and Norwegian names, and between German and Belgian names. To facilitate comparisons with the estimates with our other sources of data, we collapse the Onolytics data with the following categorization:

- **Non-Hispanic White**: English, Celtic, Central European, and Northern European names.
- **"White Ethnic"**: Eastern European, Southern European, Jewish, and Armenian names.
- **Black**: African names.
- **Hispanic/Latino**: Hispanic names.
- **Asian**: East Asian and Pacific (one category) names.
- **Indian**: Sikh and South Asian names.
- **Middle Eastern**: Muslim names.

We decided ex-ante on this categorization, blind to the results it produced. This categorization is mutually exclusive and exhaustive for the names in our dataset, except for those names Onolytics could not recognize, which we set to missing. This categorization is not intended to serve as a general-purpose racial classification scheme, but instead to best match the categories available in our other data. Unless otherwise noted, in the paper we collapse the "white ethnic" names into the "white" category.

D Coding Delegate Biographies

As argued in Section 5.2, the main threat to a causal interpretation of our results is the potential of confounding due to prior information voters have about individual delegate candidates beyond what is stated on the ballot. To address this concern, we instructed research assistants to search on Google for biographical information on every delegate in our sample. The intuition behind this strategy was that, if RAs cannot find information about a delegate holding office on Google, it is also unlikely that voters have information on the delegate. RAs used information on the delegate's county to narrow results and accepted as evidence *Chicago Tribune* articles, Wikipedia pages, LinkedIn pages, or multiple local sources, such as local newspapers.[57]

We double-coded the names using the following procedure. If a first RA did not find any evidence that a delegate candidate had served as an outside officeholder, we asked a second RA to search for information on that delegate as well. Conditional on a first RA not finding any evidence of a delegate serving as a notable official in some capacity, a second RA only found such evidence in 4 percent of cases.

If a delegate was positively identified as an officeholder, RAs recorded the office held and the counties within the congressional district that would have been affected by the officeholder. For example, a county board member is coded as affecting the county she serves. We then collapsed the offices to four categories:

- **Major Office**: U.S. House Member, Governor
- **State Legislature:** State House Member, State Senator
- **Minor Office**: All other codings
- **No Office**: No office found

Tabulations of delegates by office and race can be found in Table 1 and Appendix Table A4.

[57]The complete search instructions provided to RAs are available in the replication file.

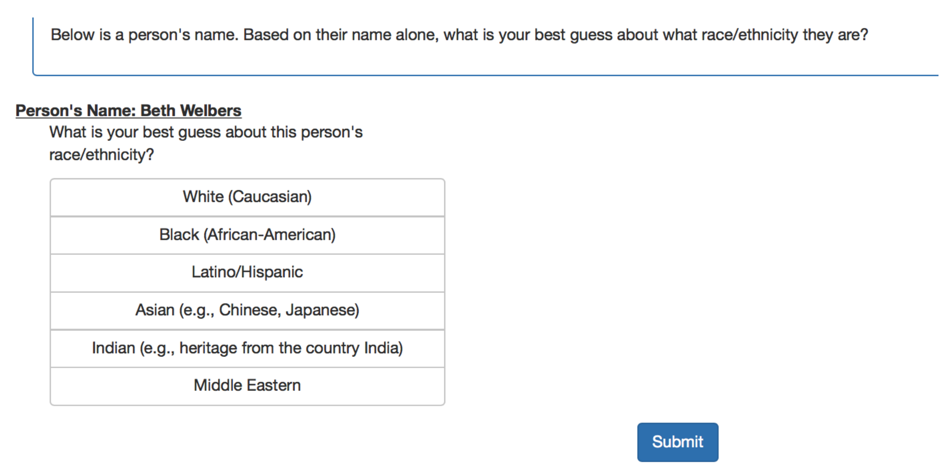
E MTurk Survey Procedure

To construct a plausible measure of how Americans would subjectively perceive the racial background of the delegate candidates' names, we hired 30 Amazon Mechanical Turk workers to record their best guess about the racial background of each name. Following Kuziemko et al. (2015), to ensure high-quality guesses, we only accepted codes from MTurk workers who (1) were located in the United States, (2) received Amazon's "master" qualification for submitting high-quality work in many previous jobs, and (3) had more than 97 percent of their prior work approved. We paid each MTurk worker 2 cents per name. Individual MTurk workers could not code the same name more than once, but could code as many names as they wished.

Figure A6 shows the interface MTurk workers used to code each name. They were told "Below is a person's name. Based on their name alone, what is your best guess about what race/ethnicity they are?" We then showed them one of the delegate candidates' names in bold font and presented them with one of seven racial categories along with the question "What is your best guess about this person's race/ethnicity?"

The presence of many names that all 30 MTurkers unanimously coded as white, many names all 30 unanimously coded as nonwhite, etc. suggests shirking (e.g., random answering behavior) was minimal. However, any shirking should bias our estimates towards zero. Appendix Section G estimates how attenuated our estimates our by the fact that a finite number of MTurk workers coded their perception of each name.

Figure A6: Coding Interface for Mechanical Turk Workers



Notes: MTurk workers recorded their perceptions using this form.

F State of the Nomination Races

Delegate outcomes are determined at the Congressional district level, with the median contest being decided by 2,541 voters. This section reviews the broader statewide and nationwide political contexts in which the four in-sample, Congressional-district-level Illinois Republican presidential primary delegate elections took place. Table A22 reports the number of delegates won by presidential candidates prior to Illinois and their prediction-market probabilities of winning the nomination and the Illinois primary, as observed the day before the primary.

Table A21: Polling and Results of Illinois Republican Presidential Primaries

2000			2008		
Candidate	Poll Av.	Vote Share	Candidate	Poll Av.	Vote Share
George W. Bush	n.a.	67.40	John McCain	38.5	47.45
John McCain	n.a.	21.54	Mitt Romney	22.8	28.60
Alan Keyes	n.a.	8.97	Mike Huckabee	14.8	16.46
Steve Forbes	n.a.	1.40	Ron Paul	7.3	5.01
2012			**2016**		
Candidate	Poll Av.	Vote Share	Candidate	Poll Av.	Vote Share
Mitt Romney	41.0	46.69	Donald Trump	36.0	38.80
Rick Santorum	31.0	35.01	Ted Cruz	29.5	30.23
Ron Paul	8.3	9.32	John Kasich	18.5	19.74
Newt Gingrich	13.3	7.98	Marco Rubio	13.5	8.74

Notes: For each year, this table reports the final pre-election Real Clear Politics state-level polling average and the official result of the 2000, 2008, 2012, and 2016 Illinois Republican presidential primaries using data from the Illinois State Board of Elections. By the primary date, four candidates remained in the race in the 2008, 2012, and 2016 primaries; in 2000, only Bush and Keyes remained, and no polling data is available. Note that these are statewide results, but that voters are pivotal with respect to the contest for delegate candidate slots within their Congressional districts.

2000. George W. Bush defeated John McCain in Illinois by a margin of 38 percentage points. Before the primary, Bush was already widely referred to as "presumptive nominee" in newspaper reports, and McCain had suspended his presidential campaign. Bush had accumulated 1,063 delegates, as compared to McCain's 237, and the Iowa Electronic Markets (IEM) gave Bush a 96-percent probability of becoming the 2000 Republican nominee the day before the primary.

2008. McCain defeated runner-up Mitt Romney in Illinois by a margin of 19 percentage points. Although McCain was not widely labeled declared the presumptive nominee, he had accumulated a slight delegate lead—95 to Romney's 83—and, according to IEM data on the day before the Illinois primary, McCain had an 87-percent probability of becoming

Table A22: State of the Race Before IL Primaries

March 20, 2000				February 4, 2008			
Candidate	Dels.	Nom.	IL	Candidate	Dels.	Nom.	IL
George W. Bush	1,063	0.96	n.a.	John McCain	95	0.87	n.a.
Steve Forbes	10	0.01	n.a.	Mitt Romney	83	0.11	n.a.
Rest of Field	7	0.01	n.a.	Rest of Field	7	0.01	n.a.
John McCain	237	0.01	n.a.	Mike Huckabee	27	0.01	n.a.
Bob Dole	0	0.00	n.a.	Fred Thompson	9	0.00	n.a.
Dan Quayle	0	0.00	n.a.	Rudy Giuliani	0	0.00	n.a.
March 20, 2012				March 14, 2016			
Candidate	Dels.	Nom.	IL	Candidate	Dels.	Nom.	IL
Mitt Romney	1,462	0.90	0.98	Donald Trump	384	0.70	0.72
Rest of Field	1	0.08	0.00	Rest of Field	7	0.18	n.a.
Rick Santorum	261	n.a.	0.03	John Kasich	44	n.a.	0.02
Ron Paul	154	0.01	0.00	Ted Cruz	324	0.15	0.28
Newt Gingrich	142	0.01	0.00	Marco Rubio	141	0.01	0.02
Herman Cain	0	0.00	n.a.	Ben Carson	8	0.00	n.a.
Rick Perry	0	0.00	n.a.				
Michele Bachmann	1	0.00	n.a.				

Notes: This table reports candidates' delegate counts and probabilities of winning the Republican presidential nomination and Illinois Republican presidential primary on the day prior to the Illinois primary in 2000, 2008, 2012, and 2016. For nominations, data come from the Iowa Electronic Markets. For the state primaries, data come from Intrade 2012) and PredictIt (2016). To the best of our knowledge, no prediction-market data are available for the 2000 and 2008 Illinois primaries. Due to the configuration of the IEM markets, note that "Rest of Field" prominently subsumes Santorum in 2012 and Kasich in 2016.

the 2008 Republican nominee, relative to Romney's 11-percent probability of becoming the nominee. Other candidates, such as Mike Huckabee, has secured some delegates and received significant shares of the Illinois primary vote but had, in the IEM data, minimal chances of becoming the 2008 Republican nominee by the day before the Illinois primary.

2012. Romney defeated runner-up Rick Santorum in Illinois by a margin of 12 percentage points. Romney was not yet the presumptive nominee but had accumulated a substantial delegate lead: 1,462 relative to Santorum's 261. IEM data from the day before the Illinois primary gave Romney a 90-percent probability of becoming the 2012 Republican nominee; similar data from Intrade gave Romney a 98-percent chance of winning the Illinois primary. Although other candidates—Ron Paul, Newt Gingrich, etc.—had won considerable numbers of delegates, IEM and Intrade data suggest they had minimal chances of becoming the 2012 Republican nominee by the day before the Illinois primary.

2016. Donald Trump defeated runner-up Ted Cruz in Illinois by 9 percentage points.

Trump was not yet the presumptive nominee but had accumulated a slight delegate lead: 384 relative to Cruz's 324. The nomination contest remained relatively open: According to IEM and PredictIt data on the day before the primary respectively, Trump had a 70-percent chance of winning the 2016 Republican nomination and a 72-percent chance of winning the Illinois primary. Ted Cruz was the next most likely to win the primary, at 28 percent; John Kasich and Marco Rubio each had 2-percent chances of winning the primary. Cruz had a 15-percent chance of becoming the nominee, and there was an 18-percent probability that another candidate—including Kasich, who was not broken out in the IEM data—would becoming the nominee. Despite having the third-most delegates, IEM data suggest Rubio had little chance of becoming the nominee; he suspended his presidential concern the day after the Illinois primary.

G Analysis of Classical Measurement Error

The 30-person sample used to produce the MTurk estimates of perceived race introduces attenuation from classical measurement error; the independent variable is sampled. Furthermore, as is suggested by the lack of perfect collinearity across our three measures of race, methods of inferring the perceived race of a delegate candidate are susceptible to measurement error relative to the latent variable to which voters respond. While attenuation is the concern for our baseline regression, differential levels of measurement error across nonwhite race categories can induce differential attenuation biases to our estimates of taste-based discrimination, invalidating inference about relative levels of discrimination among races.

In this section we present two ways of correcting our preferred estimates for these sources of attenuation bias. First, we compute test-retest reliability of the MTurk measure by bootstrapping and use this to correct the OLS MTurk coefficients in an errors-in-variables model. In particular, we bootstrap our MTurk race measure by drawing with replacement among the 30 ratings for each delegate candidate to obtain a bootstrapped average, computing the correlation between the bootstrapped MTurk race measure and the original, and then computing the average correlation over 20 runs. Second, we estimate the reliability of our PC1 race measure by Cronbach's (1951) alpha and similarly correct the estimates. Cronbach's α is a standard measure of reliability from psychometrics; intuitively, it represents the average correlation between manifest variables that measure the same latent variable. If the measures capture different latent variables (e.g., perceived vs. actual race), α may be biased downwards; however, we present it for completeness.

Table A10 reports both reliability estimates, and Table A23 reports the regression coefficients corrected using these estimates. The bootstrap reliability estimates, reported in Table suggest that measurement error from MTurk sampling is not a substantial concern, changing coefficient values only slightly in Columns 1 and 2 relative to the corresponding columns of Table 3. Under the assumption that all three variables measure the same latent construct, the adjustment for α, reported in Columns 3 and 4, would suggest that our baseline estimate of taste-based racial discrimination is attenuated by 43 percent of its true value, implying that about 18 percent, instead of 10 percent, of voters in the Illinois Republican primary do not vote for delegate candidates of their preferred presidential candidate who have names that indicate nonwhite ancestry. We interpret this as evidence that our estimates in the paper may be biased slightly towards zero but that stronger assumptions would be required to argue that they are highly biased toward zero.

One important caveat is that there appears to be heterogeneity across races in the reliability of our race measures. The entries of Table A10 confirm that reliability, measured by α, varies substantially across races: For example, our estimated reliability of the black measure is less than half that of the Hispanic/Latino measure. Columns 2 and 4 of Table A23 presents estimates for the detailed nonwhite race categories. Consistent with differential reliabilities, we find that the estimate of taste-based discrimination against black delegate candidates may be highly attenuated, whereas attenuation is a comparatively minor issue for estimates for Asian and particularly Hispanic delegate candidates. Again under the assumption that our three measures capture the same latent construct, the point estimate suggests that about 38 percent of primary voters would vote for the delegate candidates of another presidential candidate to avoid voting for a black delegate candidate of their preferred presidential candidate.

Table A23: Measurement-Error Corrected Estimates of Taste-Based Discrimination

	Bootstrap (MTurk)		Reliability (PC1)	
	(1)	(2)	(3)	(4)
Nonwhite	-0.095***		-0.111***	
	(0.024)		(0.011)	
Black		-0.056		-0.405***
		(0.083)		(0.062)
Hispanic/Latino		-0.049*		-0.128***
		(0.026)		(0.018)
Any Asian				-0.197***
				(0.034)
East Asian		-0.097***		
		(0.033)		
Indian		-0.191***		
		(0.053)		
Middle Eastern		-0.204***		
		(0.051)		

Notes: This table reports estimates, corrected for classical measurement error, of the coefficients from the baseline regressions. In all columns the dependent variable is ln(votes) for the delegate candidates. The unit of observation is the county–district-delegate. All regressions include cell-level, are weighted by the maximum number of votes a delegate candidate received in a cell, and two-way-clustered standard errors, at the delegate and cell level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.